

Vorhersage und Wahrnehmung deutscher Betonungsmuster

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu Bonn

vorgelegt von
Petra Susanne Wagner
aus
Minden in Westfalen

Bonn, 2002

Gedruckt mit Genehmigung der Philosophischen Fakultät der Rheinischen
Friedrich-Wilhelms-Universität Bonn

1. Berichterstatter: Prof. Dr. Wolfgang Hess

2. Berichterstatter: Prof. Dr. Wilfried Lenders

Tag der mündlichen Prüfung: 24.04.2002

Danksagung

Mein erster Dank gilt allen ehemaligen und gegenwärtigen Kolleginnen und Kollegen am IKP, die mich vom ersten Arbeitstag an in ein funktionierendes Team integriert und mich über die Jahre hinweg in allen möglichen Fragen und Problemen unterstützt haben, die weit über phonetische Aspekte hinausgehen.

Dr. Thomas Portele hat mich motiviert, meine anfangs sehr unausgegorenen Ideen zur Evaluierung phonologischer Vorhersagen zu verfolgen. Prof. Dr. Wolfgang Hess möchte ich insbesondere für die Kombination von großem Forschungsfreiraum und inhaltlicher Rückendeckung danken.

Meiner Familie danke ich für all die Jahre der Unterstützung und des Zuspruchs, all meinen Freunden für riesengroße Ohren zum Volljammern, spontanen Einladungen zum Essen und Ingo Gronewald für die logistische Meisterleistung, immer genügend Schokolade vorrätig zu haben.

[...] der Sprechrhythmus hat im ganzen nur wenig Formen aufzuweisen. Diese bilden ein System, das einfach und übersichtlich ist.
RUDOLF BLÜMEL, "DER NEUHOCHDEUTSCHE RHYTHMUS", 1930, S.2)

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	2
1.2	Überblick	5
2	Ein integratives Prominenzmodell	6
2.1	Prominenz als Begriff der Prosodie	7
2.2	Prominenz im Spannungsfeld zwischen Phonetik und Phonologie . .	9
2.3	Die phonetische Modellierung von Prominenz	12
2.3.1	Rein perzeptive Beschreibungsmodelle	12
2.3.2	Akustische Korrelate der Prominenz	13
2.3.3	Artikulatorische Korrelate	14
2.4	Die phonologische Modellierung von Prominenz	15
2.4.1	Einbettung von Betonung in die Prosodische Hierarchie . . .	15
2.4.2	Betonung als relative Prominenz im Äußerungskontext	17
2.4.2.1	Prominenzmuster von Zitierform und Nicht-Zitierform	20
2.4.2.2	Ein Prominenzmuster für die Modellierung von Korrekturkontrast	23
2.4.3	Prominenzmodellierungen unterschiedlicher phonologischer Theorien	30
2.4.3.1	Generative Phonologie	31
2.4.3.2	Metrische Phonologie	32
2.4.3.3	Optimalitätstheorie	35
2.4.3.4	Unifikationsphonologie	36
2.4.4	Formalisierung phonologischer Prominenzmodelle	38
2.5	Prominenz und Paralinguistik	42
2.5.1	Datenmaterial	42
2.5.2	Resultate	42
2.6	Fazit: Das Prominenzmodell	44
3	Methode	46
3.1	Was ist eine „gute“ linguistische Theorie?	46
3.1.1	Gütekriterien für phonologische Grammatiken	47

3.1.2	Formalität als notwendiger Bestandteil einer Grammatik . . .	49
3.1.3	Einfachheitskriterien	49
3.1.4	Gütekriterien der Evaluationsmethode	49
3.2	Das Empirie-Formalismus-Gefüge in der Phonologie	50
3.3	Die empirische Validierung von Betonungsvorhersagen	51
3.4	Formale Vorhersagen durch Implementierung	53
3.5	Datenmaterial: Die Bonner Prosodische Datenbank	54
3.6	Vergleichbarkeit wahrgenommener und vorhergesagter Muster . . .	54
3.7	Definition einer <i>Baseline</i>	55
3.7.1	Motivation	55
3.7.2	Baselineberechnung	56
3.7.3	Sprecherinnenvariationen durch unterschiedliche Fokussierung	57
3.8	Fazit: Überprüfung formaler Betonungsmodelle anhand perceptiver Daten	59
4	Prominenzmuster auf der Äußerungsebene	60
4.1	Vorhersage der Äußerungsprominenz im Englischen	60
4.2	Vorhergesagte und wahrgenommene Prominenz: ein erster Regelsatz	63
4.2.1	Güteüberprüfung der Vorhersagen des ersten Regelsatzes . .	64
4.3	Vergleich verschiedener Ansätze zur Generierung von Äußerungsbe- tonung	66
4.3.1	Ein wortklassenbasierter Ansatz (UHMANN 1991)	66
4.3.2	Ein syntaxbasierter Ansatz (FÉRY 1988)	66
4.3.3	Vergleich der Modelle	67
4.4	Verbesserungen der Vorhersage	68
4.4.1	Der Einfluß des Lexikons	69
4.4.2	Der Einfluß des Informationsgehalts	72
4.4.3	Der Einfluß des Rhythmus	74
4.5	Eine formale Grammatik zur Generierung deutscher Äußerungsbe- tonung	76
4.5.1	Modellierung von Regel 1 zur Vorhersage metrischer Prominenz	78
4.5.2	Modellierung von Regel 2 zur Vorhersage metrischer Prominenz	79
4.5.3	Modellierung von Regel 3 zur Vorhersage metrischer Prominenz	79
4.5.4	Modellierung von Regel 4 zur Vorhersage metrischer Prominenz	80
4.5.5	Modellierung von Regel 5 zur Vorhersage metrischer Prominenz	80
4.5.6	Modellierung von Regel 6 zur Vorhersage metrischer Prominenz	81
4.6	Fazit	82
5	Prominenzmuster auf der Wortebene	84
5.1	Monomorphematische, flektierte und derivierte Formen	85
5.1.1	Das Datenmaterial	85
5.1.2	Das Regelsystem von (JESSEN 1998)	86

5.1.2.1	Die Regeln	87
5.1.2.2	Vorverarbeitung: Abtrennung extrametrischer Suffi- xe und Resilbifizierung	88
5.1.2.3	Basisregeln	89
5.1.2.4	Regel 1: Approximantregel	90
5.1.2.5	Regel 2: Schwere-Endsilben Restriktion	90
5.1.2.6	Regeln 3a und 3b: Regeln der Paenultimabetonung .	90
5.1.2.7	Regel 4: Regel der nächsten betonbaren Silbe	91
5.1.3	Ergebnisse der Vorhersage	91
5.1.4	Fehleranalyse	92
5.1.4.1	Fehlerklasse I: leichte betonte Endsilben	92
5.1.4.2	Fehlerklasse II: schwere, unbetonte Silben	94
5.1.4.3	Fehlerklasse III: Wörter mit Antepaenultimabetonung	95
5.1.4.4	Andere Fälle	96
5.1.5	Erweiterung des Silbengewichtsmodells	97
5.1.6	Anwendung der Regeln auf die BPD	100
5.1.7	Zusammenfassung	101
5.2	Morphologisch komplexe Wörter: Komposita	102
5.2.1	„Akzentverschiebung“ in Komposita	104
5.2.2	Empirische Untersuchungen zur Akzentverschiebung in Komposita	105
5.2.3	Zusammenfassung	106
5.3	Eigennamen	107
5.3.1	Anwendung des Vorhersagealgorithmus auf Eigennamen . .	108
5.3.2	Betonungsunterschiede durch Lexikalisierung oder regelhaft?	108
5.3.3	Untersuchung des Einflusses der einzelnen Regeln	109
5.3.4	Betonung unbekannter Eigennamen	110
5.3.5	Zusammenfassung	112
5.4	Eine formale Grammatik für die Generierung deutscher Wortbetonung	112
5.4.1	Modellierung des ersten Regelapparats zur Zuweisung der Wortbetonung	113
5.4.2	Modellierung des erweiterten Regelapparats zur Zuweisung der Wortbetonung	114
5.4.3	Modellierung des Regelapparats zur Zuweisung der Wortbe- tonung von Eigennamen	115
5.5	Fazit	116
6	Einbettung in die Optimalitätstheorie	118
6.1	Formale Darstellung der Optimalitätstheorie	119
6.2	Hierarchisierung und Zuweisung der <i>Constraints</i>	124
6.3	Fazit	128

7	Zusammenfassung der Resultate	129
7.1	Motivation der Arbeit	129
7.2	Das Prominenzmodell	130
7.3	Die Methode zum Vergleich phonetischer Messungen und phonolo- gischer Vorhersagen	130
7.4	Anwendung der Methode auf die Äußerungsbetonung	132
7.5	Anwendung der Methode auf die Wortbetonung	132
7.6	Eine optimalitätstheoretische Modellierung der Ergebnisse	134
7.7	Fazit	134
8	Schlußbemerkungen und Ausblick	138
	Literaturverzeichnis	140
A	Experimentelles Material	i
A.1	Zweisilbige Nonsense-Wörter	i
A.2	Unbekannte Namen	viii
B	Extrametrische und betonte Affixe	x
B.1	Suffixe	x
B.2	Unbetonte Präfixe	x
B.3	Betonte Präfixe	x
B.4	(A(BC))-Komposita	x
C	Lebenslauf	xii

Kapitel 1

Einleitung

Es wird alles immer gleich ein wenig anders, wenn man es ausspricht.

Hermann Hesse

Angenommen, wir werden von einem Menschen, dessen Muttersprache sich von unserer unterscheidet, nach der korrekten Betonung eines bestimmten Wortes unserer eigenen Muttersprache gefragt. Die naheliegenste Reaktion ist, daß wir nun „in uns hineinzuhören“ oder uns sowie dem Fragenden das Wort laut vorsprechen, um aufgrund unserer muttersprachlichen Kompetenz die korrekte Betonung zu ermitteln. Diese Vorgehensweise hat allerdings einen Haken: Sollten wir bereits von der Gültigkeit einer bestimmten Betonung ausgehen, weil wir vielleicht eine bestimmte Regel aufgestellt haben, so werden wir dahin tendieren, daß unsere intuitiv gewonnenen Daten mit unseren Regeln übereinstimmen. Die Gültigkeit unserer intuitiv gewonnenen Urteile hat dann nur noch eingeschränkten Wert. Um eine Hypothese hinsichtlich eines korrekten Betonungsmusters bestätigen oder verwerfen zu können, benötigen wir daher möglichst viel Datenmaterial, welches auf objektivere Art und Weise gewonnen wurde, als es ein introspektives Vorgehen erlaubt. Dieser Problematik widmet sich die vorliegende Arbeit und stellt sich den folgenden zentralen Fragen:

- Welches sind die Betonungsmuster, die in bestimmten Dialogsituationen oder Äußerungssituationen mit der Übertragung bestimmter Bedeutungen verknüpft sind?
- Wie beschaffen wir objektive Daten für die Erforschung bedeutungsrelevanter Betonungsmuster?

Bei der Betonung spielen neben der Übertragung bedeutungsrelevanter Aspekte an den Hörer aber auch ästhetische Aspekte der Sprache eine Rolle. So hat eine Sprechergemeinschaft hinsichtlich der Betonung bestimmte Präferenzen, die dazu führen, daß manche Folgen betonter und unbetonter Silben bevorzugt werden. Diese Präferenzen können dazu führen, daß an einigen Stellen eigentlich betonte

Silben nicht betont werden, die Betonung auf eine andere Silbe verschoben wird, oder aber auch, daß eigentlich nicht betonte Silben betont werden. Da solche Präferenzen sprachspezifisch sind und insbesondere in sogenannten *betonungszählenden* Sprachen¹ wie dem Deutschen offenbar eine Rolle spielen, werden diese Aspekte in dieser Arbeit ebenfalls untersucht.

1.1 Motivation

The case for quantitative representations in the mind appears much stronger than the case for discrete representations of the speech signal. Progress needs to be made on formulating such representations and understanding their relationships to the qualitative representations of current phonological theories.

(PIERREHUBERT 1990, 391)

Wenn wir sprachliche Äußerungen produzieren und hören, so erscheinen uns bestimmte Wörter, Silben oder Phrasen betonter als andere. Die literaturwissenschaftliche Disziplin der Metrik hat diesem Umstand bereits seit der Antike viel Aufmerksamkeit gewidmet. Die Metrik hat Wohlgeformtheitskriterien für besonders wohlklingende Sprachrhythmen aufgestellt.

Neben der literarischen Beschäftigung mit dem Sprachrhythmus und der unterschiedlichen Betonung ist die systematische sprachwissenschaftliche Beschäftigung mit der Betonung und dem Rhythmus als Phänomen der Alltagssprache noch etwas jünger. Da die Betonung von Äußerungen im Deutschen aber auf die Bedeutung des Gesagten eine Auswirkung hat und manche Betonungen von Muttersprachlern als „falsch“ oder „ungrammatisch“ eingestuft werden, muß eine vollständige grammatische Beschreibung der deutschen Lautstruktur auch diesen Aspekt berücksichtigen.

Für eine gute linguistische Theorie gilt nun, daß sie den Formenreichtum der beschriebenen Sprache möglichst genau erfassen soll. Chomsky nennt dies *deskriptive Adäquatheit* (CHOMSKY 1965). Für Theorien, die Aussagen über die Welt machen, gilt aber zusätzlich noch die Anforderung der *Falsifizierbarkeit*. Um falsifizierbar zu sein, müssen die Vorhersagen der Theorie an objektiv erhobenen Daten getestet werden, um Vorhersagefehler aufspüren zu können. Bisher wurden die Vorhersagen (metrischer) Phonologie allerdings — wie in vielen anderen Zweigen der Linguistik auch — in erster Linie introspektiv mit Hilfe kleiner Grammatikfragmente evaluiert. Während diese Vorgehensweise für einige Teildisziplinen der Linguistik sicherlich vernünftig ist, ist sie im Bereich der Wort- und Satzbetonung problematisch. Genau

¹*Betonungszählende* Sprachen werden von *silbenzählenden* Sprachen unterschieden (PIKE 1947). In ersteren bilden die Folgen einer betonten und eventuell mehrerer unbetonter Silben jeweils ein annäherungsweise gleiches Zeitintervall. Bei letzteren ist das Zeitintervall auf Silbenebene ähnlicher. Diese sogenannte Isochronie-Hypothese konnte bisher allerdings noch nicht eindeutig experimentell nachgewiesen werden.

gesagt ist die introspektive Vorgehensweise im Bereich an dieser Stelle nicht geeignet, da

- die Intuitionen mit den tatsächlichen Ereignissen nicht immer übereinstimmen,
- es sprecherspezifische oder dialektale Variationen geben kann,
- eine Validierung anhand kleiner Sprachfragmente immer nur einen kleinen Ausschnitt einer Sprache beschreiben kann, so daß viele komplexe Interaktionen nicht berücksichtigt werden können.

Ein Nebenproblem des Testens an kleinen Fragmenten ist der Umstand, daß die Generalisierungsfähigkeit einer solchen Grammatik nie dazu dienen kann, um in tatsächlichen Applikationen eingesetzt zu werden.

Die Problematik introspektiver Evaluation soll an einem Beispiel verdeutlicht werden: Phonologische Modellierungen deutscher Betonung beschäftigen sich seit langem mit sogenannten Phänomen der *Akzentverschiebung*. Die Theoretiker gehen davon aus, daß bei zwei aufeinanderfolgenden Silben, welche die Wortbetonung tragen, die weniger stark betonte Silbe sich verschiebt, um einen *Akzentzusammenstoß* zu vermeiden. Ein klassisches Beispiel für eine solche Akzentzusammenstoßumgebung ist laut (KIPARSKY 1966) die Nominalphrase *halbtoter Mann*. Ihm zufolge stößt die betonte Silbe *tot* vom Adjektiv *halbtot* an das betonte Nomen *Mann*, verursacht einen Akzentzusammenstoß und wird deshalb von der Sprecherin² auf die Silbe *halb* verschoben (vgl. Abbildung 1.1).

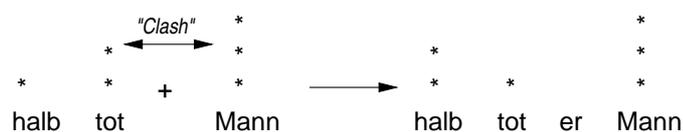


Abbildung 1.1: Vermeidung eines *Akzentzusammenstoßes* bei der Betonung der Phrase *halbtoter Mann*

Diese Analyse geht natürlich von einer Betonung der Silbe *tot* aus, wenn das Adjektiv in Isolation gesprochen würde. Legt man allerdings das am weitesten verbreitete deutsche Aussprachelexikon (MANGOLD 1990) zugrunde, so muß die Akzentverschiebung überhaupt nicht bemüht werden, um das Betonungsmuster in *halbtoter Mann* zu erklären, da hier die erste Silbe *halb* als die wortbetonungstragende Silbe betrachtet wird. Die Tatsache, daß introspektive Beurteilungen vermeintlich korrekter prosodischer Muster sehr anfällig für Autosuggestion sind, sind für Menschen, die im Bereich der Sprachsyntheseforschung aktiv sind, alltägliche Erkenntnisse, welche die Entwicklung objektiver Evaluationsverfahren von Prosodiesteuerungen erforderlich machten (SONNTAG 1999). Wenn die Introspektion als Grund-

²In dieser Arbeit ist immer von einer weiblichen *Sprecherin* sowie einem männlichen *Hörer* die Rede. Es ist selbstverständlich immer das jeweils andere Geschlecht implizit mitgemeint.

lage für die Evaluation entfällt, so muß das Datenmaterial für die Validierung phonologischer Vorhersagen durch objektive phonetische Messungen erhoben werden. Das Hauptproblem empirischer Validierungen von Theorien metrischer Phonologie bzw. von Betonungsmustern war bisher das Fehlen eines empirischen Maßes bzw. eines Annotationsschemas für das phonetische Datenmaterial. In existierenden Datenbanken wird zumeist lediglich zwischen Haupt- und Nebenakzenten unterschieden (z.B. (CELEX 1995)). Bestehende empirische Untersuchungen beschränken sich daher meist entweder auf prosodische Ebenen, die noch mit wenigen Betonungsstufen auskommen (z.B. (MENGEL 1998) zum Wortakzent). Die relativen Betonungsstufen auf Satzebene werden jedoch mit einer Unterscheidung in Haupt- und Nebenbetonung nicht erfaßt. Das Problem fehlender Annotationen hängt mit den abstrakt definierten Einheiten metrischer Repräsentationen zusammen (vgl. Kapitel 6). Dennoch sind eine Reihe von Untersuchungen zur Überprüfung metrischer Phonologie im Rahmen des Englischen durchgeführt worden (GRABE und WARREN 1995; VOGEL et al. 1995; BECKMAN und EDWARDS 1990). Diese Arbeiten sind indes kritisiert worden, da das untersuchte phonologische Material in keiner Weise phonologisch interpretiert worden sei, d.h., es sei nicht klar, inwiefern die untersuchten akustischen Muster tatsächlich Korrelate zu den phonologischen Mustern darstellen (SHATTUCK-HUFNAGEL 1995). Unklarheit herrschte insbesondere darüber, zu *welchen* akustischen Korrelaten genau die metrischen Einheiten in Beziehung zu setzen sind. Diese Unklarheit ist allerdings sehr problematisch. Soll beispielsweise in einem Produktionsexperiment untersucht werden, ob bestimmte metrische Vorhersagen wie die Akzentverschiebung auch eintreten, so muß ohne ein entsprechendes *perzeptives* Maß eine *akustische* Größe (zum Beispiel das Vorhandensein eines Pitchakzents) als Evidenz herangezogen werden. Dies geschieht ohne Wissen darüber, ob diese Größe im gegebenen Kontext tatsächlich das Korrelat zum untersuchten Phänomen darstellt. Diesen Problemen konnte nicht begegnet werden, solange keine reliablen Etikettierverfahren für metrische Einheiten vorhanden waren.

(FANT und KRUCKENBERG 1989) haben als erste für das Schwedische ein Etikettierverfahren entwickelt, welches genau diese Lücke schließt (siehe Kapitel 2). Mit Hilfe eines neuen, graduellen Prominenzbegriffes stand ein Etikettierverfahren zur Verfügung, welches wohldefinierte Schnittstellen zur akustischen sowie phonologische Beschreibungsebene zur Verfügung stellen konnte (FANT et al. 1998b; FANT und KRUCKENBERG 1999; HEUFT et al. 2000).

Mit der Erstellung größerer Datenbanken in den letzten Jahren können nunmehr die Vorhersagen phonologischer Theorien nicht mehr nur an kleinen Fragmenten überprüft werden. Um Vorhersagen an großen Datenmengen testen zu können, ist es nunmehr allerdings unausweichlich, die Vorhersagen entsprechend formal zu formulieren, so daß sie implementierbar werden.

In der vorliegenden Arbeit wird eine Methode zur Evaluation phonologischer Betonungsmodelle vorgestellt. Bei dieser Methode werden die Vorhersagen der Theorien anhand perzeptiv erhobener Prominenzurteile bewertet. Die Vorhersagen können an großen Datenmengen getestet werden, da sie zuvor implementiert werden. Die

Evaluation führt letztlich zu einer Entwicklung einer formalen Grammatik deutscher Betonungsmuster auf Wort- und Äußerungsebene.

1.2 Überblick

Das vorliegende Kapitel beinhaltet eine Einführung in den Problembereich und liefert die Motivation für die nachfolgend beschriebene Untersuchung. Im zweiten Kapitel wird ein Modell für die prosodischen Parameter der Silbentypenprominenz vorgestellt. Die Silbentypenprominenz bietet eine gute Schnittstelle zu linguistischen Betonungsmodellen, kann aber objektiv und reliabel gemessen werden. Zudem stehen Datenbanken bereit, die auf der Basis dieses Maßes annotiert wurden. Aus diesen und anderen Gründen bildet das Prominenzmodell die Grundlage für die empirischen Messungen, auf denen die Evaluationen der Vorhersagen aufbauen. Im dritten Kapitel wird eine Evaluationsmethode für linguistische Betonungsmodelle entwickelt, innerhalb welcher objektive Prominenzmessungen introspektive Betonungsanalysen ersetzen. Weiterhin fußt die Methode auf formalen Grammatikbeschreibungen, so daß die Vorhersagen faßbar werden, aber auch auf große Korpora angewandt werden können. Im vierten Teil wird die Methode verwendet, um verschiedene Ansätze zur Vorhersage der deutschen Äußerungsbetonung zu evaluieren. Die Evaluation führt zur Weiterentwicklung und formalen Beschreibung bestehender Vorhersagemodelle auf der Basis endlicher Übergangnetzwerke. Das fünfte Kapitel beinhaltet die Anwendung der Methode auf die Wortbetonung. Die Evaluation führt zur Kritik und Erweiterung der deutschen Silbengewichtshierarchie sowie zur Entwicklung eines separaten Betonungsmodells für deutsche Eigennamen. Auch die Vorhersagemodelle der Wortbetonung werden als endliche Übergangnetzwerke formal dargestellt. Anschließend werden die Erkenntnisse zur Wortbetonung in das nicht-generativ geprägte Modell der Optimalitätstheorie integriert. Das vorletzte Kapitel beinhaltet eine Zusammenfassung der Ergebnisse. Die Arbeit schließt mit einer Diskussion möglicher Anwendungsdomänen der Ergebnisse sowie einem Ausblick auf sich anschließende Forschungsfragestellungen.

Kapitel 2

Ein integratives Prominenzmodell

Calvin: I don't think, I'll go to school today.

Mom: *I think you will.*

Calvin: **I** think I won't.

Bill Watterson, Calvin and Hobbes

Wie eingangs erläutert, beschäftigt sich die vorliegende Arbeit mit wahrgenommenen Betonungsmustern sowie deren Vorhersage. Auf die wahrgenommene Betonungsstärke einer linguistischen Einheit wird in dieser Arbeit mit dem Terminus *Prominenz* Bezug genommen. Wenngleich die meisten Menschen eine intuitive Vorstellung davon haben, was mit der Betonungsstärke einer Silbe oder eines Wortes gemeint ist, treffen wir innerhalb der Prosodieforschung auf viele begriffliche Unschärfen und teilweise sehr unterschiedliche Betrachtungsweisen von Prominenz, ihrer Messung, ihrer phonetischen Ausprägung sowie ihrer Funktionen innerhalb der Kommunikationskette.

Im folgenden Abschnitt wird zunächst erläutert, inwieweit Prominenz als Teil der prosodischen Eigenschaften von Äußerungen zu betrachten ist und auf welche Dimensionen der Kommunikation (linguistische und para-/extralinguistische) die wahrgenommene Betonung Auswirkungen haben kann.

Da phonetische und phonologische Modellierungen von Betonungsmustern häufig mit gleichlautenden Begriffen arbeiten, hinter denen sich aber oft unterschiedliche Bedeutungen verbergen, werden danach beide Beschreibungsebenen voneinander abgegrenzt.

Anschließend wird beschrieben, auf welche Art und Weise Prominenz phonetisch gemessen und modelliert wird. Dieser Aspekt ist für die empirische Erhebung in der vorliegenden Arbeit von großer Relevanz.

Der nächste Teilabschnitt dieses Kapitels beschäftigt sich mit Modellierung von Prominenz in unterschiedlichen phonologischen Theorien. Es wird außerdem auf formale phonologische Beschreibungen eingegangen. Dieser Abschnitt ist für die Vorhersage von Betonungsmustern besonders relevant.

Danach wird die Relevanz von Prominenz innerhalb der Beschreibung paralinguistischer Phänomene erläutert. Es wird anhand einer empirischen Studie gezeigt, daß

mit Hilfe der vorgestellten phonetischen Meßmethode die paralinguistischen Ausprägungen von Prominenz nicht mit den linguistisch-phonologischen Betonungsmustern in Konflikt geraten.

Aufbauend auf den Untersuchungen und Analysen zu Aspekten der Prominenz, wird im letzten Teilabschnitt ein integratives Prominenzmodell vorgestellt, welches

- reliable phonetische Messungen ermöglicht,
- phonologische Betonungsmuster erfaßt,
- paralinguistische Aspekte von Betonung widerspiegelt,
- jedoch die Trennung von paralinguistischen und linguistischen Ausprägungen von Prominenz ermöglicht.

2.1 Prominenz als Begriff der Prosodie

Laut (MANGOLD 1975, 2) wurde der sehr generelle Terminus der *Emphase* in der phonetischen Forschung spätestens seit den Arbeiten von Daniel Jones immer mehr durch den Begriff *Prominenz* ersetzt. Unter *Emphase* wurde und wird die wie auch immer geartete Hervorhebung einer linguistischen Konstituente verstanden. Die Bezeichnung *Emphase* ist sehr unscharf und wurde für alle Arten phonetischer Hervorhebung sprachlicher Konstituenten benutzt. In dieser Arbeit wird der Begriff der *Emphase* dann verwendet, wenn es sich nicht um eine bedeutungsrelevante, sondern eher um eine durch Emotion oder Attitüde der Sprecherin bedingte Hervorhebung handelt. Ist vom Vorliegen von *Emphase* die Rede, so ist damit noch nicht geklärt, auf welcher phonetischen Ebene das Vorliegen von *Emphase* gemessen wurde. Im Gegensatz dazu wurde unter dem Begriff der *Prominenz* schon immer die *wahrgenommene* Betonungsstärke einer Äußerungskonstituente verstanden. Eine klassische Definition von *Prominenz* liefert (JONES 1950, 137):

Syllables occur here and there in connected speech which are heard to be more „prominent“ than others — syllables which stand out from other syllables near them in the word or sentence.

An dieser Betrachtungsweise von *Prominenz* hat sich bis heute noch nichts Grundlegendes geändert. Durch die Betonung ihrer perzeptiven Natur („which are heard ...“) wird somit die *Prominenz* zu einer phonetischen Größe, mit der *per se* weder

- eine linguistische Funktion, noch
- ein bestimmter Emotionsausdruck, noch
- eine eindeutige akustische Eigenschaft, noch
- eine bestimmte Artikulationsbewegung

verknüpft sein muß. Dennoch ist klar, daß es zu all diesen Bereichen Schnittstellen geben muß, wenn Kommunikation funktionieren soll. Die Prominenz wird in dieser Arbeit als die grundlegende Meßgröße angelegt, die in Beziehung zu allen anderen Beschreibungsebenen steht.

Wenngleich Jones die Prominenz als Eigenschaft einer Silbe versteht und auch in dieser Arbeit die Prominenz anhand von Silben gemessen wird, können selbstverständlich auch größere Bestandteile von Äußerungen wie Wörter oder Phrasen relativ zueinander unterschiedlich prominent wahrgenommen werden. Dies wird allein dadurch deutlich, daß Menschen mit der Frage „Welches Wort in dem Satz X ist stärker betont?“ etwas anfangen können. In der syntaktischen Nominalphrase „die Kartoffeln“ wird das Wort „Kartoffeln“ i.d.R. als Gesamteinheit stärker betont wahrgenommen werden als das Funktionswort „die“, obwohl ersteres eine unbetonbare Endsilbe enthält, die für sich genommen sehr wenig prominent ist (vgl. Kapitel 4). Als perzeptive Eigenschaft von Silben oder noch größeren sprachlichen Lautstrukturen ist die Prominenz ein Begriff aus der *Prosodie*, welche sich mit der

Gesamtheit sprachlicher Eigenschaften wie Akzent, Intonation, Quantität, Sprechpausen (BUSSMANN 1990, 13)

beschäftigt. Wie bereits angedeutet, ist das charakteristische Merkmal für prosodische Eigenschaften von Sprache, daß sie sich über längere Zeiträume als einen Laut bzw. ein Lautsegment erstrecken. Aus diesem Grund wird auf prosodische Größen auch häufig mit dem Begriff *Suprasegmentalia* referiert. Akustisch-prosodische Messungen beschäftigen sich typischerweise mit den Einheiten Dauer, Grundfrequenz (F_0) sowie Intensität. Dabei ist aber zu berücksichtigen, daß sich auch spektrale Signaleigenschaften, die beispielsweise die charakteristische Stimmqualität kennzeichnen, ebenfalls über einen längeren Signalabschnitt als einen Laut erstrecken. Diese werden ebenfalls als prosodische Eigenschaften betrachtet.

Nach (LAVER 1980) wirken sich prosodische Phänomene auf drei Dimensionen innerhalb der Kommunikationskette aus, nämlich auf

- die linguistische Dimension,
- die paralinguistische Dimension sowie
- die extralinguistische Dimension.

Innerhalb ihrer linguistischen Dimension transportiert eine prosodische Eigenschaft *bedeutungsrelevante* Information. Das Nichtvorhandensein dieser prosodischen Eigenschaft würde also dazu führen, daß die Äußerung anders verstanden würde¹.

¹Ich bin mir im klaren darüber, daß die Semantik eine Unterscheidung zwischen Satzbedeutung und Äußerungsbedeutung macht (vgl. z.B. (GREWENDORF et al. 1987, 387)). Ich bin mir auch ziemlich sicher, daß der gezielte Einsatz prosodischer Eigenschaften dazu führen kann, diese Ebenen voneinander zu trennen, und eine Äußerungsbedeutung von „Mir ist kalt“ in eine Äußerungsbedeutung „Schließe auf der Stelle das Fenster!“ zu verwandeln weiß. Hier ist aber nur gemeint, daß ein Bedeutungswandel auf irgendeiner Ebene, also Satz- oder Äußerungsebene, stattfinden muß.

Paralinguistische Phänomene sind nonverbale Kommunikationssignale. Typischerweise werden sie prosodisch durch den Einsatz bestimmter Stimmqualitäten wie Flüstern, Schreien etc. ausgedrückt, die das Sprachsignal modifizieren. Sie können aber auch durch Laute transportiert werden, welche nicht simultan zum Sprachsignal auftreten, wie z.B. Lachen, Schluchzen etc. Paralinguistische Signale transportieren trotz ihrer Nicht-Verbalität kommunikativen Gehalt an den Hörer über seine Stimmungslage, seine Einschätzung der Situation u.v.m.

Im Gegensatz zu paralinguistischen sind extralinguistische Signale von der Sprecherin nicht gezielt zu manipulieren, da sie auf physiologische Gegebenheiten zurückzuführen sind. Sie können sowohl konstant sein (z.B. eine sprecherinspezifische Stimmlage) oder aber temporär (wie z.B. ein nicht intendiertes Husten)². Auch extralinguistische Phänomene haben kommunikative Funktion. So werden beispielsweise unterschiedliche menschliche Stimmen von Hörern als verschieden natürlich und angenehm bewertet (SONNTAG 1999, 131). Extralinguistische Phänomene vermitteln an den Hörer auch Informationen zum Alter, Geschlecht, Rauchgewohnheiten und andere Eigenschaften der Sprecherin. Extralinguistische Betrachtungen werden in dieser Arbeit keine Rolle spielen. Beschreibungen der linguistischen und paralinguistischen Dimensionen der Prominenz werden hingegen in den Abschnitten 2.4 und 2.5 näher beleuchtet.

2.2 Prominenz im Spannungsfeld zwischen Phonetik und Phonologie

Innerhalb einer linguistischen Betrachtung ist die prosodische Ausprägung einer Äußerung Teil des *Signifiant* eines sprachlichen Zeichens. Dies beinhaltet gemäß der klassisch strukturalistischen Auffassung (DE SAUSSURE 1916) die relevanten bedeutungsunterscheidenden, materiellen Bestandteile des Zeichenkörpers wie Phoneme, Grapheme, Gesten einer Zeichensprache u.ä. Die strukturierte Beschreibung der relevanten bedeutungsunterscheidenden Parameter einer Sprache fällt in den Bereich der *Phonologie*. Die individuelle akustische und artikulatorische Realisierung der phonologischen Einheiten bzw. ihre Wahrnehmung ist Untersuchungsgegenstand der *Phonetik*.

Innerhalb der prosodischen Forschung ist die Abgrenzung zwischen phonetischer Messung und abstrakter phonologischer Beschreibung häufig schwierig, da für die Überprüfung phonologischer Hypothesen immer phonetische Messungen herangezogen werden müssen. Die traditionelle Linguistik hingegen bezieht ihre empirischen Daten häufig aus der Introspektion. Gerade im Bereich der Prosodie versagt die Introspektion als linguistisches Werkzeug allerdings (siehe Kapitel 3). Um

²Eine neue Studie von (VAN HIRTUM und BERCKMANS 2001) konnte zeigen, daß sich absichtliches Husten tatsächlich von unabsichtlichem Husten unterscheidet. Ein absichtliches Husten würde in den Gegenstandsbereich der Paralinguistik fallen.

dennoch auf Datenmaterial zurückgreifen zu können, hilft sich die Prosodieforschung daher durch die Definition von akustischen, artikulatorischen oder perceptiven *Korrelaten* zu prosodisch-phonologischen Einheiten. So sollte beispielsweise ein phonologischer Satz- oder Nuklearakzent nach (PIERREHUMBERT 1980) sich phonetisch durch eine starke Grundfrequenzbewegung im akustischen Signal bemerkbar machen. Die Korrelate sind aber selten 1:1-Entsprechungen ihrer phonologischen Kategorien. Entsprechend komplex ist die formale Modellierung der Phonetik-Phonologie-Schnittstelle (COLEMAN 1992a; COLEMAN 1992b; DIRKSEN und COLEMAN 1997; WAGNER 1997; WAGNER 1998; WAGNER und KÜHNLEIN 1998). Um eine phonologische Beschreibung zu ermöglichen, die sich näher an phonetischen Gegebenheiten orientiert, wurden deshalb bereits Vorschläge für eine Trennung in separate Produktions- und Perzeptionsgrammatiken gemacht (BOERSMA 1998). Andere phonologische Modelle sehen sogar eine Reduktion von phonologischen auf phonetische Artikulationsprozesse als Lösung dieses Abbildungsproblems (BROWMAN und GOLDSTEIN 1986; BROWMAN und GOLDSTEIN 1989). In dieser Arbeit wird aber von der Notwendigkeit einer phonologischen Beschreibungsebene ausgegangen. Die Prominenz, deren Messung in der vorliegenden Arbeit die empirische Grundlage bildet, ist ein Begriff aus der perceptiven Phonetik, welcher aber viele Schnittstellen zur linguistisch-phonologischen, aber auch zur paralinguistischen sowie den anderen phonetischen Ebenen hat. Aus diesem Umstand erfolgte eine gewisse Unschärfe hinsichtlich der Verwendung der eng verwandten Begriffe *Prominenz* („prominence“), *Betonung* („stress“) sowie *Akzent* („accent“) als *termini technici*. Ein besonders eindrucksvolles Beispiel für die begriffliche Verwirrung und eine zirkuläre Definition liefert das „Dictionary of Phonetics and Phonology“ (TRASK 1996). Hier wird „prominence“ (Prominenz) (TRASK 1996, 291) definiert als

property of a syllable which stands out of adjoining syllables by virtue of [...] phonetic characteristics. [...] The phonological use is **accent**.

„Accent“ (Akzent) (TRASK 1996, 4) wiederum wird beschrieben als

prominence attached to one syllable [...] by some phonetic means such as **stress**.

„Stress“ (TRASK 1996, 336) wird beinahe genauso wie „accent“ beschrieben, nämlich als

certain type of **prominence** which [...] is present upon certain syllables.

Die Abgrenzung von phonetischen und phonologischen Aspekten innerhalb der Begriffe ist aber nicht nur hier verwaschen. So weist (LEHISTE 1970, 119) darauf hin, daß bei (JONES 1940) nicht zwischen bedeutungsrelevanter (also phonologischer) und anders motivierter Prominenz unterschieden wird, obwohl er einen Unterschied zwischen *stress* und *prominence* mache. Sie selbst referiert mit dem Terminus *Stress* auf linguistisch signifikante, also bedeutungsunterscheidende Funktionen von Prominenz. Dieser Verwendungsweise soll hier gefolgt werden, allerdings wird statt des englischen Terminus „Stress“ der Ausdruck *Betonung* verwendet.

Der Umstand, daß der Begriff des *Akzents* ebenfalls zu der Verwirrung beiträgt, rührt daher, daß mit einem phonetischen Akzent in erster Linie eine deutlich wahrnehmbare Grundfrequenzbewegung (*Tonhöhenakzent* oder *pitch accent*) gemeint ist. Äußerungsabschnitte, auf denen Tonhöhenakzente liegen, sind i.allg. sehr prominent. Silben, die mit der Wortbetonung zusammenfallen, tragen häufig auch einen Tonhöhenakzent. Insbesondere die Höhe eines Tonhöhenakzents hat einen Einfluß auf die wahrgenommene Prominenz des so prosodisch ausgezeichneten Äußerungsabschnitts (PIERREHUMBERT 1979).

Die Form eines Tonhöhenakzents hingegen beeinflußt die Bedeutung der Äußerung über den Prominenzaspekt hinaus (KOHLE 1987; PIERREHUMBERT und HIRSCHBERG 1990). Es lassen sich verschiedene Ausprägungen von Tonhöhenakzenten feststellen. Die detaillierte Untersuchung vom komplexen Zusammenspiel zwischen Prominenz und Tonhöhenakzent steht allerdings nicht im Fokus dieser Arbeit. Wenn im folgenden von *Akzenten* die Rede ist, sind prominenzverleihende Tonhöhenakzente gemeint.

In Übereinstimmungen mit klassischen und neueren Verwendungsweisen von *Prominenz* und *Betonung* (s.o., (FANT und KRUCKENBERG 1989; TERKEN 1991; LADD et al. 1994)) kann über Prominenz zusammenfassend gesagt werden:

One of the important functions of prosody is emphasizing linguistic units in contrast to their environment. The corresponding perceptual impression is often called prominence, and stress is the pertinent linguistic feature. While stress as a linguistic feature is limited to a certain number of levels according to the phonological description of a language [...] prominence as a gradual parameter is more than the perceptual impression of linguistic stress.

(HEUFT et al. 2000, 97)

Ausgehend von diesen Überlegungen, wird Prominenz folgendermaßen definiert:

Definition 1

Prominenz bezeichnet die graduell wahrgenommene Stärke einer prosodischen Einheit, die mindestens eine Silbe umfassen muß, relativ zu ihrer Umgebung. Kommt der Prominenz innerhalb einer prosodischen Einheit eine bedeutungsrelevante Funktion innerhalb der Kommunikationskette zu, erhält sie den linguistischen Status einer Betonung.

Der Standort und die Beschreibung von Prominenz bzw. Betonung innerhalb der sprachwissenschaftlichen Disziplinen Phonetik und Phonologie wird in den folgenden Abschnitten erläutert.

2.3 Die phonetische Modellierung von Prominenz

Die Prominenz als perzeptive Einheit wird in phonetischen Modellen berücksichtigt, aber selbstverständlich ist es ein Hauptanliegen der phonetischen Wissenschaft, akustische und artikulatorische Korrelate für die perzeptiven Einheiten zu finden, um z.B. aufgrund des Vorhandenseins bestimmter akustischer oder artikulatorischer Gegebenheiten auf die entsprechende Prominenz und damit auch auf linguistisch relevante Information rückschließen zu können.

2.3.1 Rein perzeptive Beschreibungsmodelle

Das wichtigste und am weitesten verbreitete phonetische Beschreibungsmodell ist das Internationale Phonetische Alphabet oder IPA (IPA 1996). Obwohl das IPA auf der Basis artikulatorischer Parameter die phonetischen Eigenschaften gesprochener Sprache beschreibt, ist dies bei der Beschreibung der Prominenz nicht vorgesehen. Das IPA läßt die Beschreibung von drei Prominenzstufen zu, nämlich unbetont (keine Annotation von Prominenz), nebenbetont (Annotation durch ' ' vor dem Silbennukleus) sowie hauptbetont (Annotation durch ' ' vor dem Silbennukleus). Die Prominenz wird im IPA als Eigenschaft von Silben betrachtet, und annotiert wird üblicherweise die Wortbetonung. Die Tatsache, daß Hörerurteile hinsichtlich der Platzierung der Hauptbetonung in Wörtern weitgehend übereinstimmen, konnte wenigstens für das Englische von (LEA 1977; THOMPSON 1980) gezeigt werden.

Ein perzeptiv orientiertes Annotationsverfahren, welches Prominenz als graduellen Parameter betrachtet, wurde von (FANT und KRUCKENBERG 1989) eingeführt. Sie ließen gelesene Sprache (Schwedisch) von Hörern hinsichtlich der von ihnen empfundenen Prominenz auf Silbenebene annotieren. Hierzu verwendeten sie eine offene Skala von 0-30, die Hilfslinien für die empfundene Betonung der „Wortbetonung“ (bei 15) sowie der „Fokusbetonung“ (bei 30) enthielten. Das Verfahren impliziert jedoch nicht die Annahme von 30 linguistisch relevanten Betonungsstufen. Vielmehr ist entscheidend, daß das Annotationsverfahren eine feingranulierte und intervallskalierte Erfassung von Prominenzrelationen ermöglicht. Das Verfahren wurde von (HEUFT et al. 1995) auf das Deutsche und von (ELSNER et al. 1998) auf das Amerikanische Englisch übertragen und konnte für beide Sprachen aufgrund hoher Inter-Hörerkorrelationen bestätigt werden. Für alle Sprachen konnten Korrelationen der Prominenz zu akustischen Signaleigenschaften bestimmt werden, wobei insbesondere die Silbendauer sowie die Grundfrequenzbewegung als wichtige Merkmale hervorstechen (FANT und KRUCKENBERG 1989; HEUFT 1999; PORTELE 1998). (FANT und KRUCKENBERG 1999) fanden weiterhin einen signifikanten Zusammenhang zwischen dem subglottalen Luftdruck und der Prominenz.

Die Hörerbeurteilungen der wahrgenommenen Prominenz sind aber nicht in allen Kontexten klar übereinstimmend. In einer neuen Untersuchung zur Nebenbetonung in komplexen deutschen Komposita, welche Folgen hochprominenter Silben enthalten und damit sogenannte *Akzentzusammenstöße* provozieren, wichen die

Hörerurteile teilweise stark voneinander ab und zeigten noch nicht einmal in allen Fällen eindeutige Tendenzen (vgl. Abschnitt 5.2, (FISCHENBECK 2001)). Die Betonungen in diesen Kontexten stellen allerdings Randphänomene der deutschen Sprache dar, die prosodisch hochmarkiert sind. Auch berichten (ROSS und OSTENDORF 1996) für das Englische sowie (MÜLLER und HOFFMANN 2001) für das Deutsche von Problemen einer übereinstimmenden Etikettierung von Nebenbetonungen. Interessanterweise kommen Annotatoren also mit kleinen Skalen bei der Beurteilung von Betonungen nicht gut zurecht. Daher soll in dieser Arbeit dem Modellierungsansatz von Prominenz als graduellem Parameter gefolgt werden.

2.3.2 Akustische Korrelate der Prominenz

Die akustische Realisierung von wahrgenommener Prominenz ist hochkomplex und bis heute noch nicht vollständig geklärt. Bei der Aufdeckung akustischer Korrelate gilt es zudem zu berücksichtigen, daß diese sprachspezifisch sehr unterschiedlich sein können. In Sprachen, in denen die Dauer genutzt wird, um Vokalphoneme voneinander zu unterscheiden, spielt sie bei der Realisierung von Betonung eine geringere Rolle (BERINSTEIN 1979). Es ist anzunehmen, daß Ähnliches bei Tonsprachen für die Grundfrequenz gilt. Die folgenden akustisch-prosodischen Parameter stehen sprachabhängig in unterschiedlichem Maße ziemlich sicher in einer Beziehung zur wahrgenommenen Prominenz:

- Grundfrequenz
- Dauer
- Intensität
- Spektrale Parameter

Im Falle der Grundfrequenz scheint insbesondere die Höhe eines Tonhöhenakzents Einfluß auf die wahrgenommene Prominenz zu haben (PIERREHUMBERT 1979). Auch das Verhältnis vom Grundfrequenzminimum zum Grundfrequenzmaximum innerhalb eines Tonhöhenakzents (*pitch excursion*) wurde hinsichtlich seines Einflusses auf die wahrgenommene Prominenz untersucht und beeinflußt diese ebenfalls (GUSSENHOVEN und RIETVELD 1988; LIBERMAN und PIERREHUMBERT 1984; PIERREHUMBERT 1979; RIETVELD und GUSSENHOVEN 1985; 'T HART 1981). Die Steilheit des Grundfrequenzabfalls nach einem Gipfel wirkt sich ebenfalls auf die wahrgenommene Prominenz aus (BRUCE und TOUATI 1990). Die verschiedenen Einflußfaktoren wurden im Modell von (ELSNER 2000) integriert und für die automatische Erkennung von prosodischen Fokusakzenten eingesetzt. (MIXDORFF und WIDERA 2001) fanden außerdem einen Zusammenhang zwischen Silbeprominenz und Stärke des *Akzentkommandos*, welches der Modellparameter innerhalb der Grundfrequenzmodellierung nach (FUJISAKI und HIROSE 1984) ist, welcher in erster Linie für die Ausprägung eines Tonhöhenakzentes verantwortlich ist.

In der klassischen Untersuchung von (FRY 1955; FRY 1958) wurde neben der Grundfrequenzbewegung die Dauer als wichtigstes akustisches Korrelat von Betonung festgestellt. Die Dauer ist auch in den neueren Untersuchungen zum Schwedischen, Deutschen und Amerikanischen Englisch immer wieder als wichtiger akustischer Parameter für die Signalisierung von Prominenz bestätigt worden (FANT et al. 1998b; FANT und KRUCKENBERG 1999; PORTELE 1998; HEUFT et al. 2000).

Die Intensität scheint hingegen eine weniger wichtige Rolle bei der Signalisierung von Prominenz zu spielen (FRY 1955; FRY 1958), ganz ohne Einfluß ist sie allerdings nicht (BECKMAN 1986; BECKMAN und PIERREHUMBERT 1986).

Neben den klassischen akustisch-prosodischen Größen *Dauer*, *Grundfrequenz* und *Intensität* haben aber auch die spektralen Eigenschaften des Sprachsignals Einfluß auf die Prominenz. Dies wird besonders deutlich bei Reduktionsphänomenen der Sprache. Diese führen insbesondere zur Abschwächung unbetonter Vokale. Diese Reduktion führt nicht nur zu einer geringeren Dauer, sondern hat auch Auswirkungen auf die spektrale Struktur des Vokals, der sich abhängig vom Reduktionsgrad dem Zentralvokal annähert. Korrelationen zwischen dem Reduktionsgrad eines Vokals und der wahrgenommenen Prominenz der zugehörigen Silbe wurden in den Studien zu Vokalreduktion von (WIDERA et al. 1997) bestätigt. Ein weiterer spektraler Parameter, welcher im Zusammenhang mit der Prominenz immer wieder genannt wird, ist der *Spectral Tilt*, die relative Energie in verschiedenen Frequenzbereichen des Spektrums, die mit zunehmender Prominenz flacher wird (SLUIJTER und VAN HEUVEN 1996). Die Modellierung von *Spectral Tilt* ist allerdings ein noch immer ungelöstes Problem.

Trotz dem komplexen Zusammenspiel verschiedener akustischer Parameter können sprachspezifische rein akustische Modelle die Prominenz bereits relativ stabil automatisch annotieren, wenn sie auch noch hinter Modellen zurückstehen, die auf linguistische Information Zugriff haben (PORTELE 1998; STREEFKERK et al. 1998; WAGNER et al. 2000; STREEFKERK et al. 2001).

2.3.3 Artikulatorische Korrelate

Ähnlich wie bei der Bestimmung der akustischen Korrelate ergibt sich auch hinsichtlich der artikulatorischen Realisation von Betonung ein komplexes Bild. Frühe Versuche einer artikulatorischen Beschreibung beliefen sich auf allgemeine Aussagen wie den Grad des „Artikulatorischen Aufwands“ (JONES 1950), d.h., eine höhere Prominenz ist mit einem höheren artikulatorischen Aufwand verknüpft. Aufgrund des wichtigen Einflusses der Grundfrequenz ist klar, daß die Frequenz der Stimmlippenvibration ein wichtiger Faktor bei der Realisation sein muß. Ebenfalls wird die akustische Dauer klarerweise durch die Dauer der Artikulationsbewegung beeinflusst. Der Reduktionsgrad wird artikulatorisch mit dem Vorliegen eines *Target Undershoot* (LINDBLOM 1983) bzw. dem Vorliegen einer progressiven oder regressiven Koartikulation erklärt. Ein vorliegender *Target Undershoot* bedeutet, daß in reduzierten, weniger prominenten Silben die artikulatorische Zielposition von

den Artikulatoren nicht erreicht wurde. Dies kann im Extremfall bei einem Vokal zu einer totalen Zentralisierung und somit Schwa-Qualität des ursprünglichen Lautes führen. Im Falle einer Koartikulation wird der Laut durch die benachbarten Laute *gefärbt*, z.B. kann bei einem nasalen Konsonant das Gaumensegel länger als notwendig geöffnet bleiben und somit zu einer nasalen Qualität eines angrenzenden Vokals führen. Eine solche nasale Färbung wird von Hörern des Deutschen als Reduktion des Ursprungsvokals wahrgenommen (WIDERA und PORTELE 1999).

Der artikulatorische Parameter, welcher bisher am eindeutigsten mit der Prominenz in Zusammenhang gebracht wurde, ist der *subglottale Luftdruck* (vgl. u.a. (FANT und KRUCKENBERG 1999)).

2.4 Die phonologische Modellierung von Prominenz

In diesem Teilabschnitt werden zunächst die verschiedenen prosodischen Ebenen diskutiert, auf denen Prominenzunterschiede eine Rolle spielen. Dann wird gezeigt, daß die Prominenz Relevanz für die Bedeutung von Äußerungen hat und somit den Status des phonologischen Merkmals erwirbt, welches in dieser Arbeit als *Betonung* bezeichnet wird. Besonders interessant ist dabei die Entdeckung, daß ein bedeutungsrelevanter Betonungsunterschied u.U. nicht mit der lokalen Beeinflussung der Prominenz einer Einzelsilbe einhergeht, sondern daß die Prominenzen der gesamten Äußerung davon betroffen sein können. Die Modellierung von Betonung in verschiedenen phonologischen Paradigmen sowie die formale Beschreibung von Betonungsregeln werden abschließend dargestellt.

2.4.1 Einbettung von Betonung in die Prosodische Hierarchie

Bei der Betrachtung von phonologischen Arbeiten zum Thema Betonung stößt man zunächst wiederum auf leichte Unstimmigkeiten hinsichtlich der Verwendung verschiedener Begriffe, um Betonung zu kennzeichnen. Die wohl häufigste Verwirrung bietet die uneinheitliche Verwendung von *Betonung* vs. *Akzent*, welche ja bereits in Abschnitt 2.2 durch die Relevanz von Tonhöhenakzenten für besonders prominente Äußerungsabschnitte erklärt worden ist. Weiterhin auffällig ist die begriffliche Unschärfe hinsichtlich der Ebene, auf der Betonung angesiedelt ist. So kann sich eine Untersuchung mit der prominentesten Silbe in Wörtern („lexical stress“), Äußerungen („utterance stress“) oder Komposita („compound stress“) beschäftigen. Manchmal führt auch eine bestimmte Fragestellung zu einer entsprechenden Verwendungsweise. So ist von einem *Fokusakzent* meist dann die Rede, wenn davon ausgegangen wird, daß die Plazierung der prominentesten Silbe im Satz maßgeblich für eine bestimmte Satzbedeutung ist (s. Abschnitt 2.4). Einheiten, die hinsichtlich ihrer Betonungsverteilung miteinander verglichen werden, umfassen aber immer mindestens eine Silbe.

Die Erkenntnis, daß Menschen von verschiedenen Einheiten (Silben, Morpheme,

Wörter etc.) sagen können, „Einheit A ist stärker betont als Einheit B“, führte u.a. dazu, eine *prosodische Hierarchie* (NESPOR und VOGEL 1986) anzunehmen. Gemäß (WIESE 1996, 82-84) sind die für die deutsche Phonologie relevanten prosodischen Ebenen die in Abbildung 2.1 dargestellten. Wiese hält es aber für sehr wahrscheinlich, daß im Deutschen weitere relevante Ebenen (beispielsweise die Äußerungsebene) innerhalb der prosodischen Hierarchie existieren, für die sich lediglich noch keine explizite Evidenz hat finden lassen.

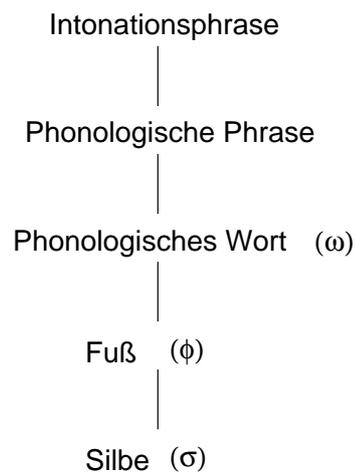


Abbildung 2.1: Für das Deutsche relevante Ebenen der prosodischen Hierarchie nach (WIESE 1996)

Eine *Silbe* ist eine intuitive phonetisch-phonologische Grundeinheit, auf welcher sich die Betonung ausprägen kann. Silben werden im Deutschen gebildet aus einem hochsonoren *Silbenkern* oder *Nukleus*, der meistens aus einem Vokal besteht, sowie weniger sonoren *Anlaut-* oder *Onsetkonsonanten* und optionalen *Auslaut-* oder *Kodakonsonanten*. Ein *Fuß* ist eine Folge einer betonten und mindestens einer unbetonten Silbe. Füße können im Deutschen aus zwei oder drei Silben bestehen (vgl. (FÉRY 1988)). Ein *phonologisches Wort* unterscheidet sich von einem *morphologischen Wort* dadurch, daß es u.U. aus zwei morphologischen Wörtern gebildet ist. Dies ist der Fall, wenn ein wenig prominentes Wort sich durch *Klitisierung* an ein anderes anbindet und mit diesem zu einer untrennbaren lautlichen Einheit verschmilzt, wie z.B. die Wörter „gibt“ und „es“ sich lautlich zu „gibt’s“ verbinden können. Eine lautliche Trennung dieser beiden Wörter in ihrer verschmolzenen Form, z.B. durch Einfügen einer Pause, würde zu einer ungrammatischen Äußerung führen. Eine *phonologische Phrase* ist eine prosodische Einheit mit einem prominenten *Nuklearakzent*, stellt aber noch keine vollständige Intonationsphrase dar. Diese ist die größte prosodische Äußerungseinheit, welche durch einen besonders prominenten *Nuklearakzent* sowie einen anschließenden *Grenzton*, also eine steigende oder fallende Grundfrequenzbewegung, verbunden mit einer Dehnung der finalen Silbe, gekennzeichnet ist. In dieser Arbeit wird die phonologische Phrase als separate Ebene der prosodischen Hierarchie ausgeklammert.

Jede phonologische Regel der Betonung muß innerhalb einer dieser Ebenen formuliert werden.

2.4.2 Betonung als relative Prominenz im Äußerungskontext

Im Gegensatz zu anderen phonologischen Merkmalen ist *Betonung* nicht binär, also nicht durch Vorhandensein oder Nichtvorhandensein gekennzeichnet. Für das Wort „Baum“ kann beispielsweise gesagt werden, daß alle Segmente des Wortes die phonologische Eigenschaft „stimmhaft“ tragen, und zwar unabhängig davon, ob noch andere, stimmlose Laute in derselben Äußerung auftreten wie das Wort Baum. Eine Betonungsstufe kann für das gleiche Wort allerdings nur in Relation zu anderen Wörtern definiert werden. Es ist nicht möglich zu sagen, dem Wort Baum komme die Betonungsstärke „3“ zu, ohne klarzustellen, welche Wörter dann die Betonungsstufen „2“, „1“ etc. haben. Eine Silbe oder ein Wort ist also durch ein Mehr oder Weniger — relativ zu anderen Wörtern oder Silben — an Betonung gekennzeichnet. Die Betonungsverhältnisse innerhalb einer Äußerung sind u.U. sehr wichtig, um bestimmte Bedeutungen zu übertragen. Zum einen können die Prominenzverhältnisse innerhalb von *Wörtern* bedeutungsrelevant — und damit phonologisch — sein, wie beispielsweise in dem phonologischen Minimalpaar *Kónstanz* vs. *Konstánz*. Des weiteren kann eine unterschiedliche Bedeutung durch Plazierung des *prosodischen Fokus* bewirkt werden, welcher sich in einer erhöhten Prominenz des fokussierten Äußerungsteils auswirkt. Tabelle 2.1 erläutert die verschiedenen Bedeutungen, welche sich im Englischen durch Verschiebung des prosodischen Fokus bzw. der prominentesten Silbe innerhalb der Äußerung ergeben können. Der Sitz des prosodischen Fokus ist dabei durch die Kursivschrift gekennzeichnet.

Fokussierung	Umschriebene Bedeutung
Peter only introduced <i>Bill</i> to Sue.	Peter hat Sue ausschließlich Bill vorgestellt.
Peter only introduced Bill to <i>Sue</i> .	Peter hat ausschließlich Sue Bill vorgestellt.
Peter only introduced <i>Bill</i> to <i>Sue</i> .	Peter hat ausschließlich Bill und Sue einander vorgestellt, sonst niemanden.

Tabelle 2.1: Unterschiedliche Satzbedeutungen, hervorgerufen durch Plazierung des prosodischen Fokus

Gemäß der Theorie von Rooth (ROOTH 1995) rückt eine prosodische Hervorhebung immer ein bestimmtes Element aus einer Reihe von kontextuell vorgegebenen Alternativen in den Vordergrund. Die Hervorhebung einer dieser Alternativen stelle die Antwort auf eine implizit oder explizit durch die Dialogsituation gestellte Frage dar. Häufig werden Fokustypen aufgrund dieser Fragen voneinander unterscheiden. Tabelle 2.2 zeigt eine Übersicht über die passenden Fragen zu verschiedenen

Antwortfokussen und die jeweilige Fokusbezeichnung, die für die Fokustypen in dieser Arbeit verwendet werden soll.

Fokustyp	Frage	Antwort mit Fokus
Weiter Fokus	Was ist passiert?	[Anna hat sich in Ben verliebt] _F .
Enger Objektfokus	In wen hat sich Anna verliebt?	Anna hat sich in [Ben] _F verliebt.
Enger Subjektfokus	Wer hat sich in Ben verliebt?	[Anna] _F hat sich in Ben verliebt.
Kontrast- oder Korrekturfokus	Anna hat sich also in Casimir verliebt?	(Nein,) Anna hat sich in [Ben] _F verliebt.

Tabelle 2.2: Verschiedene Fokustypen und ihre typischen Fragen. Die Klammerung läßt den Fokus erkennen.

Der sogenannte *weite Fokus* unterscheidet sich von den anderen Fokustypen, indem er die Antwort auf eine allgemeine Frage „Was ist passiert?“ darstellt. In der Antwort auf diese Frage ist jedes Element neu, nichts wird als bereits in den Diskurs eingeführt betrachtet und fällt somit weniger ins Gewicht. Diese Art der Betonung wird auch häufig als *Zitierform*, *Defaultbetonung* oder *all-new-Akzentuierung* bezeichnet. Einige Forscher, allen voran (BOLINGER 1972), haben bezweifelt, daß es eine solche Zitierform überhaupt gibt, weil jede Sprecherin bestimmte Elemente als besonders relevant hervorheben möchte. Das Ergebnis dieser Diskussion ist noch immer offen (siehe beispielsweise die Diskussion in (LADD 1996) sowie Seite 20ff. In dieser Arbeit wird zunächst von der Existenz einer Zitierform ausgegangen. Dies ist schon allein deshalb notwendig, weil ansonsten bei einer Vorhersage der Betonungsverhältnisse in Sätzen, die ohne Kontext gelesen wurden, die Interpretation der Sprecherin „erraten“ werden müßte.

Ein interessanter Aspekt beim Vorliegen eines weiten Fokus ist, daß dennoch ein Wort stärker hervorgehoben wird als die Restäußerung. Dieses Wort wird häufig als der *Fokusexponent* bezeichnet, der im Fokus liegende Äußerungsteil jedoch als *Fokusdomäne*. So unterscheiden sich der weite Fokus und der enge Objektfokus in Tabelle 2.2 nicht durch die Platzierung des Fokusexponenten, sondern lediglich durch die unterschiedliche Größe der Fokusdomäne. Im Falle des weiten Fokus umspannt die Fokusdomäne die gesamte Äußerung, im Falle des engen Objektfokus ist die Fokusdomäne deckungsgleich mit dem Fokusexponenten:

- weiter Fokus: [Anna hat sich in [Ben]_{Exponent} verliebt.]_{Domne}
- enger Fokus: Anna hat sich in [[Ben]_{Exponent}]_{Domne} verliebt.

Es ist noch ungeklärt, ob sich die prosodische Umsetzung dieser beiden Beispiele voneinander unterscheidet (KRAHMER und SWERTS 1998).

Ferner sollte darauf hingewiesen werden, daß die Verwendung des Wortes *Kontrastfokus*, wie es in Tabelle 2.2 eingeführt wurde, unterschiedlich verwendet wird. Manchmal wird nämlich nicht nur die Korrektur als Kontrastfokus aufgefaßt, sondern auch folgende Art der multiplen Hervorhebung als solcher gewertet:

1. Ein [altes]_F Haus hat mehr Charme als ein [neues]_F Haus.

Rein semantische Analysen wie die von (ROOTH 1995) behandeln den multiplen Fokus genauso wie den Korrekturfokus. Dennoch scheinen beide Muster sich pragmatisch voneinander zu unterscheiden: Bei einem Korrekturfokus versucht die Sprecherin, aktiv etwas zu überschreiben, von dem sie glaubt, es sei der Informationsstatus des Hörers. Ein multipler Fokus kann geäußert werden, um den Informationsstatus des Hörers auf einen neueren Stand zu bringen, ohne etwas daran ändern zu wollen. Der Umstand, daß für prosodische Hervorhebungen in erster Linie pragmatische und weniger semantische Bedingungen eine Rolle spielen, ist ebenfalls das Ergebnis der Arbeiten von (SCHMITZ et al. 2001). Die explizite Verwendung des Begriffes Kontrastfokus im Sinne einer Korrektur ist nicht neu, sondern deckt sich mit den Fokuskonzepten in (FÉRY 1988; PIERREHUMBERT und HIRSCHBERG 1990).

Die inhaltliche Hervorhebung prosodisch fokussierter Wörter kann auf der Basis von Silbeprominenzen modelliert werden. Dies konnte im Rahmen der Evaluation eines Sprachsynthesystems gezeigt werden (PORTELE und HEUFT 1997), welches eine silbeprominenzgetriebene Prosodiesteuerung enthält. Weiterhin deuten Experimente darauf hin, daß eine hohe Prominenz nicht nur eine inhaltliche Hervorhebung markiert, sondern daß der Grad der Prominenz einen Einfluß auf die Art des wahrgenommenen Fokus haben kann. Je höher die Prominenz relativ zur Umgebung ist, um so eher scheint der Eindruck einer kontrastiven oder auch korrigierenden Hervorhebung gegeben zu sein (WOLTERS und WAGNER 1998). Die Manipulation der Prominenz konnte in dieser Untersuchung allerdings den Eindruck einer Korrektur nicht reliabel modellieren, so daß der Eindruck entstand, daß die Hervorhebung relativ zu ihrer Umgebung stärker werden muß. Dies bedeutet, daß die Umgebung der fokussierten Einheit hinsichtlich ihrer Prominenz ebenfalls betroffen ist, um den einen oder anderen Fokuseindruck beim Hörer zu erzielen. Die Hypothese, daß die Prominenz einer Silbe nicht isoliert wahrgenommen wird, sondern daß es Interaktionen zwischen den Prominenzen gibt, die im Zusammenhang mit der linguistischen Struktur stehen, wurde anhand von zwei empirischen Untersuchungen getestet. Die erste Untersuchung beschäftigt sich mit Prominenzunterschieden von Äußerungen in Zitierform und Äußerungen, die nicht der Zitierform entsprechen. In der zweiten Untersuchung wurde ein Prominenzmuster isoliert, welches von Hörern als Korrekturkontrast interpretiert wird. Beide Untersuchungen weisen darauf hin, daß es für die Signalisierung bestimmter linguistischer Inhalte (hier: etwas Neues/Unvorhergesehenes wird in den Dialog eingeführt, bzw. es wird etwas korrigiert) darauf ankommt, daß die Prominenzen innerhalb der Äußerung relativ zueinander bestimmte Abstände haben müssen, um Eindrücke beim Hörer hervorzurufen, die mit den entsprechenden Bedeutungen verknüpft sind.

2.4.2.1 Prominenzmuster von Zitierform und Nicht-Zitierform

Wie bereits erwähnt (vgl. Seite 18), ist es ein noch immer debattiertes Problem, ob es einen Standardfall für die Betonungsmuster in Äußerungen gibt, welcher auf rein syntaktischer Basis vorhersagbar ist. (BOLINGER 1972) und (FUCHS 1984) sind der Ansicht, daß sich die Satzbetonung nur aufgrund semantischen oder pragmatischen Wissens vorhersagen läßt und daß andersartige Vorhersagen zum Scheitern verurteilt sind. (LADD 1996, 160ff.) hingegen argumentiert für die Annahme einer Standardbetonung. Seine Ansicht wird auch in dieser Arbeit unterstützt, und an späterer Stelle (siehe Kapitel 4) wird gezeigt werden, daß auf der Basis dieser Annahme für gelesene Sprache recht gute Vorhersagen von Betonung möglich sind. Wenn man auf semantisch-pragmatische Begriffe rekurrieren möchte, um mit der Annahme einer Standardbetonung zu leben, so sei darauf hingewiesen, daß eine solche Betonung einem „weiten Fokus“ entspricht. Semantisch-pragmatisch gesehen stellt dieser Fall eine Äußerung dar, innerhalb der die Sprecherin davon ausgeht, alles in ihr enthaltene sei „neu“ für den Hörer. Dieser Standardfall der Äußerungsbetonung soll im Folgenden als *Zitierform* bezeichnet werden.

Es sollte untersucht werden, ob sich die Werte für die prominenteste Silbe im Falle einer Zitierform von den Prominenzwerten im Falle einer Nicht-Zitierform unterscheiden. Die prominenteste Silbe sollte dort liegen, wo sich auch der Fokusexponent einer Äußerung befindet. Für die Untersuchung wurden in einer prosodischen Datenbank für das Amerikanische Englisch (ELSNER et al. 1998) die annotierten Prominenzwerte für beide Fälle untersucht. Das untersuchte Material stammt von einer Sprecherin und einem Sprecher. Die Datenbank enthält zwar gelesene Sprache, verfügt aber dennoch über reichlich Ausnahmen zur Zitierform, da das gelesene Material aus kurzen Frage-Antwort-Sequenzen besteht, welche dazu führen, daß manche Stellen eine Äußerung in Nicht-Zitierform von den Sprecherinnen erfordern. Beispielsweise enthält die Datenbank folgende Frage-Antwort-Paare:

- Beispiel für eine Zitierform in einer Frage:

Frage: Where is that BUTton?

Antwort: It's on the left hand side of the panel.

- Beispiel für eine Nicht-Zitierform in einer Frage:

Frage: Press the orange button!

Antwort: Where IS that button?

Obwohl die Betonung auf dem Verb *IS* im Falle der Nicht-Zitierform innerhalb der vorgegebenen Dialogsituation völlig vorhersagbar ist, entspricht sie nicht dem Betonungsmuster, welches eine amerikanische Muttersprachlerin in dem ersten Fall anwenden würde, in dem sie den gleichen Satz ohne Kontext äußert.

Die Äußerungen, bei denen eine Zitierform erwartbar war, wurden in der Datenbank mit der Eigenschaft $[+default]$ annotiert, die restlichen entsprechend als $[-default]$. Einen Überblick über die getesteten Werte und die Abkürzungen, mit denen auf diese Werte referiert wird, liefert die Tabelle 2.3.

maxprom	Der höchste Prominenzwert innerhalb einer Äußerung wird als <i>maxprom</i> bezeichnet. Die Silbe, die diesen Wert trägt, wird <i>maxprom-Silbe</i> genannt. Diese Silbe entspricht meist dem Ort der Realisierung einer Fokusbetonung.
meanprom	Die mittlere Prominenz innerhalb einer Äußerung mit Ausnahme des Wertes <i>maxprom</i> .
premax	Die mittlere Prominenz innerhalb einer Äußerung vor der <i>maxprom-Silbe</i>
postmax	Die mittlere Prominenz innerhalb einer Äußerung nach der <i>maxprom-Silbe</i>

Tabelle 2.3: Die untersuchten Prominenzeigenschaften innerhalb der Äußerungen und ihre Abkürzungen

Mit Hilfe statistischer Untersuchungen wurden die zwei folgenden Hypothesen getestet:

1. Es besteht ein Zusammenhang zwischen der Prominenz der *maxprom-Silbe* und der Prominenz ihrer Umgebung. Bei höheren Prominenzwerten tritt eine Deakzentuierung der Umgebung auf, die sich in einer allgemein niedrigeren Prominenz auswirken muß. Die Erwartung ist, daß mit zunehmendem *maxprom*-Wert der Wert *meanprom* niedriger wird, weil eine hohe Prominenz vermutlich besonders in Erscheinung tritt, wenn sie von relativ niedrigen Prominenzen umgeben ist.
2. Die Äußerungen, die nicht der Zitierform entsprechen, unterscheiden sich von den anderen hinsichtlich des Verhältnisses zwischen *maxprom*- und *meanprom*-Werten.

Die statistischen Untersuchungen ergaben zum einen, daß es keinerlei Korrelation zwischen der Prominenz der *maxprom-Silbe* und den Prominenzen innerhalb der Restäußerung gab. Dies bedeutet, daß eine hohe Prominenz der betontesten Silbe nicht automatisch eine Deakzentuierung der Restäußerung nach sich zieht. Anschließend wurde untersucht, ob der Abstand zwischen *maxprom*-Wert und die mittlere Prominenz der Restäußerung in einem Zusammenhang zueinander stehen. Hierbei konnte eine hohe Korrelation (Sprecher: $cc = 0.8, p < 0.0001$; Sprecherin:

$cc = 0.76, p < 0.0001$) zwischen beiden Werten gemessen werden (vgl. Abbildung 2.2). Dieses Ergebnis kann folgendermaßen interpretiert werden: Je höher der Prominenzwert der fokussierten Silbe ist, desto größer ist auch der Abstand zur mittleren Prominenz in der Restäußerung. Dies Ergebnis zeigt deutlich, daß die mittlere Äußerungsprominenz in keiner Weise vom Wert der maxprom-Silbe beeinflusst wird, sondern im Gegenteil äußerst stabil ist. Die Prominenz einer fokussierten Silbe hat offenbar keinen Einfluss auf die wahrgenommene Prominenz ihrer Umgebung, sondern kann autonom als mehr oder weniger betont wahrgenommen werden. Sprecherinnen sind dementsprechend in der Lage, ein Einzelwort oder eine einzelne Silbe innerhalb einer Äußerung hervorzuheben. Die erste Hypothese wurde somit verworfen.

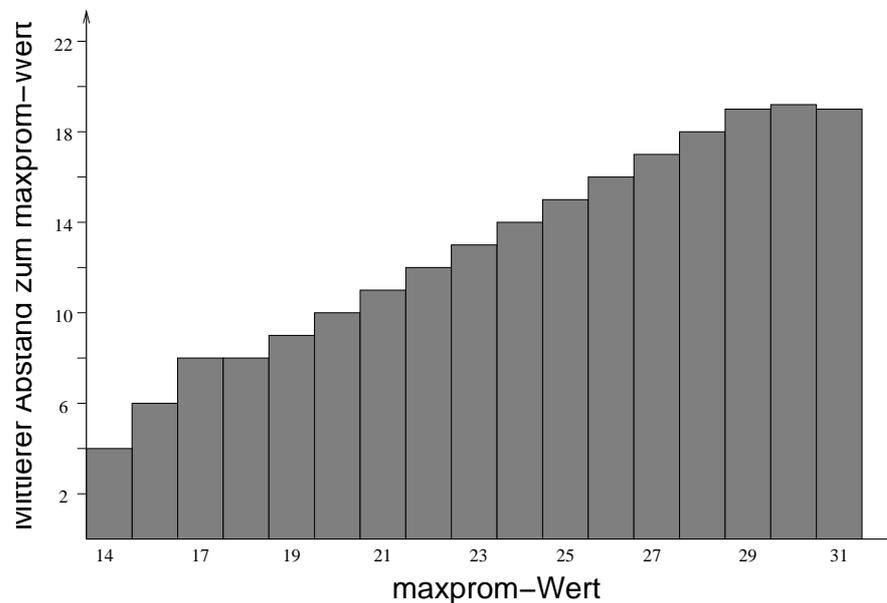


Abbildung 2.2: Dargestellt ist der Abstand zwischen der Prominenz der fokussierten Silbe und der mittleren Prominenz der Äußerungsumgebung. Prominenzwerte können innerhalb einer Äußerung die Werte 0 bis 31 annehmen. Mit zunehmender Prominenz der fokussierten Silbe steigt der Abstand. Dies bedeutet, daß die Prominenz der Umgebung relativ stabil ist.

Bei einem Vergleich der mittleren Prominenzwerte zwischen Äußerungen in Zitierform und Äußerungen ohne Zitierform konnten keine Unterschiede hinsichtlich der maxprom-Werte festgestellt werden. Fokussierte Silben, die nicht der Zitierform entsprechen, sind also nicht prominenter als andere.

Da die Differenz zwischen der prominentesten Silbe und der Restäußerung sich als äußerst stabil zeigte, wurde in den darauffolgenden Tests untersucht, ob es an dieser Stelle Unterschiede zwischen Äußerungen in Zitierform und Äußerungen ohne Zitierform gibt. Bei einem Vergleich des Abstandes zwischen dem Prominenzwert der Satzbetonung und der Prominenz der Restäußerung konnten wiederum keine Unterschiede festgestellt werden. Die Abstände waren im Durchschnitt beinahe

identisch (vgl. Abbildung 2.3). Vergleich man allerdings die mittleren Prominenzwerte jeweils VOR und NACH dem Satzakzent, so wurden Unterschiede sichtbar. In Äußerungen, die von der Zitierform abweichen, ist der Abstand zum Wert der prominentesten Silbe VOR dieser signifikant höher (t-test, Sprecher: $p < 0.0001$; Sprecherin: $p < 0.01$; vgl. Abbildung 2.4), NACH der maxprom-Silbe ist dieser Abstand allerdings signifikant geringer (t-test, Sprecher: $p < 0.05$; Sprecherin: $p < 0.05$; vgl. Abbildung 2.5).

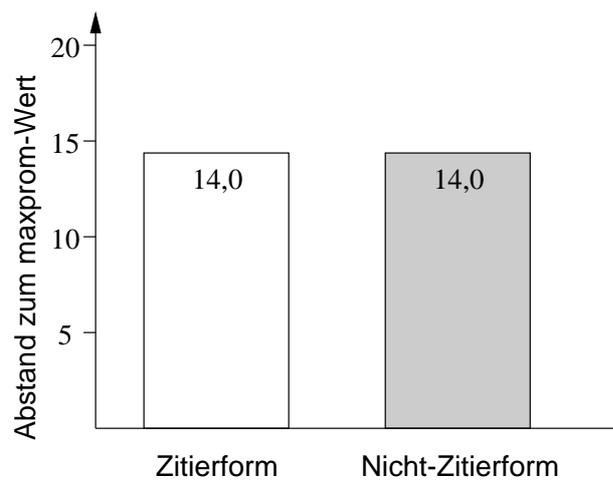


Abbildung 2.3: Dargestellt ist der Abstand zwischen der Prominenz der fokussierten Silbe und der Prominenz der Restäußerung. Beide Werte sind identisch für Zitierformen und Nicht-Zitierformen.

Insgesamt zeigen diese Untersuchungen, daß die Betrachtung der absoluten Prominenzwerte keinerlei Unterscheidungen zwischen Äußerungen in Zitierform und Nicht-Zitierform zuläßt. Die äußerungsinternen Prominenzrelationen werden jedoch durch die Plazierung der prominentesten Silben beeinflusst. Dies ist ein starkes Indiz dafür, daß linguistisch relevante Unterschiede durch die äußerungsinternen Prominenzrelationen widerspiegelt werden. Dies kann auch der Grund dafür sein, daß im Niederländischen Fokusexponenten, die nicht der Zitierform entsprechen, nur dann als betonter klassifiziert werden, wenn sie im Kontext gehört werden (KRAHMER und SWERTS 1998). Die absoluten Prominenzwerte sind hingegen nicht von den linguistisch relevanten Prominenzverhältnissen betroffen.

2.4.2.2 Ein Prominenzmuster für die Modellierung von Korrekturkontrast

Die o.a. Untersuchung konnte zeigen, daß sich Prominenzen prinzipiell eignen, um linguistisch relevante prosodische Muster zu modellieren. Die Eignung einer prominenzbasierten Beschreibung für die gezielte Modellierung eines Bedeutungsunterschiedes soll am Beispiel des Kontrastfokus dokumentiert werden (vgl. (WAGNER 1999)). Es sollte festgestellt werden, ob sich ein Prominenzmuster isolieren läßt, welches Hörer eindeutig mit einem Eindruck eines Korrekturkontrasts verknüpfen. Um

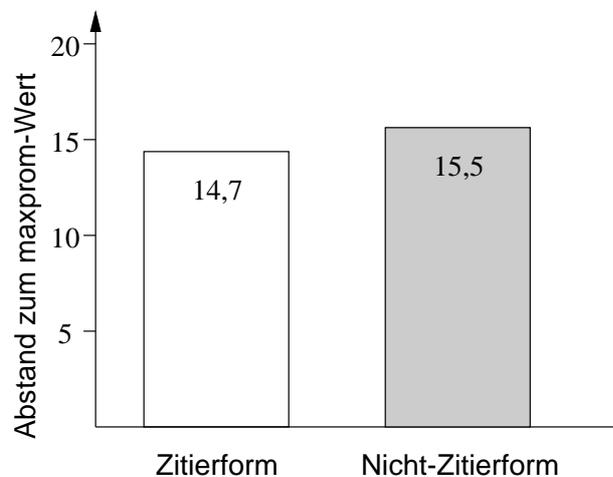


Abbildung 2.4: Dargestellt ist der Abstand zwischen der Prominenz der prominentesten Silbe und der mittleren Prominenz VOR dieser Silbe. Der Abstand ist größer bei Äußerungen, die nicht der Zitierform entsprechen.

Hinweise auf die Existenz eines solchen Musters zu erhalten, wurden in einem ersten Schritt Modellierungen von Kontrast in Phonologie und Phonetik untersucht. Obwohl die Existenz einer eindeutigen Kontrastprosodie noch umstritten ist, existieren Vorschläge für das Deutsche wie das Englische, den Kontrast mit Hilfe eines sogenannten $L + H^*$ -Tonhöhenakzents (REYELT et al. 1996a; PIERREHUMBERT und HIRSCHBERG 1990) zu modellieren. Diese Art der Beschreibung lehnt sich an das ToBI-Modell der Intonation an, das auf (PIERREHUMBERT 1980) zurückgeht und von (REYELT et al. 1996b; BAUMANN et al. 2000) an das Deutsche adaptiert wurde. Im ToBI-Modell wird die Kontur einer Intonationsphrase anhand der Tonhöhenakzente (annotiert durch ,*') sowie eines Grenztons (annotiert durch ,%') beschrieben, welche anhand der Täler (annotiert durch ,L') und Gipfel (annotiert durch ,L') charakterisiert werden kann. Andere phonologische Arbeiten schlagen für das Deutsche eine Deakzentuierung nach dem Fokusexponenten (FÉRY 1988) oder aber für das Englische und Polnische eine Operation auf dem metrischen Baum vor (vgl. Abschnitt 2.4), die als *Contrastive Relabelling* bezeichnet wird (DOGIL 1979). Das *Contrastive Relabelling* bewirkt ebenfalls eine postfokale Schwächung der Prominenz und eine gleichzeitige Stärkung des Fokusexponenten. Experimente für das Niederländische (KRAHMER und SWERTS 1998) haben gezeigt, daß die Wahrnehmung von Kontrast nur dann funktioniert, wenn die fokussierte Konstituente im Zusammenhang mit der Umgebung gehört wird. In Isolation ist eine Wahrnehmung von Kontrast offenbar nicht möglich.

Phonetische Analysen des Korrekturfokus (COOPER et al. 1985) haben gezeigt, daß sich ein Kontrast akustisch durch eine Zunahme an Dauer des fokussierten Wortes sowie einer flachen postfokalen F_0 -Kurve charakterisieren läßt. Da weder die Grundfrequenz noch die Dauer, in Isolation betrachtet, Korrelate von Korrektur

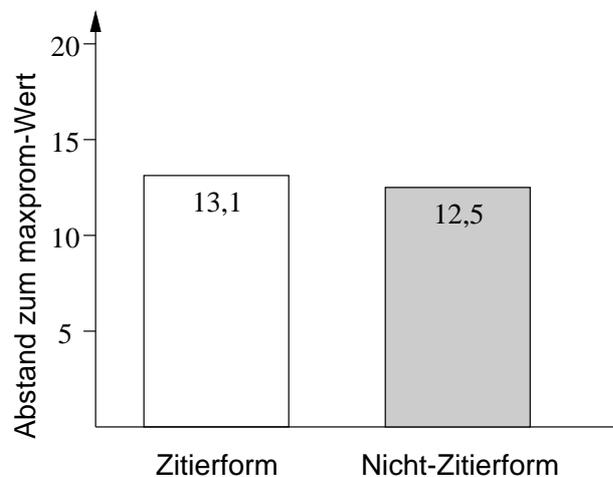


Abbildung 2.5: Dargestellt ist der Abstand zwischen der Prominenz des fokussierten Silbe und der mittleren Prominenz NACH dieser Silbe. Der Abstand ist geringer bei Äußerungen, die nicht der Zitierform entsprechen.

darstellen, ist eine einfachere quantitative Darstellung von Korrektur im phonetischen Bereich wünschenswert. Es wurde also untersucht, inwieweit dieses Desiderat durch eine Prominenzstruktur erfüllt werden kann.

Hierzu hat man sich den Umstand zunutze gemacht, daß der wohldefinierte Zusammenhang von perceptiver Prominenz und deren akustischen Korrelaten in einem Sprachsynthesystem für das Deutsche (PORTELE und HEUFT 1997) Verwendung gefunden hat. Das Synthesystem basiert ähnlich wie (FANT und KRUCKENBERG 1989) auf einer Prominenzskalierung zwischen 0 und 31 und baut auf extensive Studien zum Zusammenhang zwischen Prominenz und akustischen Korrelaten im Deutschen (HEUFT 1999) auf. Silben, die, mit hohen Prominenzwerten versehen, synthetisiert werden, tendieren dazu, als engfokussiert wahrgenommen zu werden (WOLTERS und WAGNER 1998; PORTELE und HEUFT 1997). Für eine gezielte Wahrnehmung einer Korrektur hingegen reicht es nicht aus, lediglich die Prominenz der fokussierten Silbe³ zu erhöhen. Es lag also nahe, die Prominenzen der umgebenden Silben ebenfalls zu manipulieren. Es wurden verschiedene Strategien der Prominenzmanipulation angewandt, um die beste Methode herauszufinden.

Das Experiment Es wurden drei Deklarativsätze mit intendiertem Korrekturfokus in jeweils drei unterschiedlichen Satzpositionen synthetisiert, insgesamt gab es demnach 9 Varianten unterschiedlicher Satz-Fokusposition-Paare (siehe Tabelle 2.4). Jede dieser Varianten wurde durch fünf unterschiedliche Strategien (siehe Tabelle 2.5) synthetisiert, um den Eindruck eines Korrekturfokus zu modellieren. Die Strategien waren motiviert durch die o.a. phonologischen wie phonetischen Studien zur

³Mit fokussierter Silbe ist diejenige Silbe innerhalb des fokussierten Wortes gemeint, welche Trägerin der Wortbetonung ist.

Korrektur. Die Prominenzwerte der nicht-manipulierten Äußerungskonstituenten wurden mit Hilfe des Verbmobil-Transkriptionsmoduls (STÖBER et al. 2000) generiert, welches Prominenzwerte anhand syntaktischer und lexikalischer Information berechnet, sofern keine weiteren semantischen oder pragmatischen Informationen — m.a.W. Fokusinformationen — vorliegen. Das Resultat dieses Vorgehens ist ein Prominenzmuster, welches die Prosodie der Zitierform modelliert.

<i>ENde MAI bin ICH noch im Urlaub.</i>
<i>Es würde mich freuen, wenn wir noch einen Termin ausmachen.</i>
<i>Anfang Mai hätte ich noch Zeit.</i>

Tabelle 2.4: Unterschiedliche Stimulussätze mit den verschiedenen Positionen für den Korrekturfokus (kursiv)

Strategie 1	Die fokussierte Silbe wurde auf den höchstmöglichen Prominenzwert gesetzt (31). Zusätzlich wurde die Prominenz verstärkt, indem die Dauer der fokussierten Silbe erhöht wurde. Sämtliche restlichen Silben innerhalb der prosodischen Phrase wurden deakzentuiert, indem ihre Prominenzen — falls notwendig — in einem Maß reduziert wurden, daß sie ohne Tonhöhenakzent synthetisiert wurden.
Strategie 2	Die fokussierte Silbe wurde auf den höchstmöglichen Prominenzwert gesetzt (31). Zusätzlich wurde die Prominenz verstärkt, indem die Dauer der fokussierten Silbe erhöht wurde. Sämtliche <i>postfokalen</i> Silben innerhalb der prosodischen Phrase wurden deakzentuiert, indem ihre Prominenzen in einem Maß reduziert wurden, daß sie ohne Tonhöhenakzent synthetisiert wurden.
Strategie 3	Die fokussierte Silbe wurde auf den höchstmöglichen Prominenzwert gesetzt (31). Zusätzlich wurde die Prominenz verstärkt, indem die Dauer der fokussierten Silbe erhöht wurde. Keine zusätzliche Deakzentuierung der Umgebung wurde bei der Synthese angewandt.
Strategie 4	Die fokussierte Silbe wurde auf den höchstmöglichen Prominenzwert gesetzt (31). Keine weitere Manipulation wurde angewandt.
Strategie 5	Die fokussierte Silbe wurde auf den höchstmöglichen Prominenzwert gesetzt (31), aber nicht zusätzlich in ihrer Dauer manipuliert. Sämtliche restlichen Silben innerhalb der prosodischen Phrase wurden deakzentuiert, indem ihre Prominenzen in einem Maß reduziert wurden, daß sie ohne Tonhöhenakzent synthetisiert wurden.

Tabelle 2.5: Fünf Strategien für die Synthese eines Korrekturfokus

Jeder Satz wurde zudem mit zwei unterschiedlichen Fragekontexten versehen. Einer dieser Kontexte ist genau dann wohlgeformt, wenn der intendierte Korrekturfokus vom Hörer als solchem interpretiert wird. Der zweite Frage-Antwort-Kontext sollte im Fall einer solchen Interpretation dem Hörer als nicht-wohlgeformt oder gar ungrammatisch erscheinen. Die beiden folgenden Frage-Antwort-Paare sind Beispiele für eine Korrekturantwort, welche in einem passenden (vgl. 1.) und einem unpassenden Fragekontext (vgl. 2.) präsentiert wurde:

1. passender Fragekontext für Korrektur in der Antwort:
Frage: Anfang Mai sind sie also noch im Urlaub?
Antwort: Ende Mai bin ich noch im Urlaub.
2. unpassender Fragekontext für Korrektur in der Antwort:
Frage: Ende Juni sind sie also noch im Urlaub?
Antwort: ?? Ende Mai bin ich noch im Urlaub.

Die Kontextfragen wurden von einem männlichen, phonetisch geschulten Sprecher, welcher mit dem Experiment vertraut war, in einem schalltoten Raum laut vorgelesen und aufgenommen. Er wurde instruiert, die Fragen mit einer neutralen Betonung vorzulesen, welche beim Hörer keinerlei Erwartung hinsichtlich der Antwort oder ihrer korrekten Betonung erwecken sollte. Die insgesamt 180 Frage-Antwort-Paare wurden 11 phonetisch geschulten Versuchspersonen präsentiert. Die Stimuli wurden per Kopfhörer angehört, die Fragesätze wurden zusätzlich orthographisch auf dem Bildschirm dargestellt. Die Versuchspersonen durften sich jeden synthetisierten Antwortsatz mehrfach anhören. Anschließend mußten die Versuchspersonen die Stimmigkeit eines jeden Frage-Antwort-Paares auf einer Skala von 1-6 beurteilen. Die Skala lehnt sich dabei an die deutschen Schulnoten an, d.h., eine „1“ bedeutet eine sehr gute Stimmigkeit zwischen Frage und Antwort, eine „6“ hingegen eine ungenügende. Die Skala verhindert, daß die Hörer sich für ein „unentschieden“ entscheiden, sondern verlangt eine Entscheidung zu einer eher positiven oder negativen Stimmigkeit des jeweiligen Frage-Antwort-Paares („forced choice“).

Resultate Diejenigen Frage-Antwort-Paare, welche zu dem intendierten Korrekturfokus paßten, wurden signifikant besser bewertet als die nicht passenden (Kolgomorov-Smirnov, $p < 0.001$). Außerdem konnte eine nicht unerhebliche negative Korrelation (Spearman-Rho, $\rho = -0,49$, $p < 0,01$) zwischen den passenden und den nicht-passenden Frage-Antwort-Paaren festgestellt werden. Diese Ergebnisse gelten über alle Stimuli — unabhängig von bestimmten Synthesestrategien. Abbildung 2.6 zeigt die Verteilung der Hörerbewertungen für die verschiedenen Fragekontexte.

Es war jedoch nicht möglich, einen Einzelfaktor zu bestimmen, welcher allein den Eindruck einer Korrektur beim Hörer stabil hervorrufen konnte. Die Beurteilungen für die Strategien, bei denen lediglich ein Einzelaspekt manipuliert wurde (nur Prominenzhöhung der fokussierten Silbe, nur zusätzliche Deakzentuierung der Umgebung, nur zusätzliche Dauermanipulation), wurden allesamt signifikant schlechter bewertet als die Kombinationsstrategien, unterschieden sich hinsichtlich ihrer Beurteilungen aber nicht signifikant voneinander.

Des weiteren wurde deutlich, daß eine kombinierte prä- UND postfokale Deakzentuierung zwar eine bessere, aber nicht signifikant bessere Beurteilung nach sich zieht. Weder die Hörer noch der Satz haben einen Einfluß auf die Beurteilung, lediglich wurde auffällig, daß der Korrektureindruck signifikant schlechter bewertet

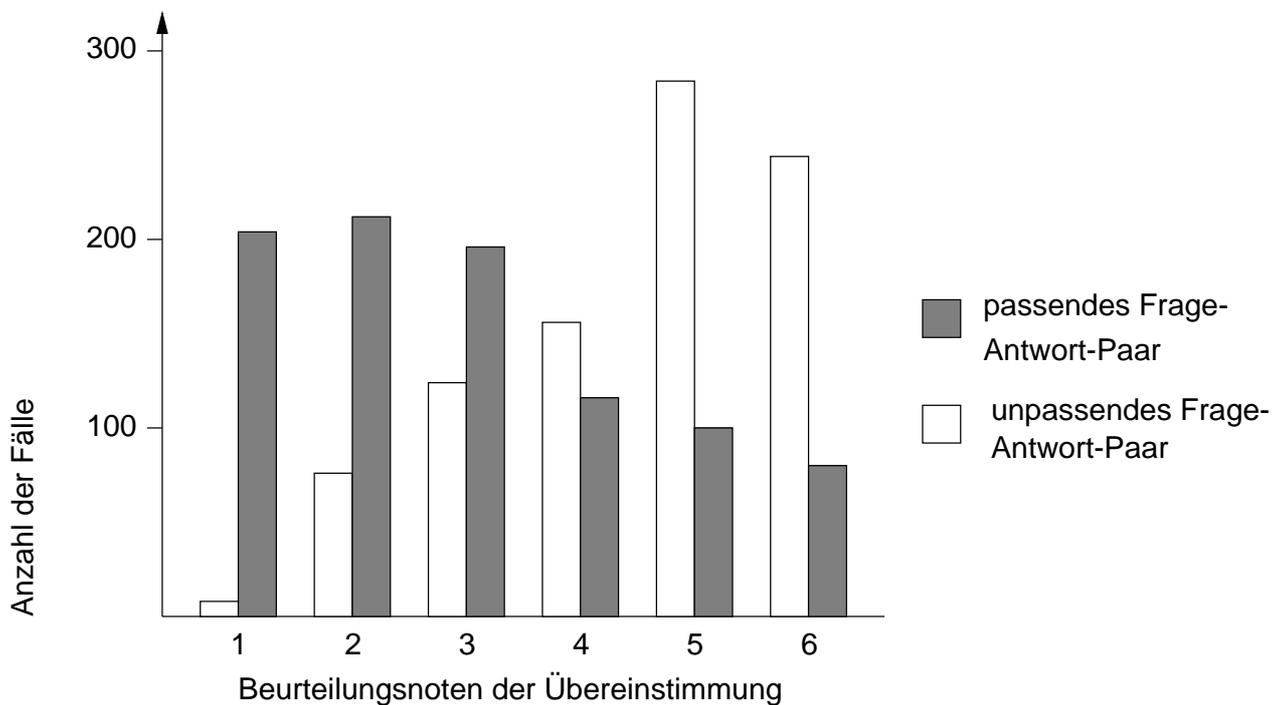


Abbildung 2.6: Unterschiedliche Hörerurteile für passende und unpassende Fragekontexte

wird, wenn der Fokus weit hinten in der Äußerung liegt. Dies ist wenig verwunderlich, wenn man bedenkt, daß die postfokale Deakzentuierung, die eine große Rolle bei der Modellierung des Korrekturindrucks spielt, in diesen Fällen zum Einsatz kommt. Außerdem überlappt in diesen Fällen häufig die Position des Korrekturfokus mit der Position der prominentesten Silbe im Falle einer Zitierform. Es ist anzunehmen, daß Sprecherinnen in Dialogen solche Fälle aktiv umgehen, indem sie zusätzlich zur prosodischen Markierung des Fokus eine syntaktische Markierung durch eine Topikalisierung (GUNDEL 1977) wählen. Evtl. verwenden Sprecherinnen in diesen Fällen aber auch noch weitere Strategien zur prosodischen Hervorhebung der korrigierten Konstituente wie eine präfokale Deakzentuierung. Äußerungen mit präfokaler Deakzentuierung wurden ja tatsächlich besser — wenn auch nicht signifikant — bewertet. Alternativ möglich ist auch eine Abgrenzung der fokussierten Konstituente durch das Einfügen einer prosodischen Phrasengrenze:

Ende Mai habe ich noch | *Zeit* ||

Welche Strategien Sprecherinnen in diesen Fällen anwenden, wurde allerdings nicht weiter untersucht.

Fazit Der Eindruck einer Korrektur kann nicht durch die Manipulation eines einzelnen Faktors erweckt werden, sondern nur durch das Zusammenspiel von stark

erhöhter Prominenz auf der hauptbetonten Silbe und einer zusätzlichen postfokalen Deakzentuierung modelliert werden. Die höchsten Prominenzwerte, welche sich mit Hilfe des Synthesystems modellieren ließen, waren nicht ausreichend, um den gezielten Eindruck beim Hörer zu erwecken. Vielmehr war eine zusätzliche Dauer-manipulation notwendig. Da die Syntheseregeln aber ursprünglich für die Modellierung einer Normprosodie innerhalb einer *Text-to-Speech*-Anwendung gedacht waren, ist dies nicht verwunderlich, da innerhalb von Syntheseanwendungen die Möglichkeit grober Ausreißer i.allg. vermieden wird. Eine Korrektur macht aber genau eine solch starke Abweichung von der Norm erforderlich und verlangt daher einen gezielten Eingriff in die eingeschränkten Modellierungsmöglichkeiten des Synthesystems.

Wichtigstes Ergebnis dieser Studie war die Bestätigung von zwei Annahmen:

- Ein (Korrektur-)Fokus ist kein lokales Phänomen, kann also vermutlich auch nicht ausschließlich durch eine bestimmte Intonationskontur oder eine bestimmte Dauer modelliert werden. Vielmehr ist es für die Wahrnehmung einer Korrektur wichtig, daß das fokussierte Element sich hinsichtlich seiner Prominenz stark vom Kontext abhebt.
- Der Bereich der Deakzentuierung, durch welche das fokussierte Element noch stärker hervorgehoben wird, ist postfokal.

2.4.3 Prominenzmodellierungen unterschiedlicher phonologischer Theorien

Die Beschäftigung der Phonologie mit dem Phänomen der Betonung ist — verglichen mit anderen Bereichen — noch verhältnismäßig jung und erfuhr ihre erste formal-systematische Darstellung innerhalb der *Generativen Phonologie* durch den Ansatz von (CHOMSKY und HALLE 1968). In den 70er Jahren erkannten (LIBERMAN und PRINCE 1977) gewisse Defizite der Darstellung und erweiterten die phonologische Modellierung der Betonung durch die Methoden der *Metrischen Phonologie*, welche später integrativer Bestandteil der *Nichtlinearen Phonologie* wurde. Die Nichtlineare oder *Autosegmental-Metrische Phonologie* ist nach wie vor tonangebend innerhalb der phonologischen Theorienbildung, allerdings wurde ihre generative, regelorientierte Prägung durch deklarative Herangehensweisen verändert. Neue deklarative phonologische Theorien sind die *Optimalitätstheorie* und verschiedene formale, deklarative Formalismen. Innerhalb dieser neueren Formalismen wurden ebenfalls Vorschläge für die Modellierung von Betonung gemacht. Die verschiedenen Herangehensweisen werden nachfolgend skizziert.

2.4.3.1 Generative Phonologie

In der Generativen Phonologie gemäß (CHOMSKY und HALLE 1968) wird die Betonung als ein *phonologisches Merkmal* von Vokalen betrachtet. Als einziges phonologisches Merkmal ist die Betonung dabei nicht binär, sondern kann sich prinzipiell in beliebig viele Stufen unterteilen. Je höher die relative Prominenz eines Vokals innerhalb der Äußerung, um so niedriger die Zahl, welche diese Betonung kennzeichnet. Diese Art der Numerierung erfolgt in Anlehnung an die Bezeichnung „primary stress“ für die prominenteste, „secondary stress“ für die zweitprominenteste Silbe eines Wortes etc., welche im IPA verwendet wird. Die Betonungsmuster einer Äußerung werden durch die Anwendung phonologischer Regeln auf die *phonologische Oberflächenstruktur* generiert. Die phonologischen Regeln haben die Form:

$$A \rightarrow B/C_D$$

Dies ist folgendermaßen zu lesen:

Der Einheit mit der Eigenschaft A wird die Eigenschaft B zugewiesen, sofern sie sich in der Umgebung von Einheiten mit den Eigenschaften C und D befindet.

Betonungsregeln können einem Vokal immer nur die Primärbetonung (markiert durch eine „1“ oberhalb des Vokals) zuweisen. Kommt eine solche Regel zum Tragen, so bedeutet das ein Herunterstufen der Betonung aller anderen Vokale der insgesamt betrachteten Einheit, auf welche die Regel angewendet wurde.

Die Regeln beziehen sich auf annotierte, geklammerte phonologische Einheiten und werden zyklisch angewendet. Das bedeutet, daß ein Regelzyklus sich zunächst auf die Inhalte der innersten Klammer bezieht. Sind sämtliche Regeln durchlaufen, werden die innersten Klammern gelöscht, und die Regeln werden auf die neuen innersten Klammerinhalte angewandt. Dies Vorgehen wiederholt sich, bis die äußerste Klammer erreicht ist. Die zyklische Regelanwendung soll exemplarisch dargestellt werden am Beispiel der segmental identischen Einheiten „ein Baum“ und „Einbaum“. Die unterschiedliche Klammer- bzw. Baumstruktur ist in Abbildung 2.7 dargestellt.



$$\left[\underset{\text{NP}}{\#} \left[\underset{\text{N}}{\#} \text{EIN} \# \right] \left[\underset{\text{N}}{\#} \text{BAUM} \# \right] \right] \quad \left[\underset{\text{N}}{\#} \left[\underset{\text{N}}{\#} \text{EIN} \# \right] \left[\underset{\text{N}}{\#} \text{BAUM} \# \right] \right] \underset{\text{N NP}}{\#}$$

Abbildung 2.7: Die Klammerstrukturen der segmental identischen Einheiten „Ein Baum“ und „Einbaum“

Folgende drei generative Regeln können nun zyklisch auf diese beiden phonologischen Strukturen angewandt werden:

1. $V \rightarrow [1V]/\#(C+)_ (C+)\#$ (Einsilbige Wörter erhalten eine primäre Betonung.)
2. $[1V] \rightarrow [1V]/_ [1V]]_N\#$ (Innerhalb eines Nomens bleibt ein Vokal Träger der primären Betonung, wenn er diese bereits trägt und einem Vokal vorausgeht, welcher ebenfalls primär betont ist. (*Compound Stress Rule*))
3. $[1V] \rightarrow [1V]/[1V]_]_{NP}\#$ (Innerhalb einer Nominalphrase bleibt ein Vokal Träger der primären Betonung, wenn dieser eine solche trägt und von einem Vokal gefolgt wird, welcher sie ebenfalls trägt. (*Nuclear Stress Rule*))

Die Anwendung der Regeln erfolgt in zwei Zyklen und führt zur einer primären Betonung auf dem Wort „Baum“ innerhalb der Nominalphrase, hingegen zur primären Betonung auf „Ein“ innerhalb des Kompositums. Jedesmal, wenn einem Vokal das Merkmal „primärbetont“ zugewiesen wird, werden die anderen betonten Vokale um eine Betonungsebene heruntergestuft:

Kompositum „Einbaum“		
$[_N\#[_N\#EIN\#]_N$	$[_N\#BAUM\#]_N\#]_N$	Zyklus 1
1	1	Regel 1
(1	1	Regeln 2/3 finden keine Anwendung)
$[_N\#\#EIN\#$	$\#BAUM\#\#]_N$	Zyklus 2
(1	1	Regel 1 findet keine Anwendung)
1	2	Regel 2
(1	2	Regel 3 findet keine Anwendung)
Nominalphrase „Ein Baum“		
$[_{NP}\#[_{DET}\#EIN\#]_{DET}$	$[_N\#BAUM\#]_N\#]_{NP}$	Zyklus 1
1	1	Regel 1
(1	1	Regeln 2/3 finden keine Anwendung)
$[_{NP}\#\#EIN\#$	$\#BAUM\#\#]_{NP}$	Zyklus 2
(1	1	Regel 1 findet keine Anwendung)
(1	1	Regel 2 findet keine Anwendung)
2	1	Regel 3

Die Zyklen werden entsprechend komplizierter, je komplexer die phrasalen und wortinternen Klammerstrukturen sind. Es gibt keine prinzipielle Obergrenze an Betonungsstufen.

2.4.3.2 Metrische Phonologie

Die *Metrische Phonologie* ist eine Theorie, welche eigens für die Beschreibung von Betonungsmustern verschiedener Sprachen entwickelt wurde (LIBERMAN und

PRINCE 1977). Die Darstellung der metrischen Phonologie in diesem Abschnitt folgt im wesentlichen den Ausführungen von (HAYES 1995). Im Gegensatz zur Generativen Phonologie wird die Betonung nicht mehr als n -näres phonologisches Merkmal verstanden, sondern ergibt sich aus einer hierarchisch organisierten rhythmischen Struktur. Die Ebenen, aus denen diese Struktur erzeugt wird, sind dabei die unterschiedlichen Ebenen der prosodischen Hierarchie einer Sprache. Auf jeder Ebene kann eine Konstituente eine andere Konstituente rhythmisch dominieren. Diese Dominanz führt jeweils zu einer stärkeren wahrnehmbaren Prominenz dieser Konstituente. Auf der prosodischen Ebene der Intonationsphrase lehnen sich die Konstituenten oft an syntaktische Konstituenten an, Konstituenten unterhalb phonologischer Wörter sind häufig morphologische Konstituenten, noch darunter liegt die Ebene der phonologischen Füße, welche sich wiederum aus Silben zusammensetzen. Ein phonologischer Fuß wird als eine Folge von einer betonten und einer oder mehreren unbetonten Silben betrachtet (vgl. Abschnitt 2.4.1). Füße können rechts- oder linksassoziativ sein, und die Anzahl der in ihnen enthaltenen Silben ist sprachspezifisch. Auch die Zusammensetzung der rhythmisch relevanten anderen Konstituenten ist abhängig von der jeweiligen Sprache. Metrische Betonungsregeln für eine bestimmte Ebene der prosodischen Hierarchie können immer nur für Einheiten der unmittelbar darunterliegenden Ebene der prosodischen Hierarchie formuliert werden (HAYES 1995, 33). Dies bedeutet, daß die phrasale Betonung nicht auf Silbenebene, sondern nur auf Wortebene definiert werden kann. Die hierarchische Konstituentenstruktur, innerhalb deren phonologische Konstituenten ihre Nachbar-konstituenten rhythmisch dominieren, wird mit Hilfe *metrischer Bäume* dargestellt. Die Verzweigungen innerhalb der Bäume unterteilen sich jeweils in einen dominierenden Teil, der mit $s(\text{trong})$ annotiert ist, sowie einen von $s(\text{trong})$ dominierten Teil, welcher mit $w(\text{eak})$ annotiert ist. Metrische Bäume können in *metrische Gitter* überführt werden mit Hilfe folgender Rechnungsvorschrift:

Die Spaltenhöhe und damit die Betonungsstärke einer jeden Silbe berechnet sich aus der Anzahl der von der Silbe dominierten und mit w annotierten Knoten, beginnend beim Schwesterknoten, plus 1.

Metrische Gitter sind eine direkte Abbildung der Silbenprominenzen, d.h., daß die Spaltenhöhe über jeder Silbe die perzeptive Prominenz der Silbe relativ zu den umgebenden Silben darstellt. Gitter sind aber nicht nur alternative Darstellungen der Zahlenwerte, wie sie aus der Generativen Phonologie bekannt sind, da sich jede Reihe auf eine bestimmte Ebene der prosodischen Hierarchie bezieht. Gitter sind außerdem nicht nur eine Reihe von Spalten, sondern sie bilden eine relationale Struktur. Gegeben zwei Gittermarkierungen x_1 und x_2 auf derselben Ebene, so muß eine der beiden folgenden Bedingungen gelten: Entweder ist eine Markierung der anderen untergeordnet, oder aber sie sind beide einer dritten Markierung untergeordnet. Aus dieser Bedingung folgt, daß auf allen Ebenen Prominenzrelationen angegeben werden können und daß es keine zwei gleich prominente hauptbetonte Silben innerhalb einer prosodischen Konstituente geben kann.

Wenn auch Gitter die Dominanzbeziehungen verschiedener Konstituenten der prosodischen Hierarchie nicht darstellen können, sind sie eher geeignet, um sogenannte „Akzentzusammenstöße“⁴ und die sich daraus ergebenden Betonungsverschiebungen („stress shift“) zu modellieren (siehe auch Abschnitte 1 und 5.2.1). Es hat Uneinigkeit dahin gehend gegeben, ob Gitter- oder Baumdarstellungen adäquater für die Modellierung rhythmischer Phänomene sind (HAYES 1983; PRINCE 1983; SELKIRK 1984). Auf diese Problematik soll hier allerdings nicht weiter eingegangen werden, zumal dies in erster Linie ein Darstellungsproblem zu sein scheint (GOLDSMITH 1990, 192), welches durch etwas komplexere Notationsformen aus der Welt geschafft werden kann (z.B. (HAYES 1995, 38f.)). Je nachdem, welche Art der Information — ob reine Prominenzrelation oder rhythmische Dominanzrelation — im Fokus der Darstellung liegt, können aber beide Darstellungsformen sinnvoll sein. Abbildung 2.8 stellt die rhythmische Struktur des Wortes „Worthypothesengraph“ mit Hilfe eines metrischen Baumes sowie eines Gitters dar.

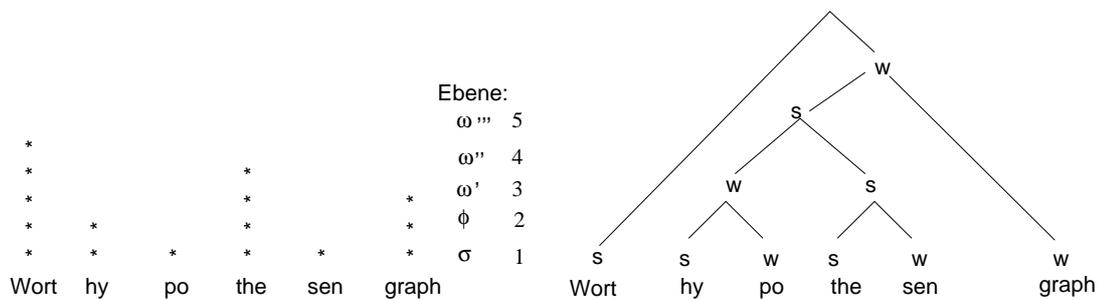


Abbildung 2.8: Die metrische Struktur des Wortes „Worthypothesengraph“ als Baum- und Gitterdarstellung

Meist wird die Metrische Phonologie im Zusammenhang mit der Autosegmentalen Phonologie genannt (GOLDSMITH 1990; LADD 1996). Letztere hat sich aus einer Beschreibung von Tonsprachen (LEBEN 1973) hin zu einer vollständigen Phonologie (GOLDSMITH 1979) entwickelt, die aber traditionell einen starken Fokus auf intonatorische Aspekte hat. Die Idee einer Kopplung von rhythmischer Struktur und Intonation wurde besonders populär durch die Arbeiten von (PIERREHUMBERT 1979), und es gilt mittlerweile als üblich, die Ebenen von sprachlichem Rhythmus und Intonationsstruktur als separate, aber interaktive Ebenen zu betrachten. Die besondere Schwierigkeit einer strikten Trennung beider Ebenen liegt insbesondere darin begründet, daß sie sich in ihren phonetischen Ausprägungen stark überlappen.

⁴Diese Bezeichnung ist irreführend, da es sich nicht notwendigerweise um das Zusammenstoßen zweier phonetischer Akzente handeln muß. Der englische Terminus „stress clash“ trifft die phonologischen Gegebenheiten besser. Da in der deutschen Literatur sich aber die Bezeichnung des „Akzentzusammenstoßes“ eingebürgert hat, wird sie an dieser Stelle ebenfalls verwendet.

2.4.3.3 Optimalitätstheorie

Die Optimalitätstheorie, die in ihren fundamentalen Zügen in (PRINCE und SMOLENSKY 1993) entwickelt wurde, ist nicht generativ geprägt. Vielmehr geht sie davon aus, daß die dem Menschen angeborene *Universalgrammatik* eine Menge universeller *Constraints* enthält, welche aus einer Menge möglicher phonologischer Formen diejenige heraussucht, welche die wenigsten *Constraints* verletzt. Die möglichen phonologischen Formen werden durch eine Generierungsfunktion GEN erzeugt. Aus dieser Menge von Kandidaten wird mit Hilfe der Funktion H-EVAL der optimale Kandidat herausgefiltert, wobei H-EVAL die universalen *Constraints* beinhaltet. Die *Constraints* sind je nach Sprache unterschiedlich hierarchisiert. Der wichtige Punkt der Optimalitätstheorie ist, daß *Constraints* verletzt werden können, sofern dadurch ein *Constraint* beachtet wird, welcher in der Hierarchie an höherer Stelle steht.

Die Optimalitätstheorie bedient sich für die Illustration von miteinander in Konflikt stehenden *Constraints* sogenannter *Tableaus*. Die linke Spalte eines *Tableaus* enthält die verschiedenen Kandidaten, die obere Zeile die verschiedenen *Constraints*, die zu berücksichtigen sind. Die Zellen darunter dienen zur Erläuterung des Evaluationsergebnisses eines jeden Kandidaten hinsichtlich des oberhalb stehenden *Constraints*. Eine leere Zelle bedeutet, daß der *Constraint* erfüllt ist, ein *-Symbol zeigt eine *Constraint*-Verletzung an. Bei skalaren *Constraints* wird der Wert des evaluierten Elements zusätzlich in der Zelle notiert. Ein Ausrufungszeichen (!) signalisiert die ausschlaggebende *Constraint*verletzung für den Ausschluß eines Kandidaten. Ein kleiner Pfeil⁵ weist auf den ausgewählten, also optimalen Kandidaten hin.

Die Vorgehensweise der Optimalitätstheorie soll anhand eines kleinen Beispiels erläutert werden. Erklärt werden soll die Wortbetonung des deutschen Wortes *subtil*. Die relevanten *Constraints* seien folgende:

- WSP: Das sogenannte *weight-to-stress*-Prinzip nach (PRINCE 1990), welches bedeutet, daß die Betonung auf eine schwere Silbe fallen soll
- FOOT-FORM-TROCHAIC: Dieser *Constraint* fordert, daß Füße, also Folgen von betonten und unbetonten Silben, eine trochäische Form haben sollen, also mit einer betonten Silbe beginnen.

Wir nehmen für das Beispiel an, daß im Deutschen der WSP-*Constraint* einen höheren Rang als FOOT-FORM-TROCHAIC trägt, daß also gemäß optimalitätstheoretischer Konvention zu schreiben ist:

WSP » FOOT-FORM-TROCHAIC

⁵In der optimalitätstheoretischen Literatur steht an dieser Stelle eigentlich ein kleines Händchen, aber ein kleiner Pfeil erfüllt m.E. denselben Zweck.

/?Eks.tre:m/	WSP	FOOT-FORM-TROCHAIC
a. /?'Eks.tre:m	*!	
b. → /?Eks.tr'e:m/		*

Tabelle 2.6: Optimalitätstheoretisches Tableau zur Illustration der Vergabe von Wortbetonung

Die erste Spalte in Tabelle 2.6 verdeutlicht die Evaluation des Kandidaten a., des Wortes *extrem* mit der Betonung auf der ersten Silbe. Der Kandidat erfüllt nicht den *WSP-Constraint*, da die zweite Silbe einen Langvokal enthält und außerdem geschlossen ist, so daß sie hinsichtlich ihres Silbengewichts der ersten Silbe überlegen ist (vgl. Abschnitt 5.1.5). Allerdings wird der zweite *Constraint* nicht verletzt. Kandidat b. hingegen verletzt zwar den *FOOT-FORM-TROCHAIC-Constraint*, nicht aber den in der Hierarchie höher stehenden *WSP-Constraint*. Die Verletzung des *WSP* von Kandidat a. ist ausschlaggebend für die Auswahl von b. als optimalem Kandidaten. Würde an dieser Stelle keine Auswahl getroffen werden können, so würden weitere *Constraints*, die in der Hierarchie niedriger angesiedelt sind, die Entscheidung fällen.

2.4.3.4 Unifikationsphonologie

Wie die Optimalitätstheorie ist auch die *Unifikationsphonologie* deklarativ geprägt. Sie versucht demnach, ohne eigenschaftsverändernde Regeln auszukommen. Statt dessen wird einer unterspezifizierten grammatischen Beschreibung unter Beachtung bestimmter *Constraints* Information hinzugefügt. Der unifikationsgrammatische Formalismus wurde ursprünglich für die Beschreibung syntaktischer Strukturen entwickelt und wird in verschiedenen Grammatikformalismen wie der GPSG (GAZDAR et al. 1985) oder HPSG (POLLARD und SAG 1987; POLLARD und SAG 1994) verwendet. Die starre Unterscheidung zwischen Regeln und Repräsentationen wird bei diesen Grammatiktypen aufgelöst, indem linguistische Regeln selbst als partielle Beschreibungen linguistischer Objekte betrachtet werden. Lexikalische Repräsentationen, die nicht vorhersagbar sind, werden mit partiellen Repräsentationen, die vorhersagbare Strukturbeschreibungen sind und somit „Regeln“ entsprechen, vereinigt. Bei dieser Operation müssen bestimmte *Constraints* berücksichtigt werden. Die Vereinigung kann entweder mittels *Konkatenation*, also syntagmatischer Verkettung, oder *Unifikation*, also eher paradigmatischer Anreicherung, erfolgen. Der Vorgang der Regelanwendung ist nicht geordnet, und die Repräsentationen können nicht verschiedenen linguistischen Ebenen (wie z.B. einer Oberflächen- und Tiefenstruktur) zugeordnet werden. Die linguistischen Objekte werden als *Attribut-Wert-Paare* dargestellt. Phonologische Arbeiten im Rahmen des Unifikationsformalismus bestehen z.B. von (SCOBIE 1991; COLEMAN 1992a; HÖHLE 1998). Anwendungsbezogene Arbeiten im Rahmen der Sprachsynthese existieren von (LOCAL 1994; DIRKSEN und COLEMAN 1997; WALTHER und KRÖGER 1994; OGDEN et al. 1999). Eine unifikationsbasierte Modellierung der Betonung in Anlehnung an das Sprachproduktions-

modell von (LEVELT 1989) wurde im SYNPHONICS-Projekt von (ABB und LEBETH 1992; ABELN und GÜNTHER 1993) durchgeführt. Die Modellierung berücksichtigte die Interaktion von Syntax, Semantik und Phonologie. Die phonologische Struktur des SYNPHONICS-Lexikoneintrags nach (ABB et al. 1993, 28) für das Lexem „Hans“ ist in Abbildung 2.9 dargestellt. Hinter den Attributen SYN bzw. SEM verbirgt sich die Spezifikation der syntaktischen bzw. semantischen Werte, die aus Gründen der Übersichtlichkeit an dieser Stelle nicht berücksichtigt wurden.

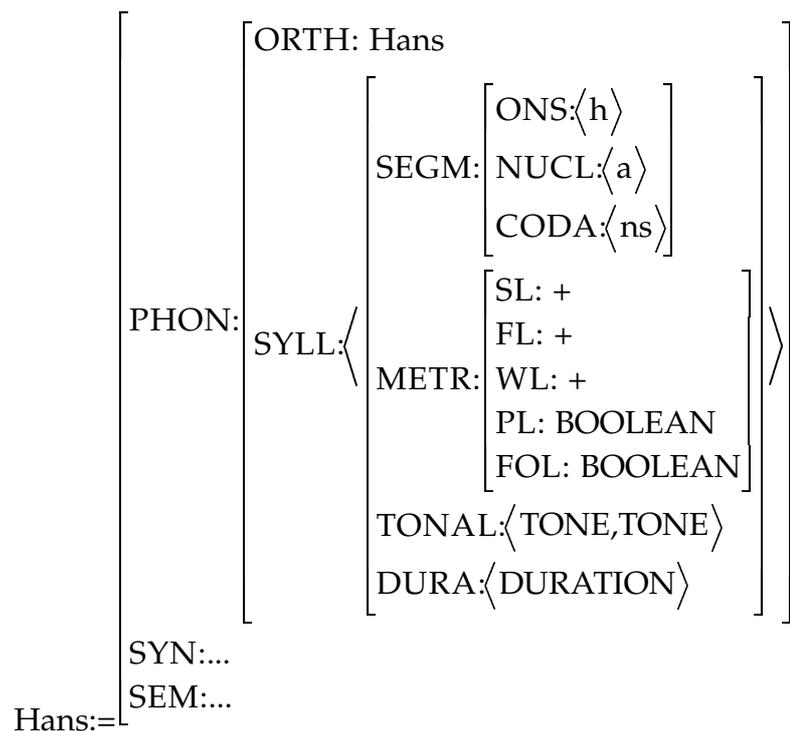


Abbildung 2.9: Die phonologische Struktur des Lexikoneintrags „Hans“ in Attribut-Wert-Notation

Abbildung 2.9 ist folgendermaßen zu lesen: Die phonologische Struktur besteht aus ihrer orthographischen Form und ihrer Silbenstruktur. Die Silbenstruktur ist eine Liste (dargestellt durch spitze Klammern), die im Fall des einsilbigen Wortes Hans lediglich ein Element enthält. Die Silbe ist durch verschiedene Attribute spezifiziert, welche die segmentale (SEGMENTAL), metrische (METRICAL), tonale (TONAL) und Dauerstruktur (DURATION) beschreiben. Die segmentale Struktur wird durch verschiedene Listen von Segmenten dargestellt, welche den Onset, Nukleus und die Silbenkoda enthalten. Die metrische Struktur enthält Attribute, welche auf den verschiedenen Ebenen der prosodischen Hierarchie nun Boolesche Werte tragen. Die Ebenen entsprechen auch den Ebenen eines metrischen Gitters oder Baumes. Ein positiver Wert auf einer Ebene ist identisch mit einem Eintrag auf der entsprechenden Ebene eines metrischen Gitters. Im vorliegenden Falle sind positive Werte für die Silbenebene (SL), Fußebene (FL) sowie Wortebene (WL) vergeben worden, da es

sich um die Repräsentation des Lexems handelt. Die Werte auf Phrasenebene (PL) und Fokusebene (FOL) sind noch nicht spezifiziert. Diese Ebenen werden im Äußerungskontext durch Unifikation mit partiellen Beschreibungen von „Regeln“ angereichert. Eine solche „Regel“ ist laut (ABB et al. 1993, 29) die „Fokusakzent-Regel“ (vgl. Abbildung 2.10). Diese besagt, daß die Silbe eines semantisch fokussierten Wortes mit einem positiven Eintrag auf der WL-Ebene positive Werte auf der phrasalen und fokalen Ebene erhält. Somit erhält diese Silbe die höchste Prominenz innerhalb der Äußerung. Außerdem wird dieses Wort mit einem Tonhöhenakzent H* versehen. Letzterer entspricht der ToBI-Notation nach (PIERREHUMBERT 1980). Die tonale Ebene (TONAL) ist jedoch noch nicht völlig ausspezifiziert, da die fokussierte Silbe u.U. noch Trägerin des Grenztones wird, sofern sie am Ende der Äußerung steht. In diesem Fall muß noch abhängig vom Satzmodus ein steigender oder fallender Grenzton spezifiziert werden. Die Fokusakzent-Regel wurde in (GÜNTHER 1999) für die verschiedenen Fokustypen (vgl. auch Abschnitt 2.4.2) noch genauer formuliert.

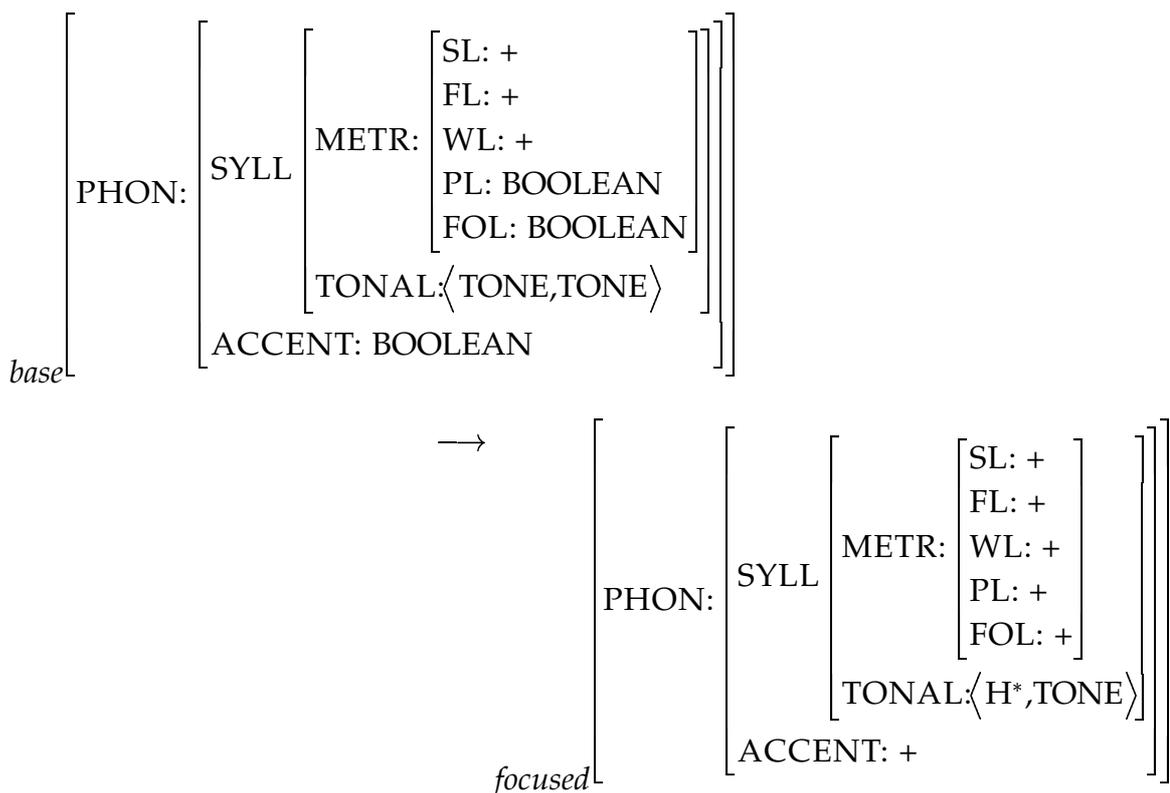


Abbildung 2.10: Die Fokusakzent-Regel nach (ABB et al. 1993)

2.4.4 Formalisierung phonologischer Prominenzmodelle

Phonologische Regelsysteme werden formal üblicherweise als reguläre Grammatiken ausgedrückt und unter Verwendung endlicher Automaten modelliert. Die

Anwendung von endlichen Automaten in der Phonologie geht auf die ältere Entdeckung von (JOHNSON 1972) zurück, daß die kontextsensitiven Regeln der Generativen Phonologie mit Hilfe regulärer Grammatiken des Typs 3 der Chomsky-Hierarchie⁶ beschreibbar sind, sofern darauf verzichtet wird, dieselbe Regel auf ihre Ausgabe anzuwenden. Diese Entdeckung führte, nachdem sie von (KAPLAN und KAY 1994)⁷ wiederaufgegriffen worden war, zu einer rapiden Entwicklung von Beschreibungen phonologischer und morphologischer Regelsysteme als endliche Übergangnetzwerke (z.B. (KOSKENNIEMI 1983)), wobei das Eingabeband in Anlehnung an die klassische Generative Phonologie normalerweise der *phonologischen Tiefenstruktur* oder lexikalischen Ebene, das Ausgabeband der *phonetischen Oberflächenstruktur* entspricht. Wenn die Grammatik die Anwendung mehrerer phonologischer Regeln in einer bestimmten Reihenfolge vorsieht, so kann jede Regel als eigenes endliches Übergangnetzwerk formuliert werden. Das Ausgabeband des ersten Übergangnetzwerkes bildet die Eingabe für das nachfolgende Netzwerk usw. Die intermediären Automaten einer solchen Folge von *kaskadierten* Übergangnetzwerken können durch den Vorgang der *Komposition* in einen einzelnen Automaten überführt werden. Ein solcher Automat besteht demnach lediglich aus einer einzigen phonologischen Tiefenstruktur sowie einer Oberflächenstruktur (siehe Abbildung 2.11). Für einen detaillierten Überblick zu diesem Thema siehe (CARSON-BERNDSEN 1993, 46-60) sowie (KARTTUNEN 1993).

Die Beschreibungen mit Hilfe endlicher Automaten werden als *deklarative* Beschreibung betrachtet, da die früheren generativen Regeln durch die Angabe kontextueller *Constraints* dargestellt werden, die innerhalb des Automaten als Übergänge zwischen Zuständen spezifiziert sind. Neben der Beschreibung linearer Phonologie mit Hilfe endlicher Automaten wurden auch verschiedene Versuche unternommen, endliche Automaten für die Beschreibung *nichtlinearer* oder *multilinear*er phonologischer Strukturbeschreibungen zu nutzen. Wichtige Arbeiten auf diesem Forschungsgebiet sind dabei die von (KAY 1987), (KORNAI 1991), (BIRD und ELLISON 1992; BIRD und ELLISON 1994), (WIEBE 1992) sowie (CARSON-BERNDSEN 1991; CARSON-BERNDSEN 1993). (GIBBON 1987; GIBBON 2001) verwendet außerdem endliche Automaten für die Beschreibung phonologischer Tonstrukturen von westafrikanischen Sprachen.

(BIRD 1992) liefert außerdem einen Vorschlag für die Integration endlicher Automaten in die unifikationsbasierten phonologischen Grammatiken, indem er das PHON-Attribut erweitert, dessen Wertspezifikation in der klassischen HPSG als Liste von Phonemen erfolgt. Bei Bird geschieht die Wertspezifikation mit Hilfe regulärer Ausdrücke bzw. endlicher Automaten.

Meist sind segmentale morphophonologische Phänomene mit Hilfe endlicher Automaten modelliert worden, aber auch die Formulierung von Betonungsregeln ist

⁶Für eine Einführung in die Chomsky-Hierarchie sowie die Automatentheorie siehe (HOPCROFT und ULLMAN 1979)

⁷Das entsprechende Papier ist zwar erst 1994 veröffentlicht worden, war aber bereits seit geraumer Zeit in der Forschungsdiskussion.

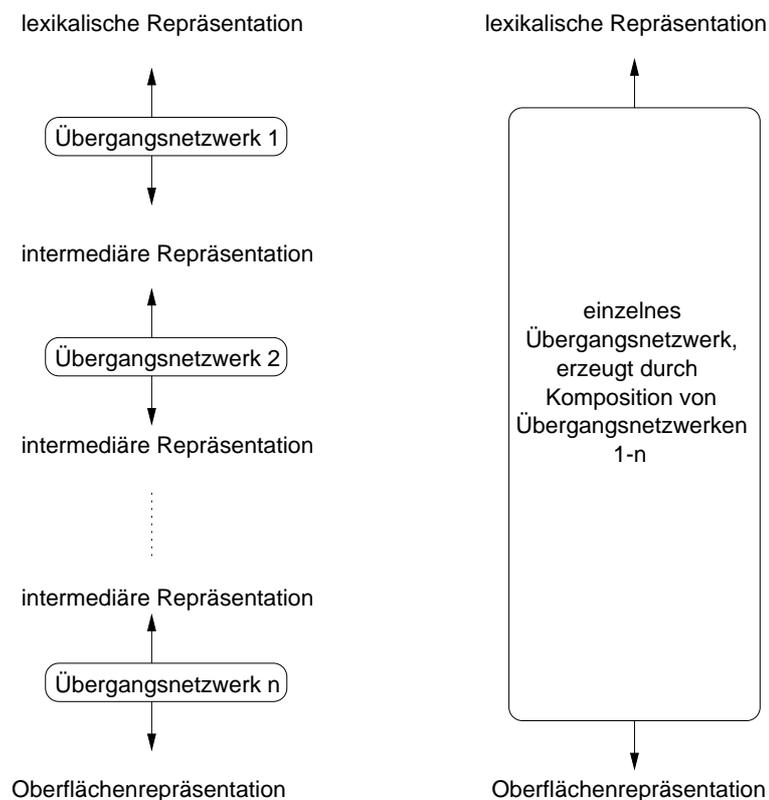


Abbildung 2.11: Die Möglichkeit der Überführung einer Folge von kaskadierten in ein einziges Übergangnetzwerk nach (KARTTUNEN 1993)

ohne weiteres möglich. In Abbildung 2.12 ist ein endliches Übergangnetzwerk dargestellt, welches eine Folge von Silben mit Spaltenhöhen metrischer Gitter einliest, wobei die Spaltenhöhen als Zahlenwerte kodiert sind. Das Ausgabeband enthält die gleiche Silbenfolge mit entsprechend der Regel veränderter metrischer Struktur. Der Automat fügt der letzten Silbe, welche mindestens einen Schlag auf der Gitterebene 3 besitzt, einen weiteren Schlag hinzu. Wenn man annimmt, daß die Ebene 3 im metrischen Gitter der Wortebene entspricht, so stellt dieser Automat eine einfache metrische Formulierung der *Nuclear Stress Rule* (vgl. Abschnitt 2.4.3.1) dar. Auf den Übergängen sind die Betonungsstärken der Silben in der Eingabe und in der Ausgabe abgebildet. Es ist zu beachten, daß dieser Automat nur Äußerungen als wohlgeformt erkennen würde, die mindestens *eine* Betonung auf der Ebene des prosodischen Wortes enthalten.

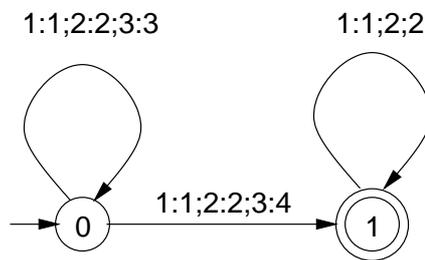


Abbildung 2.12: Vereinfachte Modellierung der *Nuclear Stress Rule* mit Hilfe eines endlichen Übergangnetzwerkes. Die Zahlenwerte stellen die Spaltenhöhe der Einträge in metrischen Gittern dar.

Die formale Definition eines endlichen Übergangnetzwerkes, welcher Operationen auf metrischen Gittern modelliert, lautet wie folgt.

Definition 2

Ein endliches Übergangnetzwerk zur Darstellung metrischer Strukturen (EÜMS) kann dargestellt werden als ein Sechstupel $EMS = \langle Q, \Sigma, \Delta, \delta, q_0, F \rangle$,

und es sei

$Q :=$ eine endliche, nichtleere Menge von Zuständen

$\Sigma :=$ ein endliches Eingabealphabet, bestehend aus Attribut-Wert-Spezifikationen von Silben, z.B. $\langle \text{WEIGHT:HEAVY}; \text{PROMINENCE:3} \rangle$

$\Delta :=$ ein endliches Ausgabealphabet, bestehend aus Attribut-Wert-Spezifikationen von Silben

$\delta :=$ eine Abbildungsfunktion von $Q \times \Sigma$ in endliche Teilmengen von $Q \times \Delta^*$

$q_0 \in Q :=$ ein Startzustand

$F \subseteq Q :=$ eine Menge von Endzuständen

Es existieren ebenfalls Vorschläge zur Darstellung der hierarchisierten *Constraints* der Optimalitätstheorie auf der Basis konkatenierter endlicher Automaten (KARTTUNEN 1998). Auch (ELLISON 1994b) macht einen Vorschlag zur Darstellung von optimalitätstheoretischen *Constraints* auf der Basis endlicher Übergangnetzwerke. Seine Modellierung sieht vor, auf dem Eingabeband den jeweiligen Kandidaten einzulesen, während auf dem Ausgabeband die *Constraint*verletzungen ausgegeben werden. Durch eine anschließende Minimierung der *Constraint*verletzungen des Ausgabebandes könne der optimale Kandidat ermittelt werden. Diese Beschreibung entspricht im wesentlichen der Auswahl eines optimalen Pfades, wie er aus der datenbasierten Sprachsynthesetechnologie (CAMPBELL 1996; BEUTNAGEL und CONKIE 1999; STÖBER et al. 1999) bekannt ist. Hierbei wird ebenfalls aus einer Reihe von für die Synthese zur Verfügung stehenden Kandidaten die optimale Kette von Kandidaten berechnet, indem für jeden Übergang zwischen zwei potentiellen Kandidaten sowie jedem Kandidaten selbst *Kosten* zugewiesen werden. Die Auswahl der optimalen Kandidatenfolge geschieht über die Minimierung der Gesamtkosten. Unklar bei einer Formalisierung der Optimalitätstheorie ist allerdings noch

die Generierung der Kandidatenmenge. Da die Optimalitätstheorie insgesamt keinen generativen Ansatz liefert, baut diese Arbeit nicht auf ihren Ergebnissen auf. In Kapitel 6 wird jedoch ein Vorschlag für eine optimalitätstheoretische Modellierung der Ergebnisse geliefert.

2.5 Prominenz und Paralinguistik

Wie in Abschnitt 2.1 erwähnt, wirken sich die prosodischen Eigenschaften von Äußerungen auf die linguistische, paralinguistische und extralinguistische Dimension aus. Gegenstand dieser Arbeit ist ein Vergleich der linguistischen Dimension von Prominenz mit tatsächlich wahrgenommener Prominenz. Bei der Wahrnehmung von Prominenz spielt es aber keine Rolle, ob die Prominenz auf linguistische oder paralinguistische Äußerungseigenschaften zurückzuführen ist. Wenn diese beiden Ebenen aber nicht klar voneinander abzugrenzen sind, könnte dies zu Problemen führen. In einem Vergleich emphatischer und neutraler englischer Sprache wurde daher untersucht, inwieweit die paralinguistische und linguistische Ausprägung von Prominenz sich überlappen. Um dies festzustellen, wurde, aufbauend auf der Untersuchung zum Verhältnis von Prominenz und Zitierform (vgl. 2.4.2.1), untersucht, ob die linguistisch zu motivierenden Prominenzverhältnisse in paralinguistisch unterschiedlichen Kontexten erhalten bleiben (WAGNER und PORTELE 1999). Als Untersuchungsgegenstand wurde die *Emphase* gewählt, die hier nicht im Sinne von *Kontrastfokus* oder *prosodischem Fokus* zu verstehen ist, sondern als rein paralinguistisches Phänomen (vgl. Seite 7). Wie die phonetische Ausprägung der Emphase aussieht, ist noch unklar, aber intuitiv kann jedes Wort emphatisch produziert werden. Zu untersuchen ist insbesondere, ob es Interaktionen zwischen der emphatisch motivierten und der linguistisch motivierten Prominenz einer sprachlichen Äußerung gibt.

2.5.1 Datenmaterial

Es wurden im wesentlichen die in Abschnitt 2.4.2.1 beschriebenen Annotationen der Amerikanischen Prosodischen Datenbank (ELSNER et al. 1998) verwendet. Die Datenbank wurde ergänzt mit Markierungen [+*emphatic*] für solche Äußerungen, in denen die Sprecherinnen explizit aufgefordert waren, das zu lesende Material mit Emphase zu sprechen. Hierbei wurden die Leser instruiert, in allen Fällen lediglich ein einzelnes Wort innerhalb der Äußerung emphatisch hervorzuheben.

2.5.2 Resultate

Für nicht-emphatische Sprache hatte sich bei einem Vergleich zwischen der neutralen Zitierform (vgl. Seiten 18ff.) und den markierten Fällen kein signifikanter Unterschied bzgl. des Prominenzwertes der fokussierten Silbe gezeigt. Unterschiedliche

Muster hatten lediglich bei Betrachtung der Prominenz vor bzw. nach der fokussierten Silbe isoliert werden können. Es wurde aber in der Untersuchung auch deutlich, dass Sprecherinnen durchaus gewollt eine Silbe stärker betonen können, ohne daß dies einen direkten Effekt auf die Prominenz der umgebenden Silben haben muß. Vergleicht man hingegen die emphatischen mit den neutral gesprochenen Äußerungen, ergibt sich ein anderes Bild. Innerhalb der emphatischen Äußerungen reichte die mittlere Prominenz der Äußerung näher an die Prominenz der (emphatischen) maxprom-Silbe heran (t-test, $p < 0.0001$, vgl. Abbildung 2.13).

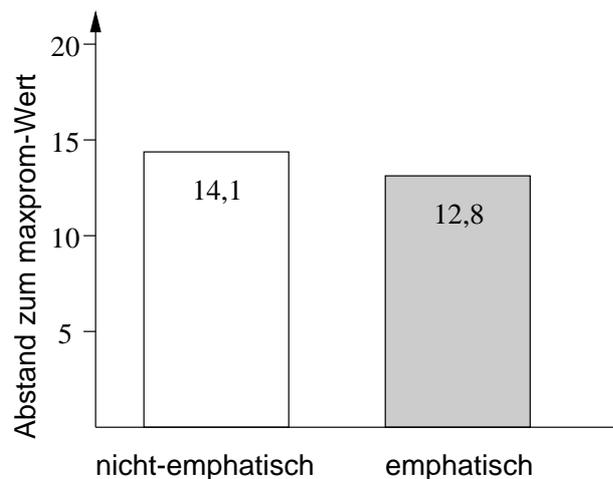


Abbildung 2.13: In emphatischen Äußerungen ist die Prominenz der Äußerung näher an der Prominenz der fokussierten Silbe.

Diese Tendenz war für Silben sowohl vor als auch nach der maxprom-Silbe stabil. Die emphatischen fokussierten Silben waren, verglichen mit den neutralen fokussierten Silben, allerdings nicht signifikant prominenter. Lediglich in emphatischen Äußerungen, die außerdem noch nicht der Zitierform entsprachen, war die maxprom-Silbe prominenter (t-test, $p < 0.0001$). Außerdem zeigte sich in den nicht-neutralen emphatischen Äußerungen noch ein größerer Abstand zwischen der Prominenz der maxprom-Silbe zu den premax-Werten (t-test, $p < 0.0001$). Dieser Effekt spiegelt das Muster wider, welches für nicht-emphatische Äußerungen gefunden wurde (siehe Abschnitt 2.4.2). Es scheint demnach der Fall zu sein, daß vom Sprecher intendierte Emphase zu einer allgemeinen Erhöhung des Prominenzniveaus über die gesamte Äußerung führt. Da die maxprom-Silbe in Äußerungen ohne Zitierform relativ zu ihrer Umgebung prominenter ist als in der Zitierform, führt dies in emphatischer Umgebung zusätzlich zu einer stärkeren Prominenz der fokussierten Silbe.

Emphatische Äußerungen zeichnen sich nicht durch einen einzelnen extrem hohen Prominenzwert aus, obwohl in fast allen Fällen die Sprecherinnen instruiert waren, lediglich ein Einzelwort innerhalb der Äußerung emphatisch auszusprechen. Der linguistisch zu motivierende prominenzrelevante Unterschied zwischen Zitier-

form und Nicht-Zitierform bleibt auch in emphatischen Äußerungen erhalten. Im Fall einer markierten Nicht-Zitierform wird auch die fokussierte Silbe prominenter wahrgenommen als in den neutralen Äußerungen. Dies kann damit zusammenhängen, daß bei einem Nicht-Anheben die linguistisch relevanten Prominenzverhältnisse gefährdet wären. Der Umstand, daß in emphatischen Äußerungen mit Zitierform dieser Effekt nicht zu beobachten ist, könnte damit zusammenhängen, daß es dem Ökonomieprinzip der Sprache zuwiderliefe; solange also die Verständlichkeit nicht leidet und der Eindruck der Emphase durch eine allgemeine Anhebung des Prominenzniveaus gegeben ist, ist eine weitere Anhebung auf der hauptbetonten Silbe demnach nicht notwendig. Sie ist bereits sehr prominent und muß nicht noch prominenter gebildet werden, da der Eindruck der Emphase nicht an ihr allein hängt (vgl. Abbildung 2.14).

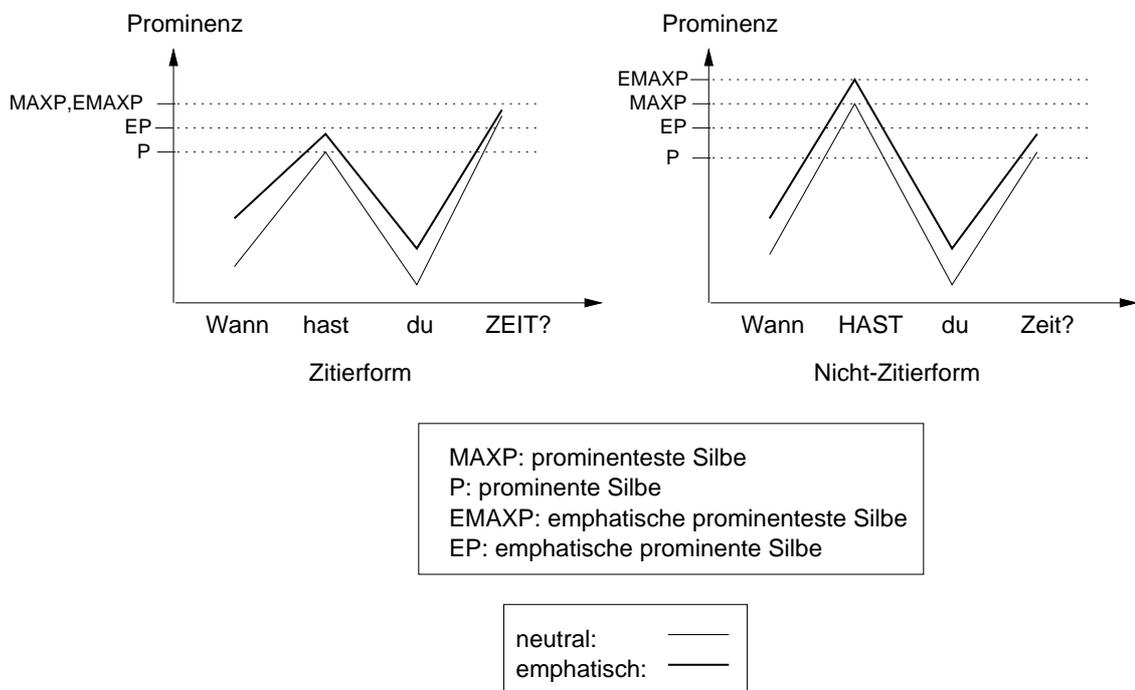


Abbildung 2.14: Prominenzunterschiede zwischen hauptbetonter Silbe und ihrer Umgebung von emphatischen und neutralen Äußerungen in Zitierform und Nicht-Zitierform

2.6 Fazit: Das Prominenzmodell

Prominenz wurde definiert als eine perzeptive Eigenschaft von Silben, die sich aber auch auf höheren Ebenen der prosodischen Hierarchie niederschlägt. Die perzeptive Messung von Prominenz als gradueller, intervallskalierter Wert ist möglich und konnte für mehrere Sprachen nachgewiesen werden. In der linguistischen Repräsentation sind jedoch kategoriale Unterscheidungen von Prominenz relevant. Die-

se spiegeln sich in den relativen, äußerungsinternen Prominenzwerten wider. Eine Äußerung, die der Zitierform entspricht, hat andere Prominenzrelationen als eine, die der Zitierform nicht entspricht. Nicht das absolute Prominenzniveau ist hierbei für den einen oder anderen Eindruck entscheidend, sondern die Prominenzabstände innerhalb der Äußerung. Auch eine Äußerung, die eine Korrektur enthält, trägt ein ganz spezifisches Prominenzmuster, wenn der Eindruck der Korrektur eindeutig aus ihr hervorgehen soll. Für phonologische Unterscheidungen ist also das Wissen um die relativen Prominenzen innerhalb der Äußerung entscheidend. Phonologische Theorien, welche die Betonungsverhältnisse modellieren, versuchen, diese relativen Prominenzverhältnisse darzustellen und vorherzusagen. Metrische Gitter bieten dabei 1:1-Abbildungen der bedeutungsrelevanten Prominenzverhältnisse. Die Regeln der Phonologie können auf der Basis endlicher Übergangnetzwerke formalisiert werden. Eine identische linguistische Struktur kann aber mit unterschiedlichen Sprechstilen einhergehen, welche beispielsweise paralinguistische Informationen über den emotionalen Zustand der Sprecherin oder ihre Einschätzung der Situation verraten. Möchte die Sprecherin ihren Worten besonderen Nachdruck verleihen, so kann sie dies mittels emphatischer Sprache tun. Es konnte gezeigt werden, daß Emphase als Merkmal der gesamten Äußerung zu verstehen ist, da sie sich auf das Prominenzniveau der gesamten Äußerung auswirkt. Eine solche Betrachtung von Emphase als Äußerungseigenschaft deckt sich mit Modellen der automatischen Emotionserkennung (BATLINER et al. 2000) sowie mit Arbeiten aus dem Bereich der Konversationsanalyse (SELTING 1994). Die linguistisch relevanten äußerungsinternen Prominenzrelationen hingegen bleiben auch in emphatischer Sprache erhalten. Der Grad der Emphase, welcher sich auf das Prominenzniveau einer Äußerung auswirkt, muß aber nicht als kategoriales Merkmal verstanden werden. Gerade bei der empirischen Erfassung von Sprechstilcharakteristika ist daher die Verwendung eines intervallskalierten Prominenzmaßes die richtige Entscheidung. Sicherlich ist die Unterscheidung in kategoriale linguistische Parameter und paralinguistische Parameter häufig schwierig. So werden diskursstrukturelevante prosodische *Register* (MAYER 1997) in manchen Modellen u.U. als verschiedene Sprechstile subsumiert. Auch konnte die hier vorgestellte Studie lediglich einen kleinen Teilaspekt von paralinguistischen Phänomenen behandeln. Es ist durchaus denkbar, daß sich andere Sprechstile wesentlich von der Emphase unterscheiden. Antworten, in welchem Ausmaß dies der Fall ist und wieweit das hier vorgestellte integrative Prominenzmodell auf weitere Phänomenbereiche übertragbar ist, muß Gegenstand zukünftiger Forschung bleiben.

Kapitel 3

Methode

Man lerne den Stil aus dem Sprechen und nicht Sprache aus künstlichem Stil.

J.G. von Herder

In Kapitel 1 wurde erklärt, daß ein Ziel dieser Arbeit die Entwicklung einer Evaluationsmethode für bestehende phonologische Modelle sein soll. Um eine Evaluation vorzunehmen, sollen in diesem Kapitel zunächst Gütekriterien definiert werden, anhand deren die Evaluation vorgenommen werden kann. Dies erfolgt in Abschnitt 3.1 in Anlehnung an Arbeiten von Noam Chomsky. Es wird präzisiert, inwieweit sich seine Überlegungen auf prosodische Fragestellungen übertragen lassen und an welchen Stellen Probleme auftreten. In Abschnitt 3.2 wird erläutert, wie die Methode der introspektiven Grammatikalitätsbeurteilung der klassischen Linguistik durch objektivere Verfahren ersetzt werden kann. Zusätzlich wird besprochen, inwieweit sich Möglichkeiten der Formalisierung phonologischer Theorien in die hier vorgeschlagene Methode integrieren lassen. Die anschließenden Abschnitte widmen sich konkret den Anforderungen geeigneten empirischen Datenmaterials für die geplanten Evaluationen von linguistischen Betonungsmustern.

3.1 Was ist eine „gute“ linguistische Theorie?

Laut (CHOMSKY 1957) ist die Bereitstellung von Evaluationskriterien genau das, was eine linguistische Theorie ausmacht:

Given a corpus and given two proposed grammars G_1 and G_2 , the theory must tell us which is the better grammar of the language from which the corpus is drawn. (CHOMSKY 1957, 51)

und ferner

[...] there are three main tasks in the kind of program for linguistic theory that we have suggested:

1. state precisely adequacy criteria
2. characterize the form of grammars in a general and explicit way
3. analyse and define the notion of simplicity that you intend to use in choosing among grammars which are of the proper form.

(CHOMSKY 1957, 53-54)

Eine Grammatik muß also, um vollständig zu sein, auch Maße bereitstellen, die uns erlauben zu sagen, wann sie einer alternativen Grammatik überlegen oder unterlegen ist. Diesen Vorgaben Chomskys soll bei der Festlegung von Gütekriterien für phonologische Grammatiken weitgehend gefolgt werden.

3.1.1 Gütekriterien für phonologische Grammatiken

Hier werden die Gütekriterien für unsere phonologischen Grammatiken festgelegt, um erkennen zu können, welche Defizite in den später betrachteten Grammatiken vorliegen und wie sie verbessert werden können, um eine bessere Vorhersage von Betonung zu ermöglichen. Laut (CHOMSKY 1965, 24-26) sind die relevanten Gütekriterien für Grammatiken zum einen die (a) *deskriptive Adäquatheit*, also der Grad der Übereinstimmung mit den intuitiven Grammatikalitätsurteilen einer Muttersprachlerin, sowie die (b) *explanatorische Adäquatheit*, womit der Grad der psychologischen Erklärbarkeit für diese intuitiven Urteile gemeint ist. Diese Arbeit beschäftigt sich zunächst lediglich mit dem Kriterium der deskriptiven Adäquatheit, da diese eine notwendige (wenn auch nicht hinreichende) Voraussetzung für die Entwicklung einer explanatorisch adäquaten Grammatik ist.

Im Bereich der deskriptiven Adäquatheit ist zu fragen, ob es für sie eine Obergrenze oder *Baseline* gibt, die überhaupt erreichbar ist, da sich Menschen sowohl hinsichtlich ihrer Grammatikalitätsurteile als auch ihres Sprachgebrauchs voneinander unterscheiden. In diesem Zusammenhang ist weiterhin zu erwähnen, daß im Bereich der Prosodieforschung die Erhebung introspektiver Grammatikalitätsurteile sehr problematisch ist. Der Umstand, daß introspektive Prosodiebeurteilungen stark voneinander abweichen können und daß die Intuitionen oft nur am Rande mit tatsächlich produzierten Mustern zu tun haben, wurde bereits in Kapitel 1 erwähnt. Daher bleibt nichts weiter übrig als für einen „Ersatz“ für introspektiv gewonnene Daten zu sorgen. Dieser findet sich in den über mehrere Hörer gemittelten perceptiven Beurteilungen von tatsächlich geäußerten Betonungsmustern. Die Entscheidung für einen Evaluationsansatz, basierend auf Performanzdaten, ist gerechtfertigt, da viele Sprecherinnen keinen reliablen intuitiven Zugang zur Grammatikalität von Betonungsmustern haben. Die Erhebung von Perzeptionsdaten wird ferner unterstützt durch die Tatsache, daß Hörer sich sehr einig sind, wie eine Sprecherin etwas *betont hat*, obwohl sie möglicherweise sehr unterschiedliche intuitive Auffassungen davon haben können, wie eine Äußerung zu *betonen ist* (vgl. Abschnitt 3.3). Die Problematik von intuitiven Grammatikalitätsurteilen hat zu einer erhöhten Popularität von empirisch geprägten Forschungsprogrammen geführt, welche

unter den Namen *Experimentalphonologie* oder *Laboratory Phonology* bekannt geworden sind (vgl. Abschnitt 3.2). Diese Forschungsrichtung zielt auf die experimentelle Falsifikation von phonologischen Vorhersagen und ist stark von modernen wissenschaftsphilosophischen Theorien (POPPER 1934) geprägt. Die klassische Evaluationsmethode Chomskys wird daher in dieser Arbeit durch die Überlegungen der *Laboratory Phonology* ergänzt (vgl. Abbildung 3.1).

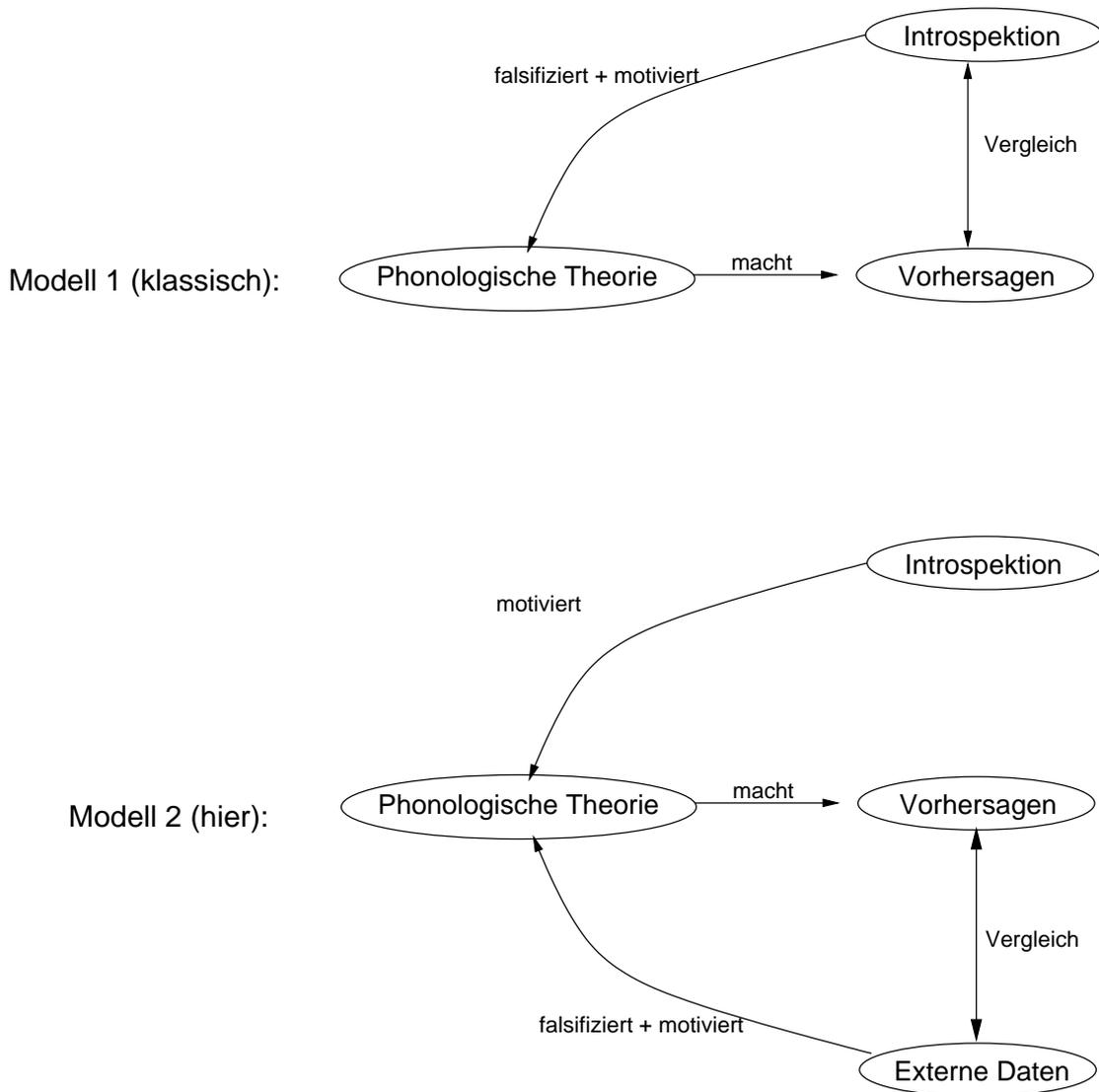


Abbildung 3.1: Die hier verwandte Methode im Vergleich zur traditionell linguistischen Herangehensweise

3.1.2 Formalität als notwendiger Bestandteil einer Grammatik

Ein weiterer wichtiger Punkt bei der Evaluation von Grammatiken ist laut Chomsky die präzise Formulierung derselben (vgl. Seite 47). Hier wird bereits die Forderung an eine formale Grammatikbeschreibung deutlich, welche sich in der computerlinguistischen Forschung als Paradigma weitgehend durchgesetzt hat. Eine formale Darstellung ist unverzichtbar, wenn Inkonsistenzen und Redundanzen erkannt werden sollen. In der Phonologie haben sich formale Beschreibungsmodelle noch nicht annähernd so stark durchgesetzt wie in den Bereichen der Syntax oder Semantik. An dieser Stelle muß daher viel nachgearbeitet werden, um eine Evaluation der bestehenden Grammatiken zu ermöglichen (vgl. Kapitel 4 und 5). Die Formalisierung soll sich an der in Abschnitt 2.4 entwickelten Formalisierung von Betonungs*constraints* anlehnen, die ja im Einklang mit neueren Trends der Phonologie steht. Die Forderung einer Formalisierung bildet einen ganz wesentlichen Bestandteil der hier vorgeschlagenen Evaluationsmethode (vgl. Abbildung 3.1).

3.1.3 Einfachheitskriterien

Als letzten Punkt fordert Chomsky die Definition eines Einfachheitskriteriums („*notion of simplicity*“), mit dessen Hilfe die bessere von zwei Grammatiken mit der gleichen Vorhersagegüte ausgewählt werden kann. Ein wichtiges Einfachheitskriterium für phonologische Grammatiken wird in (CHOMSKY und HALLE 1968, ix) definiert. Demnach ist jede Grammatik vorzuziehen, welche die Anzahl der aufzuziehenden Ausnahmefälle durch die Formulierung von generellen Regeln minimieren kann.

Ein weiteres klassisches Einfachheitskriterium, welches in dieser Arbeit in der Evaluation Berücksichtigung finden soll, geht auf den mittelalterlichen Scholastiker William von Ockham (1285-1347/49) zurück und ist als *Ockhams Rasiermesser* bekannt geworden. Es lautet

non sunt multiplicanda entia praeter necessitatem,

und wird i.allg. als das Gebot verstanden, daß eine Theorie keinerlei redundante Regeln enthalten soll.

3.1.4 Gütekriterien der Evaluationsmethode

Bei der Evaluation von Grammatiken sollen demnach die folgenden Gütekriterien beachtet werden:

1. Die Grammatik muß maximale deskriptive Adäquatheit anstreben, wobei auf introspektive Grammatikalitätsurteile zugunsten von objektiv erhobenem Datenmaterial verzichtet werden soll. Die Grammatik mit einer besseren deskriptiven Adäquatheit ist einer anderen vorzuziehen.

2. Die Grammatik soll keine redundanten Regeln enthalten, aber alle Regeln, welche die Anzahl der Ausnahmefälle minimieren, zulassen.
3. Um Widersprüchlichkeiten und Redundanzen erkennen zu können sowie die Vorhersagen auf großen Datensätzen zu ermöglichen, muß die Grammatik formal notiert werden und implementierbar sein.

3.2 Das Empirie-Formalismus-Gefüge in der Phonologie

Neuere, empirisch geprägte phonologische Forschungsprogramme bieten einen geeigneten methodischen Rahmen, um den o.a. Gütekriterien gerecht zu werden. Eines der neuen Forschungsprogramme ist die sogenannte *Experimental Phonology*. Gemäß (OHALA und JAEGER 1986, 4) sollte eine Experimentalphonologie ein Zusammenspiel von formalen Theorien sein, deren Vorhersagen experimentell überprüft werden können. Dabei sei der Anspruch an die Formalisierung einer Theorie insbesondere dadurch motiviert, daß sie das Erfassen von Vorhersagen erleichtere. In neueren formal-linguistischen Theorien wird das Zusammenspiel von Formalismus und Empirie ähnlich betrachtet. Hierbei wird innerhalb des Formalismus-Empirie-Gefüges dem Formalismus ein maßgebliches Gewicht zugebilligt. Die Argumentation läuft darauf hinaus, daß nur mit Hilfe einer formalen Theorie eine genaue Vorhersagenbildung möglich ist. Ohne diese Vorhersagen seien aber auf der Basis der Empirie lediglich rein deskriptive Theorien möglich, die einen eingeschränkten Wert besäßen (COLEMAN 1998).

Aus ähnlichen Überlegungen anderer Forschergruppen entwickelte sich das Forschungsprogramm der *Laboratory Phonology* (BECKMAN und KINGSTON 1990). Innerhalb dieses Forschungsprogramms wird die Frage der Abbildung zwischen diskreten symbolischen Einheiten der phonologischen Repräsentation und den phonetischen Repräsentationen angegangen. Ferner wird nach möglichst passenden Modellen in beiden Domänen geforscht, um aus verschiedenen Perspektiven die Muster gesprochener Sprache möglichst genau beschreiben zu können (BECKMAN und KINGSTON 1990, 1). Man kann sagen, daß sich die Überlegungen der *Laboratory Phonology* und der *Experimental Phonology* gegenseitig ergänzen, da ein Zusammenspiel von theoretischen Vorhersagen und empirischen Überprüfungen der Vorhersagen schnell die Frage nach der Abbildung zwischen beobachteten phonetischen Strukturen und symbolischen phonologischen Strukturen nach sich zieht.

Da eine Messung von phonologischen Lautstrukturen eigentlich immer nur über den „Umweg“ der Phonetik möglich ist, gehen Forscher daher auch dazu über, direkt phonetische Korrelate von linguistischen Entitäten wie *Fokus* oder *Satzmodus* zu formulieren. Auf eine Beschreibungsebene der Phonologie soll in dieser Arbeit allerdings nicht verzichtet werden, zumal dabei die Gefahr besteht, Generalisierungen über phonetische Entitäten zu postulieren, die mit dem beschriebenen gram-

matischen Phänomen lediglich in zufälligem Zusammenhang stehen: Eine Pause im Sprachsignal kann z.B. nicht von sich aus als syntaktische Phrasengrenze interpretiert werden, auch wenn beide Phänomene häufig simultan auftreten, denn Pausen sind auch anderen Orts in gesprochener Sprache anzutreffen. Eine *Interpretation* der Pause als prosodische Phrasengrenze muß auf der Ebene der Phonologie geschehen. Die Analyse der Beziehung zu anderen linguistischen Ebenen (wie die Beziehung zur syntaktischen Phrasengrenze) ist ein weiterer Schritt.

The relevance of a phonological level of description — an abstraction that mediates between meaningful units and acoustic/articulatory parameters — is taken for granted in any useful interpretation of instrumental findings. (LADD 1992)

Auf eine phonologische Beschreibung sollte also nicht verzichtet werden, um die Realisation von linguistischen Phänomenen, z.B. des prosodischen Fokus oder der Wortbetonung, zu untersuchen. Eine Klärung ihres Verhältnisses zu phonetischen Größen ist allerdings unumgänglich, um empirisch auf festem Boden zu stehen.

3.3 Die empirische Validierung von Betonungsvorhersagen

Wenn phonologische Theorien, welche sich mit Betonung beschäftigen, nicht mehr rein introspektiv evaluiert werden sollen, so muß geklärt werden, wie metrisch-phonologische Beschreibungen zu phonetischen Messungen in Beziehung gesetzt werden können. Das Hauptproblem empirischer Validierungen von Theorien metrischer Phonologie war bisher das Fehlen eines empirischen Maßes bzw. eines Annotationsschemas für phonetisches Datenmaterial. In existierenden Datenbanken wird meist lediglich zwischen Haupt- und Nebenbetonung im Wort unterschieden (z.B. im CELEX-Korpus (CELEX 1995)). Bestehende empirische Untersuchungen an Korpora beschränken sich daher zumeist auf solche prosodischen Ebenen, deren Beschreibung noch mit wenigen Betonungsstufen auskommt. Die relativen Betonungsstufen auf Satzebene werden jedoch mit einer Unterscheidung in Haupt- und Nebenbetonung nicht erfaßt. Das Problem fehlender Annotationen hängt mit den abstrakt definierten Einheiten metrischer Repräsentationen zusammen (vgl. 2.4). Trotz dieser Probleme sind eine Reihe von Untersuchungen zur Überprüfung metrischer Phonologie im Rahmen des Englischen durchgeführt worden (GRABE und WARREN 1995). Weitere Arbeiten, ebenfalls zum Englischen, existieren von (VOGEL et al. 1995; BECKMAN und EDWARDS 1990). Diese Arbeiten sind allerdings auch kritisiert worden, da das untersuchte phonetische Material in keiner Weise phonologisch interpretiert worden ist, d.h., es ist nicht klar, ob die Untersuchungen tatsächlich die gesuchten Muster perceptiv erfassen (SHATTUCK-HUFNAGEL 1995). Unklarheit herrscht insbesondere darüber, zu welchen akustischen Korrelaten die

metrischen Einheiten in Beziehung zu setzen sind. Dies ist in mehrererlei Hinsicht problematisch. Soll beispielsweise in einem Produktionsexperiment untersucht werden, ob bestimmte metrisch-phonologische Vorhersagen auch eintreten, so muß ohne ein entsprechendes *perzeptives* Maß eine *akustische* Größe (zum Beispiel das Vorhandensein eines Akzents) als Evidenz herangezogen werden. Dies geschieht ohne Wissen darüber, ob diese Größe im gegebenen Kontext tatsächlich das Korrelat zum untersuchten phonologischen Phänomen darstellt. Diesen Problemen konnte nicht begegnet werden, solange keine reliablen Etikettierverfahren für metrische Einheiten vorhanden waren.

Das in Abschnitt 2.3 vorgestellte Annotationsverfahren von (FANT und KRUCKENBERG 1989) schließt genau diese Lücke. Im Gegensatz zu einer binären Auffassung von perzeptiver Prominenz (in etwa analog zu der Verwendung *akzentuiert* vs. *nicht-akzentuiert*) betrachten sie Prominenz als ein graduell wahrnehmbares Phänomen. Die Tatsache, daß die Prominenzannotationen sich ebenfalls eignen, um linguistisch relevante Betonungsunterschiede zu beschreiben, wurde in Abschnitt 2.4 anhand experimenteller Erhebungen nachgewiesen. Die Gründe für eine Anwendbarkeit dieses Etikettierverfahrens für das Ziel einer Evaluation phonologischer Vorhersagen liegen somit auf der Hand:

The prominence scaling fulfills the need of a descriptive level interfacing prosodic phonological categories and acoustic correlates. (FANT et al. 1998b)

Daß insbesondere die Vorhersagen der metrischen Phonologie anhand des etikettierten Materials untersucht werden können, sollte aufgrund folgender Ähnlichkeiten zwischen perzeptiven Prominenzetiketten und metrischen Gitterstrukturen klar sein:

- Beide drücken im Äußerungsgefüge die relative Prominenz der Silben zueinander aus.
- Die Strukturen der metrischen Phonologie sind ordinalskaliert. Aufgrund der Art und Weise der Annotation auf einer offenen Skala können phonetische Prominenzmuster auch (z.B. für die Untersuchung paralinguistischer Phänomene) als intervallskaliert betrachtet werden.

Es existieren selbstverständlich ebenfalls Unterschiede zwischen beiden Größen:

- Prominenzen sind perzeptive Meßwerte, die einer produzierten Äußerung zugeordnet wurden. Somit enthalten sie auch eine Bewertung von Betonungsmustern, die aufgrund von paralinguistischen Einflüssen in dieser Form produziert wurden.
- Metrische Strukturen sind Vorhersagen von Betonungsverhältnissen, die lediglich auf der Basis linguistischer Strukturen erfolgen.

- Die etikettierten Prominenzstrukturen haben wohldefinierte akustische Korrelate.
- Die metrischen Gitter haben wohldefinierte Schnittstellen zu linguistischen Ebenen wie Wort und Satz, Fokus, syntaktische Struktur etc.

Einige dieser Unterschiede, wie die Bezüge zu linguistischen Kategorien auf der einen Seite, die Bezüge zu akustischen Größen auf der anderen Seite, sind natürlich wünschenswert. Es bleiben aber einige Probleme für einen Vergleich: Zum einen bleibt die Frage offen, wie die verschiedenen Skalen (0–31 innerhalb der Etikettierung; ca. 0–5 bei metrischen Modellen) zueinander in Beziehung gesetzt werden können. Ein störender Einfluß von paralinguistischen Phänomenen auf die Prominenzverhältnisse linguistisch relevanter Muster konnte jedoch durch entsprechende Voruntersuchungen ausgeschlossen werden (vgl. Abschnitt 2.5). Es wird daher davon ausgegangen, daß eine Evaluation von Vorhersagen metrischer Phonologie anhand perzeptiv erhobener Prominenzmuster eine verlässliche Methode darstellt.

3.4 Formale Vorhersagen durch Implementierung

Um die Vorhersagen einer Theorie sowie ihre Konsistenz zu überprüfen, eignen sich Implementierungen der Regel- bzw. *Constraints*sätze. Die ungenügende empirische Datenlage hat allerdings ebenfalls die Weiterentwicklung und Präzisierung formaler phonologischer Modelle gehemmt. So wird beispielsweise eine vollständige metrische Beschreibung der wortinternen Prominenzunterschiede innerhalb des Modells von (GÜNTHER 1999, 194) abgelehnt:

Ebenso wird in der hier vorliegenden Modellierung des prosodischen Merkmals Akzent bei Kompositabildung keine zusätzliche Akzentabstufung vorgenommen. Für eine korrekte Modellierung, die sowohl die phonologische Spezifizierung als auch die phonetische Realisierung berücksichtigt, kann auf keine gesicherte Datengrundlage zurückgegriffen werden.

An dieser Stelle wird einmal mehr das wichtige Zusammenspiel von phonologischer Theorienbildung und empirischer Grundlage deutlich, welche im vorgestellten methodologischen Ansatz verfolgt wird.

Die computerlinguistischen Modellierungen phonologischer Regeln und *Constraints* gehen unterschiedliche Wege. Das Aufzeigen der *Lernbarkeit* von Regeln oder *Constraints* ist ein wichtiges Anliegen von phonologischen Theorien, die kognitive Adäquatheit für sich in Anspruch nehmen. Die selbsternannte Nähe der Optimalitätstheorie zum Konnektionismus (PRINCE und SMOLENSKY 1993, 200ff.) legt z.B. Lernverfahren auf der Basis neuronaler Netze zur Modellierung optimalitätstheoretischer *Constraints* nahe. Um die Lernbarkeit von Betonungsregeln zu zeigen, wurden maschinelle Lernverfahren eingesetzt (DRESHER und KAYE 1990) und führten

zu einer Kritik des *Prinzipien- und Parameteransatzes* („principles and parameters approach“) (DAELEMANS et al. 1994). Diese Theorie, welche Spracherwerb als Setzen sprachspezifischer Parameter in einem System von Universalprinzipien betrachtet, hat Probleme mit der Erklärung von Irregularitäten in Betonungssystemen, welche mittels maschineller Lernverfahren besser beschrieben werden konnten.

Als computerlinguistische Methode für die Darstellung generativer Phonologien haben sich Formalisierungen unter Zuhilfenahme *Endlicher Automaten* durchgesetzt (vgl. Abschnitt 2.4.4). Die Möglichkeit der Modellierung von Betonungsregeln innerhalb dieser Herangehensweise wurde bereits in Abschnitt 2.4 verdeutlicht und wird die Basis für die formale Beschreibung der Vorhersagen in den folgenden Kapiteln bilden.

3.5 Datenmaterial: Die Bonner Prosodische Datenbank

Das in den letzten Abschnitten entwickelte Evaluationsmodell fußt auf perzeptiv annotiertem Datenmaterial. Die in den folgenden Kapiteln durchgeführte Untersuchung erfolgt auf der Basis der Prominenzannotationen in der *B(onner) P(rosodischen) D(atenbank)* (HEUFT et al. 1995). Basierend auf der Methode von (FANT und KRUCKENBERG 1989) wurde jede Silbe in einer prosodischen Datenbank geleseener Sprache hinsichtlich ihrer perzeptiven Prominenz etikettiert. Die Datenbank enthält 227 Einzelsätze und drei kurze Geschichten, welche jeweils von drei Sprecherinnen (2 weibliche, 1 männlicher Sprecher) gelesen wurden. Insgesamt besteht die Datenbank aus 10661 Silben. Die wahrgenommene Prominenz einer jeden Silbe wurde von drei Hörern ohne vorgegebene Skalierung gemäß dem eigenen Empfinden visuell beurteilt. Die Urteile wurden anschließend in Werte auf einer Skala zwischen 0 und 31 überführt. Es konnten hohe Korrelationen bei den Prominenzbeurteilungen zwischen den verschiedenen Hörern nachgewiesen werden (vgl. Seite 12). Um sprecherspezifische Effekte bei den Beurteilungen zu minimieren, beziehen sich alle weiteren Analysen auf den Median der Prominenzurteile aller Hörer. Ferner enthält die Datenbank Annotationen bzgl. der Position und Stärke prosodischer Phrasengrenzen, der Wortklasse, einer Unterscheidung zwischen Lehnwörtern und nativen Wörtern, der Wortbetonung (basierend auf der prominentesten Silbe eines jeden Wortes), Morphemgrenzen sowie einer Reihe von weiteren segmentalen und suprasegmentalen phonetischen Annotationen.

3.6 Vergleichbarkeit wahrgenommener und vorhergesagter Muster

Bevor ein Vergleich zwischen vorhergesagten (phonologischen) und wahrgenommenen (phonetischen) Mustern vorgenommen wird, muß die Vergleichbarkeit beider Größen sichergestellt werden. Hierzu muß nochmals festgehalten werden, daß

die ursprünglichen Prominenzgrößen als intervallskaliert betrachtet werden können, da sie auf einer freien Skala eingetragen worden sind. Jede Intervallskala kann in eine niedrigere Ordinalskala überführt werden. Ebenso kann jedes metrische Gitter als ordinalskaliert interpretiert werden, da für jede Silbe innerhalb eines metrischen Gitters entschieden werden kann, ob sie in Relation zu ihrer Umgebung stärker oder schwächer ist. Für die Möglichkeit eines Vergleichs spricht zudem die allgemeine Annahme, daß metrische Gitter modellhafte Darstellungen perzeptiver Prominenzmuster sind (vgl. 2.4). Daher sollte die Hypothese eines starken Zusammenhangs zwischen vorhergesagten und wahrgenommenen Prominenzrängen sich genau dann bewahrheiten, wenn die phonologischen Regeln zutreffend sind. Ist hingegen dieser Zusammenhang nicht herstellbar, so müssen die getesteten Regeln als falsifiziert betrachtet werden.

3.7 Definition einer *Baseline*

3.7.1 Motivation

Es ist nicht sinnvoll davon auszugehen, daß eine metrisch-phonologische Grammatik Vorhersagen zu machen in der Lage ist, welche für jede potentielle Sprecherin und jede Äußerung in jeder Dialogsituation zutreffend sind. Vielmehr müssen Variationen berücksichtigt werden. Diese Variationen können teils dialektalen oder sprecherspezifischen Ursprungs sein. So scheint es teilweise systematische Abweichungen zwischen dem Hochdeutschen und dem österreichischen Dialekt zu geben.¹ Eine Studie zur Betonung von Komposita bestätigt diese systematischen Abweichungen (DOLESCHAL 1988). Des weiteren gibt es noch sprecherspezifische Variationen, die schwer dialektalen Gebieten zuzuordnen sind (*Kómpost* vs. *Kompóst*, *Fáchhochschule* vs. *Fachhóchschule*). So fand (BENWARE 1987) in einer Studie zu dreigliedrigen Komposita teilweise hohe Abweichungen zwischen Sprecherinnen, welche nicht dialektal motiviert werden konnten. Weitere Variationen können mit unterschiedlichen semantischen Interpretationen verschiedener Sprecherinnen zusammenhängen, welche zu unterschiedlichen Fokussierungen führen können. Auch die Einbettung einer Äußerung in den situativen Kontext kann u.U. zu einer stärkeren Betonung von neuen vs. kontextuell gegebenen (u.a. (CHAFE 1976)) Äußerungsteilen führen. Bei letzteren ist jedoch anzumerken, daß die prosodische Markierung eines solchen pragmatischen Fokus zumindest gemäß einschlägiger semantisch-phonologischer Theorien (u.a. (PIERREHUMBERT und HIRSCHBERG 1990)) vorher-sagbar ist. Wenn dem so ist und ein prosodisch markierter Fokus durch den Kontext vorherzusagen ist, dann sollte eine phonologische Theorie diese Variation ebenfalls mitberücksichtigen können. Es ist durchaus denkbar, daß die meiste Variation zwischen verschiedenen Sprecherinnen durch unterschiedliche Fokussierungsstrategien hervorgerufen wird. Es ist eine interessante Frage, wieviel Variation die Spre-

¹Vgl. hochdeutsch: *Káffee*, *Sákkó* vs. österreichisch: *Kaffée*, *Sakkó*.

cherinnen hinsichtlich unterschiedlicher Fokussierungen in gleichen Umgebungen treffen. Sollte ein Großteil der Betonungsvariation auf systematische unterschiedliche Fokussierungsstrategien zurückführbar sein, müssen diese im Modell berücksichtigt werden. Es wird in den folgenden Abschnitten zunächst die *Baseline* auf der Basis der Sprecherinnenvariationen berechnet. Anschließend wird überprüft, wie groß der Anteil an möglicherweise unterschiedlichen Fokussierungen war, der zu den Variationen geführt hat.

3.7.2 Baselineberechnung

Zunächst soll als Baseline der von der Theorie erwartbaren Vorhersagegüte die Korrelation der Sprecherinnen untereinander definiert werden. Grundlage der Berechnungen waren alle Äußerungen, in denen die Silbenanzahl pro Wort identisch und der segmentale Gehalt vergleichbar waren.

Innerhalb der Bonner Prosodischen Datenbank ist die durchschnittliche Übereinstimmung der drei Sprecherinnen hoch und liegt bei einem Wert von (Spearman-Rho, $\rho = 0.78$, $p < 0.001$). Die Korrelationen zwischen den verschiedenen Sprecherinnen rangierte zwischen $\rho = 0.72$ and $\rho = 0.82$. Abbildung 3.2 illustriert die verschiedenen Prominenzurteile für die drei Sprecherinnen für die gleiche Äußerung.

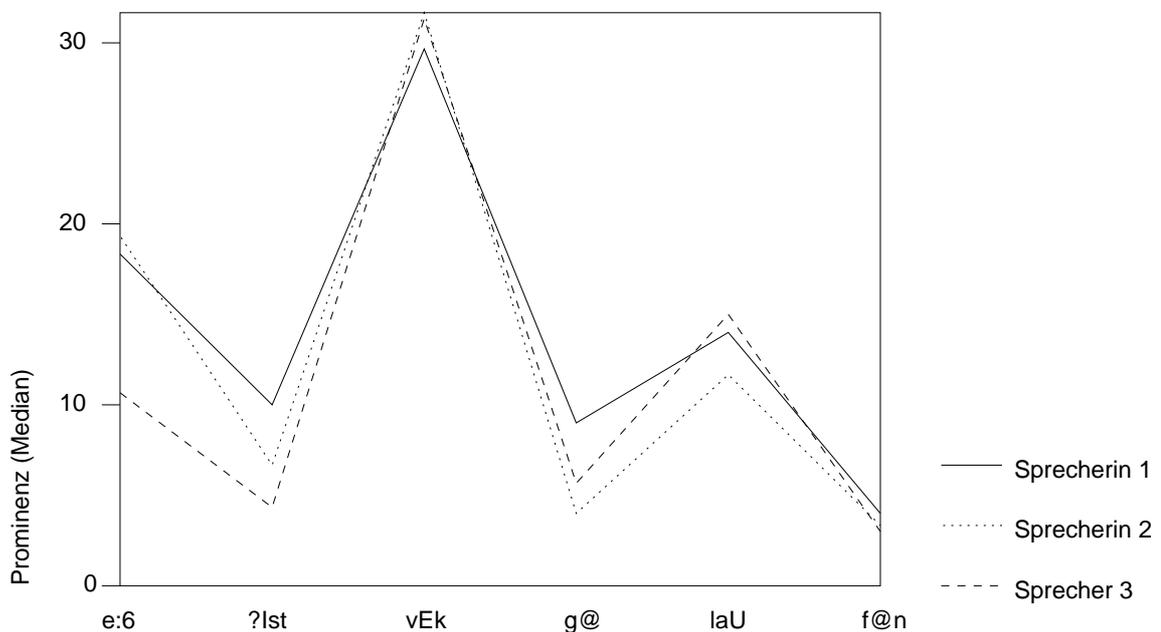


Abbildung 3.2: Prominenzurteile für die Äußerung „Er ist weggelaufen“ von drei verschiedenen Sprecherinnen

3.7.3 Sprecherinnenvariationen durch unterschiedliche Fokussierung

Da es möglich ist, daß die Sprecherinnenvariation in großem Maße durch verschiedene Fokussierungen hervorgerufen wird, wurden alle Äußerungen untersucht, innerhalb derer die Sprecherinnen keine sehr hohe Korrelation untereinander haben (hier: alle Äußerungen mit mittlerer Korrelation von $\rho < 0.8$). Der Anteil dieser Äußerungen an der Gesamtzahl der Äußerungen lag bei 19,5%. Eine geringe Korrelation kann nur bei 3,8% der Äußerungen eindeutig auf eine unterschiedliche Fokussierung der Sprecherinnen bei gleichem gegebenen Kontext zurückgeführt werden. Alle anderen kommen zustande durch

- abweichende rhythmische Muster bei Aufzählungen, u.a. von Telefonnummern (vgl. zu diesem Problem auch (BAUMANN und TROUVAIN 2001)),
- durch abweichende Wortbetonungen (z.B. warúm vs. wárum),
- unterschiedliche prosodische Phrasierungen („Ach so|die||“ vs. „Ach so die||“),
- unterschiedlich starke Betonung von Wörtern, ohne daß die Fokussierung davon betroffen wäre.

Dies zeigt, daß jedenfalls in gelesener Sprache unterschiedliche Fokussierungen der Sprecherinnen nicht die Hauptursache für Sprecherinnenvariationen hinsichtlich der Prominenzmuster darstellen, sondern lediglich *eine* Möglichkeit von vielen Variationsmöglichkeiten innerhalb der Sprachproduktion. Daher erhalten die Fokussierungen in dieser Arbeit keinerlei Sonderbehandlung, sondern werden lediglich als ein Aspekt der Variation von Sprecherinnen gesehen. Auffällig ist, daß der Datenbanksprecher, der über eine Schauspielausbildung verfügt, häufiger hinsichtlich der produzierten prosodischen Fokusstrukturen von den anderen Sprecherinnen abweicht (vgl. Abbildungen 3.3 und 3.4). Die Ausbildung läßt vermuten, daß sich dieser Sprecher stärker mit einer möglichen unterschiedlichen Interpretation des Textes sowie den entsprechenden Betonungen beschäftigt hat. Die inhaltliche Auseinandersetzung mit einem Text kann demnach sicherlich zu verschiedenen Betonungen führen, sie scheint allerdings in gelesener Sprache nicht die Hauptursache für abweichende prosodische Realisierungen darzustellen. Die Tatsache, daß die beiden schauspielerischen Laien (von denen eine professionelle Sprecherin ist) jedoch stark miteinander übereinstimmen, ist ein weiteres Argument für das Vorhandensein einer Zitierform, welche in Abschnitt 2.4.2.1 bereits empirisch gestützt werden konnte.

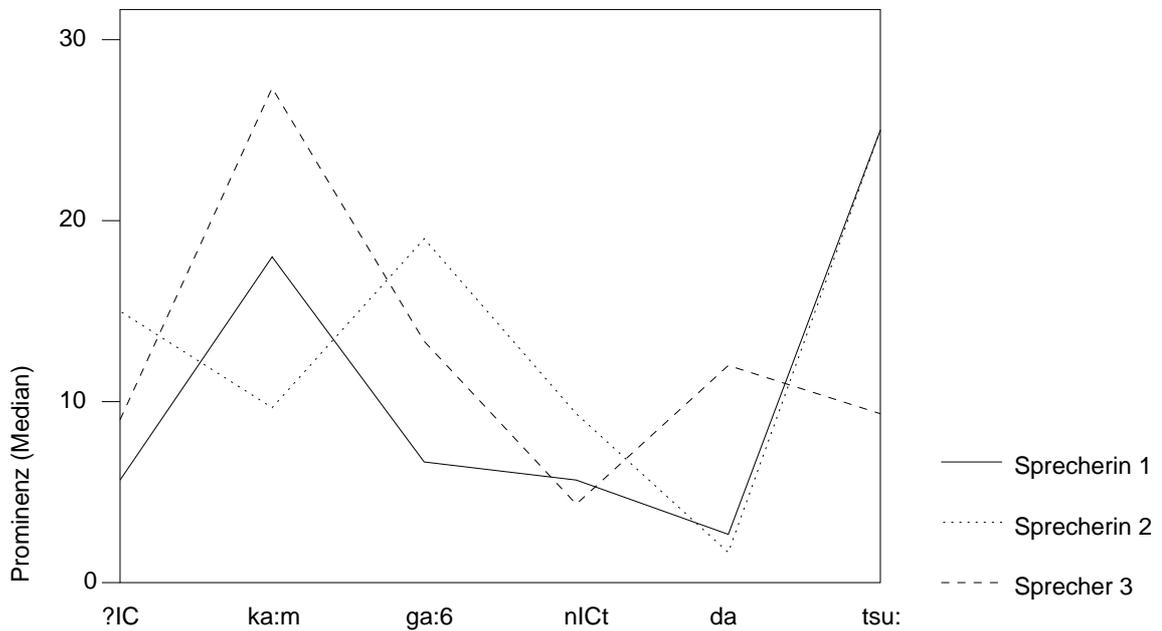


Abbildung 3.3: Prominenzmuster verschiedener Antworten auf die Frage „Wieviele Äpfel hast du geklaut?\": **Ich kám gar nicht dazu vs. Ich kam gár nicht dazu**

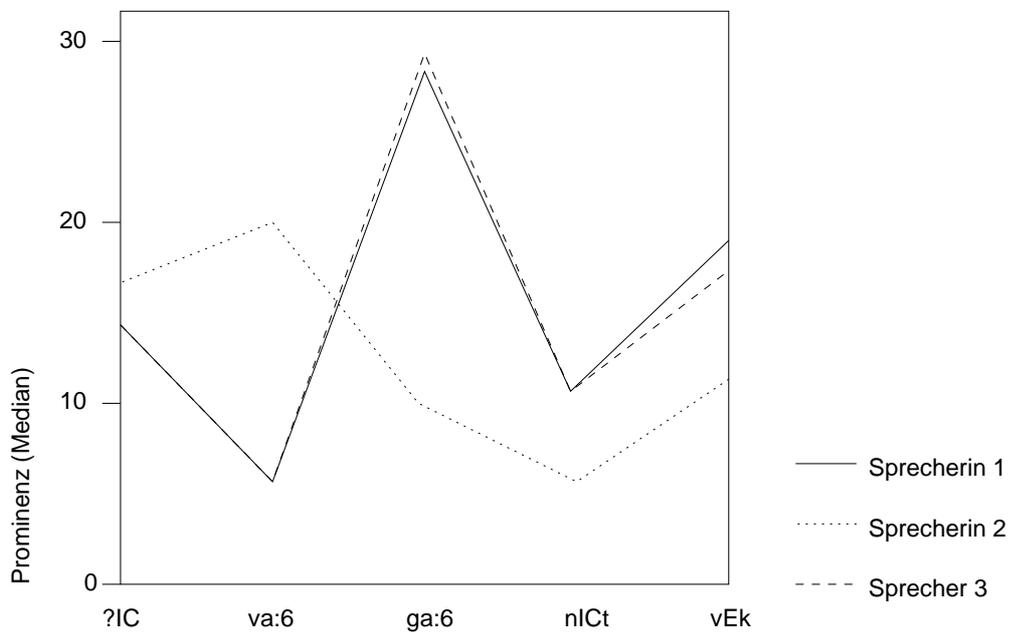


Abbildung 3.4: Prominenzmuster verschiedener Antworten auf die Frage „Wieviele Tage warst du weg?\": **Ich wár gar nicht weg vs. Ich war gár nicht weg**

In dieser Arbeit wird davon ausgegangen, daß die vorhergesagten Prominenzmuster einer phonologischen Grammatik eine ähnliche Korrelation zu den wahrge-

nommenen Mustern aufweisen sollten wie die Muster verschiedener Sprecherinnen untereinander. Im folgenden wird die Inter-Sprecherinnenkorrelation daher als *Baseline* für die Güte einer phonologischen Grammatik zugrunde gelegt.

3.8 Fazit: Überprüfung formaler Betonungsmodelle anhand perzeptiver Daten

In diesem Kapitel wurde auf der Basis der Anforderungen Chomskys an linguistische Theorien ein Evaluationsmodell von metrischer Phonologie entwickelt, welches perzeptiv erhobene Prominenzdaten als empirisches Datenmaterial nutzt. Die Vorhersagen der phonologischen Modelle sollen mit Hilfe endlicher Übergangnetzwerke formalisiert werden und durch ihre Implementierung automatische Vorhersagen liefern, so daß große Datenmengen getestet werden können. Die Inter-Sprecherinnenkorrelationen bilden die *Baseline* der Untersuchung, also das Gütemaß, welches von einer phonologischen Vorhersage im Idealfall erreicht werden sollte.

Als Fazit der Forschungslage innerhalb der formal geprägten Phonologie läßt sich zusammenfassend sagen, daß auf der einen Seite erst jetzt durch die Entwicklung eines verlässlichen Annotationsverfahrens für Prominenzstrukturen eine Grundlage für empirische Untersuchungen der Vorhersagen metrischer Modelle vorliegt. Auf der anderen Seite ermöglichen nur computergestützte Verfahren die Untersuchung größerer Korpora, wie sie für über Fallstudien hinausgehende Erhebungen notwendig sind. Ferner bietet nur die Methode der Implementierung die formale Evaluierungsmöglichkeit einer vorliegenden Theorie, eine Anforderung, welche in anderen Bereichen der Linguistik mittlerweile einen Standard darstellt:

[...] since formality demands (or at least promotes) a certain kind of explicitness and (in theory, if not always in practice) clarity of representation, the formalist programme offers phonologists the possibility of pursuing certain debates in a rigorous and impersonal fashion, thus allowing the prospect of avoiding chasing non-issues. (COLEMAN 1998, 13)

Kapitel 4

Prominenzmuster auf der Äußerungsebene

Selbst wenn er [die Wörter] nicht aussprach, hörte er sie, manchmal schienen es ihm sogar, als schallten sie. Sobald man sie aus der Reihe der Sätze löste, in die sie gehörten, bekamen sie, falls man dafür empfänglich war, etwas Angsterregendes, eine Fremdheit, über die man besser nicht nachdachte, weil sonst die ganze Welt ins Wanken geriet.

Cees Nooteboom, *Allerseelen*

Dieses Kapitel handelt von der Vorhersage von Prominenzmustern im Äußerungskontext. Hierzu werden zunächst verschiedene phonologische Modellierungen miteinander verglichen, von denen einige die Äußerungsprominenz auf der Basis rein syntaktischer, andere auf der Basis semantischer oder lexikonbasierter Eigenschaften vorherzusagen suchen (vgl. auch (WAGNER 2000a; WAGNER 2000b)). Die verschiedenen Ansätze werden sukzessive anhand der in Kapitel 3 erarbeiteten Methode überprüft. Die Ergebnisse werden zuletzt in ein optimiertes Vorhersagemodell integriert, welches anschließend auf der Basis endlicher Übergangnetzwerke formal dargestellt wird.

4.1 Vorhersage der Äußerungsprominenz im Englischen

Das wichtigste und einflußreichste Modell für die Generierung von Prominenzrelationen auf Satz- bzw. Äußerungsebene stammt von (SELKIRK 1984). Ihr Modell bedient sich dabei der Gitterdarstellungen der Metrischen Phonologie. Sie liefert Abbildungsvorschriften für die Überführung der syntaktischen Struktur auf die metrische und auch tonale Struktur der Äußerung. Die tonale Ebene soll an dieser Stelle jedoch ausgeklammert werden, da sich Prominenzrelationen phonologisch durch

die metrischen Gitter abbilden lassen und die tonale Struktur sich sowieso im metrischen Gitter durch erhöhte Prominenzen niederschlägt. Wird durch die syntaktische Analyse nämlich ein Ton — z.B. ein phrasenfinaler Grenzton oder fokusbedingter Tonhöhenakzent — vorgegeben, so muß die mit diesem Ton verknüpfte Silbe metrisch verstärkt, also prominenter werden. Diese Interaktion zwischen metrischer und tonaler Darstellung geht bereits auf (PIERREHUMBERT 1980) zurück.

Laut (SELKIRK 1984) ist die rhythmische Struktur einer Äußerung rein syntaktisch bestimmbar, die Semantik hat keinerlei direkten Einfluß. Aus der syntaktischen Konstruktion ergibt sich bei ihr simultan die rhythmische Struktur. Die phonologischen Regeln im eigentlichen Sinne erhalten diese rhythmische Struktur anschließend als Eingabe. Einen hohen Stellenwert bei der Vergabe von rhythmischen Strukturen ist bei Selkirk das *Prinzip der rhythmischen Alternation*, welches auf eine Alternanz metrisch starker und schwacher Silben zielt, wie sie auch für das Deutsche beobachtet werden konnte (HEUFT 1999, 65). Laut (WIESE 1996, 256) stützt sich das Prinzip auf folgende Einschränkungen, die für metrische Gitter gelten und auch *Euphonieregeln* genannt werden:

- *Jede starke Position auf einer metrischen Ebene n sollte von wenigstens einer schwachen Position auf dieser Ebene gefolgt werden.*
- *Einer schwachen Position auf einer Ebene n kann höchstens eine schwache Position auf dieser Ebene vorangehen.*

Rhythmische Alternation wird im metrischen Gitter durch Hinzufügung oder Tilgung bzw. Verschiebung von rhythmischen Prominenzeinheiten oder *Schlägen* („Beats“)¹ erreicht. (WIESE 1996) bemerkt allerdings, daß dies Prinzip vermutlich keine strenge universelle Einschränkung ist, sondern eher eine Beschreibung der präferierten Strukturen darstellt. Diese Vermutung Wieses wird durch neuere Ergebnisse aus der Sprachsyntheseforschung gestützt (EOM et al. 1999), die bemerkt haben, daß streng rhythmische Sprache von Menschen eher als „unnatürlich“ und „maschinenhaft“ empfunden wird. Eine gewisse Unregelmäßigkeit scheint daher durchaus den Normalfall darzustellen.

Ein wichtiger Punkt in Selkirks Modellierung ist die Behandlung von Funktionswörtern. Diese erhalten keinerlei Schlag auf der Gitterebene des prosodischen Wortes, da für sie angenommen wird, daß sie mit den ihnen vorausgehenden Wörtern klitisch verbunden sind. Mit Hilfe dieser Klitisierung erklärt Selkirk die Möglichkeit der Reduktion von Funktionswörtern, wie in *Mary's, isn't, he'll*. Solche Klitisierungen sind auch im Deutschen möglich (z.B. *wie geht's?*, vgl. Seite 16). Um zu erklären, wann diese Klitisierung bzw. bei welchen Funktionswörtern sie eintritt, entwickelt sie einen recht komplexen Mechanismus, der sogenannte „stumme Halbschläge“ nach Inhaltswörtern annimmt, welche eine Klitisierung blockieren, unter bestimmten Voraussetzung aber verschoben werden können. Ferner sind Funktionswörter

¹Ein Schlag entspricht dabei einem Eintrag auf einer Ebene eines metrischen Gitters.

am Phrasenende nicht reduzierbar. Ob Klitisierung rein syntaktisch oder morphologisch determiniert ist, bleibt laut (WIESE 1996, 271) noch offen. Interessant ist, daß Funktionswörter in Selkirks Modell keinen Schlag auf der Wortebene erhalten und zusätzlich deakzentuiert werden können, so daß Strukturen wie in Abbildung 4.1 möglich werden.

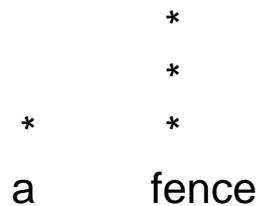


Abbildung 4.1: Metrisches Gitter für die Nominalphrase „a fence“ gemäß (SELKIRK 1984)

Eine solche Darstellung entspricht aber nicht dem streng relationalen Konzept von metrischen Repräsentationen, wie sie mit der Baumstruktur ausgedrückt wird. Diese drückt zwar aus, daß Inhaltswörter prominenter als Funktionswörter sind, macht aber keinerlei quantitative Angaben über den Prominenzabstand. Die Gitterdarstellung nach Abbildung 4.1 impliziert aber, daß Inhaltswörter eine intrinsische Prominenz haben. Diese intrinsische Prominenz ist nicht nur höher als die eines davorstehenden Funktionswortes, sondern es muß offenbar zwischen beiden Wörtern ein bestimmter „Prominenzabstand“ eingehalten werden. Dies bedeutet,

- daß Prominenzen auf der Ebene der Phonologie eine intrinsische Stärke haben
- und daß diese Prominenzen auf der Ebene des Lexikons festgesetzt werden.

In dem in Abschnitt 2 definierten Prominenzmodell wird nicht von intervallskalierten linguistischen Kategorien ausgegangen. Dies impliziert jedoch nicht, daß es keine wortklassenabhängige Prominenz geben kann, die sich aus dem Lexikon ergibt, anstatt auf der Basis der syntaktischen Struktur einer Äußerung motiviert zu sein. Eine Ordinalskalierung ergibt sich aus der Prominenz relativ zu anderen wortklassenspezifischen Prominenzwerten, und zwar unabhängig davon, ob diese im betrachteten Äußerungsabschnitt auftauchen oder nicht.

Selkirks Regelsatz für die Vorhersage von Wort- und Phrasenakzent im Englischen sieht leicht vereinfacht folgendermaßen aus:

Selkirks Prominenzregel 1

Auf der ersten Ebene des metrischen Gitters erhalten alle Silben einen Halbschlag.

Selkirks Prominenzregel 2

*Schwere Silben² erhalten einen zusätzlichen Schlag auf der zweiten Ebene (Fußebe-
ne), sofern sie am Wortanfang stehen.*

²Bei (SELKIRK 1984) sind schwere Silben als solche definiert, die mindestens einen Langvokal oder einen Kodakonsontanten haben. Zur Silbengewichtsproblematik im Deutschen siehe Abschnitt 5.1.5.

Selkirks Prominenzregel 3

Es werden rhythmisch alternierende Schläge von links nach rechts eingeführt.

Selkirks Prominenzregel 4

Die Position des letzten Grundschlages innerhalb einer Wortwurzel wird verstärkt und erhält einen Schlag auf der dritten Ebene (Wortbetonung). Funktionswörter erhalten keinen Schlag auf der dritten Ebene.

Selkirks Prominenzregel 5

In Komposita wird zyklisch der Wortakzent der linken Konstituente erhöht. (Compound Stress Rule)

Selkirks Prominenzregel 6

Die letzte Silbe innerhalb einer Phrase, die Trägerin der Wortbetonung und kein Funktionswort ist (also einen Schlag mindestens auf der dritten Ebene trägt), wird verstärkt. Die phrasenfinale Stärkung darf jedoch nicht ein vorhergehendes Wort an Prominenz übertreffen, welches fokussiert ist. Letztere Einschränkung ist notwendig, da bei (SELKIRK 1984) Tonhöhenakzente auf der Basis syntaktisch motivierter Fokusregeln vergeben werden. Da sich ein Tonhöhenakzent wie o.a. auf die Prominenz einer Silbe auswirkt, darf die phrasenfinale Stärkung nicht die Fokussierung neutralisieren, indem sie die fokussierte Konstituente an Prominenz übertrifft. (Nuclear Stress Rule)

Dieser Regelsatz bildete den Ausgangspunkt für verschiedene generative Vorhersagemodelle der Äußerungsbetonung für das Deutsche. Zwei dieser Modellierungen sollen in den nun folgenden Abschnitten evaluiert werden.

4.2 Vorhergesagte und wahrgenommene Prominenz: ein erster Regelsatz

Als Ausgangspunkt für weitere Studien sollte untersucht werden, ob sich regelhaft vorhergesagte Prominenzmuster überhaupt mit wahrgenommenen in einen Zusammenhang bringen lassen. Hierfür wurde ein Regelsatz implementiert, welcher weitgehend auf den Vorhersageregeln von (UHMANN 1991, 69) basiert. Dieser Regelsatz schien besonders geeignet, da er

- Prominenzen bis hinauf zur prosodischen Äußerungsebene vorhersagt,
- auf Regeln basiert, über die innerhalb der phonologischen Forschung weitgehende Übereinstimmung herrscht,
- sehr einfach ist.

Der Regelsatz enthält keinerlei elaborierte Behandlung der deutschen Wortbetonung, da dies auch nicht das Thema von Uhmans Arbeit war. Der hier dargestellte Regelsatz stellt eine vereinfachte Version der Regeln Uhmans dar, da sie die Satzprominenz auf der Basis von vorher detektierten Fokusexponenten vorhersagt. Eine Annotation von Fokusexponenten in einer größeren Datenbank macht allerdings eine automatische tiefe semantische Analyse erforderlich. Diese stellt heutzutage noch ein fast unlösbares Problem dar, weshalb an dieser Stelle mit einem vereinfachten Ansatz gearbeitet wird. Da es sich bei den Sätzen allerdings um gelesene Sprache handelt, welche nur teilweise mit Kontext dargeboten war, kann in der Regel von einer „Zitierform“ des gelesenen Materials ausgegangen werden. Der erste Regelsatz sieht folgendermaßen aus:

Prominenzregel 4.1

Jede Silbe erhält einen Basisschlag.

Prominenzregel 4.2

Jede Silbe, deren Nukleus nicht aus einem reduzierten Vokal (ə oder ʊ) oder einem silbischen Konsonanten (n, m, ŋ oder l) besteht, erhält einen Schlag.

Prominenzregel 4.3

Die erste nichtreduzierte Silbe jeden nativen Wortes erhält einen Schlag, die letzte nichtreduzierte Silbe eines jeden Lehnwortes erhält einen Schlag (Zuweisung der Wortbetonung).

Prominenzregel 4.4

Jede Silbe, der die Wortbetonung zugewiesen wurde und die Teil eines Inhaltswortes ist, erhält einen Schlag (vereinfachte inhaltliche Betonung von potentiellen Fokusdomänen).

Prominenzregel 4.5

Die letzte Silbe innerhalb der Intonationsphrase, die Wortbetonung trägt und Teil einer potentiellen Fokusdomäne ist, erhält einen Schlag (entspricht der Nuclear Stress Rule).

4.2.1 Güteüberprüfung der Vorhersagen des ersten Regelsatzes

Die Regeln wurden implementiert und dazu verwendet, für jede Äußerung innerhalb der BPD ein Prominenzmuster zu generieren. Eine Korrelationsberechnung zeigte, daß beide Maße substantiell aber nicht hoch miteinander korrelieren ($\rho = 0.64$, $p < 0.01$). Dieses Resultat bestätigte die Annahme, daß die in der BPD gemessenen Prominenzmuster prinzipiell durch metrische Phonologie modelliert werden können, da ansonsten bei einem solch einfachen Vorhersagealgorithmus kein Zusammenhang zu erwarten wäre.

Um die größten Schwächen innerhalb der Vorhersage zu bestimmen, wurde eine qualitative Evaluation der Vorhersagefehler durchgeführt. Hierfür wurde für jede

Phrase die Spearman-Rho-Korrelation bestimmt. Alle Phrasen, die eine Korrelation von $\rho < 0.6$ aufwiesen, wurden hinsichtlich der Vorhersagefehler untersucht. Die Evaluation zeigte, daß nahezu alle Fehler erklärbar waren (92%). Hauptursache für schlechte Vorhersagen war eine falsche Wortbetonung (35%), welche durch die stark vereinfachte Regel 4.3 nicht überrascht. Ein weiterer Hauptgrund für falsche Vorhersagen waren Wörter, die eine höhere Prominenz aufwiesen, als ihnen zugewiesen worden war (25%). Hierbei ist besonders auffällig, daß bestimmte Klassen von Funktionswörtern anfällig für eine höhere Prominenz zu sein scheinen, unter ihnen Affirmativpartikeln, Demonstrativ- und Interrogativpronomina. Dies ist ein Indiz dafür, daß die Unterscheidung in Funktions- vs. Inhaltswörter zu vereinfacht ist und daß bestimmte Funktionswörter intrinsisch bereits höhere Prominenz aufweisen als andere. In einigen Fällen interpretierten die Sprecherinnen die Äußerung in einer Art und Weise, welche zu einer Hervorhebung auf einem Wort führte, welches nicht das letzte Inhaltswort innerhalb der Äußerung war. Letztere würde zu einer Betonung führen, welche einem sogenannten „weiten Fokus“, der Zitierform oder einer Defaultbetonung entspricht (vgl. Seite 18). Die Sprecherinnen wichen aber in 17% der Fälle mit einer schlechten Vorhersage von diesem Muster ab. Weitere 10% der schlechten Vorhersagen scheinen auf das Konto fehlender Euphonieregeln zu gehen, welche rhythmische Lücken verhindern, indem bei längeren Folgen eigentlich unbetonter Silben Schläge hinzugefügt werden. Klare Fälle von Vermeidungen sogenannter „Akzentzusammenstöße“ durch das Verschieben von Betonungen (vgl. Seite 3 und Abschnitt 5.2.1) konnten nicht gefunden werden. Ein interessantes Ergebnis war, daß in 5% der fehlerhaften Fälle die letzte schwere Silbe einer Äußerung prominenter war als vorhergesagt, auch dann, wenn sie nicht Trägersilbe der Wortbetonung war. In einigen Fällen erschien selbst eine unbetonbare Silbe äußerungsfinal als recht prominent. Dies weist darauf hin, daß die Domäne der „Nuclear Stress Rule“ nicht nur Auswirkungen auf die letzte Wortbetonungssilbe hat, sondern einen weiter gehenden Einfluß ausübt. Ein Überblick über die Ergebnisse der qualitativen Evaluation findet sich in Tabelle 4.1.

Falsche Wortbetonung	35%
Funktionswörter wurden prominenter empfunden als vorhergesagt (gilt für Affirmativpartikeln und Demonstrativpronomina)	25%
Unvorhergesehener prosodischer „enger“ Fokus	17%
Nicht aufgefüllte rhythmische Lücken	10%
Letzte schwere Silbe (nicht Trägerin der Wortbetonung) prominenter als vorhergesagt	5%
Unklar	8%

Tabelle 4.1: Gründe für schlechte Vorhersagen beim Verwenden eines stark vereinfachten Regelsatzes der Prominenz

Insgesamt weisen die Ergebnisse des ersten, einfachen Regelsatzes darauf hin, daß die wahrgenommenen Prominenzmuster durchaus mit Hilfe phonologisch motivierter Regeln vorhersagbar sind. Dennoch bleiben eine Reihe von Unklarheiten hinsichtlich einer präziseren Vorhersage bestehen. Klar ist, daß ein Großteil der Vorhersagefehler durch fehlende Wortbetonungsregeln erklärbar sind. Da in diesem Abschnitt aber nicht die Wortbetonungsregeln im Fokus der Untersuchung stehen, sondern vielmehr die Vorhersagbarkeit der Betonungsmuster auf Äußerungsebene, werden in den nachfolgenden Untersuchungen die Wortbetonungen anhand der wahrgenommenen Prominenzmuster bestimmt. Auf diese Weise können zwei verschiedene Herangehensweisen für die Generierung der Äußerungsbetonung im Deutschen miteinander verglichen werden. Beide sind stark von den Arbeiten in (SELKIRK 1984) zum Englischen geprägt. Im ersten Ansatz wird die metrische Struktur auf der Basis der syntaktischen Phrasierung generiert. Im zweiten Ansatz ist die Eigenschaft eines Wortes, potentieller Fokusexponent (vgl. Seite 2.4.2) zu sein, ausschlaggebend für die Berechnung der Äußerungsbetonung.

4.3 Vergleich verschiedener Ansätze zur Generierung von Äußerungsbetonung

4.3.1 Ein wortklassenbasierter Ansatz (UHMANN 1991)

Schon der erste Regelsatz in Tabelle 4.1 bezog sich auf den Algorithmus in (UHMANN 1991). Um das Problem der mangelhaften Wortbetonungsvorhersage zu umgehen und die Tauglichkeit des Ansatzes unabhängig von der Wortbetonung zu prüfen, wurden die Wortbetonungen (vgl. Regel 4.3) anhand der Prominenzannotationen in der Datenbank durch einen einfachen *Look-Up* bestimmt. Der hier dargestellte Ansatz wird natürlich Uhmans Ansprüchen nur teilweise gerecht, weil die (potentiellen) Fokusexponenten lediglich auf der Basis einer Unterscheidung zwischen Inhalts- und Funktionswörtern bestimmt werden, anstatt den semantischen Fokus der Äußerung zu bestimmen.

4.3.2 Ein syntaxbasierter Ansatz (FÉRY 1988)

Die Regeln dieses Ansatzes stimmen bis zur Ebene der Wortbetonung (Regeln 4.1 bis 4.3) mit denen in Abschnitt 4.3.1 überein. Die bisherigen Regeln 4.4 und 4.5 wurden durch nachfolgend erläuterte Modifizierte Prominenzregel 4.5 ersetzt. Basierend auf den Ausführungen von (FÉRY 1988), wurde nunmehr die Äußerungsbetonung anhand syntaktischer Phrasierung in der Datenbank nach folgender Vorschrift bestimmt:

Modifizierte Prominenzregel 4.5

(a) Die am weitesten rechts stehende Konstituente von zwei oder mehr syntakti-

schen Schwesterknoten erhält wenigstens einen Schlag mehr als ihre Schwesterkonstituenten.

(b) Schritt (1) muß für zwei syntaktische Ebenen berechnet werden, wobei zwei Ebenen unterhalb des Startknotens begonnen wird.

Die ursprüngliche Regel Férys bezog sich auf sämtliche syntaktischen Ebenen. Diese Regel wurde jedoch vereinfacht, um die — in diesem Fall manuellen — Betonungszuweisungen zu ermöglichen. Da die allermeisten syntaktischen Analysen aus nicht mehr als drei Ebenen bestanden, kommt dieser Einschränkung keine größere Bedeutung zu. Die syntaktischen Analysen erfolgten mit Hilfe des HPSG-basierten Parsers von (MÜLLER 1996). In vielen Fällen waren die Äußerungen syntaktisch ambig, so daß der Parser mehrere Vorschläge zur syntaktischer Phrasierung lieferte. In diesen Fällen wurde diejenige Phrasierung ausgewählt, die der Interpretation der Annotatorin am ehesten entsprach. Abbildung 4.2 illustriert die Vorhersage der Satzbetonung auf der Basis syntaktischer Strukturen.

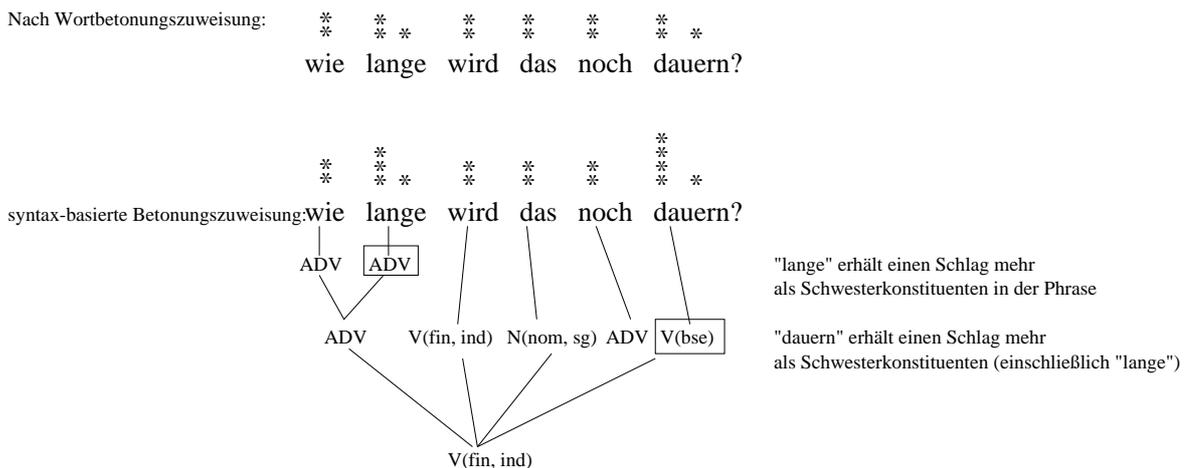


Abbildung 4.2: Generierung der Äußerungsbetonung anhand syntaktischer Strukturen

4.3.3 Vergleich der Modelle

Die Vorhersagen beider Modelle wurden mit einem weiteren Modell verglichen, welches lediglich das Silbengewicht sowie die aus der Datenbank entnommenen Wortbetonungen berücksichtigte (Regeln 4.1 und 4.2 plus Wortbetonungsbestimmung durch *Look-Up*). Diese Modellierung macht also keinerlei Annahmen über eine spezielle Äußerungsbetonung. Bei einem Vergleich beider Modelle mit dieser vereinfachten Vorhersage zeigten beide eine Verbesserung. Dennoch zeigt die wortklassenbasierte Vorhersage Uhmans gegenüber dem rein syntaktisch basierten Ansatz klare Vorteile. In vielen Fällen stimmen die Vorhersagen beider Modelle

überein. Mögliche Gründe für überlappende Vorhersagen liegen in der „Nuclear Stress Rule“, welche zwar syntaktisch motiviert ist, jedoch von Uhmans wie Féry gleichermaßen berücksichtigt wird. Weiterhin tauchen die potentiellen Fokusexponenten Uhmans häufig am rechten Rand einer syntaktischen Phrase auf. Dieser Umstand liefert weitere Gründe für sich deckende Vorhersagen. Dennoch scheint es wichtiger zu sein, Inhaltswörter und damit potentielle Fokusexponenten prosodisch hervorzuheben als die syntaktische Phrasierung zu berücksichtigen. Es ist anzunehmen, daß die syntaktische Phrasierung stärker mit der prosodischen Phrasierung zusammenhängt als mit der Prominenz der einzelnen Konstituenten. Diese Annahme wird gestützt durch Arbeiten, die auf der Basis syntaktischer Annotationen erfolgreich die prosodische Phrasierung vorhersagen können (HENRICI 2001; SCHWEITZER und HAASE 2000). Tabelle 4.2 zeigt eine Übersicht über die Korrelationen der verschiedenen Regelsätze mit den wahrgenommenen Prominenzwerten. Abbildungen 4.3 und 4.4 zeigen die wahrgenommenen und vorhergesagten Prominenzen der beiden Modelle anhand des Beispielsatzes „Sprechen Sie bitte langsam und deutlich!“.

Regeln basierend auf:	Korrelation vorhergesagter und wahrgenommener Prominenz:
Silbengewicht, Wortbetonung	$\rho = 0.64, p < 0.001$
Silbengewicht, Wortbetonung, syntaktische Phrasierung	$\rho = 0.72, p < 0.001$
Silbengewicht, Wortbetonung, potentielle Fokusexponenten	$\rho = 0.76, p < 0.001$

Tabelle 4.2: Vergleich von drei Modellen zur Vorhersage der Prominenzen auf Äußerungsebene

Wenn man die wahrgenommenen Prominenzwerte für Sprecherin 1 der Datenbank zugrunde legt, so ist die Vorhersage mit Hilfe der Regeln Uhmans bereits recht nah an der spezifizierten *Baseline*-Korrelation von $\rho = 0.78$, wie sie in Abschnitt 3.7 definiert wurde. Dennoch liegt die Korrelation noch immer wesentlich unterhalb der *Baseline*, wenn man sämtliche Annotationen der Datenbank zugrunde legt. Da eine phonologische Beschreibung möglichst allgemein sein sollte, ist die weitere Verbesserung der Vorhersageregeln demnach notwendig, sofern das phonologische Modell nicht nur sprecherspezifisch greifen soll.

4.4 Verbesserungen der Vorhersage

Ausgehend von dem bereits vielversprechenden Algorithmus in Anlehnung an die Arbeiten von (UHMANN 1991) sollen systematisch Verbesserungen durchgeführt werden. Da die Vorhersage auf der Basis der Wortklasse bereits der syntaktischen

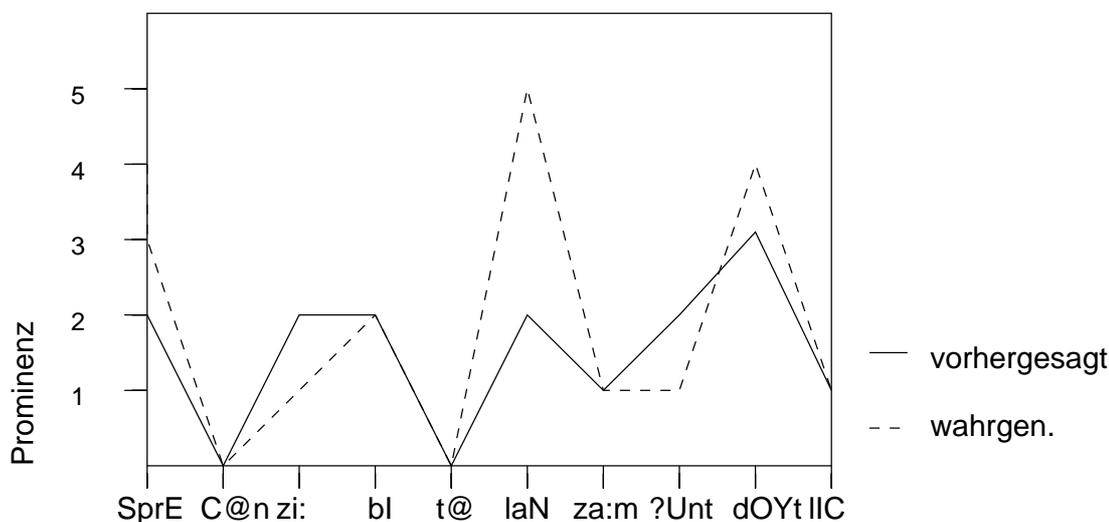


Abbildung 4.3: Vorhergesagte und wahrgenommene Prominenz bei Verwendung eines rein syntaktisch basierten Vorhersagemodells

Analyse gegenüber erfolgreicher war, soll untersucht werden, inwieweit der Einfluß des Lexikons noch weiter ausgenutzt werden kann. Die dahinterstehende Idee ist, daß bestimmte Wörter wichtiger für eine erfolgreiche Kommunikation sind. Wieder sollen zwei Ansätze, die den Einfluß des Lexikons modellieren, hinsichtlich ihrer Vorhersagekraft miteinander verglichen werden. Der erste trifft Vorhersagen lediglich auf der Basis der Wortklasse, wohingegen der zweite auf der Basis der Wortfrequenz die Betonungsstufe ermittelt. Es ist jedoch zu erwähnen, daß keiner der beiden Ansätze aus phonologischen Arbeiten stammt. Jedoch ist es offenbar für eine Verbesserung der phonologischen Modelle notwendig, Ideen aus anderen Quellen zu berücksichtigen.

4.4.1 Der Einfluß des Lexikons

Die Einsicht, daß die Prominenz einer Silbe davon abhängt, welcher Wortklasse sie zugehörig ist, ist nicht neu. Bereits (BEHN 1912, 56) konnte mit Experimenten an phonetisch geschulten Versuchspersonen fünf Betonungsstufen für deutsche Silben isolieren. Obwohl diese Betonungsstufen introspektiv ermittelt wurden und es somit fraglich ist, inwieweit diese Erhebung mit perzeptiven Daten vergleichbar ist, sind Behns Ergebnisse nicht weit von neueren Untersuchungen entfernt. Allerdings unterscheidet auch er in erster Linie zwischen Inhalts- (Begriffs-) und Funktionswörtern (Beziehungswörtern). Behn begründet die unterschiedlichen Betonungsstufen mit einem höheren „Sinnwert“ von Inhaltswörtern. Diese Beobachtung findet sich auch in der neueren Literatur wieder, die davon ausgeht, daß Inhaltswörter zu meist den informativeren Teil einer Äußerung ausmachen und außerdem seltener verwendet werden (siehe Abschnitt 4.4.2). Daher muß die Sprecherin insbesondere

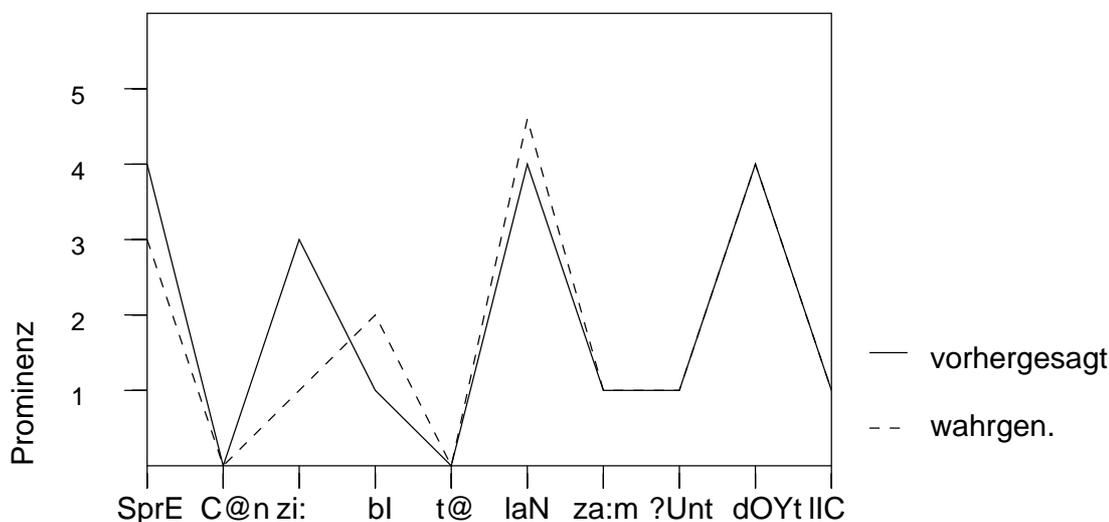


Abbildung 4.4: Vorhergesagte und wahrgenommene Prominenz bei Berücksichtigung potentieller Fokusdomänen

die Übertragung der Inhaltswörter sicherstellen, damit die Äußerung beim Hörer ankommt und die Kommunikationsübertragung erfolgreich ist.

Arbeiten zur wortklassenspezifischen Betonung finden sich zumeist im Umfeld der sprachtechnologischen Forschung, u.a. von (FANT und KRUCKENBERG 1989) für das Schwedische, (KLATT et al. 1987) für amerikanisches Englisch sowie (WIDERA et al. 1997), (STÖBER et al. 2000; MIXDORFF und JOKISCH 2001; MÜLLER und HOFFMANN 2001) und (SCHWEITZER und HAASE 2000) für das Deutsche. Die hier untersuchte Modellierung bezieht sich in erster Linie auf die Ergebnisse von (WIDERA et al. 1997). In ihrer Studie analysierten sie nicht nur die wortklassenabhängige Prominenz, sondern ebenfalls den Einfluß auf die Prominenzen benachbarter Wortklassen. Ihre Ergebnisse spiegeln die bereits mehrfach genannte Intuition wider, daß Inhaltswörter inhärent prominenter als Funktionswörter sind. Allerdings lassen die Ergebnisse eine feinere Klassifikation zu und stellen die einfache Dichotomie Inhalts- vs. Funktionswort sehr in Frage. Die Untersuchungen von (WIDERA et al. 1997) stützen sich auf eine 30-stufige Prominenzskala. Da in der Phonologie lediglich relative Prominenzen wichtig sind, wurden die Ergebnisse vereinfacht und auf eine 5-stufige Prominenzskala abgebildet (vgl. Tabelle 4.3). Die folgende Regel ersetzten die Regeln 4.3 und 4.4 in Abschnitt 4.1, welche den im Wort hauptbetonten Silben von Funktionswörtern einen, denen von Inhaltswörtern zwei zusätzliche Schläge zuwiesen:

Modifizierte Prominenzregel 4.4

Die prominenteste Silbe eines Wortes erhält einen Prominenzwert entsprechend der Prominenzkategorie des Wortes, dessen Teil sie ist.

Wortklasse		Prominenzwert
<i>Inhaltswörter</i>	<i>Funktionswörter</i>	
Nomina, Numeralia, Eigennamen		5
Adverbien, Adjektive		4
Verben	Demonstrativpronomina, Fragepronomina	3
	Modal-/Auxiliarverben, Affirmativ-/Negationspartikeln	2
	Artikel, Konjunktionen, Subjunktionen, Präpositionen	1

Tabelle 4.3: Wortklassenabhängige inhärente Prominenzwerte

Weiterhin fanden (WIDERA et al. 1997) heraus, daß inhärent hochprominente Wörter sehr stabile Prominenzwerte haben, wohingegen weniger prominente Wörter hinsichtlich ihrer Prominenz abhängig von ihrer Umgebung oder Äußerungsposition veränderlich sind. So weichen Nomina selten von ihrer inhärenten Prominenz ab. Artikel jedoch sind offenbar recht anfällig für gewisse Abweichungen. Es findet sich die Tendenz, daß Funktionswörter zu Beginn prosodischer Phrasen prominenter sind, was eine Art Gegengewicht zur klassischen *Nuclear Stress Rule* darstellt, da diese das letzte Inhaltswort einer Äußerung in seiner Prominenz anhebt. Die Prominenzstabilität inhärent hochprominenter Wörter liefert weitere Evidenz für die Wichtigkeit dieser Wörter innerhalb der Kommunikationsübertragung. Die äusserungsinitial erhöhte Prominenz von Funktionswörtern wurde ebenfalls in den Vorhersagealgorithmus integriert, indem jedem äusserungsinitialen Funktionswort ein zusätzlicher Schlag zugewiesen wurde:

Zusätzliche Prominenzregel 4.6

Die prominenteste Silbe eines äusserungsinitialen Funktionswortes erhält einen zusätzlichen Schlag.

Die *Nuclear Stress Rule* wurde ebenfalls unter Berücksichtigung der Wortklasse neu definiert:

Modifizierte Prominenzregel 4.5

Die prominenteste Silbe innerhalb des letzten Inhaltswortes einer prosodischen Phrase erhält einen zusätzlichen Schlag, außer es handelt sich um ein Verb. Verben erhalten einen zusätzlichen Schlag, wenn ihnen nicht ein Nomen, Adjektiv oder Adverb folgt.

Die wortklassenbasierten Modifikationen des Regelsatzes stellten sich als äußerst erfolgreich heraus. Eine Vorhersage, die lediglich sehr einfache phonologische Sil-

benmerkmale (Akzentuierbarkeit der Silbe) und die wortklassenspezifischen Prominenzwerte (Regel 4.4) berücksichtigt, resultiert in einer Korrelation zwischen vorhergesagten und wahrgenommenen Prominenzwerten, die bereits nahe an die in Abschnitt 3.7 definierte Baseline heranreicht und die vorherigen Analysen übertrifft. Die phraseninitiale Schlaghinzufügung auf Funktionswörtern verbessert die Ergebnisse weiter, ebenso die wortklassenabhängig definierte *Nuclear Stress Rule*. Die höchsten Korrelationen zwischen vorhergesagten und wahrgenommenen Prominenzen zeigten sich für die (professionelle) Sprecherin 1 ($\rho = 0.81, p < 0.01$). Tabelle 4.4 liefert einen Überblick über die Ergebnisse für verschiedene Regeln.

In der Vorhersage verwendete Regeln	Korrelation zwischen vorhergesagter und wahrgenommener Prominenz (Baseline: $\rho = 0.78$)
Regeln 4.1, 4.2 (Akzentuierbarkeit der Silbe)	$\rho = 0.58, p < 0.01$
Regeln 4.1, 4.2, 4.4 (zusätzlich: Wortklassenspezifische Prominenzzuweisung)	$\rho = 0.77, p < 0.01$
Regeln 4.1, 4.2, 4.4, 4.6 (zusätzlich: Hinzufügung eines Zusatzschlags auf äußerungsinitiale Funktionswörter)	$\rho = 0.78, p < 0.01$
Regeln 4.1, 4.2, 4.4, 4.6, 4.5 (zusätzlich: <i>Nuclear Stress Rule</i>)	$\rho = 0.784, p < 0.01$

Tabelle 4.4: Korrelationen zwischen wahrgenommenen und vorhergesagten Prominenzen für wortklassenspezifische Regelsysteme

4.4.2 Der Einfluß des Informationsgehalts

Die ausschlaggebende Motivation für die Beachtung der Wortklasse bei der Vergabe der Prominenz war der Gedanke, daß Inhaltswörtern innerhalb der Kommunikationskette ein höherer Informationsgehalt zukommt. Dies hat zur Folge, daß diesen Wörtern von der Sprecherin eine höhere Prominenz zugewiesen wird. Es ist auch der Fall, daß Funktionswörter insgesamt eine höhere Frequenz im Sprachgebrauch haben, u.a. deshalb, weil sie eine geschlossene Klasse darstellen und kontextuell vermutlich besser vorhersagbar sind. Ihr informativer Gehalt bzw. ihr *semantisches Gewicht* (MONAGHAN 1993) steht aus diesem Grund hinter dem von Inhaltswörtern zurück. Es ist aber durchaus einsichtig, daß es auch innerhalb von Wörtern ein und derselben Wortklasse ähnliche Unterschiede gibt. Beispielsweise kann man sich den folgenden Satz vorstellen, wobei das *< noise >* für einen nicht verstehbaren Äußerungsteil steht.

Bundeskanzler < noise > hat in einer Rede seine Politik der ruhigen Hand verteidigt.

Wenn man davon ausgeht, daß der Hörer dieses Satzes in den letzten Jahren in Deutschland gelebt und vom öffentlichen Leben ansatzweise Notiz genommen hat, wird er vermutlich keine Schwierigkeiten haben, den unverständlichen Teil mit dem Eigennamen *Schröder* aufzufüllen. Würde der Satz jedoch lauten:

Das Staatsoberhaupt der Malediven, Präsident < noise >, hat auf die unmittelbare Bedrohung seines Landes durch die gobale Erwärmung hingewiesen.

hätten wohl die meisten Hörer Probleme, auf Anhieb den unverständlichen Äußerungsteil mit einem Namen zu ergänzen. Die Sprecherin dieses Satzes würde aufgrund ihres Weltwissens vermutlich den geringen Bekanntheitsgrad des maledivischen Staatsoberhauptes in Betracht ziehen und dafür Sorge tragen, daß sein Name zum Hörer durchdringt. Erhöhte Prominenz stellt nur eine von mehreren Möglichkeiten dar, dies Ziel sicherzustellen. Dennoch ist es anzunehmen, daß Sprecherinnen auch von dieser Möglichkeit Gebrauch machen. Auf der Basis dieser Überlegung wurde eine Vorhersage der Prominenzen anhand von Auftretenshäufigkeiten deutscher Wörter in großen Sprachkorpora durchgeführt. Die Annahme hierbei war, daß Wörter mit einer hohen Frequenz wenig Neuinformation liefern, eher kontextuell erschlossen werden können und somit weniger prominent sind, wohingegen selten auftretende Wörter für den Hörer einen hohen Informationsgehalt haben und prominenter sein sollten.

Um die wortfrequenzabhängige Prominenz zu bestimmen, wurde für alle Wörter der BPD eine Häufigkeitsklasse anhand des Normalisierungsverfahrens von (QUASTHOFF 1998) bestimmt. In einer Datensammlung aus verschiedenen deutschen Zeitungskorpora wurde für jedes Wort eine Häufigkeitsklasse ermittelt. Die Ermittlung einer Häufigkeitsklasse anstatt einer absoluten Frequenz stellt ein Normalisierungsverfahren dar, welches zum Ziel hat, daß die Häufigkeitsklasse unabhängig von der Korpusgröße relativ stabil bleiben sollte. Die Häufigkeitsklasse wird für jedes Wort relativ zur Auftretensfrequenz des häufigsten deutschen Wortes im Korpus, „der“, ermittelt. Die Häufigkeitsklasse h eines Wortes drückt dabei aus, daß das Wort „der“ $\frac{2^h}{e}$ mal so häufig vorkommt wie das Wort selbst. Die Wortfrequenzen für alle Wörter der BPD wurden mit Hilfe des „Deutschen Referenzkorpus“ (ENDRES 2000) bestimmt. Anschließend erfolgte die Bestimmung der Häufigkeitsklassen mit Hilfe der Berechnungsvorschrift

$$h(\text{Wort}) = \ln(\text{Frequenz}(\text{Wort})/\text{Frequenz}(\text{der})).$$

Die so ermittelten Häufigkeitsklassen spiegeln eine Unterteilung in Funktionswörter (niedrige Häufigkeitsklassen) und Inhaltswörter (hohe Häufigkeitsklassen) wider. Die Hoffnung, daß diese feinere Aufteilung zu einer besseren Vorhersage der Prominenzen führt, ließ sich allerdings nicht bestätigen. Die Vorhersagen auf der Basis der Häufigkeitsklassen korrelierten mit den wahrgenommenen Werten lediglich mit $\rho = 0.63, p < 0.001$, was im Verhältnis zu den wortklassenspezifischen Analysen eine erhebliche Verschlechterung darstellt. Eine Prominenzvorhersage, die von

grammatischen Kategorien abstrahiert, scheint demnach keinen Ersatz zu bieten. Immerhin ist aber denkbar, daß wenigstens wortklassenintern die Häufigkeitsklassen eine feinere Vorhersage der Prominenz ermöglichen. Hierfür wurden innerhalb einzelner Wortklassen die Korrelationen zwischen Häufigkeitsklassen und den wahrgenommenen Prominenzwerten ermittelt. Diese Analyse ergab jedoch allenfalls leichte Tendenzen, die für einen Zusammenhang zwischen Häufigkeitsklasse und Prominenz sprachen. Die Korrelationen waren aber zu geringfügig, um hieraus Generalisierungen ableiten zu können.

Es ist jedoch möglich, daß die Hypothese vom Zusammenhang zwischen Häufigkeitsklasse und Prominenz eines Wortes nicht komplett fallengelassen werden muß. Da Kommunikation immer situations- und domänengebunden stattfindet, kann eine domänenspezifische Häufigkeitsklasse u.U. bessere Ergebnisse liefern. So ist in einer Dokumentation über die beliebtesten Ferenziele der Welt das Wort „Malediven“ sicherlich eher zu erwarten als das Wort „Dortmund“, obwohl letzteres bezogen auf das Deutsche Referenzkorpus eine niedrigere Häufigkeitsklasse besitzt. In einer Äußerung wie

Auch Dortmund zieht in diesem Jahr viele Touristen an.

könnte aufgrund der ungewöhnlichen Information das Wort „Dortmund“ eine hohe Prominenz zur Folge haben. Eine domänenabhängige Wortfrequenz könnte u.U. zu einer verbesserten Prominenzvorhersage beitragen, domänenunabhängig ist aber eine Verbesserung nicht ersichtlich.

4.4.3 Der Einfluß des Rhythmus

In Abschnitt 4.1 wurde das *Prinzip der Rhythmischen Alternation* als Wohlgeformtheitskriterium für Prominenzmuster beschrieben. In Abschnitt 4.1 wurde bereits bemerkt, daß falsche Prominenzvorhersagen in vielen Fällen durchaus mit dem Nichtbeachten dieses Prinzips erklärbar sind. Daher wurde untersucht, inwieweit die Einführung von Euphonieregeln die Vorhersage verbessern könnten. Ausgegangen wurde hierbei von dem wortklassenspezifischen Vorhersagealgorithmus, welcher in Abschnitt 4.4.1 entwickelt wurde.

Die Methode einer wortklassenspezifischen Prominenzzuweisung löst das Problem des sogenannten Akzentzusammenstoßes in vielen Fällen automatisch mit, da es selten der Fall ist, daß zwei Wörter gleicher Prominenzklasse aufeinanderstoßen. Allerdings fanden (WIDERA et al. 1997) Evidenz für eine Vermeidung ähnlicher Prominenzen auch auf Wortebene: Werden zwei Funktionswörter derselben Wortklasse nacheinander geäußert, so wird das erste tendenziell prominenter wahrgenommen, handelt es sich jedoch um zwei Inhaltswörter, so wird das zweite von beiden prominenter wahrgenommen. Diese (hier stark vereinfachte Darstellung) fand Berücksichtigung in der folgenden Euphonieregel zur Vermeidung von gleichen Prominenzen auf Wortebene:

Zusätzliche Prominenzregel 4.7

Die prominenteste Silbe eines Funktionswortes erhält einen zusätzlichen Schlag, wenn das Wort von einem Wort gleicher Prominenzklasse gefolgt wird. Die prominenteste Silbe eines Inhaltswortes erhält einen zusätzlichen Schlag, wenn ein Wort gleicher Prominenzklasse vorausgeht.

Diese Regel führte allerdings zu einer leichten Verschlechterung der Ergebnisse und zu einem Absinken der Korrelation unter das *Baseline*-Niveau ($\rho = 0.77, p < 0.001$). Dies zeigt, daß die wortklassenspezifischen Prominenzen relativ zueinander eingehalten werden müssen. Daß die Arbeiten in (WIDERA et al. 1997) mit einer feineren Prominenzskalierung arbeiteten, weist wiederum darauf hin, daß die *absolute Prominenz* durchaus beeinflußt werden kann, solange die linguistisch relevanten Muster nicht durcheinanderkommen (vgl. Abschnitt 2.6). Solche Effekte scheinen demnach eher sprechstil- oder sprecherspezifische Charakteristika darzustellen, die aber keine linguistische Relevanz im engeren Sinne besitzen.

Wenn auch die wortklassenspezifischen Prominenzniveaus nicht stark von Euphonieprinzipien betroffen zu sein scheinen, so verhindert der Algorithmus es dennoch nicht, daß Silben gleicher Prominenz aufeinander folgen. Das „Prinzip der rhythmischen Alternation“ bevorzugt allerdings Folgen starker und schwacher Silben (binäres Betonungsmuster), wenn auch (FÉRY 1986) die Ansicht vertritt, eine betonte Silbe werde im Deutschen auch genauso häufig von zwei unbetonten Silben gefolgt (ternäres Betonungsmuster). Die Metrische Phonologie kennt zwei phonologische Prozesse, welche rhythmisch alternierende Folgen betonter und unbetonter Silben erzeugen sollen. Der erste Prozeß ist die sogenannte *Schlaghinzufügung* (vgl. Abbildung 4.5). Diese verhindert längere Sequenzen unbetonter Silben. Beim zweiten Prozeß handelt es sich um die *Schlagtilgung* (vgl. Abbildung 4.6). Diese sorgt für eine Verhinderung von Folgen mehrerer betonter Silben (*Stress Clash*).

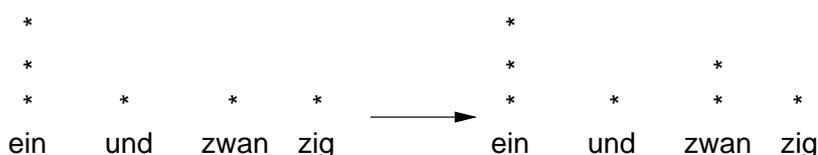


Abbildung 4.5: Der phonologische Euphonieprozeß der Schlaghinzufügung zur Vermeidung rhythmischer Lücken

Um zu bestimmen, wie diese Prinzipien im Deutschen angewendet werden, wurden diejenigen Fälle nochmals untersucht, in denen der erste, einfache Vorhersagealgorithmus (vgl. Abschnitt 4.1) Schwächen zeigte, die auf Verletzung des Euphonieprinzips hinwiesen. Diese Analyse ergab, daß die allermeisten Verletzungen des Euphonieprinzips in zu langen Folgen mehrerer unbetonter Silben bestanden. Diese traten zumeist in Numeralia auf, welche aus vielen Silben bestehen. Die hier bestehende Lücke in der Vorhersage wird auch deutlich, wenn auf der Basis der bisher besten Vorhersage eine Korrelation zwischen vorhergesagter und wahrge-

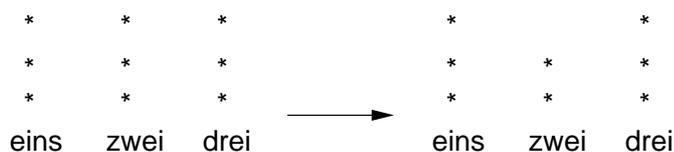


Abbildung 4.6: Der phonologische Euphonieprozeß der Schlagtilgung zur Vermeidung von Betonungszusammenstößen

nommener Prominenz für denjenigen Teil der Datenbank berechnet wird, welcher viele polysyllabische Numeralia enthält. Diese Korrelation liegt nämlich nur bei $\rho = 0.76, p < 0.001$ und somit unterhalb der angestrebten *Baseline*. Offensichtlich ist in diesen Wörtern zusätzlich zur Wortbetonung die Vergabe einer Nebenbetonung durch das Prinzip der Schlaghinzufügung vonnöten. Dies geschah mit Hilfe folgender zusätzlicher Regel, welche auf die Ausgabe des bisher besten Ergebnisses angewendet wurde:

Zusätzliche Prominenzregel 4.8

Wenn mehr als zwei betonbare Silben niedriger Prominenz (Prominenzniveau 1) aufeinander folgen, so füge der zweiten und jeder übernächsten Silbe dieser Folge einen Schlag hinzu.

Diese Vorschrift erzeugt binäre Betonungsmuster bei Folgen von drei oder mehr betonbaren Silben ohne Wortakzent. Folgen von zwei unbetonten Silben werden jedoch akzeptiert, so daß in einigen Fällen auch ternäre Betonungsmuster entstehen. Nach Anwendung der Regel steigt auch die Korrelation innerhalb des kritischen Teils der Datenbank auf das *Baseline*-Niveau an ($\rho = 0.78, p < 0.001$). Die Verbesserung ist für alle Sprecherinnen klar ersichtlich. Abbildungen 4.7 und 4.8 zeigen einen Vergleich von vorhergesagter und wahrgenommener Prominenz vor und nach der Regelanwendung für die Teiläußerung „Ist der einundzwanzigste August.“

Eine Reanalyse der Vorhersagefehler ergab keine Evidenz für eine rhythmisch zu begründende *Schlagtilgung* auf Äußerungsebene (siehe aber Abschnitt 5.2). Daher wurde beschlossen, diese zunächst nicht zusätzlich zu modellieren.

4.5 Eine formale Grammatik zur Generierung deutscher Äußerungsbetonung

Die Analysen für eine erfolgreiche Vorhersage der Prominenz auf Äußerungsebene ergaben abschließend folgendes Regelwerk für die Generierung metrischer Gitter.

Bestätigte Prominenzregel 1

Jede Silbe erhält einen Basisschlag.

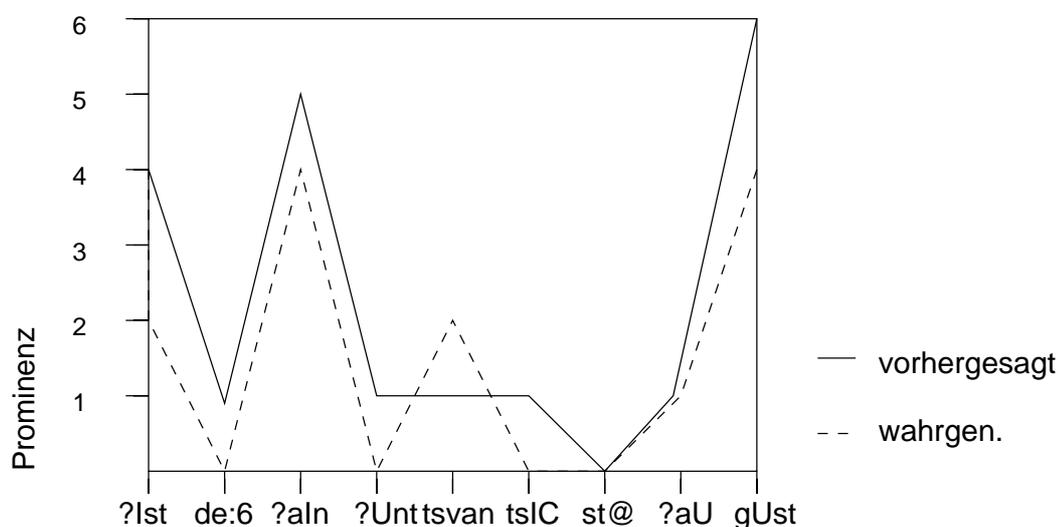


Abbildung 4.7: Vorhersage ohne Schließen rhythmischer Lücken

Bestätigte Prominenzregel 2

Jede Silbe, deren Nukleus nicht aus einem reduzierten Vokal (/ə/ oder /ɐ/) oder einem silbischen Konsonanten (/ŋ/, /m/ oder /ŋ/) besteht, erhält einen Schlag.

Bestätigte Prominenzregel 3

Die prominenteste Silbe eines Wortes erhält einen Prominenzwert entsprechend der Prominenzkategorie des Wortes, dessen Teil sie ist.

Bestätigte Prominenzregel 4

Die prominenteste Silbe innerhalb des letzten Inhaltswortes einer prosodischen Phrase erhält einen zusätzlichen Schlag, außer es handelt sich um ein Verb. Verben erhalten einen zusätzlichen Schlag, wenn ihnen nicht ein Nomen, Adjektiv oder Adverb folgt.

Bestätigte Prominenzregel 5

Die prominenteste Silbe eines äußerungsinitialen Funktionswortes erhält einen zusätzlichen Schlag.

Bestätigte Prominenzregel 6

Wenn mehr als zwei betonbare Silben niedriger Prominenz (Prominenzniveau 1) aufeinander folgen, so füge der zweiten und jeder weiteren übernächsten Silbe dieser Folge einen Schlag hinzu.

Die Formalisierung der metrischen Regeln erfolgt in Form von konkatenierten endlichen Übergangnetzwerken gemäß Definition 2. Das Ausgabeband des ersten Automaten ergibt das Eingabeband des zweiten Automaten usw.

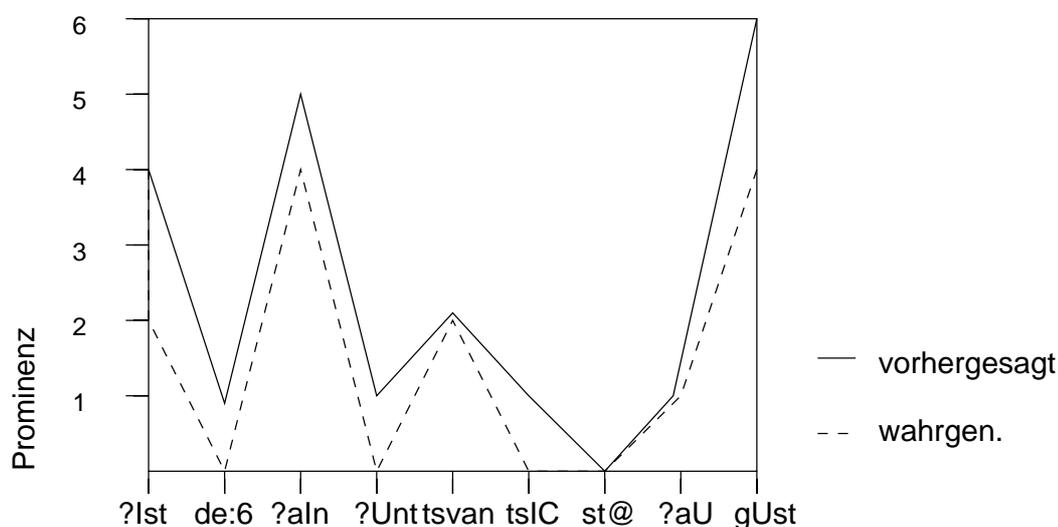


Abbildung 4.8: Vorhersage mit Schließen rhythmischer Lücken

4.5.1 Modellierung von Regel 1 zur Vorhersage metrischer Prominenz

In dieser Regel wird jeder Silbe ein Grundschlag zugewiesen.

Das **Eingabealphabet** für den ersten Automaten besteht aus Silben, wobei jede Silbe mit einem Tripel bestehend aus den Attributen \langle Akzentuierbarkeit; Wortbetonung; Wortklasse \rangle annotiert ist. Die Attribute *Akzentuierbarkeit*, *Wortbetonung* können die Werte 0 oder 1 annehmen. Das Attribut *Wortklasse* kann die skalaren Werte *ART(ikel)*, *KONJ(unktion)*, *SUBJ(unktion)*, *PRAEP(osition)*, *MOD(alverb)*, *AUX(iliarverb)*, *NEG(ationspartikel)*, *AFF(irmationspartikel)*, *V(erb)*, *DEM(onstrativpronomen)*, *W – Frage(pronomen)*, *ADV(erb)*, *ADJ(ektiv)*, *N(omen)*, *E(igen)N(ame)* oder *NUM(eralium)* annehmen.

Das **Ausgabealphabet** des ersten Automaten besteht aus Silben, die mit einem Quadrupel annotiert sind, welches die Attribute \langle Akzentuierbarkeit, Wortbetonung, Wortklasse, Prominenz \rangle enthält. Das Attribut *Prominenz* kann in diesem Fall nur den Wert 0 annehmen. Im Prinzip stellt dieser *Constraint* nur die Lizenzierung des Prominenz-Attributs als Eigenschaft von Silben dar. Alle Zustandsmengen sind identisch und enthalten jeweils den Zustand q_0 . Das Übergangsnetzwerk wird illustriert in Abbildung 4.9. Werte, die nicht in die Entscheidung einfließen, sind in dieser und den folgenden Abbildungen aus Übersichtsgründen nicht spezifiziert.

<Akzentuierbarkeit;Wortbetonung;Wortklasse>
 <Akzentuierbarkeit;Wortbetonung;Wortklasse,Prominenz:0>

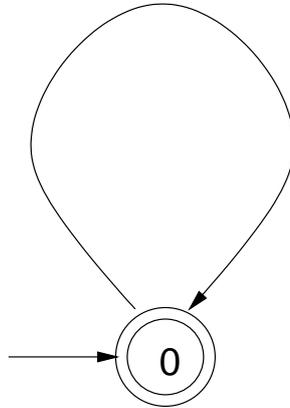


Abbildung 4.9: Ein Übergangnetzwerk für die Zuweisung von Basisschlägen

4.5.2 Modellierung von Regel 2 zur Vorhersage metrischer Prominenz

In dieser Regel wird jeder akzentuierbaren Silbe ein weiterer Schlag zugewiesen. Das **Eingabealphabet** definiert sich durch das **Ausgabealphabet** von Regel 1. Im **Ausgabealphabet** dieses Automaten können die *Prominenz*-Attribute die Werte 1 oder 2 annehmen. Der Automat hat wiederum nur den einen Zustand q_0 . Die Wertezuweisung für die verschiedenen Eingabebedingungen wird illustriert in Abbildung 4.10.

<Akzentuierbarkeit:1;Prominenz:0>:<Akzentuierbarkeit:1;Prominenz:2>
 <Akzentuierbarkeit:0;Prominenz:0>:<Akzentuierbarkeit:1;Prominenz:1>

Abbildung 4.10: Die Zuweisung von Zusatzschlägen auf betonbare Silben

4.5.3 Modellierung von Regel 3 zur Vorhersage metrischer Prominenz

Diese Regel weist Silben, die die Wortbetonung tragen, einen wortklassenspezifische Prominenzwert zu.

Das **Eingabealphabet** definiert sich durch das **Ausgabealphabet** von Regel 2. Das **Ausgabealphabet** dieses Automaten unterscheidet sich von seinem Eingabealphabet dadurch, daß das *Prominenz*-Attribut die Werte 2, 3, 4, 5 oder 6 annehmen kann. Der Automat hat wiederum nur den einen Zustand q_0 . Die Wertezuweisung für die verschiedenen Eingabebedingungen wird illustriert in Abbildung 4.11.

<Wortbetonung:0;Prominenz:1>:<Prominenz:1>
 <Wortbetonung:0;Prominenz:2>:<Prominenz:2>
 <Wortbetonung:1;Wortklasse:<DET|CONJ|SUBJ|PREP>;Prominenz:2>:<Prominenz:2>
 <Wortbetonung:1;Wortklasse:<MOD-V|AUX|NEG-P|AFFIRM)>;Prominenz:2>:<Prominenz:
 <Wortbetonung:1;Wortklasse:<V|DEM-P|WH-P>;Prominenz:2>:<Prominenz:4>
 <Wortbetonung:1;Wortklasse:<ADJ|ADV>;Prominenz:2>:<Prominenz:5>
 <Wortbetonung:1;Wortklasse:<N|EN|NUM>;Prominenz:2>:<Prominenz:6>

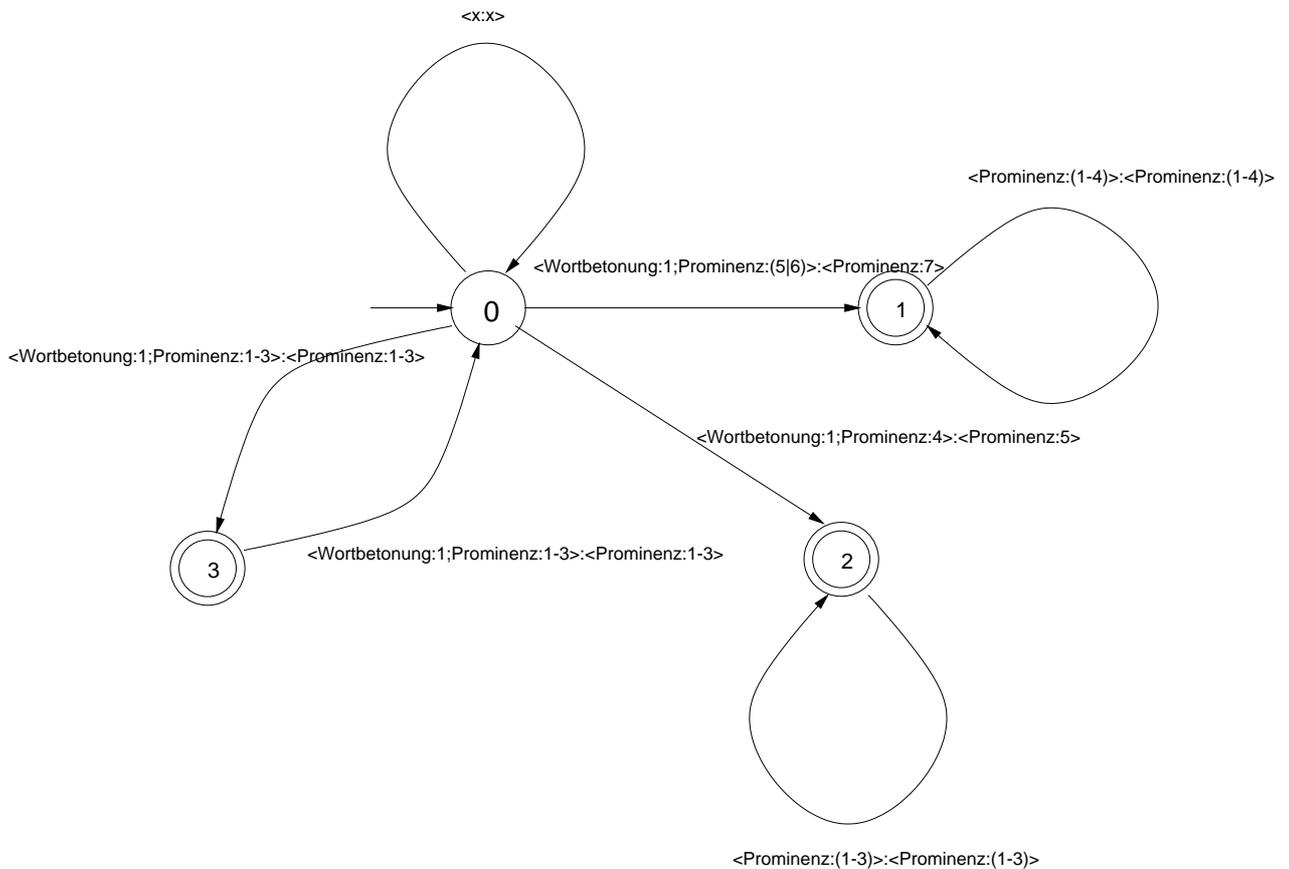
Abbildung 4.11: Die Zuweisung von wortklassenspezifischen Prominenzwerten auf Silben, die die Wortbetonung tragen

4.5.4 Modellierung von Regel 4 zur Vorhersage metrischer Prominenz

Diese Regel weist hauptbetonten Silben des letzten Inhaltswortes in der Äußerung unter bestimmten Voraussetzungen einen weiteren Schlag (Nuclear Stress) zu. Das **Eingabealphabet** definiert sich durch das **Ausgabealphabet** von Regel 3. Das **Ausgabealphabet** dieses Automaten unterscheidet sich von seinem Eingabealphabet dadurch, daß das *Prominenz*-Attribut von Silben der Form $\langle 1, 1, V \vee DEM - P \vee WH - P \vee ADV \vee ADJ \vee N \vee EN \vee NUM, prominenz \rangle$ den Wert 5, 6 oder 7 annehmen kann, jeweils einen Schlag mehr als das korrespondierende Eingabeattribut. Der Automat hat einen Startzustand q_0 , sowie drei weitere Zustände q_1 , q_2 und q_3 , die gleichzeitig Endzustände sind. Das Übergangsnetzwerk wird illustriert in Abbildung 4.12. Es ist zu beachten, daß dieser Automat nicht nur solche Äußerungen akzeptiert, in denen mindestens ein Inhaltswort enthalten ist. Äußerungen, die lediglich aus der Negationspartikel „nein“ bestehen, würden vom Automaten akzeptiert, aber einfach nicht manipuliert.

4.5.5 Modellierung von Regel 5 zur Vorhersage metrischer Prominenz

Diese Regel weist hauptbetonten Silben von äußerungsinitialen Funktionswörtern einen zusätzlichen Schlag zu. Das **Eingabealphabet** definiert sich durch das **Ausgabealphabet** von Regel 4. Das **Ausgabealphabet** dieses Automaten unterscheidet sich von seinem Eingabealphabet dadurch, daß das *Prominenz*-Attribut von Silben der Form $\langle 1, 1, V \vee ART \vee KONJ \vee SUBJ \vee PRAEP \vee NEG \vee AFF \vee MOD - V \vee AUX, prominenz \rangle$ einen Wert 4 statt 3 bzw. 3 statt 2 annimmt. Der Automat hat einen Startzustand q_0 und zwei weitere Zustände q_1 und q_2 , welche gleichzeitig Endzustände sind. Das Übergangsnetzwerk wird illustriert in Abbildung 4.13. In der Ausgabe ist aus Gründen der Übersichtlichkeit lediglich das Prominenzattribut dargestellt.

Abbildung 4.12: Ein Übergangnetzwerk für die Modellierung der *Nuclear Stress Rule*

4.5.6 Modellierung von Regel 6 zur Vorhersage metrischer Prominenz

Diese Regel sorgt für einen alternierenden Rhythmus bei Folgen mehrerer unbetonter Silben.

Das **Eingabealphabet** definiert sich durch das **Ausgabealphabet** von Regel 5. Das **Ausgabealphabet** dieses Automaten unterscheidet sich von seinem Eingabealphabet dadurch, daß das *Prominenz*-Attribut von Silben der Form $\langle 1, 0, \text{wortklasse}, 1 \rangle$ den Wert $\langle 1, 0, \text{wortklasse}, 2 \rangle$ annehmen kann. Der Automat hat einen Startzustand q_0 , fünf Endzustände q_3, q_4, q_5, q_6, q_7 sowie zwei weitere Zustände q_1, q_2 . Beginnt die Äußerung mit zwei oder mehr aufeinanderfolgenden unbetonten Silben, so wird die erste, die vierte und jede weitere gerade Silbe betont. Ansonsten wird in Folgen von drei oder mehr unbetonten Silben die zweite und jede weitere gerade Silbe betont. Dies bedeutet, daß in den meisten Fällen binäre Betonungsmuster bevorzugt werden, ternäre Betonungsmuster aber am Anfang einer Äußerung zugelassen sind. Das Übergangnetzwerk wird illustriert in Abbildung

<Wortbetonung:0;Prominenz:(1-3)>:<Prominenz:(1-3)>

x:x

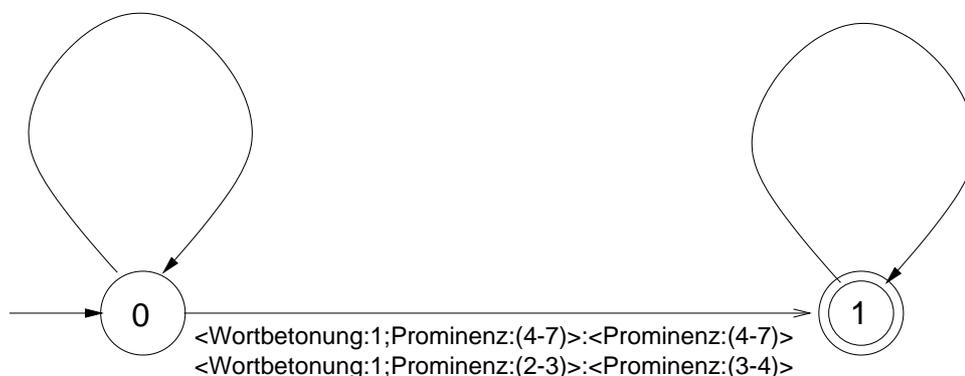


Abbildung 4.13: Ein Übergangsnetzwerk für die Zuweisung von Zusatzschlägen auf äusserungsinittiale Funktionswörter

4.14. Zur besseren Übersicht sind in der Abbildung lediglich die Werte des Prominenzattributs dargestellt. Diejenigen Übergänge, bei denen alle Prominenzwerte akzeptiert werden, die nicht den Wert 1 haben, sind mit “!” annotiert. Diese Werte werden auf dem Ausgabeband nie verändert.

4.6 Fazit

In diesem Abschnitt wurde ein verbessertes Vorhersagemodell für Prominenzen auf Äußerungsebene entwickelt. Den Ausgangspunkt bildete dabei die Theorie von (SELKIRK 1984). Die Prominenz auf Wortebene wird dabei als wortklassenspezifische lexikalische Eigenschaft betrachtet. Zwei Adaptionen dieser Theorie auf das Deutsche wurden miteinander verglichen. Hierbei stellte sich heraus, daß eine wortklassenspezifische Vorhersage der Äußerungsprominenzen einer rein auf syntaktischen Phrasen basierenden Vorhersage überlegen ist. Euphonierregeln spielen offenbar bei der Betonung auf Äußerungsebene lediglich eine marginale Rolle, insbesondere für die Vermeidung von Akzentzusammenstößen konnte keine Evidenz gefunden werden. Das Hinzufügen von Schlägen zur Vermeidung rhythmischer Lücken führte allerdings durchaus zu einer Verbesserung der Vorhersage. Einige der phonologischen Regeln konnten damit durchaus bestätigt werden. Um eine befriedigende Vorhersage von Prominenzen zu erreichen, wurde das Vorhersagemodell auf der Basis sprachtechnologischer sowie korpuslinguistischer Forschungsergebnisse verbessert. So wurde die ursprüngliche Unterteilung in Funktions- und Inhaltswörter durch eine feinere Klassifikation ersetzt. Die Vermutung über einen Zusammenhang zwischen Prominenz und Frequenz in deutschen Referenzkorpora konnte hingegen nicht erhärtet werden. Die bestätigten Prominenzregeln führten zu einer Vorhersage, die im Bereich der in Abschnitt 3.7 festgelegten Korrelations-Baseline lag. Abschließend wurden die Vorhersageregeln mit Hilfe konkatenierter endlicher Über-

gangsnetzwerke formal dargestellt.

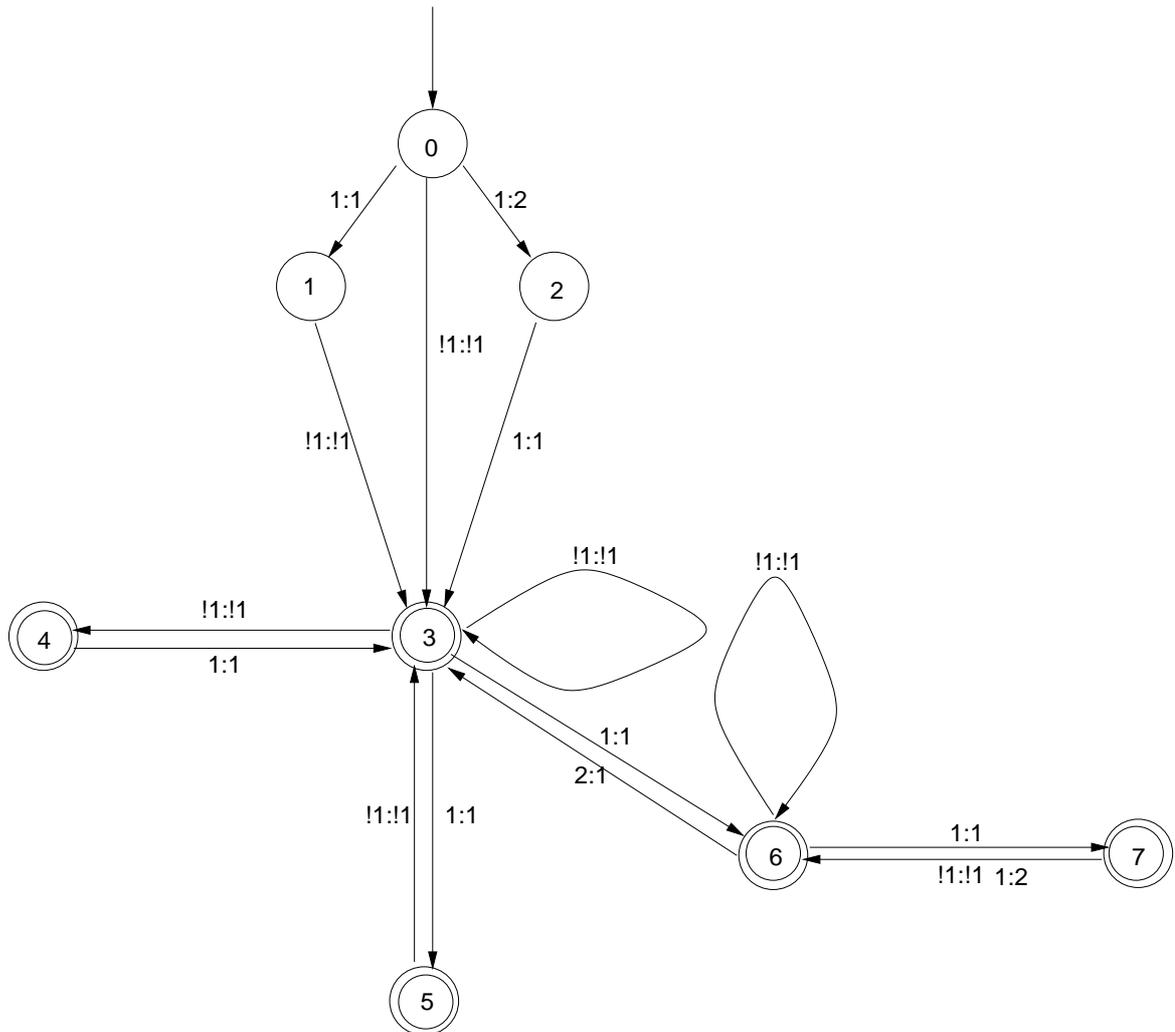


Abbildung 4.14: Ein Übergangnetzwerk zur Vermeidung rhythmischer Lücken

Kapitel 5

Prominenzmuster auf der Wortebene

Raistlin setzte sich mit dem Buch hin und brütete über den Spruch. Er sprach jeden Buchstaben laut aus, dann sagte er jedes Wort laut und wiederholte die Wörter, bis sie für ihn richtig klangen, wie ein Sänger mit perfektem Gehör seine Laute stimmt. Das machte er sehr gut und war ziemlich stolz auf sich, bis er an das siebente Wort kam. Das siebente Wort des Spruches war eines, das er noch nie gehört hatte. Es konnte auf mehrere verschiedene Weisen betont werden, was ihm jedesmal eine andere Bedeutung verlieh. Welche Aussprache war wohl die richtige?

Margaret Weis, Die Zauberprüfung

Im letzten Abschnitt wurde die in Kapitel 3 entwickelte Methode verwendet, um ein Vorhersagemodell für die Prominenz auf Äußerungsebene zu evaluieren und anschließend zu erweitern. In diesem Kapitel soll die Methode zusätzlich auf die Vorhersage von Wortbetonungen angewendet werden. Da die Datenbank, auf der die vorherige Evaluation aufbaute, lediglich über ein sehr eingeschränktes Lexikon verfügt (996 *Types*, 2312 *Tokens*), welche nicht den Formenreichtum deutscher Wortformen repräsentativ abdeckt, wird hierfür auf zusätzliches Datenmaterial in Form eines manuell etikettierten Lexikons zurückgegriffen. Selbstverständlich ist die in Kapitel 3 ermittelte *Baseline* auf diese Art von Datenmaterial nicht anwendbar. Es wird daher zunächst vereinfachend angenommen, daß sich die Sprecherinnen einer Sprache i.d.R. einig sind, auf welche Silbe die Hauptbetonung eines Wortes fällt, wenn es auch für diese Annahme Gegenbeispiele gibt (vgl. Abschnitt 5.2). Der Abschnitt 5.1 behandelt zunächst die Vorhersage der Wortbetonung auf Wörtern, die aus einem ggf. flektierten und/oder derivierten freien Morphem bestehen. Die Darstellung basiert auf den Arbeiten, die in (WAGNER 2001) präsentiert wurden. In Abschnitt 5.2 wird die phonologische Regel der „Akzentverschiebung“ überprüft, die im Deutschen laut phonologischer Analysen in erster Linie beim Aufeinandertreffen zweiter betonter Silben in Komposita auftritt. Es wurde bereits darauf hingewiesen, daß die intuitiv begründeten phonologischen Vorhersageregeln oft weit von den tatsächlich wahrzunehmenden Betonungsmustern abweichen. Dies könnte gerade den Bereich der Akzentverschiebung betreffen (vgl. Abbildung 1.1). Daher

wird an dieser Stelle experimentell überprüft, ob die Vorhersagen der metrischen Phonologie die tatsächliche Wahrnehmung der Hörer widerspiegelt. In Abschnitt 5.3 wird diskutiert, inwieweit das metrische Modell der Wortbetonung auf Eigennamen übertragen werden kann, da die Betonung von Eigennamen im Deutschen ein phonologisch noch weitgehend ungeklärtes Problem darstellt.

5.1 Monomorphematische, flektierte und derivierte Formen

Dieser Abschnitt behandelt die Güte von Vorhersageregeln für die deutsche Wortbetonung. Komposita werden vorerst aus der Untersuchung ausgeklammert, da für diese Sonderregeln gelten (vgl. Abschnitt 5.2). Die Studie stützt sich in erster Linie auf die Regeln, wie sie von (JESSEN 1998) aus verschiedenen phonologischen Arbeiten zusammengetragen wurden. Zunächst wird das Datenmaterial vorgestellt (vgl. Abschnitt 5.1.1), anschließend werden die zur Untersuchung stehenden Regeln eingeführt und getestet (vgl. Abschnitte 5.1.2 und 5.1.3). Aus den Ergebnissen ergeben sich Zweifel an der zunächst verwendeten Kategorisierung in *leichte* und *schwere* Silben. Aus diesem Grund wurde als Grundlage für weitere Forschung in dieser Richtung ein Vortest durchgeführt, um das Silbengewichtsmodell kritisch zu überprüfen (vgl. Abschnitt 5.1.5). Die Ergebnisse werden in Abschnitt 5.1.7 kurz zusammengefaßt.

5.1.1 Das Datenmaterial

Um die Vorhersagen der Wortbetonung an einem größeren Korpus testen zu können, wurde zusätzlich zur BPD auf die Verbmobil-Wortliste (LÜNGEN et al. 1998) zurückgegriffen. Diese Wortliste ist vergleichsweise klein, insbesondere verglichen mit einem großen zur Verfügung stehenden Aussprachelexikon wie dem CELEX-Korpus (CELEX 1995). Allerdings hat die Verbmobil-Wortliste den entscheidenden Vorteil, daß sie komplett manuell annotiert wurde. Das CELEX-Korpus hingegen ist wenigstens partiell automatisch transkribiert worden. Eine Evaluation auf der Basis von CELEX würde daher u.U. auf einen Vergleich verschiedener Vorhersagealgorithmen hinauslaufen, was jedoch nicht Zweck der hier beschriebenen Untersuchung ist. Vor der Untersuchung wurden alle Komposita unter Zuhilfenahme der Morphemgrenzen automatisch aus der Liste entfernt, da die Betonung von Komposita nach anderen Regeln verläuft als die affigierter freier Morpheme (vgl. Abschnitt 5.2). Die resultierende Wortliste bestand aus 5385 Vollformen. Nach der Entfernung der Komposita enthielt die Liste noch immer eine große Anzahl morphologisch komplexer Vollformen, bestehend aus jeweils einem freien Wurzelmorphem und einem oder mehreren Flexions- oder Derivationssuffixen. Auch eine Vielzahl von Präfixen verblieben in der Wortliste. Jeder Eintrag in der Wortliste wird durch seine orthographische Form, seine phonetische Oberflächenform inklusive Haupt-

und evtl. Nebenbetonungen sowie durch Morphem- und Silbengrenzen dargestellt (vgl. Abbildung 5.1).

```
' `Agypten ?E.g'Yp.t+@n
' `Ahnliches ?'E:n.+lI.C#+@s
' `Amter ?'Em.t#+6
' `Anderung ?'En.d@.r+UN
' `Anderungen ?'En.d@.r+U.N#+@n
.
.
```

Abbildung 5.1: Fragment aus der verwendeten Wortliste

5.1.2 Das Regelsystem von (JESSEN 1998)

Die Untersuchung soll sich in erster Linie auf die in (JESSEN 1998) zusammengetragenen Regeln stützen, da diese einen repräsentativen Überblick der im Rahmen der metrischen Phonologie erarbeiteten Betonungsregeln bieten. Zwar stützt sich Jessen nicht ausschließlich auf die Arbeiten zur Metrischen Phonologie, sondern berücksichtigt auch die generativ geprägten Regelsysteme von (WURZEL 1980) oder (KIPARSKY 1966). Allerdings sind diese Ergebnisse auch von metrisch geprägten Phonologen (FÉRY 1986; GIEGERICH 1984a; GIEGERICH 1984b; GIEGERICH 1985; WIESE 1996) rezipiert und entsprechend verändert ebenfalls modelliert worden. Der von Jessen gegebene Überblick erscheint für die hier dargelegte Studie des weiteren besonders geeignet, da er die Gültigkeit seiner Regeln an eine hohe empirische Abdeckung knüpft, wie es in der vorliegenden Arbeit ebenfalls geschieht.

Die Zuweisung deutscher Wortbetonung auf geeignete Trägersilben verläuft bei Jessen „von rechts nach links“, d.h., zuerst wird die letzte Silbe eines Wortes hinsichtlich ihrer Fähigkeit, die Wortbetonung zu tragen, analysiert. Falls diese nicht in Frage kommt, wird die vorletzte Silbe analysiert u.s.w. Weiterhin unterscheidet Jessen in seinen Regeln zwischen *leichten* und *schweren* Silben. Eine Unterscheidung zwischen diesen beiden Silbentypen hat eine lange Tradition in der phonologischen Literatur, es besteht aber eine große Unstimmigkeit, wann genau eine Silbe tatsächlich *schwer* bzw. *leicht* ist (vgl. die Diskussion in (MENGEL 1998)). In dieser Arbeit wird zunächst die Unterscheidung gemäß der Definition von (FÉRY 1998, 3) vorgenommen:

Open syllables, which always have a tense vowel in their nucleus, as well as syllables with a lax vowel and a single closing consonant or glide, are light. Syllables with a tense vowel and a closing consonant and those with a lax vowel and two closing consonants (CVCC) are heavy.

In der sogenannten CV-Notation werden ungespannte Vokale¹ als *V*, gespannte Vokale² und Diphthonge³ als *VV* und jeder Konsonant als *C* dargestellt. Gemäß dieser Vorgehensweise erhalten die in Tabelle 5.1 abgebildeten Wörter ihre *schwer-leicht*-Muster.

Wort (orth.)	IPA-Transkription	CV-Struktur	schwer/leicht-Muster
Kompost	k'ɔm.pɔst	CVC.CVCC	leicht.schwer
Zebra	ts'e:bra:	CVV.CVV	leicht.leicht
Wagnis	v'a:k.nɪs	CVVC.CVC	schwer.leicht
extrem	ʔɛks.tr'e:m	CVCC.CCVVC	schwer.schwer

Tabelle 5.1: Zuweisung der Silbeneigenschaften *schwer* und *leicht*

Für gewöhnlich werden unbetonbare Silben⁴ noch „leichter“ als leichte Silben eingestuft, so daß sich in puncto Silbengewicht folgende Hierarchie aufstellen läßt:

1. **schwa-Silben:** ə, ɐ, ɪ, ɪ̯, ʊ, ʊ̯ als Nukleus
2. **leichte Silben:** C⁺V[V|C]⁵,
3. **schwere Silben:** C⁺V[V|C]C⁺

Die Komplexität des Silbenanlautes spielt bei der Zuweisung des Silbengewichts keinerlei Rolle.

5.1.2.1 Die Regeln

Die Vorhersage der Betonung erfolgte nach deren Implementierung also „von rechts nach links“ durch die Anwendung der nachfolgend erläuterten Regeln.

¹Dies sind im Deutschen die Vokale a, ɛ, ɪ, ʏ, ɔ, ʊ, œ.

²Dies sind im Deutschen die Vokale ai, ɛi, ei, ii, yi, oi, ui, øi.

³Dies sind im Deutschen die Diphthonge ay, ɔy, əy.

⁴Dies sind schwa-Silben sowie Silben mit silbentragenden Nasalen oder Lateralen wie in *tragen* (/tr'a:ɡŋ/) oder *Hagel* (/h'a:ɡl/).

⁵Das hochgestellte „+“ soll ein Kleene-Plus darstellen, wie es aus der Notation regulärer Ausdrücke bekannt ist. Ein „+“ bedeutet also, daß das Zeichen, dem es folgt, mindestens einmal vorkommen muß, aber beliebig oft vorkommen darf. Die Folgen in eckigen Klammern stellen obligatorische Bestandteile des Ausdrucks dar, die entweder die Form *C* oder *V* annehmen können.

5.1.2.2 Vorverarbeitung: Abtrennung extrametrischer Suffixe und Resilbifizierung

Es wird allgemein angenommen (u.a. (FÉRY 1986)), daß eine große Anzahl von Suffixen sowie einige Präfixe im Deutschen generell unbetonbar sind. Diese Affixgruppe ist beinahe ausschließlich germanischen Sprachursprungs und darf bei einer Vorhersage der Betonung nicht berücksichtigt werden. Aus diesem Grund werden in einem ersten Schritt zur Betonungsbestimmung genau diese Affixe vom Restwort abgetrennt. Häufig erfordert dieser Prozeß eine Resilbifizierung des Restwortes, die wiederum Einfluß auf die Silbengewichte des Wortes haben kann (vgl. Tabelle 5.2). Daher geschieht die Bestimmung der Silbengewichte erst im Anschluß an die Abtrennung unbetonbarer Affixe.

Beschreibungsebene	Ursprüngliche Struktur	Nach Affixtrennung und Resilbifizierung
orthographisch	geniales	genial
morphologisch	genial+es	genial
phonetisch	gen.j'a:ləs	gen.j'a:l
Silbengewicht	leicht-leicht-schwa	leicht-schwer

Tabelle 5.2: Beeinflussung des Silbengewichts durch Affixabtrennung

Die Liste von abzutrennenden Affixen wurde wie die Regeln selbst nach den Angaben von (JESSEN 1998) erstellt und anhand eigener Analysen ergänzt. Tabelle 5.3 enthält die Liste der unbetonbaren abgetrennten Suffixe. Diese Liste ist allerdings orthographisch. An der phonetischen Oberfläche jedoch kann ein und dieselbe orthographische Form signifikante Variation anzeigen (vgl. Tabelle 5.4). Solche Abweichungen kommen einerseits durch morphophonologische (regelmäßige) Alternation, andererseits aber auch durch abweichende Transkriptionsstrategien zustande. Die Liste der abzutrennenden Suffixe muß daher beträchtlich erweitert werden (s. Anhang B).

-bar, -be, -ber, -chen, -el, -elnd, -elst, -elt, -em, -en, -end, -ens, -ent, -ents, -er, -el, -erl, -ern, -erst, -ert, -et, -es, -est, -ge, -gen, -gung, -haft, -heit, -ig, -igt, -ig, -igs, -in, -ir, -is, -isch, -keit, -lein, -ler, -lig, -ling, -los, -mäßig, -men, -nen, -ner, -nis, -sam, -schaft, -sche, -se, -sel, -ste, -stem, -sten, -te, -tel, -tem, -ten, -tes, -tig, -tum, -tüm, -ung, -werk, -wärts, -zeug, -zig

Tabelle 5.3: Extrametrische Suffixe des Deutschen

Orthographische Form	Phonetische Variante	Beispiel
zig	tsɪç	ʔ'am.tɪç
zig	tsɪg	ʔ'am.tɪ.gə
er	ɐ	b'au.ɐ
er	əɐ̯	b'au.əɐ̯
er	ər	b'ɔy.ə.rɪn

Tabelle 5.4: Transkriptionsvarianten orthographisch identischer Suffixe

Ähnlich wie extrametrische Suffixe sind einige Präfixe im Deutschen nie betont. Folgende dieser Präfixe wurden entsprechend ebenfalls vor der Anwendung der eigentlichen Betonungsregeln vom Restwort abgetrennt:

be-, er-, ent-, ge-, hinter-, inter-, ver-, zer-

Ferner existieren im Deutschen einige Präfixe, die die Wortbetonung an sich ziehen. Diese wurden ebenfalls berücksichtigt:

ab-, an-, auf-, bei-, dar-, ein-, empor-, fort-, her-, hin-, los-, mit-, nach-, nieder-, weg-, weiter-, vor-, zu-

Einige weitere Präfixe können sowohl betont als auch unbetont auftreten, z.B.:

durch-, hinter-, un-, unter-, voll-, wieder-, um-, über-

Eine Entscheidung über Betontheit läßt sich aber nicht anhand des isolierten Auftretens als Infinitivform in einer Wortliste treffen:

„Die Konkurrenz will das Projekt *dú*rchsetzen.“ vs. „Die Konkurrenz will das Projekt *durchsét*zen.“

Erkennbar sind betonte Präfixe erst an ihrer Abtrennbarkeit im Äußerungskontext:

„Die Konkurrenz *setzt* das Projekt *durch*.“ vs. „Die Konkurrenz *durchsetzt* das Projekt.“

Die hierdurch entstehenden Vorhersagefehler werden in Kauf genommen und gehen in die später berechnete Fehlerrate mit ein.

5.1.2.3 Basisregeln

Die Grundregeln der Vorhersage lauten folgendermaßen:

1. Silben, die ein schwa oder einen silbischen Konsonanten als Silbennukleus enthalten, werden niemals betont.
2. Wenn nach der Abtrennung der unbetonbaren Affixe sowie der erfolgten Re-silbifizierung das verbliebene Wort einsilbig ist, so wird dieser Silbe trivialerweise die Wortbetonung zugewiesen.

5.1.2.4 Regel 1: Approximantregel

Wenn die (leichte) Endsilbe eines Wortes keinen Silbenanlaut besitzt und die Paenultimasilbe auf einen hohen Vokal⁶ auslautet, wird dieser Vokal vermutlich als Approximant realisiert. In diesem Fall kann der Vokal zum Silbenanlaut der Endsilbe werden und darf demzufolge nicht gleichzeitig Träger der Wortbetonung sein. Daher rutscht in diesen Fällen die Wortbetonung auf die Antepaenultima.

$$\begin{aligned} \text{Júlia:} & /j'u:li.a/ \rightarrow [j'u:l.ja] \\ \text{Jáguar:} & /j'a:gu.a:ɾ/ \rightarrow [j'a:gu.a:ɾ] \end{aligned}$$

5.1.2.5 Regel 2: Schwere-Endsilben Restriktion

Endet ein Wort auf einer schweren Silbe, so wird dieser Silbe die Wortbetonung zugewiesen. Dies ist auch dann der Fall, wenn die Paenultima ebenfalls schwer ist. Beispiele für die Korrektheit dieser Regel finden sich in den folgenden Wörtern (schwere Silbe ist fettgedruckt):

E.le.fánt: leicht.schwa.schwer
ex.trém: schwer-schwer
Pro.spékt: leicht-schwer

5.1.2.6 Regeln 3a und 3b: Regeln der Paenultimabetonung

Kommt keine der o.a. Regeln zur Anwendung, so wird die Wortbetonung der Paenultimasilbe zugewiesen, vorausgesetzt, diese ist betonbar und fällt somit nicht unter die Einschränkung der Basisregeln. Für die Paenultimaregel finden sich im Deutschen zahlreiche Beispiele. Dies ist vermutlich auch der Grund dafür, daß die Paenultimabetonung häufig als das *Default*-Muster der deutschen Wortbetonung angesehen wird. Folgende Beispiele unterstützen diesen Standardfall:

Al.go.ríth.mus, Co.lo.rá.do, Ká.ffee, Kú.ckuck

Die so beschriebene (und implementierte) Regel stellt eine Vereinfachung von zwei Regeln aus (JESSEN 1998) dar, nämlich die „Restriktion der finalen Schwasilbe“ sowie die „Restriktion der geschlossenen Endsilben“. Beide führen zu einer Betonung der Paenultima und sind im hier vorgestellten Regelsystem implizit enthalten. Sie sollen im folgenden skizziert werden:

Regel 3a: Restriktion der finalen Schwasilbe Wenn die Endsilbe unbetonbar ist, fällt die Wortbetonung auf die Paenultima, sofern diese betonbar ist. Diese Regel wird empirisch gestützt durch folgende Wörter (unbetonbare Silbe ist fettgedruckt):

Ka.lén.der, A.le.xán.der, Man.da.rí.ne

⁶Dies sind im Deutschen u, u, i, l, y, ʏ.

Regel 3b: Restriktion der geschlossenen Endsilbe Diese Restriktion besagt, daß eine Paenultimasilbe, die einen Kodakonsonanten besitzt, ebenfalls die Betonung an sich zieht. Ausnahme sind (hier) Wörter, die eine schwere Endsilbe besitzen und somit Regel 2 unterliegen.

5.1.2.7 Regel 4: Regel der nächsten betonbaren Silbe

Wenn keine der o.a. Regeln zur Anwendung kommt, so fällt die Wortbetonung auf die nächste betonbare Silbe zur Linken. Diese Regel wird gestützt durch Wörter mit leichter Endsilbe und unbetonbarer schwa-Silbe als Paenultima:

Kár.ne.val, cé.te.ra

Eine solche Regel wird von (JESSEN 1998) nicht explizit vorgeschlagen, aber da jedem Wort eine Betonung zugewiesen werden muß, wurde sie dennoch eingeführt.

5.1.3 Ergebnisse der Vorhersage

Wenn alle Regeln, die oben erläutert wurden, verwendet werden, so beträgt die Fehlerrate der Vorhersage 4,17%. Diese Ergebnisse entsprechen in ihrer Genauigkeit solchen Vorhersagen, die auf der Basis neuronaler Netze getroffen wurden (vgl. hierzu die Ergebnisse von (HAIN und ZIMMERMANN 2001)). Im Gegensatz zu diesen Ergebnissen haben sie aber den Vorteil einer wesentlich einfacheren und eleganten Modellierung, die zudem explizite Aussagen über die Struktur der Sprache macht. Gemäß dem Einfachheitskriterium, welches in Abschnitt 3.1.3 aufgestellt wurde, ist die regelbasierte Modellierung daher vorzuziehen.

Alle aufgestellten Regeln führen zu einer Verbesserung der Vorhersage und können als relevant bestätigt werden. Läßt man die Approximantregel (Regel 1) weg, steigt die Fehlerrate leicht an auf 4,32%. Dies ist nicht verwunderlich, da dieses Muster im Deutschen nur recht selten vorkommt. Läßt man hingegen die Regel der schweren Endsilbe weg (Regel 2), so steigt die Fehlerrate dramatisch an und erreicht 9,22%. Die Paenultimaregel hingegen ist weniger einflußreich, indem ihr Weglassen die Fehlerrate lediglich auf 5,52% ansteigen läßt. Die Regel der nächsten betonbaren Silbe (Regel 5) führt zu seiner geringfügigen Verbesserung (4,73% Fehlerrate bei Weglassen). Von einem linguistischen Standpunkt, der auf eine möglichst hohe empirische Abdeckung zielt, ergibt sich somit folgende Hierarchie hinsichtlich des Einflusses der verschiedenen Regeln:

1. Restriktion der schweren Endsilbe
2. Paenultimaregel
3. Regel der nächsten betonbaren Silbe
4. Approximantregel

Es ist zu beachten, daß die Regeln keiner eingehenden Prüfung unterzogen worden sind, inwieweit sie miteinander in Konflikt geraten können. Eine solche Hierarchie macht noch keinerlei Aussagen über die Relevanz der einzelnen Regeln für die kognitive Verarbeitung bzw. ihren Status innerhalb der sprachlichen *Kompetenz*, da die hier verfolgte Methode zunächst eine möglichst einfache deskriptiv adäquate Darstellung verfolgt. Dennoch sollten diese Zahlen bei einer möglichen weitergehenden phonologischen Modellierung, die ferner auch auf *explanatorische Adäquatheit* zielt, berücksichtigt werden.

5.1.4 Fehleranalyse

Eine Fehleranalyse dient der Aufdeckung systematischer Lücken und Vorhersagefehler in der Regelmenge und sollte demnach zu einer Verfeinerung bzw. Vereinfachung des ursprünglichen Regelsystems beitragen. Außerdem kann es Anhaltspunkte liefern, welche Wörter tatsächliche Ausnahmefälle sind, die regelhaft nicht modelliert werden können und daher in ein Ausnahmelexikon aufgenommen werden sollten. Ausnahmefälle werden entsprechend dem in Abschnitt 3.1.3 definierten Einfachheitskriterium genau dann angenommen, wenn eine Regel eigens für diesen Fall zugeschnitten werden muß oder aber in den überwiegenden Fällen zu einer falschen Vorhersage und damit zu einem größeren Vorhersagefehler führt. Nach einer Untersuchung der Vorhersagefehler wurden diejenigen Fehler aus der Liste entfernt, die ohne weiteres morphosyntaktisches Wissen nicht adäquat modelliert werden können. In diese Kategorie fallen beispielsweise Wörter, die mit dem Präfix „unter“ beginnen, für welches es sowohl betonte als auch unbetonte Varianten gibt (unterbieten vs. únterschieben). Die restlichen fehlerhaft vorhergesagten Wörter wurden in drei Hauptklassen unterteilt:

- *Fehlerklasse I*: Wörter, die leichte, aber betonte Endsilben enthalten
- *Fehlerklasse II*: Wörter, die schwere, aber unbetonte Endsilben enthalten
- *Fehlerklasse III*: Wörter, in denen die Betonung auf die Antepaenultimasilbe fällt

Die aufgetretenen Fehler werden in den folgenden Abschnitten näher beleuchtet.

5.1.4.1 Fehlerklasse I: leichte betonte Endsilben

Die häufigsten Vorhersagefehler fallen in diese Klasse, weshalb die Suche nach systematischen Vorhersagelücken hier besonders gerechtfertigt erscheint. Die Gruppe kann wiederum in zwei Unterklassen aufgeteilt werden:

- *Fehlerklasse Ia*: Wörter, die auf einem Langvokal oder Diphthong enden, aber keine Silbenkoda haben (-VV)

- Fehlerklasse Ib: Wörter, die auf einer Silbe mit ungespanntem Vokal und einem einzelnen Kodakonsonanten enden (-VC)

Fehlerklasse Ia: -VV In diese Klasse fallen folgende Wörter:

Etát, Tabú, Kopíe, Regíe, Filét, Café, Idée, apropós, Parté

Mit Ausnahme des Wortes *Tabu*⁷ sind alle Wörter dieser Liste Lehnwörter französischen Ursprungs und haben ihre Originalbetonung beibehalten. Es erscheint nicht sinnvoll anzunehmen, daß Wörter, die auf einem Langvokal enden, regelhaft endbetont sind, da die Liste der Wörter, die dieses Muster aufweisen, aber nicht endbetont sind, wesentlich länger ist als die Liste derer, die Endbetonung haben, z.B.:

Káffee, Káro, Túba, Áuto, Dónau, Fóto, Lótto, Kíno, Kómma, palétti, Úni, Colorádo, ...

Außerdem weist das Deutsche einige Lehnwörter französischen Ursprungs auf, welche zwar auf -VV enden, aber sowohl endbetont als auch paenultimabetont ausgesprochen werden:

Bistro, Filet, Büro

Von einem diachronen Standpunkt aus betrachtet, befinden sich diese Wörter offensichtlich noch in einem Prozeß der Assimilierung an das deutsche prosodische System. Diese Variation zwischen Lehnwörtern mit Ursprungsbetonung und angepaßter Betonung führte sogar zu einem prosodischen Minimalpaar im (Hoch-)Deutschen⁸:

Káffee vs. Café

Zusammenfassend weisen all diese Indizien darauf hin, daß es nicht erforderlich ist, die Endbetonung auf Langvokal endender Wörter als Regelfall anzunehmen. Das Wort *Partei* fällt hierbei allerdings in zweierlei Hinsicht aus dem Rahmen. Zum einen existierte es bereits im Mittelhochdeutschen als Lehnwort (vgl. (KLOSA et al. 2001)), so daß es ausreichend „Zeit“ für eine Assimilierung der Prosodie gehabt hätte. Des weiteren endet es im Gegensatz zu allen anderen Ausnahmen auf einen Diphthong. Die naheliegende Erklärung hierfür ist, daß Diphthonge insgesamt die Wortbetonung eher anziehen als gespannte Monophthonge. Diese Hypothese wurde im Anschluß an die Untersuchung der Wortbetonungsregeln in einem Produktionsexperiment überprüft (vgl. Abschnitt 5.1.5). Eine weitere mögliche Erklärung für die abweichende Wortbetonung ist eine Analogiebildung seitens der Sprechergemeinschaft zum betonten deutschen Nominalisierungssuffix *-(r)ei* (*Mogel-éi, Fische-réi*). Letztere Vermutung hat allerdings lediglich den zweifelhaften Status einer *ad hoc*-Erklärung.

⁷Das Wort *Tabu* ist auf dem Umweg über das Englische aus der polynesischen Sprache Tonga in das Deutsche entlehnt worden (vgl. (KLOSA et al. 2001)).

⁸In einigen oberdeutschen Dialekten wie z.B. dem Österreichischen ist dieses Minimalpaar nicht vorhanden. Dort liegt die Wortbetonung für beide orthographisch und semantisch unterschiedlichen Formen auf der finalen Silbe.

Fehlerklasse Ib: -CVC Beispiele für Wörter, die in diese Klasse fallen, sind die folgenden:

Eß.zétt, Ho.tél, A.príl, Pro.to.kóll, Ma.dríd, ka.pútt, kom.plétt, Pro.grámm

Diese Fehler sind schwieriger zu klassifizieren als die vorherigen. Dennoch ist auffallend, daß bei allen Beispielen die Endsilbenstruktur komplexer ist als die der Paenultima. Die Paenultima weist bei den meisten Beispielen im Gegensatz zur finalen Silbe keinen schließenden Kodakonsonanten auf. Bei *Eßzett* und *komplett* ist die Paenultima zwar geschlossen, aber der Silbenanlaut der finalen Silbe ist komplexer. Dies könnte ein Hinweis darauf sein, daß im Gegensatz zur gängigen phonologischen Betrachtung der Silbenanlaut bei der Verteilung des Silbengewichts und somit für die Verteilung der Wortbetonung durchaus eine Rolle spielt. Auch diese Hypothese wird in dem späteren Produktionsexperiment überprüft.

Die zweite Beobachtung ist, daß die Wörter mit Endbetonung die Tendenz haben, auf einen Lateral /l/, einen alveolaren Plosiv /t/ oder einen Nasal (hier /m/) zu enden. Laterale und Nasale sind innerhalb der Sonoritätshierarchie (vgl. (JESPERSEN 1904)) gleich unterhalb der an der Spitze stehenden Vokalen angesiedelt. Sie können in unbetonten Silben anstelle von Vokalen sogar den Silbennukleus bilden.

Da die Metrische Phonologie davon ausgeht, daß die metrisch starken Segmente innerhalb von Silben aufgrund ihrer Sonorität die Vokale sind (vgl. (SELKIRK 1982)), ist der Schluß, daß Silben mit hochsonoren Kodakonsonanten schwerer als solche mit weniger sonoren Kodakonsonanten sind, einleuchtend. Da die Wortbetonung wiederum im Zusammenhang mit dem Silbengewicht erklärt wird, kann somit nicht nur die Anzahl, sondern auch die Artikulationsart der Kodakonsonanten für die Vorhersage der Wortbetonung eine Rolle spielen. Wenn der Effekt eines hochsonoren Kodakonsonanten mit einem komplexen Anlaut kombiniert wird, könnte dies zu einer ausreichenden Verstärkung des Silbengewichts führen, welche die Wortbetonung an sich zieht und z.B. das Betonungsmuster in *April* erklären kann. Auch diese Hypothese wird in dem anschließend erläuterten Produktionsexperiment zu prüfen sein.

Im Gegensatz zu Nasalen und Lateralen sind Plosive innerhalb der Sonoritätshierarchie an unterster Stelle angesiedelt, so daß es nicht offensichtlich ist, wie das Vorliegen eines Plosivs in der Silbenkoda diese Silbe „schwerer“ machen sollte, auch wenn die Wörter *komplett, Eßzett* und *kaputt* diese Vermutung nahelegen.

5.1.4.2 Fehlerklasse II: schwere, unbetonte Silben

Wörter, die eine unbetonte schwere Endsilbe aufweisen, sind zwar selten, dennoch sollte geklärt werden, inwieweit dieses Muster systematisch erklärt werden kann:

*Arbeit, Áttentat, Héimat, Stándard, Ántwort, Héirat, Zéppelin, Báltimore, Cócktail,
Dóllar, Jápan, Níkolaus, Kónstanz, Síemens*

Wenn man die englischen Lehnwörter *Baltimore*, *Cocktail* und *Dollar* unter der Annahme ignoriert, daß sie die ursprüngliche Betonung beibehalten haben, bleibt noch immer eine relativ lange Liste problematischer Fälle übrig. Bei den verbleibenden Wörtern ist auffällig, daß es sich häufig um Eigennamen handelt (*Japan*, *Nikolaus*, *Konstanz*, *Siemens*). Eigennamen sind bekannt für eine hohe Variabilität hinsichtlich ihrer Betonung (BERG 1997), was in Abschnitt 5.3 noch weiter diskutiert wird. Da sich das Wort *Zeppelin* ebenfalls von einem Eigennamen ableitet, wird es auch ausgeklammert. Es verbleiben die folgenden Wörter mit Erklärungsbedarf:

Stándard, Árbeit, Héimat, Ántwort, Héirat, Áttentat

Diese Liste ist schwieriger zu erklären. Für das englische Lehnwort *Standard* (KLOSA et al. 2001) kann analog zu Lehnwörtern französischen Ursprungs angenommen werden, daß es seine Ursprungsprosodie beibehalten hat. *Hei-rat*, *Hei-mat* und *Arbeit* waren in früheren Sprachperioden des Deutschen allesamt Kompositaformen (KLUGE 1995; BRAUN 1993). Da die Betonung in zweigliedrigen Komposita im Deutschen regelhaft auf die linke Konstituente fällt (vgl. Abschnitt 5.2), könnte dies darauf hinweisen, daß sich in diesen Formen zwar das Betonungsmuster, nicht aber die morphologische Wortstruktur erhalten hat. Das *Ant-* in *Antwort* hatte in früheren Varianten des Deutschen den Status eines betonten Nominalisierungspräfix. In seiner Bedeutung verhielt es sich analog zum (unbetonten) Verbalisierungspräfix *ent-* (KLUGE 1995). Es tritt heutzutage allerdings nur noch selten auf, z.B. in *Antlitz*. Zudem ist bei *Antwort* auffällig, daß die zweite Silbe *wort* im Deutschen noch immer als freies Morphem existiert. Dies macht die Annahme, daß sich das Betonungsmuster in *Antwort* als eine Art Pseudokompositum erklären läßt, um so wahrscheinlicher. Letzteres liefert auch eine mögliche Erklärung für die Betonung des französischen Lehnwortes *Attentat*, da die Silbe *-tat* ebenfalls als freies Morphem existiert. Dies kann von der Sprechergemeinschaft als ein deutliches Indiz zum Vorliegen eines Kompositums und zur Auswahl einer entsprechenden Betonung gewertet werden. Bei (MANGOLD 1990) wird für *Attentat* alternativ auch die Betonung auf der letzten schweren Silbe zugelassen (*Attentát*). Dieses Muster entspricht der phonologisch vorhergesagten Betonung.

Mögliche Erklärungen sind an dieser Stelle allerdings lediglich durch Indizien motiviert und stellen keineswegs den Anspruch unumstößlicher Wahrheit. Die Wörter stellen lediglich Ausnahmen einer Regel dar, die sich in der Vorhersage als sehr erfolgreich herausgestellt hat und deshalb nicht grundsätzlich in Frage gestellt werden kann.

5.1.4.3 Fehlerklasse III: Wörter mit Antepaenultimabetonung

Eine signifikante Anzahl von Wörtern tragen die Betonung auf der Antepaenultimasilbe, obwohl die letzten beiden Silben leicht sind und somit die Betonung laut Vorhersage auf der Paenultima landet:

Ýpsilon, Mínum, Rísiko, Ámeise, Rókoko, Amérika, Wàshington, Flórida, Jerúsalem, Fílofax, Músical, Términál, Níkolaus

Wiederum sind einige der Wörter klar als nicht assimilierte meist englische Lehnwörter erkennbar (*Fílofax, Términál, Músical*). Auch die große Anzahl von Eigennamen in den Ausnahmefällen ist auffällig (*Amérika, Flórida, Wàshington, Jerúsalem*) und liefert ein weiteres Indiz für den Sonderstatus von Eigennamen bei der Betonungszuweisung (siehe Abschnitt 5.3). Die übriggebliebenen Wörter der Liste weisen sämtlich eine offene Paenultimasilbe mit meist kurzem Vokal auf. Die Fälle reichen aber nicht aus, um eine explizite Regel hinzuzufügen, welche diese Fälle erfaßt, da es zu viele Gegenbeispiele gibt, m.a.W. Formen, welche eine betonte Paenultimasilbe der gleichen Struktur aufweisen (z.B. *Aréna, Ikéa, Mikádo*). Daher scheint es zunächst sinnvoll, diese Fälle vorläufig als Ausnahmen zu behandeln. Das Wort *Ámeise* kann ähnlich wie zuvor die Wörter *Áttentat* und *Ántwort* als Pseudokompositumsform erklärt werden, da das Wort *Meise* im Deutschen wie auch *Tat* und *Wort* als freies Morphem vorkommt. Diese Erklärung wird gestützt für gelegentliche humorvolle Wortschöpfungen wie *B-Meise*. Für *Rísiko* sowie *Mínum* kann u.U. angenommen werden, daß sie ihre italienisch-lateinische Ursprungsbetonung beibehalten haben. Diese Vermutung wird dadurch gestützt, daß Lehnwörter lateinischen Ursprungs häufig eine CV-Silbe statt einer CVV-Silbe aufweisen. Eine CV-Silbe ist offenbar nicht in der Lage, die Betonung am rechten Wortrand an sich zu ziehen und rutscht daher auf die erste betonbare Silbe eines Wortes. Für das Wort *Rókoko* existieren laut (MANGOLD 1990) genausoviele Betonungsmöglichkeiten wie Silben, wenngleich die Initialbetonung der häufigste Fall zu sein scheint. Hierbei kann analog zu den Wörtern lateinischen Ursprungs von einer superleichten CV-Silbenstruktur ausgegangen werden, die offenbar nicht die Betonung vorzeitig an sich ziehen kann. Die Hypothese, daß CV-Silben sich im Deutschen ähnlich zu schwa-Silben verhalten, indem sie nie bzw. lediglich am linken Wortrand betont werden können, ist allerdings noch experimentell.

5.1.4.4 Andere Fälle

In einem Fall fällt die Betonung sogar noch weiter an den linken Wortrand, nämlich bei *Cábrioleť*. Es gibt eine halbwegs plausible Erklärung für diese Abweichung von der Paenultimaregel. Die erste ist, daß das Wort *Cabrio* eine sehr geläufige Kurzform für das längere *Cabriolet* ist. Würde die Approximantregel (Regel 1) zur Anwendung kommen, so könnte das Muster *Càbrio* richtig vorhergesagt werden. Wenn die Sprecherin das *-let* als extrasilbisches Morphem empfindet, so wäre hier eine mögliche Erklärung. Diese hat allerdings ihre Tücken, da die Anwendung der Approximantregel eigentlich nicht möglich ist, denn die Phonemfolge /brj/ verstößt gegen die phonotaktischen Beschränkungen des Deutschen (vgl. (CARSON-BERNDSEN 1998)) und ermöglicht somit keine alternative Silbifizierung in Ca-brio-let.

5.1.5 Erweiterung des Silbengewichtsmodells

Gerade die Unstimmigkeiten in den Vorhersagen für Wörter, die auf zwei leichte Silben enden (vgl. Abschnitt 5.1.4.1), führten zu der Fragestellung, ob die nach unserer ursprünglichen Klassifikation als „leicht“ eingestuft Silben (vgl. Seite 86) die Wortbetonung in finaler Position vielleicht eher anziehen, wenn sie

- auf einem Langvokal enden (-VV)?
- auf einem Diphthong enden (-Vv)?
- einen komplexen Silbenanlaut haben (mindestens CC-)?
- auf einem (hochsonoren) Nasal, Lateral oder aber einem alveolaren Plosiv enden?
- eine vorangehende leichte Paenultimasilbe ohne Silbenauslaut aufweisen?

Es geht dabei insgesamt um eine Erweiterung der Klassifikation der Silbengewichtshierarchie in „leichte“, „schwere“ und unbetonbare Silben. Eine solche Erweiterung ist bereits von (KENSTOWICZ 1994) im Rahmen der Optimalitätstheorie vorgeschlagen worden. Die o.a. Fragestellungen dienen als Hypothesen in einem Produktionsexperiment. 15 phonetisch ungeschulte deutsche Muttersprachler wurden hierfür gebeten, zweisilbige Nonsense-Wörter vorzulesen, welche aus leichten Silben unterschiedlicher CV-Struktur bestanden. Die insgesamt 44 Stimuli wurden aus den Silben gebildet, die in Tabelle 5.5 abgebildet sind.

erste Silbe	letzte Silbe
zɔm, zo:, zɔY, ʔo:	bu:, ba:, baɪ, bu:ɣ, bɪn, bɪl, bat, bɔk, brit, brɪl, ʃpɪt

Tabelle 5.5: Stimulizusammensetzung bei Produktionsexperiment in IPA-Notation

Die Stimuli sollten zwar die Varianten abdecken, die zu überprüfen waren, aber dennoch nicht zuviel Unterschiede auf segmentaler Ebene beinhalten, um die Einflußvariablen möglichst konstant zu halten. Außerdem wurden die Stimuli so ausgewählt, daß es eine möglichst klare graphemische Entsprechung für sie gibt, da die Versuchspersonen phonetisch ungeschult und somit auch nicht in der Lage waren, phonetisch transkribierte Äußerungen zu lesen. Die Stimuli wurden in sehr komplexe Sätze eingebettet, die für jeden Stimulus variierten (siehe Anhang A.1). Die Sätze hatten zwei Funktionen. Zum einen sollten sie verhindern, daß die Versuchspersonen sich zu Beginn des Experiments für eine Betonungsstrategie entschieden und diese für jedes Wort wiederholten, wie es bei einem gleichbleibenden Kontext u.U. der Fall gewesen wäre. Außerdem sollten die Satzkontexte die Versuchspersonen davon abhalten, sich zu sehr auf die Aussprache des unbekanntes Wortes

im Satz zu konzentrieren. Die resultierenden Sätze wurden von einer professionellen Sprecherin zunächst auf ihre Lesbarkeit hin überprüft und teilweise vereinfacht. Die Aufnahmen erfolgten im Tonstudio des IKP Bonn. Jede Versuchsperson bekam die Instruktion, die Sätze ohne viel Nachdenken laut zu lesen und sich gedanklich ein Bild von der im Satz dargestellten Situation zu machen. Nach jedem gelesenen Satz mußten die Versuchspersonen auf einer 4-Punkte-Skala von „sehr gut“ bis „schlecht“ bewerten, wie gut das unbekannte Wort zu der im Satz beschriebenen Situation paßt. Diese Aufgabe sollte die Versuchspersonen aber lediglich davon abhalten, sich auf die Betonung oder Aussprache der unbekanntesten Wörter zu konzentrieren und sich gedanklich eher mit der semantischen Ebene der Äußerungen zu beschäftigen. In die Sätze waren auch jeweils relativ lange andere Wörter eingebaut, wohingegen die eigentlichen Stimuli mit ihrer zweisilbigen Struktur verhältnismäßig einfach waren. Dies sollte die Versuchspersonen ebenfalls davon abhalten, sich zu viele Gedanken über die korrekte Aussprache der entsprechenden Stimuli zu machen. Die langen sehr unterschiedlichen Sätze sollten ferner dazu dienen, einen Listeneffekt zu vermeiden, bei dem die Versuchsperson die prosodische Struktur vom vorhergehenden Satz imitiert. Die resultierenden Betonungsmuster der Versuchspersonen wurden von einer phonetischen Expertin perzeptiv bestimmt.

Die Ergebnisse der Testpersonen bestätigen insgesamt die Präferenz einer Betonung der ersten Silbe im Kontext zweier leichter Endsilben ($\chi^2, p < 0.0001$). Endsilben, die auf einem Langvokal enden, und Silben, die auf einem nicht-sonoren Konsonanten auslauten, konnten die Betonung nur in Ausnahmefällen an sich ziehen. Bei Wörtern mit finalen Diphthongen betonte die Mehrzahl der Versuchspersonen zwar nach wie vor signifikant häufiger ($\chi^2, p < 0.0001$) die erste Silbe (vgl. Abbildung 5.2), aber das Betonungsmuster fiel nicht so deutlich zugunsten der ersten Silbe aus wie in den Wörtern, die auf einen Langvokal enden (vgl. Abbildung 5.3). Die Hypothese, daß Diphthonge schwerer als Langvokale sind, konnte demnach nicht eindeutig bestätigt werden, wenn auch unterstützende Indizien vorliegen.

Neben Wörtern mit finalen Diphthongen wurden auch Wörter mit einer offenen vorletzten Silbe häufig auf der letzten Silbe betont (vgl. Abbildung 5.4). Obwohl auch hier zahlenmäßig die Erstbetonung überwiegt, ist keine signifikante Präferenz für eine Betonung auf der ersten Silbe nachweisbar. In dem Moment, wo eine geschlossene vorletzte Silbe gegeben ist, liegt die Betonung hingegen eindeutig auf der vorletzten Silbe ($\chi^2, p < 0.0001$). Die Hypothese, daß geschlossene Silben schwerer sind als solche ohne Auslaut, wurde dadurch zunächst gestützt. Die finale Silbe wird nur dann signifikant bevorzugt betont, wenn die letzte Silbe auf einen Sonoranten auslautet ($\chi^2, p < 0.05$, vgl. Abbildung 5.5). Beim Vorliegen einer offenen Paenultimasilbe wird dieses Muster deutlicher ($\chi^2, p < 0.0001$, vgl. Abbildung 5.6). Die Hypothese, daß ein alveolarer Plosiv im Silbenauslaut die Betonung an sich ziehen kann, konnte nicht bestätigt werden.

Die Bevorzugung der Endbetonung wird noch deutlicher, wenn die finale Silbe einen komplexen Anlaut hat, der also aus mehreren Konsonanten besteht. Allerdings ist der Unterschied zum Muster in Abbildung 5.6 nicht signifikant, so daß die

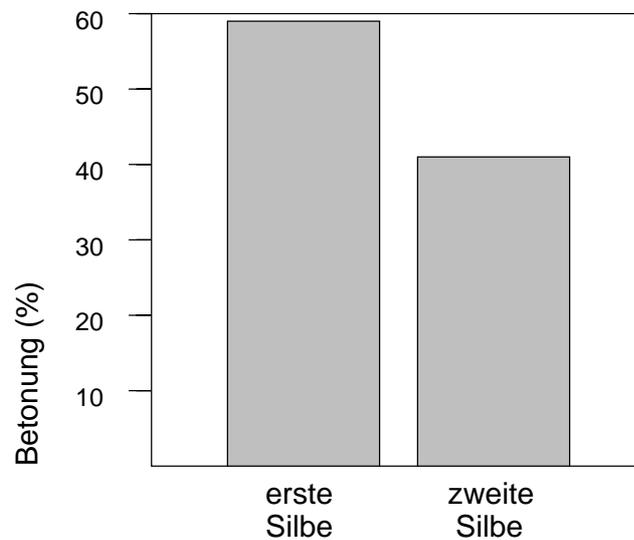


Abbildung 5.2: Betonungsverhältnisse bei zweisilbigen Wörtern mit finalem Diphthong

Hypothese, daß der Silbenanlaut einen Einfluß auf das Silbengewicht hat, zunächst nicht bestätigt werden kann.

Vom alleinigen Vorhandensein eines Sonoranten in der Silbenkoda läßt sich jedoch nicht zuverlässig auf eine Betonung der letzten Silbe schließen. Das Zusammenspiel mit einer sehr leichten offenen Paenultimasilbe ist erforderlich, um die Endbetonung zu bewirken. Andernfalls bleibt die Betonung der vorletzten Silbe das bevorzugte Muster seitens der Sprecherinnen.

Insgesamt deuten die Ergebnisse darauf hin, daß der Silbenanlaut nicht die ausschlaggebende Rolle bei der Beeinflussung des Silbengewichts spielt, sondern eher die Art bzw. das Vorhandensein des konsonantischen Auslauts. Wenn man annimmt, daß die Betonung im Deutschen auf die schwerere Silbe fällt, kann man demnach die Silbengewichtshierarchie ergänzen, so daß sich im Vergleich zur Hierarchie auf Seite 87 folgendes Bild ergibt:

1. **schwa-Silben:** ə, ɐ, ɪ, ʊ, ɨ, ʉ, l als Nukleus, evtl. auch CV (vgl. Seite 96)
2. **sehr leichte Silben:** C⁺VV
3. **leichte Silben:** CVC_{obstruent}
4. **halbschwere Silben:** CVC_{sonorant}
5. **schwere Silben:** C⁺V[V|C]C⁺

Eine zusätzliche Regel, welche Silben mit sonorem Auslaut als betonungsanziehend kategorisiert, sofern die vorletzte Silbe „wesentlich leichter“ ist, kann daher Betonungsmuster wie in *April* vorhersagen und ihren Status als Ausnahmen relativieren.

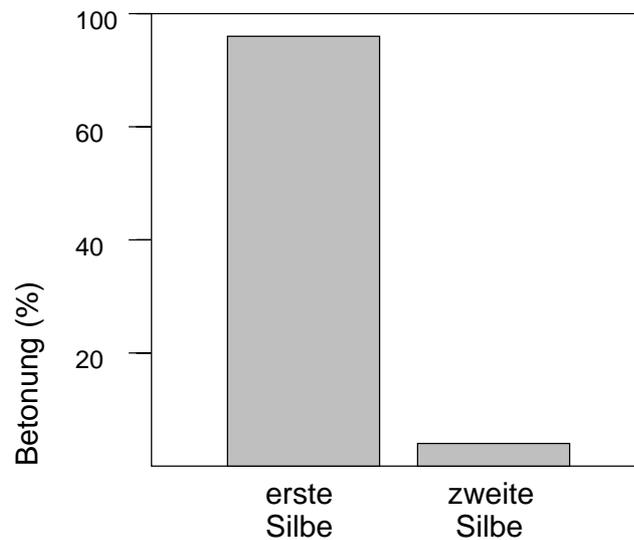


Abbildung 5.3: Betonungsverhältnisse bei zweisilbigen Wörtern mit finalem Langvokal

Diphthonge können im Vergleich mit Langvokalen vorerst nicht systematisch als „schwerer“ betrachtet werden. Auch alveolare Plosive haben keinen Sonderstatus innerhalb der Klasse der Obstruenten, indem sie das Silbengewicht so beeinflussen, daß die Betonung beeinflußt wird.

5.1.6 Anwendung der Regeln auf die BPD

Da im Sinne der in Kapitel 3 entwickelten Methode der Erfolg der erarbeiteten Regeln anhand der *Baseline* zu überprüfen ist, die sich an der Übereinstimmung der verschiedenen Sprecherinnen orientiert, wurde überprüft, wie gut der in Kapitel 4 getestete Vorhersagealgorithmus funktioniert, wenn man ihn um die Wortvorhersageeregeln erweitert. Dort wurde die Wortbetonung zunächst anhand der Annotationen in der Datenbank vergeben, um eine Gütebewertung der Vorhersageregeln auf der Äußerungsebene zu ermöglichen. Natürlich wird die Vorhersagequalität, verglichen mit den Ergebnissen, die auf manuell bestimmter Wortbetonung basierten, schlechter. Die Korrelation zwischen vorhergesagter und wahrgenommener Prominenz liegt nur noch bei $\rho = 0.7$, $p < 0.001$. Dennoch konnte die Vorhersage im Vergleich zum ersten Regelsatz ($\rho = 0.64$, $p < 0.001$) entscheidend verbessert werden. Dieser verwendete lediglich eine sehr vereinfachte Wortbetonungszuweisung. Interessanterweise gilt der Verlust der Vorhersagegenauigkeit anders als bei den Regeln für die Vorhersage auf Äußerungsebene im gleichen Maße für alle Sprecherinnen. Dies deutet darauf hin, daß die vereinfachende Annahme, daß die Wortbetonung wesentlich weniger sprecherspezifisch ist als die Äußerungsbetonung, tatsächlich zutrifft. Ein rein regelbasierter Ansatz ohne Ausnahmelexikon, der rein auf der Basis der bekannten metrischen Regularitäten basiert, funktioniert nicht voll-

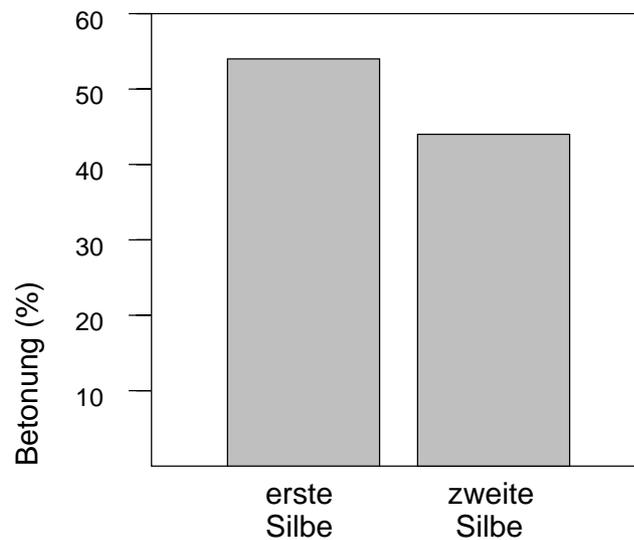


Abbildung 5.4: Betonungsverhältnisse bei zweisilbigen Wörtern mit offener Paenultima

ständig befriedigend.

5.1.7 Zusammenfassung

Im großen und ganzen konnten die Generalisierungen der Metrischen Phonologie durchaus bestätigt werden. Einige der Regeln konnten allerdings entscheidend vereinfacht bzw. zusammengefaßt werden. Das Ergebnis ist ein überschaubares System für die Vorhersage deutscher Wortbetonung. Der Ansatz ist recht erfolgreich, sofern eine Vorverarbeitung vorgenommen wird, bei der eine Reihe von Flexionssuffixen abgetrennt wird.

Die Regeln konnten gemäß ihrer Relevanz hierarchisiert werden. Das hervorstechendste Merkmal ist hierbei der auffällige Einfluß des Silbengewichts auf die Betonung, so daß die vorliegende Untersuchung die Annahme, daß das Deutsche eine quantitätssensitive Sprache ist, bestätigen konnte, wenn es hierzu auch noch gegenteilige Auffassungen gibt (z.B. (WIESE 1996)). Dies ist um so mehr der Fall, da die vorliegende Untersuchung nicht nur zeigen konnte, daß eine quantitätssensitive Vorhersage erfolgreich ist, sondern auch, daß die Anzahl der richtigen Vorhersagen die Gegenbeispiele bei weitem übertrifft. Wörter, die auf zwei leichte Silben enden, zeigen das instabilste Betonungsmuster im Deutschen. Ein Produktionsexperiment konnte eine Tendenz aufzeigen, daß die Silbengewichtshierarchie an dieser Stelle erweitert werden sollte, um bessere Vorhersagen zu erzielen. Letztendlich ist auffällig, daß die Betonung von Eigennamen häufig aus dem Rahmen dieses Regelwerks fällt. Auch funktioniert der Vorhersagealgorithmus insgesamt nicht fehlerfrei und verschlechtert die Vorhersage des Vorhersagealgorithmus auf Äußerungsebene bis weit unter die festgelegte *Baseline*. Dies kann zweierlei bedeuten:

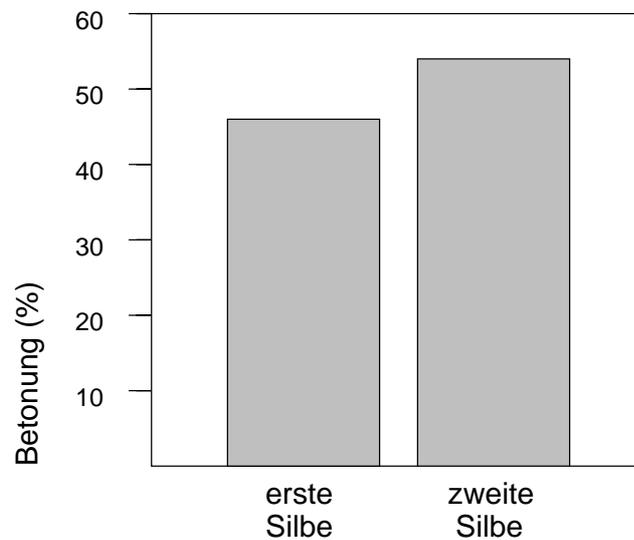


Abbildung 5.5: Betonungsverhältnisse bei zweisilbigen Wörtern mit finaler Silbe, die auf einem Sonoranten auslautet

- Entweder ist die Wortbetonung im Deutschen stärker lexikalisiert als allgemein angenommen,
- oder es sind noch nicht alle Regularitäten entdeckt worden.

Das Produktionsexperiment konnte Hinweise liefern, daß sich durchaus Präferenzen der Betonung zeigen, die nicht innerhalb der üblichen Klassifikationen erklärt werden können. Die Erweiterung der Silbengewichtstheorie könnte auch im Rahmen einer Mora-basierten Phonologie erklärt werden, wie sie beispielsweise von (LAHIRI und KOREMAN 1988) und (GUSSENHOVEN 2000) für das Niederländische vorgeschlagen wurde. Gegen die Annahme, daß es weitere Regularitäten gibt, spricht allerdings der Umstand, daß Vorhersagealgorithmen, die auf der Basis neuronaler Netze operieren und in der Lage sind, äußerst komplexe Abhängigkeiten und Zusammenhänge zu modellieren, keine besseren Vorhersagen erzielen als das hier vorgestellte Regelsystem. Daher erscheint es als fraglich, ob die nicht-vorhersagbaren Fälle nicht einfach als Ausnahmen gewertet werden sollten.

5.2 Morphologisch komplexe Wörter: Komposita

Eine einschlägige linguistische Kategorie, für die die Metrische Phonologie Betonungsregeln aufgestellt hat, sind Komposita. Diese weisen interessante Betonungsvariationen auf. Unter einem Kompositum versteht man die Zusammensetzung von zwei freien Morphemen, wie z.B. *Hand+Tuch=Handtuch*. Komposita können im Deutschen prinzipiell beliebig komplex werden, wie das berühmte Beispiel der

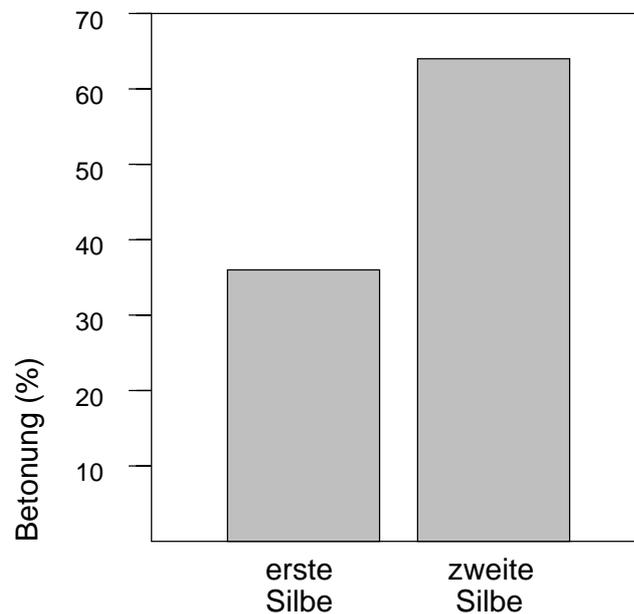


Abbildung 5.6: Betonungsverhältnisse bei zweisilbigen Wörtern mit offener Paenultimasilbe und finaler Silbe, die auf einem Sonoranten auslautet

Donau-dampf-schiff-fahrts-kapitäns-witwe zeigt. Außerdem sind sie äußerst produktiv, da Fachsprachen und Zeitungstexte immer neue Zusammensetzungen erfordern oder auch herausfordern. Die Betonung von zweigliedrigen Komposita ist regelhaft auf der ersten Konstituente. Gemäß den Regeln der metrischen Phonetik (z.B. (FÉRY 1986)) erhält also diejenige Silbe des ersten Morphems, die Trägerin der Wortbetonung ist, einen zusätzlichen Schlag. Bei dreigliedrigen, also aus drei freien Morphemen zusammengesetzten Komposita, ist die Vorhersage schwieriger. In sogenannten ((AB)C)-Komposita, bei denen die ersten beiden Morpheme ein eigenes Kompositum ausmachen, fällt die Hauptbetonung auch auf das erste Teilkompositum (z.B. (*Milch-kaffee*)-*trinker*). Bei sogenannten (A(BC))-Komposita ist die Betonungsvorhersage schwieriger. Gemäß (GIEGERICH 1985) fällt die Hauptbetonung in solchen Fällen dann auf die erste Konstituente, wenn die beiden letzten Morpheme als Kompositum einen sehr hohen Grad an Lexikalisierung erreicht haben, wie es für *Denkmal* oder *Sparkasse* in den (A(BC))-Komposita (*Industrie*-(*Denkmal*)) und (*Stadt*-(*Spar-kasse*)) der Fall ist. In einigen (A(BC))-Komposita fällt die Hauptbetonung aber auch auf die zweite Konstituente, wie z.B. in (*Städte*-(*Schnellverkehr*)), (*Landes*-(*Sport-bund*)). Für diese Zusammensetzungen gibt es aber kein allgemeingültiges Gesetz, sondern teilweise sogar sprecherspezifische Varianten, die sich nicht als systematische dialektale Abweichungen erklären lassen, da sie zwischen den Versuchspersonen stark streuen (BENWARE 1987).

5.2.1 „Akzentverschiebung“ in Komposita

Aufgrund unterschiedlicher Möglichkeiten der Wortbetonung, welche in (A(BC))-Komposita vorliegen, eignen sie sich dazu, die in der Metrischen Phonologie postulierte Eurhythmierregel der *Akzentverschiebung*⁹ zu überprüfen, da sie große Gefahr laufen, gegen die in Abschnitt 4.1 vorgestellten Euphonieprinzipien, also einer Alternation von starken und schwachen Silben, zu verstoßen. Für metrische Gitter hat (SELKIRK 1984) die folgende Akzentverschiebungsregel aufgestellt:

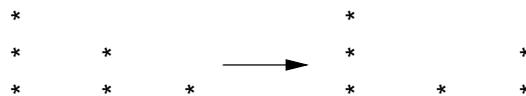


Abbildung 5.7: Die Akzentverschiebungsregel zur Vermeidung von Akzentzusammenstößen

Das Akzentverschiebungsphänomen wurde bereits von (KIPARSKY 1966) für das Deutsche als relevant postuliert. Nun gab es auf Äußerungsebene keine Belege, die die Relevanz von Akzentverschiebungsphänomenen im Deutschen untermauern konnten (vgl. Abschnitt 4.4.3). Auf Wortebene ist diese Frage allerdings noch unbeantwortet. In (A(BC))-Komposita, in denen die einzelnen Morpheme aus einsilbigen Wörtern bestehen, liegen klassische Umgebungen für die Anwendung einer Akzentverschiebungsregel vor, wenn die Hauptbetonung auf die erste Silbe, die A-Konstituente, fällt und an die hauptbetonte Silbe des BC-Kompositums direkt angrenzt (vgl. Abbildung 5.8). Gemäß den Vorhersagen der metrischen Phonologie müßte sich demnach im Kompositum *Welt-spar-tag* die Nebenbetonung des Gesamtkompositums von dem Morphem *Spar* auf das Morphem *Tag* verschieben, obwohl dies nicht der Betonungsstruktur des Kompositums *Spartag* in Isolation entspricht.

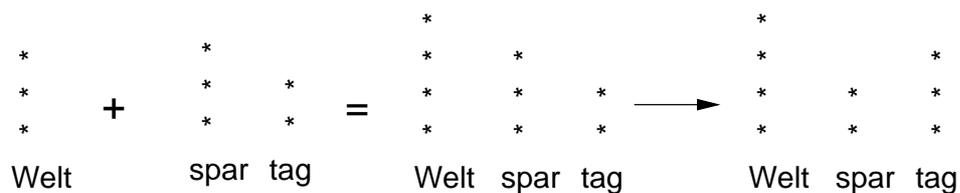


Abbildung 5.8: Akzentzusammenstoß im (A(BC))-Kompositum *Weltspartag*

⁹Auf die englischen Termini wie *stress clash* für einen Akzentzusammenstoß sowie *stress shift* für eine Akzentverschiebung wird in dieser Arbeit aufgrund der hohen Gebräuchlichkeit der deutschen Begriffe verzichtet. Dennoch ist die englische Verwendungsweise in der Tat passender, da in Kapitel 2 der Terminus *Akzent* für Tonhöhenakzente, also eindeutig akustische Phänomene reserviert wurde. Ein *Akzentzusammenstoß* ist aber auf einem metrischen Gitter formuliert, muß also nicht unbedingt mit einem Tonhöhenakzent verknüpft sein.

5.2.2 Empirische Untersuchungen zur Akzentverschiebung in Komposita

(MENGEL 1998) konnte anhand perzeptiver Untersuchungen mit dreisilbigen synthetisierten Pseudo-Komposita eine Hörertendenz zur Wahrnehmung der Nebenbetonung auf der letzten Silbe zeigen, auch wenn die Grundfrequenz über die Äußerung konstant gehalten wurde. Aufbauend auf diesen Ergebnissen, führte (FISCHENBECK 2001) eine Reihe weiterer Perzeptionsexperimente durch, wobei sie allerdings mit natürlichsprachlichen Stimuli und empirisch belegten Komposita, entnommen aus dem *Wortschatz-Lexikon* der Universität Leipzig (QUASTHOFF 1998), arbeitete. Ihr Stimulusmaterial bestand aus 26 dreisilbigen und 9 viersilbigen Wörtern (vgl. Anhang B.4). In den verwendeten viersilbigen Wörtern war lediglich das letzte Morphem zweisilbig und endete auf einer unbetonten Silbe (*Stadt.spar.kas.se*), so daß keine wortinternen unbetonten Silben eine rhythmische Alternation erzeugen und den Akzentzusammenstoß verhindern konnten. In einem ersten Experiment verwendete sie als Stimuli Aufnahmen von einer professionellen Sprecherin, welche die Komposita jeweils eingebettet in dem Trägersatz „Er sagte <...> sei das richtige Losungswort“ produziert hatte. Alle von der Sprecherin produzierten Komposita trugen ihre Hauptbetonung deutlich auf der ersten Silbe. Anschließend wurden phonetisch geschulte sowie ungeschulte Versuchspersonen (insgesamt 21) gebeten zu bewerten, ob sie die Nebenbetonung auf der zweiten oder dritten Silbe wahrgenommen hatten. Diese Aufgabe erwies sich als schwierig, insbesondere für die phonetisch ungeschulten Versuchspersonen. Diese Schwierigkeit ist vermutlich darauf zurückzuführen, daß die erste Silbe von der Sprecherin durch die paradigmatische Einbettung in den gleichbleibenden Trägersatz beinahe kontrastiv hervorgehoben worden war. Dies führte offenbar zu einer Deakzentuierung sämtlicher nachfolgender Äußerungsbestandteile (vgl. auch Abschnitt 2.4.2.2). Trotz der Schwierigkeit der Versuchspersonen legen die Ergebnisse nahe, daß in der überwiegenden Zahl der Fälle (66%) keine „Akzentverschiebung“ wahrgenommen wurde. In dreisilbigen Komposita fällt diese Tendenz noch deutlicher aus (Wahrnehmung der Nebenbetonung auf der zweiten Silbe in 72%). Bei viersilbigen Komposita gibt es keine signifikante Präferenzen für die zweite oder dritte Silbe.

Aufgrund der Beurteilungsprobleme im ersten Perzeptionsexperiment wurde ein weiterer Versuch durchgeführt. Hierbei nahmen lediglich phonetisch geschulte Versuchspersonen teil (insgesamt 19), da deren Beurteilungen sich im vorherigen Experiment als konsistent erwiesen hatten. Als Sprecher wurden jedoch zwei phonetische Laien gewählt, die nicht in die Fragestellung eingeweiht waren und bei denen deshalb davon ausgegangen werden konnte, daß sie nicht hinsichtlich einer bestimmten Betonung voreingenommen waren. Als Trägersätze fungierten diesmal die Satzkontexte, aus denen die Komposita ursprünglich entstammten. Um den Versuchspersonen die Chance zu geben, zwei Stimuli u.U. auch als „gleich stark betont“ beurteilen zu können, konnten sie die wahrgenommene Betonung anhand eines Reglers einstellen, der intern auf eine 30-stufige Prominenzskala abgebildet

wurde, ohne daß diese für die Versuchspersonen sichtbar war. Die beiden zentralen Tendenzen des Vorgängerexperiments konnten bestätigt werden, nämlich daß eine Akzentverschiebung eher die Ausnahme denn die Regel ist und lediglich bei viersilbigen Wörtern überwiegend auftritt. Das vorherrschende Betonungsmuster bei (A(BC))-Komposita mit vermeintlichen Akzentzusammenstößen bleibt die Folge *hauptbetont-nebenbetont-unbetont*. In den fast ausnahmslos viersilbigen Wörtern, in denen ein „Akzentverschiebung“ wahrgenommen wurde, blieb diese Wahrnehmung auch dann recht gut erhalten, nachdem die A-Konstituente vom Restwort abgeschnitten wurde. Dies weist darauf hin, daß die Wahrnehmung der Verschiebung nicht allein auf die Hörererwartung zurückzuführen ist, sondern auch ihre Entsprechung auf der artikulatorischen und akustischen Ebene haben muß. (FISCHENBECK 2001) schließt aus ihren Ergebnissen, daß die generelle Annahme einer „Akzentverschiebung“ beim Aufeinandertreffen zweier wortbetonungstragender Silben für das Deutsche so nicht aufrechterhalten werden kann. Eine Akzentverschiebung trete nur dann auf, wenn die wortinterne Struktur die Bildung eines zweiten *phonologischen Fußes* zuließe. Dies sei genau dann der Fall, wenn nach der im Falle eines „Akzentverschiebung“ nebenbetonten Silbe noch mindestens eine weitere Silbe folge, da phonologische Füße im Deutschen mindestens aus einer betonten gefolgt von einer unbetonten Silbe gebildet würden (vgl. auch (FÉRY 1988) sowie Abschnitt 5.3). Eine Folgestudie mit Hilfe delexikalisierte Sprache (z.B. nach dem Verfahren von (SONNTAG und PORTELE 1998)) könnte diese Frage beantworten, ohne auf synthetische Sprache ausweichen zu müssen. Inwieweit eine Wortbetonungsvorhersage die phonologische Fußstruktur mitberücksichtigen muß, ist eine weitere offene Frage, eine stärkere Einbeziehung ist aber vermutlich sinnvoll. Gemäß den Arbeiten von (EISENBERG 1991) kommt eine Vorhersage der Wortbetonung sogar ganz ohne die Einbeziehung des Silbengewichts aus, wenn sie die phonologische Fußstruktur entsprechend berücksichtigt.

5.2.3 Zusammenfassung

Wie schon für die Äußerungsebene gezeigt werden konnte (vgl. 4.4.3), scheinen die Euphoniereregeln im Deutschen eher selten zur Anwendung kommen. Offensichtlich angewandt werden sie lediglich dann, wenn ein wohlgeformtes alternierendes metrisches Gitter erzielt werden kann, wie es viersilbige Komposita durch eine Verschiebung des Nebenakzents auf die dritte Silbe ermöglichen. Trotz der hier dargestellten Ergebnisse konnte (MENGEL 1998) eine Wahrnehmung von Akzentverschiebungen auch in dreisilbigen Pseudokomposita feststellen. Es ist durchaus denkbar, daß seine Verwendung von synthetisierter Sprache sowie die Tatsache, daß es sich bei den Stimuli nicht um reale Komposita gehandelt hat, zu diesen Ergebnissen beigetragen haben. Vielmehr handelte es sich um Permutationen der Silben *koch, reis, mal*, die zusätzlich delexikalisiert waren. Die (BC)-Konstituenten dieser Permutationen bilden aber keine erkennbaren realen Komposita, die eine sinnvolle semantische Einheit bilden, wie bei (*Haupt-*)*Bahnhof*, (*Welt-*)*Spartag*, (*Stab-*)*Hochsprung*. Für letz-

tere (BC)-Konstituenten weiß der Hörer natürlich, daß sie regelhaft auf der ersten Konstituente betont werden. Dies ist für die von Mengel gewählten Silbenzusammensetzungen natürlich nicht unbedingt der Fall. Eventuell hat der Hörer also doch die Tendenz, eine alternierende Betonung wahrzunehmen, wenn er nicht *top down* eine andere Betonung erwartet.

5.3 Eigennamen

In Abschnitt 5.1 war insgesamt auffällig, daß eine große Anzahl der nicht erklärbaren Ausnahmen von den Regeln in Eigennamen auftreten. Dies kann zum einen bedeuten, daß die Betonung von Eigennamen *frei*, also nicht vorhersagbar ist, so wie es in einigen Sprachen für die Betonung insgesamt der Fall ist (z.B. im Russischen). Diese Annahme ist einleuchtend, da bei Eigennamen von einem höheren Grad der Abweichung von den sprachlichen Normen ausgegangen werden kann, was auch zu Problemen bei Verfahren der automatischen Transkription führt (MENDEL 1995). Die unvorhersagbaren Abweichungen werden schon allein dadurch deutlich, daß ortsfremde Muttersprachler häufig andere Betonungsmuster für Städtenamen verwenden als Ortsansässige, ein Umstand, der schon seit langer Zeit bekannt ist (vgl. die Literaturdiskussion in (KAUFMANN 1959, 35)). Interessanterweise zeigen aber viele von den auftretenden Eigennamen ein Betonungsmuster auf der ersten oder wenigstens auf der Antepaenultima-Silbe. Die Betonung der ersten Silbe wird traditionell als der „germanische Wurzelakzent“ (vgl. z.B. (WURZEL 1970)) betrachtet. (KAUFMANN 1959, 34ff.) findet Evidenz für eine sprachgeschichtlich ältere Betonung der initialen Silbe, welche mit einer sprachgeschichtlich jüngeren Tendenz der Betonung des rechten Wortrandes konkurriert. Die Tendenz der Verschiebung zum rechten Wortrand sei im Deutschen stärker vorhanden als in anderen germanischen Sprachen wie beispielsweise dem Englischen (vgl. engl.: *pá.per*, *Á.pril*, *ná.ture* mit dt.: *Pa.pier*, *A.pril*, *Na.túr*). Diese Beobachtung deckt sich mit einer korpuslinguistischen Studie jüngeren Datums von (BERG 1997), der im Gegensatz zum Englischen bei deutschen Nicht-Eigennamen eine starke Tendenz zur Betonung des rechten Wortrandes feststellen konnte. Für Eigennamen fremdsprachlichen Ursprungs hingegen fand er folgendes heraus:

Names with non-initial stress in the original language tend to retract their stress in both languages. The inverse pattern is less frequent. Words with initial stress in the original language are relatively reluctant to undergo stress advancement.

(BERG 1997, 11)

Diese Beobachtung führt zu der Frage, ob Eigennamen regelhaft abweichende Betonungsmuster aufweisen. Die Tatsache, daß Eigennamen die einzige größere Quelle für mehrsilbige Wörter germanischen Ursprungs darstellen, läßt die Annahme zu, daß das Deutsche entgegen den Annahmen der jüngeren Metrischen Phonologie

doch einen Unterschied zwischen der Betonung von Lehnwörtern und nativen Wörtern macht. (BERG 1997) stellt fest, daß insbesondere nicht-native Eigennamen, deren phonologische Struktur mit der deutschen Phonotaktik kompatibel ist, tendenziell eher Initialbetonung aufweisen. Die neuere metrische Phonologie (z.B. (WIESE 1996) gegenüber (WURZEL 1970)) geht hingegen meist von einem einzigen System für die Betonung aller Wörter des Deutschen aus. Es scheint so zu sein, daß eine Untersuchung von Eigennamen hier Licht ins Dunkel bringen kann.

5.3.1 Anwendung des Vorhersagealgorithmus auf Eigennamen

Zunächst wurde die Tauglichkeit des in Abschnitt 5.1 erarbeiteten Vorhersagealgorithmus an Vornamen getestet. Zu diesem Zweck wurde auf das Eigennamenaus-sprachelexikon ONOMASTICA (MENGEL und WIRTH 1995) zurückgegriffen. Dieses Lexikon ist zwar problematisch, da die Transkriptionen in erster Linie mit Hilfe automatischer Verfahren erstellt wurden, aber da mehrere Algorithmen verwendet wurden, kann dennoch von einer relativ hohen Genauigkeit der Transkriptionen ausgegangen werden. Um das Problem dennoch zu entschärfen, wurden lediglich diejenigen Transkriptionen verwendet, die eine durchschnittliche Fehlerrate $< 2\%$ aufweisen. Es wurden in der Untersuchung lediglich Vornamen betrachtet, da sowohl Straßen- als auch Ortsnamen eine gewisse morphologische Komplexität aufweisen, die vermutlich einen starken Einfluß auf die Betonung hat (z.B. *Schlumpfhäusen* vs. *Schlümpf-dorf*, vgl. auch (KAUFMANN 1959)). Bei Vornamen kann hingegen von einem einfachen Wortbildungsgefüge ausgegangen werden. So werden lexikalisierte Doppelnamen jedenfalls nicht regelhaft wie Komposita mit der Betonung auf dem ersten Glied realisiert (*Eva-María*, *Karlhéinz*). Vor dem Vergleich der vorhergesagten mit der annotierten Betonung wurde die Vornamenliste manuell auf fehlerhafte Einträge korrigiert, die eindeutig keine Vornamen darstellen (z.B. *Airport*, *Altdeutsche*, *aktuell*). Hinsichtlich der Betonung wurde keine Korrektur der annotierten Daten vorgenommen. Übrig blieb eine Liste von insgesamt 6867 Einträgen. Die Fehlerrate des Vorhersagealgorithmus stieg im Vergleich mit der Vorhersage auf der Verbmobil-Wortliste (vgl. Seite 91) dramatisch an und lag bei 18,32%. Es konnte also nochmals bestätigt werden, daß sich die Betonung von Eigennamen von der anderer Wörter stark unterscheidet. Selbst bei Berücksichtigung des Anteils an ausländischen Namen in der Liste ist dieser Einbruch der Vorhersagegüte nicht befriedigend zu erklären. Gründe hierfür und Ansätze für eine Sonderbehandlung von Eigennamen werden in den folgenden Abschnitten ausführlich diskutiert.

5.3.2 Betonungsunterschiede durch Lexikalisierung oder regelmäßig?

Eine qualitative Analyse an ca. 30% der fehlerhaft vorhergesagten Vornamen ergab, daß der größte Anteil von Fehlern auf das Konto der Betonung schwerer Endsilben ging (ca. 75%). Diese Gruppe enthielt zum überwiegenden Teil Namen deutschen

Ursprungs, aber auch slawische und englische Namen. Die Paenultimabetonungsregel führte in erster Linie bei Eigennamen südeuropäischer, romanischer Sprachen zu fehlerhaften Vorhersagen, da diese Namen häufig eine Antepaenultimabetonung aufweisen. Diese Fehler machen ca. 11% aus. In dieser Gruppe fanden sich kaum Namen germanischen Ursprungs. Allerdings war die Paenultimaregel erfolgreich bei Namen, die auf eine unbetonbare Silbe auslauten (*Adèle, Pauline, Irène*). Weitere 7,5% der Fehler sind durch eine finale Betonung in Eigennamen französischen Ursprungs erklärbar (z.B. *Anétt, Aláin, Chantál, Bernadétte*). Für diese Namen wurde laut der Paenultimaregel eine Betonung auf der vorletzten Silbe vorhergesagt. Dies deckt sich mit dem Problem, welches in Abschnitt 5.1.2.6 bereits für nicht-assimilierte französische Lehnwörter berichtet wurde. Die restlichen 6,5% Fehler gehen auf das Konto von der Anwendung der Approximantregel, welche ja verwendet wurde, um die Betonung von *Jú.lia* und *Dániel* zu erklären, die sich aber nicht auf *María* und *Elías* anwenden läßt. Namen wie *Amadéus, Bartholom'äus, Nivéa* zeigen zudem deutlich, daß sich die Regel tatsächlich auf hohe Vokale in der vorletzten Silbe beschränkt. Am auffälligsten ist aber sicherlich, daß insbesondere die Namen germanischen Ursprungs keine Endbetonung finaler schwerer Silben aufweisen, obwohl diese Regel bei Nicht-Eigennamen einen sehr wichtigen Einfluß hatte. Wenn auch viele der initialbetonten Eigennamen wie Komposita anmuten (z.B. *Adelheid, Dago-bert*), so kann die Betonung der auf eine schwere Silbe auslautenden Namen *Ármin, Álex, Béatrix, Bóries, Gérald* wohl kaum mit einer Kompositumsbetonung erklärt werden. Da die Betonung schwerer Endsilben sprachhistorisch jünger ist, liefert dies eine mögliche Erklärung für die allgemeine Tendenz, daß die Regel bei — insbesondere alten — Eigennamen nicht greift. (BERG 1997) konnte ebenfalls feststellen, daß die Initialbetonung in Eigennamen sehr konservativ ist. Die Vermutung, daß Eigennamen tendenziell eher initialbetont sind als Nicht-Eigennamen, wird außerdem bekräftigt durch vorhandene Minimalpaare des Deutschen wie *Áugust vs. Aug.úst* oder *Ró.man vs. Ro.mán*. Die Annahme der neueren Metrischen Phonologie, die nicht von einem Vorhandensein einer Initialbetonungsregel ausgeht, scheint durch die Analyse von Eigennamen zumindest stark angreifbar.

5.3.3 Untersuchung des Einflusses der einzelnen Regeln

Nachdem festgestellt werden konnte, daß der Regelsatz der deutschen Wortbetonung nicht ohne weiteres auf Eigennamen übertragbar ist, wurde der Einfluß der einzelnen Regeln systematisch überprüft. Die einschneidendste Verbesserung konnte durch ein Weglassen der *Restriktion schwerer Endsilben* erzielt werden. Die Fehlerrate sank dadurch auf 12,5%. Die Approximantregel hingegen erwies sich als wichtig, wenn sie auch bei einigen Namen zu Fehlern führte. Ihr Weglassen verschlechterte die Fehlerrate wieder auf 13%. Auch die *Restriktion der finalen Schwasilbe*, die als Teil der allgemeinen Paenultimabetonung zu sehen ist, erwies sich als sinnvoll (Verbesserung der Fehlerrate auf 11%). Eigennamen mit mehr als drei Silben sind selten initialbetont. Dies ist erklärbar, da in diesen Namen mehr als ein Fuß untergebracht

werden muß, denn Füße können im Deutschen nicht mehr als drei Silben enthalten. Bei einem viersilbigen Wort müssen demnach zwei Füße vorhanden sein, und da Füße im Deutschen linksköpfig sind, hat ein viersilbiges Wort zwangsweise die Folge *betont-unbetont-betont-unbetont*. In diesen Fällen haben Sprecherinnen meist die Tendenz, den letzten Fuß stärker zu betonen, was somit auf eine Paenultima-betonung des Gesamtwortes hinausläuft (*Ge.no.véva, Alexándra*). Es gibt allerdings auch Gegenbeispiele zu dieser Regel, wie die zusammengesetzten Namen *Hánnelore, Ánnemarie* zeigen. Ob diese Muster auf eine Kompositumsbetonung hinweisen, ist unklar, da andere zusammengesetzte Namen ein solches Muster nicht aufweisen: *Annaléna, Hans-Christían*. Eine regelhafte Paenultimabetonung wie bei der Vorhersage für Nicht-Eigennamen ist nicht nachweisbar. Eine durchgängige Initialbetonungsregel von dreisilbigen Wörtern, wie sie viele Namen vermuten lassen (*Úrsula, Ánnika, Bárbara, Ádomeit*), erwies sich allerdings ebenfalls nicht als erfolgreich, da die Paenultimabetonung auch hier relativ häufig ist (z.B. *Amánda, Elvíra, Beátus, Alfónso*). In Namen, in denen die Paenultima allerdings unbetonbar ist, fällt die Betonung auch dann auf die Antepaenultima, wenn die letzte Silbe schwer ist: *Á.del.heit, Én.gel.ber*t. Bei Beachtung dieser Ausnahme kann der Vorhersagefehler auf 9,5% gesenkt werden. Sicherlich ist auch diese Fehlerrate noch relativ hoch, aber angesichts des stark individuellen Charakters von Eigennamen sowie des hohen Anteils nicht assimilierter Namen fremdsprachlichen Ursprungs ist nicht die gleiche Vorhersagegenauigkeit zu erwarten wie bei Nicht-Eigennamen. Dennoch ist auffällig, daß die Tendenz der Betonung finaler starker Silben sich in jedem Fall nicht auf Namen erstreckt und somit ein wortklassenspezifisches Phänomen ist. Eine durchgängige Initialbetonung von dreisilbigen Namen konnte nicht belegt werden, aber bei Zweisilbern führt das Ignorieren der Betonung schwerer Endsilben zu einer korrekten Vorhersage auf der ersten Silbe.

5.3.4 Betonung unbekannter Eigennamen

In den vorherigen Abschnitten konnte eine Tendenz zu einer stärkeren Initialbetonung bei Eigennamen und einem Nichtbeachten der *Schwere-Endsilben-Restriktion* gezeigt werden. Es sollte anhand unbekannter Wörter untersucht werden, welches Muster deutsche Muttersprachler generell bevorzugen. Hierfür wurde ein Produktionsexperiment durchgeführt. Als Stimuli wurde auf eine Reihe fiktiver Namen aus der deutschen Übersetzung des Fantasy-Romans „Der Herr der Ringe“ (TOLKIEN 1996) zurückgegriffen. Diese Namen sind sämtlich nach den Regeln der deutschen Phonotaktik gebildet. Sie wurden in einen fiktiven Dialog eingebettet, wobei durch den Kontext klar wurde, daß die einzelnen Wörter teilweise Eigennamen, teilweise jedoch Bezeichnungen für Gegenstände waren (siehe Anhang A.2). An dem Produktionsexperiment nahmen 5 Versuchspersonen teil, die mit dem Roman (und somit auch mit den entsprechenden Namen) vertraut waren, aber auch 6 Versuchspersonen, denen diese Namen unbekannt waren. Die Hypothesen waren die Folgenden:

1. Versuchspersonen bevorzugen die Initialbetonung, wenn die Wörter als Namen im Kontext eingebettet sind.
2. Versuchspersonen bevorzugen die Paenultimabetonung, wenn zwei Füße im Wort untergebracht werden können, das Wort also mehr als drei Silben hat.
3. Die Versuchspersonen, die mit den Namen bereits vertraut sind, tendieren eher zu einer Betonung schwerer Endsilben als andere.

Die zu lesenden Wörter waren die Folgenden:

*Cirdan, Palantiri, Anduril, Elrond, Aragorn, Fingolfin, Beutlin, Legolas, Bombadil,
Gandalf*

Der Text wurde von den phonetisch ungeschulten Versuchspersonen laut vorgelesen. Bei einigen Wörtern ist die Graphem-Phonem-Beziehung nicht eindeutig und kann sich in unterschiedlichen Silbengewichtsrelationen auswirken. Das Graphem „Legolas“ kann z.B. als /le:go:l̩as/ oder /le:go.la:s/ interpretiert werden. Im zweiten Fall endet das Wort auf einer schweren Silbe, was zu einem Anziehen der Wortbetonung führen kann. Daher wurde zusätzlich zur Realisierung der Betonung für jede Äußerung eine Transkription der Äußerung durchgeführt. Auf eine graphemische Darstellung der Wörter konnte aufgrund des fehlenden phonetischen Vorwissens der Versuchspersonen nicht verzichtet werden. Die Ergebnisse zeigen wie die vorherigen Eigennamenanalysen, daß schwere Endsilben die Betonung nicht stärker anziehen als leichte Endsilben. Zweisilbige Wörter sind beinahe ausschließlich initialbetont, Dreisilber zeigen gelegentlich, wenn auch selten, eine finale Betonung. Viersilber werden stark bevorzugt paenultimabetont, was die Hypothese (2) bestätigt. Unbekannte Wörter werden insgesamt bevorzugt initialbetont, unabhängig von ihrem Auftauchen im Text als Name oder Nicht-Name. Hypothese (1) muß somit verworfen werden. Der Grad der Bekanntheit scheint allerdings zu einer Tendenz der finalen Betonung zu führen, so daß Hypothese (3) nicht widerlegt werden konnte. Dies ist eventuell dadurch erklärbar, daß diejenigen Versuchspersonen, die mit dem Wort nicht vertraut waren, es aufgrund der Unbekanntheit wie einen Namen behandelt haben und zunächst initial betonten. Falls eine alternative Betonung (wie die Paenultimabetonung in Viersilbern) möglich war, so wurde diese bevorzugt. Diejenigen Versuchspersonen, die mit den Namen bereits vertraut waren, konnten sich entscheiden, die allgemeinen Betonungsregeln anzuwenden oder die Tendenz der Initialbetonung von Namen anzuwenden. Die wenigen Daten, auf welche die Studie baut, lassen allgemeinere Schlüsse oder reliable statistische Untersuchungen allerdings nicht zu. Die Initialbetonung scheint aber ein Muster zu sein, für sich genügend empirische Evidenz findet, um sie nicht völlig aus einer phonologischen Grammatik zu tilgen. Eine Übersicht der Ergebnisse ist in den Tabellen 5.6 dargestellt.

Bedingung	initial	antepaenult	paenult	final
schwere Endsilbe	76%	-	-	24%
leichte Endsilbe	68%	5%	22%	5%
Zweibilber	90%	-	-	10%
Dreibilber	72%	-	4%	24%
Vierbilber	-	20%	80%	-
Namen	72%	-	-	28%
Nicht-Namen	72%	4%	8%	16%
bekannte Namen	56%	-	17%	27%
unbekannte Namen	82%	3%	6%	9%

Tabelle 5.6: Die Ergebnisse des Produktionsexperiments zu Eigennamen

5.3.5 Zusammenfassung

Die Tendenz, in Eigennamen eher die erste Silbe zu betonen, konnte wenigstens für unbekannte bzw. wenig bekannte Wörter in einem Produktionsexperiment überprüft werden. Dies spricht für das Vorhandensein einer Regel, die nicht im Apparat der metrischen Phonologie für das Deutsche aufgenommen wurde. Der sogenannte „germanische Wurzelakzent“ (vgl. (BERG 1997)) ist demnach wenigstens bei der Aussprache von Eigennamen durchaus vorhanden und bei der Aussprache von unbekanntem Wörtern vorherrschende Praxis. Durch den hohen Anteil an Eigennamen, für die dieses Prinzip aber nicht gilt, kann auch hier lediglich von einer Tendenz gesprochen werden. In dem Moment, wo die Sprecherinnen aber in der Lage sind, einen weiteren Fuß in einem Wort unterzubringen, was bei vierbilbigen Wörtern der Fall ist, tendieren sie zu einer Paenultimabetonung. Die Betonung finaler schwerer Silben konnte nur in Ausnahmefällen beobachtet werden und scheint ein wortklassenspezifisches Phänomen zu sein, welches für die Betonungsrealisierung von Eigennamen weitaus weniger wichtig ist als für Nicht-Eigennamen. Allerdings gibt es auch hier ganze Ausnahmeklassen wie beispielsweise Städtenamen wie *Berlín*, *Eutín*, *Schwerín*, die an der Gesamtzahl von untersuchten Wörtern zwar lediglich einen kleinen Prozentsatz ausmachen, aber für sich ein regelmäßiges Muster darstellen.

5.4 Eine formale Grammatik für die Generierung deutscher Wortbetonung

Die Analysen für eine erfolgreiche Vorhersage der metrischen Struktur auf Wortebene ergaben zunächst ein Regelwerk für Nicht-Eigennamen, welches der Regel zur Zuweisung der wortklassenspezifischen Prominenz aus Abschnitt 4 vorgeschaltet

wird. Das Regelwerk erwartet als Eingabe die vorverarbeitete Silbenkette des Wortes, d.h., die unbetonbaren Affixe (vgl. Tabelle 5.3) wurden vom Wort abgetrennt, und das Restwort wurde anschließend silbifiziert.

Wortbetonungsregel 1

In einsilbigen Wörtern ist die eine Silbe Trägerin der Wortbetonung.

Wortbetonungsregel 2

Unbetonbare Silbe tragen niemals die Wortbetonung.

Wortbetonungsregel 3

Schwere Endsilben werden betont.

Wortbetonungsregel 4

Hat ein Wort mehr als zwei Silben und beginnt die vorletzte Silbe mit einem hohen Vokal, so fällt die Betonung auf die Antepaenultima.

Wortbetonungsregel 5

Ansonsten wird die Paenultimasilbe betont, sofern diese betonbar ist.

Wortbetonungsregel 6

Ist die Paenultimasilbe unbetonbar, fällt die Betonung auf die Silbe, die betonbar ist und dem rechten Wortrand am nächsten liegt.

Die Formalisierung der metrischen Regeln erfolgt in Form eines endlichen Übergangnetzwerks. Die dargestellten Automaten gehen immer von mehrsilbigen Wörtern aus. Der triviale Fall eines einsilbigen Wortes kann durch einen in der Reihe der kaskadierten Netzwerke vorgeschalteten Automaten abgefangen werden.

5.4.1 Modellierung des ersten Regelapparats zur Zuweisung der Wortbetonung

Das **Eingabealphabet** für den Automaten besteht aus Silben, wobei jede Silbe mit einem Tripel, bestehend aus den Attributen $\langle \textit{Onset}; \textit{Nucleus}; \textit{Coda} \rangle$, annotiert ist. Die Attribute werden durch Listen spezifiziert, welche die entsprechenden Segmente enthalten. Aus dieser Information leitet sich das Silbengewicht ab, welches für die Zuweisung der Betonung notwendig ist.

Das **Ausgabealphabet** des Automaten besteht aus Silben, die Wertspezifikationen für das Attribut *Wortbetonung* enthalten, welches die Werte 0 oder 1 annehmen kann. Das Übergangnetzwerk wird illustriert in Abbildung 5.9. Aus Gründen der Übersichtlichkeit sind als Abkürzungen für die entsprechenden *Nucleus*- und *Coda*-Wertkonstellationen die Attribute *heavy* bzw. *light* bzw. *schwa* angegeben. Als Abkürzung für die hohen Vokale *i, ɪ, y, ʏ, u, ʊ* wird der Wert *high* notiert, als Abkürzung für die entsprechenden anderen Vokale *middlelow*. Als Abkürzung für eine beliebige Eingabe wird ein *x* notiert.

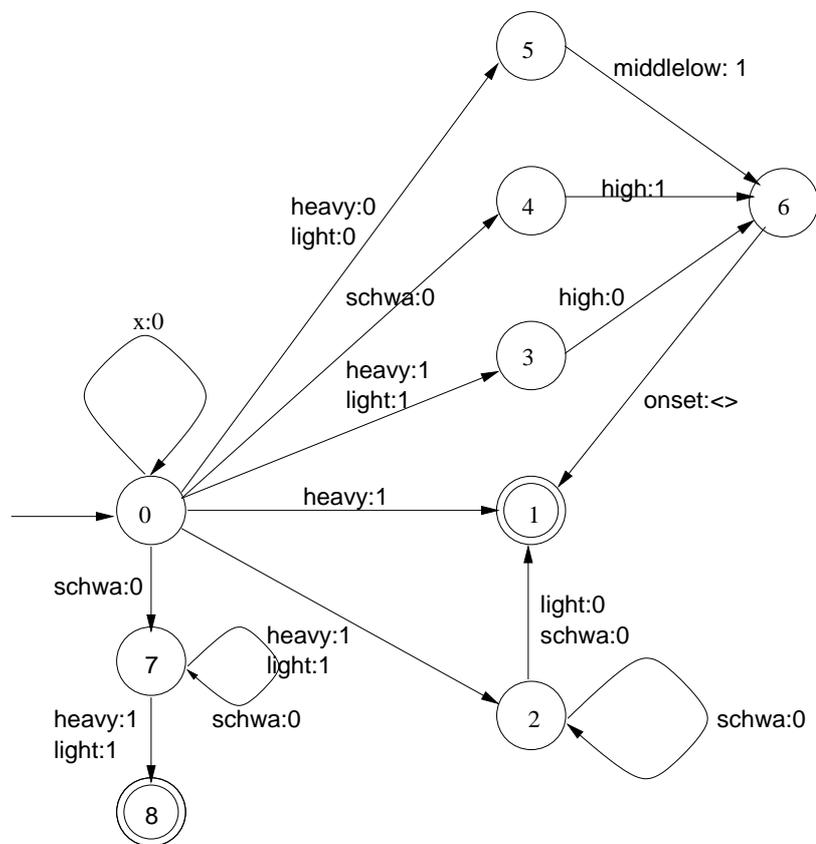


Abbildung 5.9: Ein Übergangnetzwerk für die Zuweisung von Wortbetonung

5.4.2 Modellierung des erweiterten Regelapparats zur Zuweisung der Wortbetonung

Ein Produktionsexperiment (vgl. Abschnitt 5.1.5) lieferte Argumente für eine Erweiterung der deutschen Silbengewichtshierarchie, da somit bestimmte Betonungsmuster erklärt werden konnten, die andernfalls als Ausnahmen betrachtet werden müßten. Die Ergebnisse lassen sich mit folgender Zusatzregel formulieren:

Wortbetonungsregel 7

Bei Wörtern, die auf zwei eigentlich „leichten“ Silben enden, ziehen auf einem Sonoranten endende Endsilben die Betonung an sich.

Das Übergangnetzwerk zur Modellierung der erweiterten Betonungsgenerierung entspricht hinsichtlich seines *Ein- und Ausgabealphabets* den Spezifikationen aus Abschnitt 5.4.1. In der Illustration (vgl. Abbildung 5.10) werden zur Identifikation der zusätzlichen relevanten Silbengewichtsstufen die Abkürzungen *CVV*, *CV_{son}*, sowie *CV_{obs}* verwendet. Letztere können natürlich in einer detaillierteren Darstellung durch die Wertspezifikationen der Attribute *Nucleus* und *Coda* ausgedrückt

werden.

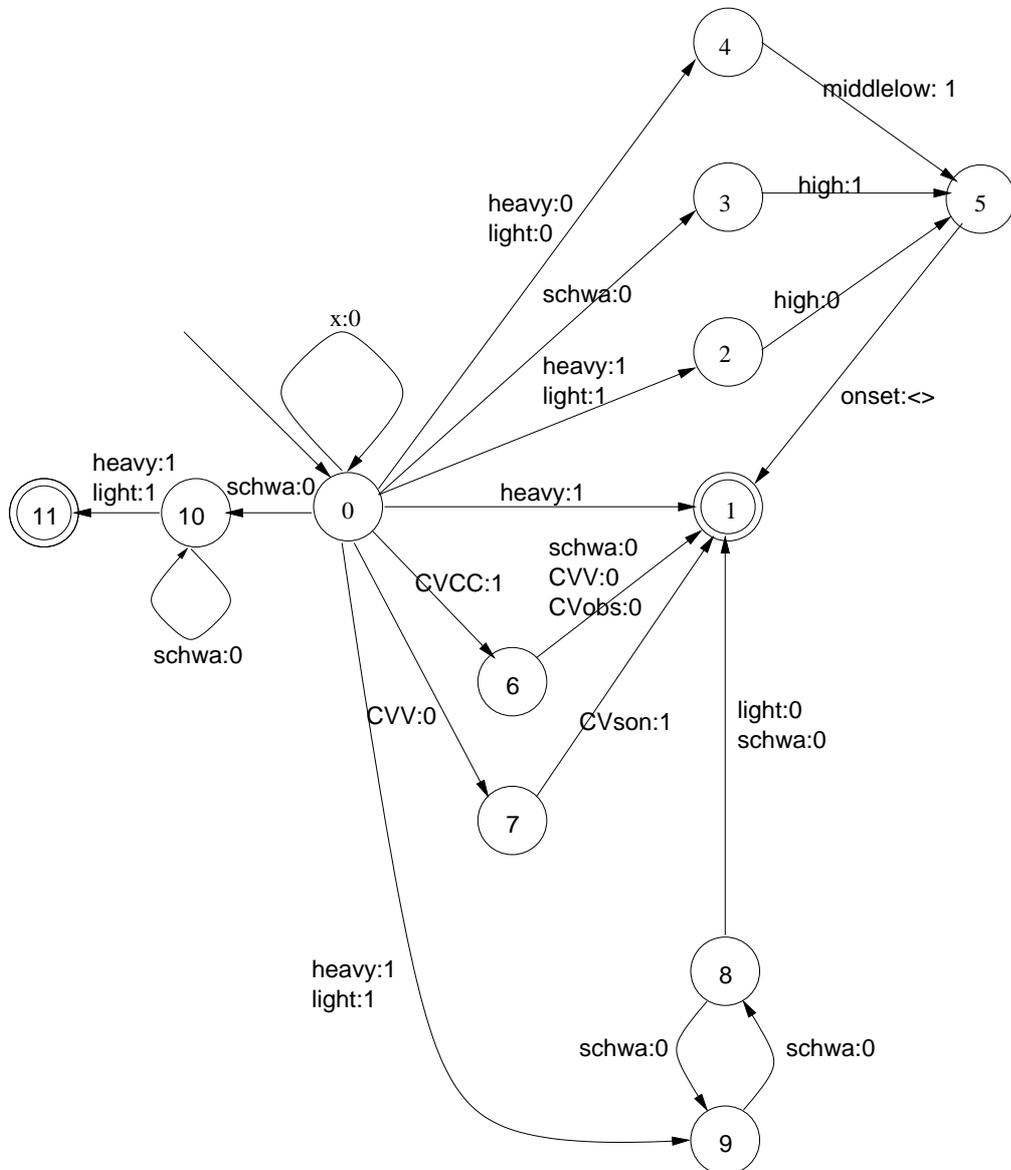


Abbildung 5.10: Ein erweitertes Übergangnetzwerk für die Zuweisung von Wortbetonung

5.4.3 Modellierung des Regelapparats zur Zuweisung der Wortbetonung von Eigennamen

In Abschnitt 5.3 wurde deutlich, daß die Betonung von Eigennamen im Deutschen offenbar anders funktioniert als die Betonung von Nicht-Eigennamen, indem die Betonung schwerer Endsilben nicht gilt. Ansonsten konnten zwar Tendenzen für

eine Betonung der ersten Silbe bei alten deutschen Namen gefunden werden, vorherrschend war diese Tendenz allerdings lediglich bei der Betonung unbekannter Namen, so daß die Initialbetonung in der Modellierung der Betonung von Eigennamen nicht berücksichtigt wird. Das Übergangsnetzwerk zur Modellierung (vgl. Abbildung 5.11) der erweiterten Betonungsgenerierung entspricht hinsichtlich seines *Ein- und Ausgabealphabets* den Spezifikationen aus Abschnitt 5.4.1. Da die Silbengewichte bei der Betonungsvorhersage von Eigennamen keine Rolle spielen, sind die Erweiterungen der Silbengewichtshierarchie an dieser Stelle nicht mehr vertreten. Relevant sind lediglich die Unterscheidungen in betonbare und unbetonbare Silben. Die Approximantregel bleibt allerdings bestehen.

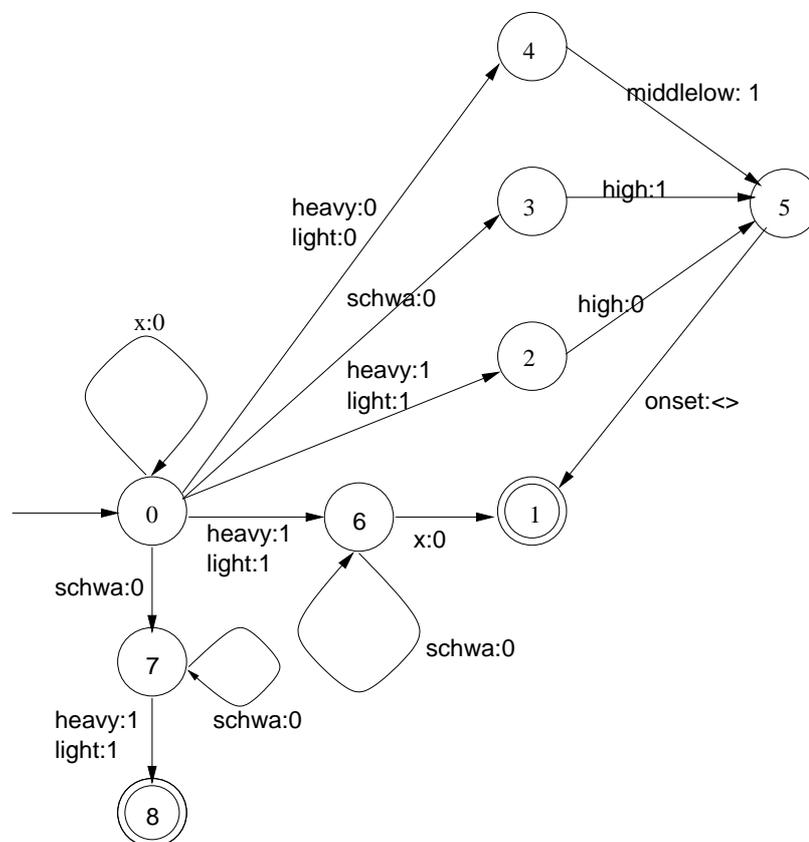


Abbildung 5.11: Ein Übergangsnetzwerk für die Zuweisung der Betonung von Eigennamen

5.5 Fazit

In diesem Abschnitt sind eine Reihe von phonologischen Vorhersageregeln empirisch überprüft worden. Die Vielzahl der phonologischen Regeln konnte bestätigt werden. Indem mehrere Regeln, welche auf eine Betonung der vorletzten Silbe ei-

nes Wortes hinausliefen, zusammengefaßt wurden, konnte das Regelwerk vereinfacht werden. Da gemäß der in Kapitel 3 festgelegten Einfachheitskriterien diejenige Grammatik zu bevorzugen ist, welche mit weniger Regeln auskommt, bedeutet dies einen Fortschritt. Mit Hilfe einer Erweiterung der Silbengewichtshierarchie auf der Basis eines Produktionsexperiments können außerdem einige vermeintliche Ausnahmefälle der Betonung vorhergesagt werden. Die Vorhersagegüte erzielte Ergebnisse im Bereich von Vorhersagen auf der Basis Neuronaler Netze. Ferner wurde von einer Arbeit berichtet, in der die phonologische Regel der Akzentverschiebung anhand mehrerer Perzeptionsexperimente überprüft wurde. Die Akzentverschiebung scheint im Deutschen nicht die Relevanz zu besitzen, welche ihr von der Phonetik bisher zugewilligt wurde, wenn sie auch bei viersilbigen (A(BC))-Komposita tendenziell eintritt. Zum Schluß wurde die Vorhersage der Betonung an Eigennamen getestet. Hierbei konnte festgestellt werden, daß sich Eigennamen insbesondere dadurch auszeichnen, daß in ihnen schwere Endsilben nicht die Betonung an sich ziehen. Der Vorhersagealgorithmus konnte aber trotz dieser Anpassung nicht die Vorhersagegüte erreichen wie bei Nicht-Eigennamen. Die Regelwerke für Eigennamen und Nicht-Eigennamen wurden mit Hilfe endlicher Übergangnetzwerke formalisiert, so daß sie in eine formale Grammatikbeschreibung integriert werden können.

Kapitel 6

Einbettung in die Optimalitätstheorie

optimal: [lat.-nlat.]: sehr gut, bestmöglich, beste, Best-
Duden Fremdwörterbuch

In den vorherigen Kapiteln wurden generative phonologische Regelsätze hinsichtlich ihrer empirischen Adäquatheit überprüft, indem sie implementiert und ihre Vorhersagen an größeren Datensätzen überprüft wurden. Die Formalisierungen als endliche Übergangnetzwerke sind dabei direkte Abbildungen der generativen Regeln der Metrischen Phonologie. Neuere Arbeiten innerhalb der Phonologie gehen allerdings andere Wege. Hier wird nicht — wie in Abschnitt 2.4.3.3 bereits erläutert —, ausgehend von einer phonologischen Tiefenstruktur, mit Hilfe von Regeln eine phonologische Oberflächenstruktur generiert. Statt dessen wird, ausgehend von allen möglichen Oberflächenformen, anhand von hierarchisch angeordneten *Constraints* diejenige Form herausgesucht, welche die wenigsten *Constraints* verletzt. Die Menge der möglichen phonologischen Formen wird mit Hilfe der Funktion GEN erzeugt. Die optimale Form muß dabei nicht notwendigerweise diejenige sein, die keinen einzigen *Constraint* verletzt. Jedoch ist sie diejenige Form, welche die *wenigsten* wichtigen *Constraints* verletzt. Dieses Kapitel ist der Frage gewidmet, wie sich die Erkenntnisse der Untersuchungen in den letzten Abschnitten in den Formalismus der Optimalitätstheorie integrieren lassen. Dabei soll der Anspruch der formalen Darstellung einer Grammatik, wie er in Abschnitt 3 formuliert wurde, nicht außer acht gelassen werden. Darum ist es zunächst notwendig, sich mit Möglichkeiten der Formalisierung der Optimalitätstheorie zu beschäftigen. Dies wird in Abschnitt 6.1 geschehen. Die folgenden Abschnitte beschreiben, wie sich diese Erkenntnisse auf eine optimalitätstheoretische Modellierung der Wortbetonung übertragen lassen. Zunächst wird in Abschnitt 6.1 ein Algorithmus vorgestellt, welcher die Generierungsfunktion GEN für Wortbetonungen als endlichen Automaten erzeugt und wie sich optimalitätstheoretische *Constraints* in diesen Kandidatenautomaten integrieren lassen. Anschließend wird ausgehend von den empirischen Analysen zur Wortbetonung eine *Constraint*hierarchie aufgestellt. Auf der Basis dieser Hierarchie werden den Übergängen in den Kandidatenautomaten auf dem Ausgabeband Kosten zugewiesen, je nachdem, ob ein *Constraint* erfüllt ist oder nicht.

Das Kostenmaß ergibt sich aufgrund der Stellung des *Constraints* in der Hierarchie, m.a.W. je wichtiger die Einhaltung des *Constraints*, um so mehr Kosten verursacht seine Verletzung. Nachdem die Kosten verteilt wurden, kann der optimale Kandidat durch Aufsummierung der Kosten ermittelt werden, die sich beim Durchlaufen eines Pfades zwischen Start- und Endknoten im Automaten ergeben. Derjenige Pfad, der die wenigsten Kosten erzeugt, entspricht auch dem optimalen Kandidaten.

6.1 Formale Darstellung der Optimalitätstheorie

Bei der Formalisierung optimalitätstheoretischer *Constraints* wird hier von den Überlegungen von (ELLISON 1994b) und darauf aufbauenden Arbeiten (WALTHER 1996; FRANK und SATTÀ 1998) ausgegangen. Dieses Modell wurde zwar von (KARTTUNEN 1998) kritisiert, da es ein explizites "Zählen" der Constraintverletzungen benötige, welches einen "offline"-Prozeß darstelle und nicht direkt in die Grammatik integriert sei. An dieser Stelle wird dennoch auf Ellisons Verfahren zurückgegriffen, da

- es eine übersichtliche Darstellung der einzelnen *Constraints* ermöglicht.
- auch in der orthodoxen Optimalitätstheorie Constraintverletzungen registriert und kognitiv verarbeitet werden müssen, wenn sie auch nicht in der hier dargestellten Art aufsummiert werden.

Ob eine Abbildung nach (KARTTUNEN 1998) psychologisch adäquater ist, kann an dieser Stelle nicht beantwortet werden. Allerdings wurde die Notwendigkeit des Zählens auch bereits an Darstellungen der metrischen Phonologie kritisiert. Eindeutige psycholinguistische Evidenz, die gegen eine Integration solcher Prozesse in eine Grammatik spricht, konnte nicht entdeckt werden.

Gemäß (ELLISON 1994a) benötigt die Formulierung optimalitätstheoretischer *Constraints* mit Hilfe endlicher Übergangnetzwerke die folgenden Teilschritte:

1. Zunächst muß die Kandidatenmenge, also die Menge der nach optimalitätstheoretischen Gesichtspunkten zu bewertenden Klasse phonologischer Strukturen, als endlicher Automat spezifiziert werden.
2. Anschließend müssen die jeweiligen *Constraints* als endliche Übergangnetzwerke formuliert werden, wobei die *Constraints* auf die Werte 0 („harmonisch“) oder -1 („disharmonisch“) abgebildet werden.
3. Dann müssen die Produkte des Kandidatenautomaten mit den *Constraint*-automaten gebildet werden, beginnend mit dem Übergangnetzwerk, welches den hierarchisch höchsten *Constraint* modelliert.

4. Um den optimalen Pfad (also den optimalen Kandidaten) durch den Automaten) zu bestimmen, müssen die Harmoniewerte auf den Übergängen addiert werden. Der Pfad mit der maximalen Harmonie beschreibt gleichzeitig den optimalen Kandidaten.

Zunächst muß also eine Generierungsfunktion GEN bestimmt werden, welche sämtliche möglichen Betonungsmuster für eine Folge von Silben oder Segmenten beschreibt. Geht man idealisierenderweise davon aus, daß es pro Wort genau eine Silbe geben muß, welche Trägerin der Wortbetonung ist, so muß GEN einen Automaten generieren, der zwischen Start- und Endknoten genausoviele Pfade hat, wie das beschriebene Wort Silben enthält. In jedem Pfad wird jeweils eine andere Silbe betont. Der endliche Automat, der die möglichen Betonungsmuster des Wortes *E.le.fant* beschreibt, ist in Abbildung 6.1 beschrieben. Dieser Automat kann auf genau drei Pfaden zwischen Start- und Endknoten durchlaufen werden, nämlich:

- Pfad 1: $0 - 1_{betont} - 2_{betont} - 3_{betont}$ (*é.le.fant*)
- Pfad 2: $0 - 1_{unbetont} - 2_{betont} - 3_{betont}$ (*e.lé.fant*)
- Pfad 3: $0 - 1_{unbetont} - 2_{unbetont} - 3_{betont}$ (*e.le.fánt*)

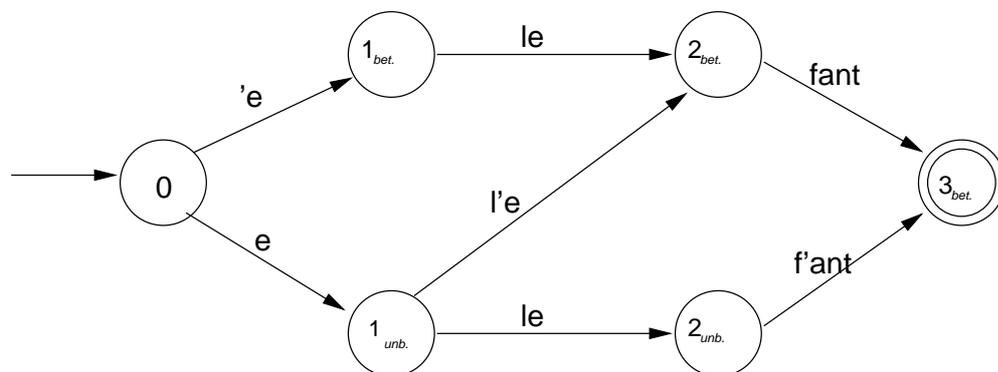


Abbildung 6.1: Die Kandidaten für die Betonung des Wortes *Elefant*, modelliert als Pfade in einem endlichen Automaten

Die Generierungsfunktion GEN, welche für ein mindestens zweisilbiges, silbengeparstes Wort die Kandidatenmenge für die Betonungszuweisung beschreibt, kann mit folgendem Algorithmus beschrieben werden:

1. Gegeben sei ein phonologisches Wort ω bestehend aus den Silben $\sigma_1, \sigma_2, \dots, \sigma_n$. Die Indizes 1 bis n entsprechen der jeweiligen Position der Silbe in ω .
2. Definiere einen Startknoten 0. und Pfade zu zwei weiteren Knoten 1_{betont} sowie $1_{unbetont}$. Annotiere den Pfad zum Knoten 1_{betont} mit einer betonten Variante von σ_1 sowie den Pfad zum Knoten $1_{unbetont}$ mit einer unbetonten Variante von σ_1 .

3. Definiere Zustände 2_{betont} bis n_{betont} und erzeuge jeweils einen Pfad von Zustand 1_{betont} zu 2_{betont} u.s.w., bis ein Pfad von $n-1_{betont}$ zu n_{betont} führt. Markiere n_{betont} als den Endzustand. Annotiere jeden Pfad mit der unbetonten Variante der Silbe, die mit dem Zielknoten des Pfades in ihrer Indexzahl übereinstimmt.
4. Definiere nun Zustände $2_{unbetont}$ bis $n-1_{unbetont}$ und, beginnend bei Zustand $m=1_{unbetont}$, schaffe Pfade zwischen jedem unbetonten Zustand und dem betonten Folgeknoten $m+1_{betont}$. Annotiere jeden Pfad mit der betonten Variante der Silbe σ_{m+1} .
5. Beginnend bei $m=1_{unbetont}$, schaffe einen Pfad zwischen jedem unbetonten Knoten zum Knoten $m+1_{unbetont}$, wobei $m < (n-1)$ sein muß. Annotiere jeden so entstandenen Pfad mit der unbetonten Variante der Silbe σ_{m+1} .

Um die optimalitätstheoretischen *Constraints* abbilden zu können, muß als nächstes ein Kandidatenautomat mit endlichen Übergangsnetzwerken verknüpft werden, welche genau diese *Constraints* darstellen. Geht man zunächst davon aus, daß ein *Constraint* bei seiner Einhaltung auf den Wert 0 („harmonisch“) sowie bei einer Verletzung auf den Wert -1 („nicht harmonisch“) abbildet, kann ein einfaches Übergangsnetzwerk formuliert werden, welches im Fall einer betonten (aber eigentlich unbetonten) schwa-Silbe der Eingabe den Wert -1 auf dem Ausgabeband zuweist, alle anderen Eingaben aber als harmonisch durchgehen läßt. In Abbildung 6.2 ist ein solches Übergangsnetzwerk dargestellt, wobei eine betonte schwa-Silbe mit '@' dargestellt ist, jede andere Eingabe wird mit x abgekürzt.

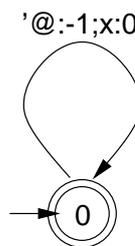


Abbildung 6.2: Darstellung eines optimalitätstheoretischen *Constraints* als Übergangsnetzwerk

Man kann nun das Produkt bilden aus dem Automaten, welcher die Kandidatenmenge beschreibt, und dem Übergangsnetzwerk, welches die Zuweisung der Harmoniewerte erledigt. Das Produkt dieser beiden Automaten akzeptiert nur solche Eingaben, die von beiden Automaten auch separat akzeptiert werden. Jede Kombination von Übergängen, die sowohl vom Kandidatenautomaten K als auch vom Harmonieautomaten H eingelesen werden können, also eine gültige Eingabe in $K \cap H$ darstellt, definiert einen Übergang im Produktautomaten. Das Produkt des in Abbildung 6.1 dargestellten Kandidatenautomaten und des in Abbildung 6.2 dargestellten Harmonieautomaten ergibt das Übergangsnetzwerk aus Abbildung 6.3.

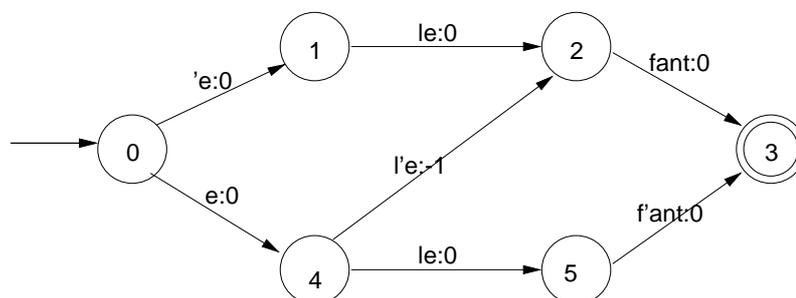


Abbildung 6.3: Produkt eines Kandidatenautomaten und eines Harmonieautomaten

Die Gesamtharmonie eines Kandidaten kann berechnet werden, indem für jede akzeptierte Eingabe die korrespondierenden Harmoniewerte der Ausgabe addiert werden. Je niedriger der Gesamtharmoniewert, um so weniger harmonisch ist der entsprechende Kandidat im Eingabeband. Der Zahlenwert des Gesamtharmoniewertes spiegelt die Häufigkeit der Constraintverletzung wider. Im Automaten aus Abbildung 6.3 würde dies zu einem Ausschluß des Kandidaten *E.l'e.fant* führen. Für eine Entscheidung zwischen den beiden verbleibenden Kandidaten sind natürlich weitere *Constraints* zu berücksichtigen. Genauso wie die Harmonie relativ zu einem Einzelconstraint berechnet werden kann, kann auch die Harmonie mit einer relativ zu einer geordneten Constrainthierarchie geordneten Liste natürlicher Zahlen gemessen werden. Hierfür muß die Liste der natürlichen Zahlen eine 1:1-Entsprechung in der Constrainthierarchie haben, d.h., der erste Zahlenwert der Liste beschreibt den Harmoniewert für den dominantesten Constraint, der zweite Zahlenwert den zweitdominantesten *Constraint* u.s.w. Der Zahlenwert gibt dabei wieder, wie häufig der mit dem Listenplatz verknüpfte *Constraint* von dem Kandidaten verletzt wurde. Ein Kandidat mit der Harmoniewert-Liste $\langle -2, -1 \rangle$ verletzt den dominantesten *Constraint* zweimal, den untergeordneten *Constraint* hingegen einmal. Um den optimalen Kandidaten auf der Basis dieser Listen zu bestimmen, müssen Listen dieser Form natürlich miteinander verglichen werden. (ELLISON 1994b) stellt fest, daß gemäß der Optimalitätstheorie bei einem Vergleich der beiden Harmoniewert-Listen $\langle -5, -2, -20 \rangle$ und $\langle -5, -3, -1 \rangle$ die erste der optimalen Form näher ist. Dies entspricht der strikten Einhaltung der optimalitätstheoretischen Constrainthierarchie. Laut dieser ist der Umstand, daß der erste Kandidat einen niedrigrangigeren *Constraint* (an dritter Stelle in der Liste) sehr viel häufiger verletzt, irrelevant, da der Ausschluß des zweiten Kandidaten bereits auf der Basis des höherrangigen *Constraints* an zweiter Stelle in der Harmoniewert-Liste erfolgt ist. M.a.W. kann eine einzelne Verletzung eines höherrangigen *Constraints* nicht durch die Befolgung vieler niedrigrangiger *Constraints* ausgeglichen werden. Die Art der Modellierung optimalitätstheoretischer *Constraints* weist viele Gemeinsamkeiten zu Verfahren auf, wie sie in der korpusbasierten Sprachsynthese in den letzten Jahren populär geworden sind (CAMPBELL 1996; STÖBER et al. 2000). Die-

se Synthesysteme bestimmen die optimale Kette von Syntheseeinheiten, indem ein gerichteter Graph über alle der Verkettung zur Verfügung stehenden Syntheseeinheiten gebildet wird. Dieser Graph entspricht in wesentlichen Merkmalen dem Automaten in Abbildung 6.1, da auch hier die Kombination möglicher Kandidaten abgebildet wird. Aufgabe des Synthesealgorithmus ist nun, die für das Syntheseziel optimale Kette zu selektieren, indem der optimale Pfad durch den Graphen gefunden wird. Ähnlich wie bei der Optimalitätstheorie werden auch hierbei Disharmoniewerte oder *Kosten* vergeben.¹ Die Gesamtkosten oder der Disharmoniewert für den jeweiligen Pfad werden ebenfalls durch eine Addition der Kosten berechnet, wobei sich ebenfalls der optimale Pfad durch eine Minimierung der Gesamtkosten ergibt. Eine weitere Ähnlichkeit ist der Umstand, daß es unterschiedlich gewichtete Kosten gibt, also Kosten, deren Mißachtung weitaus schwerer wiegen als andere.

Anders als in der optimalitätstheoretischen Modellierung sind die Kosten aber nicht binär. Dort konnten lediglich die Anzahl der Verletzungen eines jeden *Constraints* gezählt werden. Die Anzahl der Verletzungen eines höherwertigen *Constraints* führte zum Ausschluß eines Kandidaten auf dieser Ebene. Der Kandidat spielte bei weiteren Vergleichen keine Rolle, egal wie optimal hinsichtlich niederwertiger *Constraints* er sich verhält. Die Gewichtung der *Constraints* in der Einheitenselektion, wie sie aus der Synthese bekannt ist, geht anders vor: Hier entscheidet die Gewichtung über die Schwere der Constraintverletzung, d.h., die Verletzung eines bestimmten *Constraints* kann dreimal so viel Gewicht haben als ein weniger wichtiger Constraint. Eine Constrainthierarchie wird daher über die Gewichtung eingeführt. Diese Gewichtung führt aber dazu, daß die Constraintverletzungen eben nicht mehr pro *Constraint* gezählt werden müssen, sondern daß die Gesamtharmonie direkt abgelesen werden kann. Diese Art der Vorgehensweise wird der Optimalitätstheorie aber nur bedingt gerecht, da eine Verletzung eines hochrangigen *Constraints* durchaus durch die Befolgung vieler niedrigrangiger *Constraints* ausgeglichen werden kann. Soll die strenge Hierarchie, wie sie von der Optimalitätstheorie gefordert wird, eingehalten werden, so muß dies durch eine entsprechend hohe Gewichtung hochrangiger *Constraints* berücksichtigt werden. Ob diese Einschränkung für die phonologische Theorienbildung allerdings absolut notwendig ist, wird an dieser Stelle offengelassen.

Abgesehen davon bietet diese Art der Berechnung des optimalen Pfades eine elegante formale Modellierung optimalitätstheoretischer Modelle. Im folgenden Abschnitt wird geprüft, ob sich die vorher gewonnenen Erkenntnisse zur Wortbetonung auf diese Art und Weise nach optimalitätstheoretischen Gesichtspunkten modellieren lassen.

¹Bei der Synthese wird zusätzlich zwischen Einheitenkosten und Übergangskosten unterschieden. Die Einheitenkosten stellen dabei die Entfernung der betrachteten Einheit zu einer Zieleinheit dar. Diese Art der Kosten ist bei der Phonologie natürlich nicht relevant, da es keine Referenzform gibt.

6.2 Hierarchisierung und Zuweisung der *Constraints*

In diesem Abschnitt wird geprüft, ob sich die im vorherigen Kapitel gewonnen Regularitäten der Wortbetonung mit Hilfe der Berechnung von Kosten in einem Übergangnetzwerk, welches die Kandidaten darstellt, optimalitätstheoretisch modellieren lassen. Es ist zunächst zu beachten, daß die hier vorgeschlagene Modellierung deutscher Wortbetonung sich nicht orthodox an die in der Optimalitätstheorie vorgesehene *Constraint*-Menge hält. Vielmehr stehen die Regularitäten des Kapitels 5 auf dem Prüfstand. Hierbei ist insbesondere die Frage interessant, ob der Einfluß der einzelnen Regeln auf die empirische Adäquatheit der Vorhersagen sich auch in der optimalitätstheoretischen *Constraint*-Hierarchie widerspiegelt.

Wenn man die *Constraint*-Verletzungen beim Automaten in Abbildung 6.3 berücksichtigt, so konnte hier der Kandidat mit einer betonten unbetonbaren Silbe ausgeschlossen werden. Der Umstand, daß schwa-Silben nie betont werden, war Teil der Basisregeln für die Zuweisung der Wortbetonung (vgl. Seite 113) und auch fundamentaler Bestandteil der Berechnung der Äußerungsbetonung (vgl. Seite 64). Es gab niemals Anlaß, diese Einschränkung in Frage zu stellen. (FÉRY 1998) findet in ihrer Untersuchung deutscher Wortbetonung keinen einzigen Fall, der dieser Basisregel zuwiderliefe. Daher kann dieser Einschränkung höchste Priorität zugewiesen werden, und Verletzungen dieses *Constraints* sollten daher bei den Kandidaten mit hohen Kosten versehen werden. An oberster Stelle in der *Constraint*-Hierarchie steht demnach folgender *Constraint*:

Optimalitätstheoretischer *Constraint* 1

Unbetonbare Silben dürfen nicht betont werden.

Um eine Entscheidung zwischen den beiden übriggebliebenen Formen *É.le.fant* und *E.le.fánt* treffen zu können, muß darüber hinaus aber noch ein weiterer *Constraint* eingefügt werden. Hierfür kommt ein *Constraint* in Frage, der sich an der *Schwere-Endsilben-Restriktion* orientiert (vgl. Seite 113). Diese Regel hat sich zumindest bei Nicht-Eigennamen von großer Wichtigkeit für eine korrekte Vorhersage herausgestellt, und sie wurde nur in seltenen Fällen verletzt. Daher steht dieser *Constraint* in der Hierarchie an nächster Stelle

Optimalitätstheoretischer *Constraint* 2

Schwere Endsilben müssen betont werden.

Aufgrund dieser Hierarchisierung werden betonten schwa-Silben eines Kandidaten höhere Kosten zugewiesen (hier: „2“) als unbetonten schweren Endsilben (hier: „1“). Das entsprechende Übergangnetzwerk für den Beispielkandidaten *Elefant* ist in Abbildung 6.4 dargestellt. Das Netzwerk weist dem Pfad, der den Kandidaten *E.le.fánt* beschreibt, nun als den optimalen Kandidaten aus, der keinerlei Kosten auf dem Ausgabeband verursacht:

- Pfad 1: 0 – 1 – 2 – 3 (*é.le.fant*) verursacht die Kosten $0+0+1=1$.

- Pfad 2: 0 – 4 – 2 – 3 (*e.lé.fant*) verursacht die Kosten $0+2+1=3$.
- Pfad 3: 0 – 4 – 5 – 3 (*e.le.fánt*) verursacht die Kosten $0+0+0=0$ und ist damit optimal.

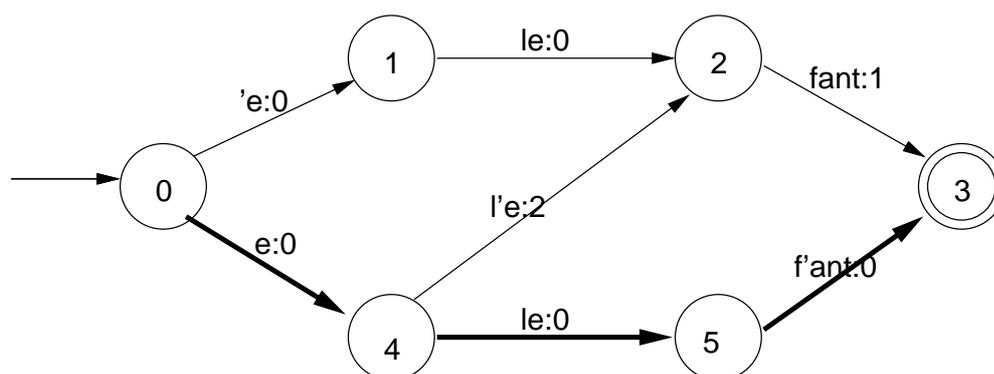


Abbildung 6.4: Ein Übergangsnetzwerk, in welchem das Betonungsmuster *Elefánt* als optimaler Kandidat beschrieben wird

Dieses Übergangsnetzwerk würde auch dann noch den richtigen Kandidaten als optimal auswählen, wenn *Constraint 1* und *Constraint 2* gleich stark gewichtet wären. Dies liegt aber daran, daß der durch Pfad 2 beschriebene Kandidat zwei *Constraints* verletzt. In Fällen, wo *Constraints* in Konkurrenz zueinander stehen, ist deren Hierarchisierung unumgänglich. Dies wird anhand eines Beispiels erläutert. Hierzu wird ein weiterer *Constraint* in die Constrainthierarchie eingeführt, welcher sich an Wortbetonungsregel 5 auf Seite 113 orientiert. Diese Regel stellte sich bei der Vorhersage der Wortbetonung als weniger wichtig heraus, scheint aber dennoch Relevanz zu besitzen und kommt bei Wörtern, die auf eine leichte Silbe auslauten, zum Zug (z.B. *Libéllé*, *Casíno*, *Kaléndér* ...).

Optimalitätstheoretischer *Constraint 3*

Betonbare Paenultimasilben müssen betont werden.

Würden nun *Constraint 2* und *3* gleich gewichtet, so könnte man in einem Wort wie *subtil*, welches auf eine schwere Endsilbe auslautet, aber eine betonbare vorletzte Silbe hat, nicht mehr die korrekte Endbetonung als optimal bestätigen. Daher muß der *Constraint 3* in der Hierarchie niedriger gewichtet werden und darf im Falle einer Verletzung nicht so viele Kosten verursachen wie *Constraint 2*. Die Anwendung auf das Beispielwort *subtil* wird in Abbildung 6.5 illustriert.

Einen Spezialfall innerhalb der Betonungsmuster bilden Wörter, die auf der Antepaenultimasilbe betont werden, wie z.B. *Kár.ne.val*. Der generative Regelansatz konnte diese Fälle vorhersagen, da

1. der letzten Silbe nur im Falle einer vorliegenden schweren Silbe die Betonung zufiel.
2. nur betonbare Silben die Wortbetonung erhalten konnten.

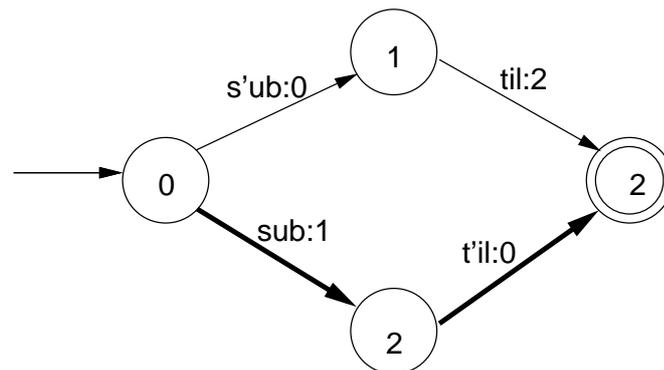


Abbildung 6.5: Ein Übergangnetzwerk, in welchem das Betonungsmuster *subtíl* als optimaler Kandidat beschrieben wird

Dies wurde auch als *Antepaenultima*regel in das Regelsystem mitaufgenommen (vgl. Wortbetonungsregel 6 auf Seite 113). War also die letzte Silbe leicht und die vorletzte Silbe unbetonbar, so fiel die vorhergesagte Wortbetonung auf die Antepaenultima-silbe. Um diese Regularität optimalitätstheoretisch zu modellieren, müssen einem Kandidaten mit einer leichten, aber betonten Endsilbe ebenfalls Kosten zugewiesen werden. Ansonsten könnte nicht zwischen den Kandidaten *Kár.ne.val* und *Kar.ne.vál* unterschieden werden (vgl. Abbildung 6.6), da beide Pfade keinerlei Kosten erzeugen würden. Hierzu wird der *Constraint 2* entsprechend erweitert.

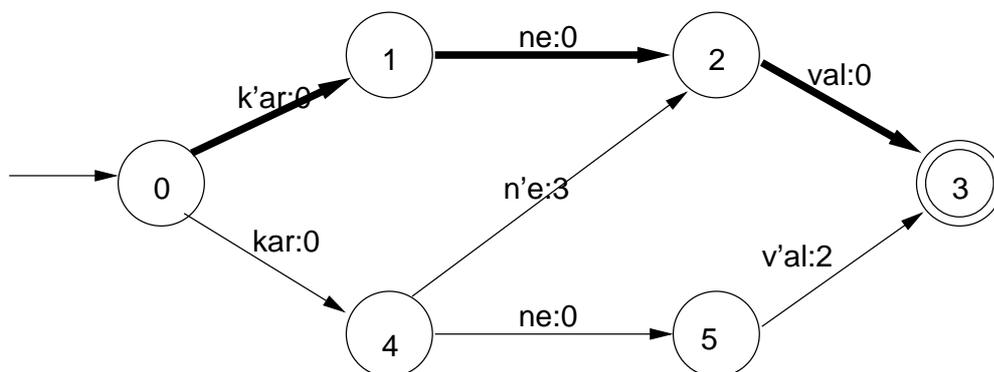


Abbildung 6.6: Ein Übergangnetzwerk, in welchem das Betonungsmuster *Kárneval* als optimaler Kandidat beschrieben wird

Modifizierter Optimalitätstheoretischer Constraint 2

Schwere Endsilben müssen betont werden, aber leichte Endsilben dürfen nicht betont werden.

Innerhalb dieses *Constraintsystems* ist noch nicht die Wortbetonungsregel 4 von Seite 113 berücksichtigt. Diese führte zu einer Betonung auf der Antepaenultimasilbe, wenn die vorletzte Silbe auf einem hohen Vokal endete und die letzte Silbe keinen Anlautkonsonanten enthielt. Um diese abbilden zu können, muß ein weiterer *Constraint* eingeführt werden, der sich in Kosten niederschlägt. Innerhalb der *Constrainthierarchie* steht er an gleicher Stelle mit dem *Constraint 2*. Ansonsten könnte das Betonungsmuster in *Jú.li.a* nicht korrekt vorhergesagt werden, da die Kosten für die Form *Ju.lí.a* höher sein müssen als für die Form *Jú.li.a* (vgl. Abbildung 6.7):

Optimalitätstheoretischer Constraint 4

Vorletzte Silben, die auf einem hohen Vokal auslauten, dürfen nicht betont werden, wenn die letzte Silbe keinen Anlaut hat.

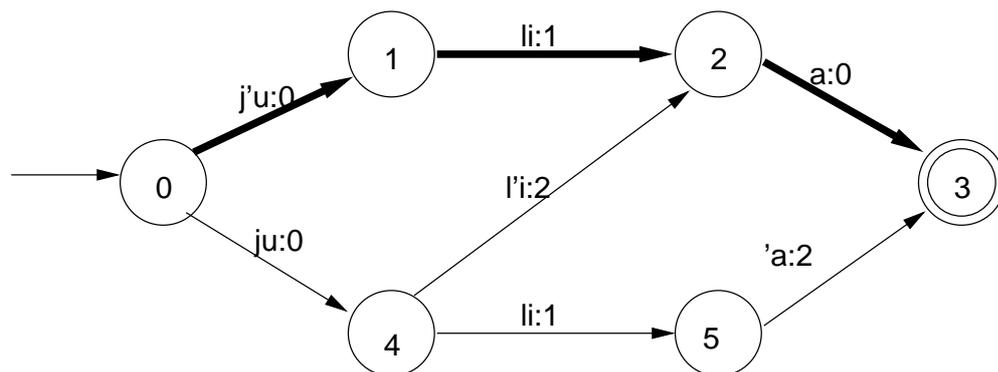


Abbildung 6.7: Ein Übergangnetzwerk, in welchem das Betonungsmuster *Júlia* als optimaler Kandidat beschrieben wird

Insgesamt ergibt sich daher folgende optimalitätstheoretische *Constraint-Hierarchie*, nach welcher die Kosten in absteigender Reihenfolge vergeben werden:

Constraint 1 >> *Constraint 2*, *Constraint 4* >> *Constraint 3*

Interessanterweise spiegelt die *Constrainthierarchie* in den meisten Punkten die Reihenfolge wider, die sich aus der Analyse des Einflusses der einzelnen Regeln ergab (vgl. Abschnitt 5.3.3). Einzige Ausnahme bildet der *Constraint 4*. Dieser konnte — als Regel formuliert — nur wenig zur Verbesserung der Wortbetonungsvorhersage beitragen. Dies liegt aber in erster Linie an dem seltenen Vorkommen der Konstellation, die zur Anwendung der Regeln bzw. des *Constraints* führt. Weiterhin interessant ist die Beobachtung, daß es keine Notwendigkeit gibt, die Antepaenultimaregel (vgl.

Seite 113) explizit optimalitätstheoretisch zu modellieren. Die Auswahl der vorvorletzten Silbe als Trägerin der Wortbetonung bei entsprechender Konstellation ergibt sich bereits aus der Verletzung der anderen *Constraints*. Allerdings mußte ein *Constraint* hierfür erweitert werden, so daß sich das Regelsystem insgesamt sehr stark in der Formulierung der *Constraints* widerspiegelt.

6.3 Fazit

In diesem Kapitel wurde zunächst eine formale Modellierung von Wortbetonung auf der Basis endlicher Übergangnetzwerke vorgestellt, die den Anforderungen der Optimalitätstheorie gerecht wird. Es wurde ein Algorithmus angegeben, mit dem eine Kandidatenmenge als endlicher Automat dargestellt werden kann. Anschließend wurden *Constraints* formuliert, mit deren Hilfe den einzelnen Übergängen innerhalb der Übergänge in den Automaten *Disharmoniewerte* oder *Kosten* vergeben wurden. Die Summe der Kosten auf dem Ausgabeband ergibt die Gesamtkosten für den jeweiligen Kandidaten, der durch das Eingabeband beschrieben wird. Der optimale Kandidat kann durch Minimierung der Gesamtkosten ermittelt werden. Ausgehend von den generativen Regeln, welche für die Bestimmung der Wortbetonung in Kapitel 5 ermittelt wurden, wurden *Constraints* abgeleitet. Um die korrekten Kandidaten bestimmen zu können, wurden den jeweiligen *Constraints* bei deren Mißachtung unterschiedliche Kosten zugewiesen. Diese Kosten ergaben eine Constraint-Hierarchie, innerhalb welcher den hochrangigen *Constraints* die meisten Kosten zugewiesen werden. Diese Hierarchie entsprach in wesentlichen Aspekten den Ergebnissen der Analysen zum Einfluß der einzelnen Regeln auf die Korrektheit der vorhergesagten Wortbetonung. Dies zeigt, daß die in Kapitel 3 vorgestellte Methode durchaus Ergebnisse liefert, welche auch für optimalitätstheoretische Modellierungen geeignet sind. Eine Implementierung der hier gezeigten Automaten unter Zuhilfenahme des Algorithmus für GEN und die Annotierung der Pfade mit Hilfe der *Constraints* könnte aber ohne weiteres dazu genutzt werden, um auf der Basis optimalitätstheoretischer *Constraints* Vorhersagen zu leisten, da ein generativer Algorithmus GEN vorliegt, welcher die Kandidatenmenge beschreibt. Insgesamt kann daher gesagt werden, daß

- den methodologischen Anforderungen aus Kapitel 3 auch im Sinne einer optimalitätstheoretischen Betrachtung mit Hilfe der vorgestellten Formalisierung entsprochen werden kann.
- die Ergebnisse der empirischen Analysen sich auch in eine optimalitätstheoretische Modellierung integrieren lassen.

Kapitel 7

Zusammenfassung der Resultate

An dieser Stelle werden die Ergebnisse der einzelnen Kapitel nochmals zusammengefaßt und zueinander in Beziehung gesetzt, um ein Gesamtresümee ziehen zu können. Einen Gesamtüberblick über die zentralen Punkte und Ergebnisse der Arbeit liefert Abbildung 7.1.

7.1 Motivation der Arbeit

Gegenstand von Kapitel 1 war die Motivation der vorliegenden Arbeit. Der dort dargestellte Ausgangspunkt für die weiteren Überlegungen war die Beobachtung, daß ein großes Problem phonologischer-prosodischer Modelle im Bereich der mangelhaften Validität introspektiver Verfahren liegt. Es wurde gefolgert, daß aussagekräftige Modelle in diesem Bereich anhand objektiv gewonnenen empirischen Datenmaterials evaluiert werden müssen. Ein weiterer Punkt der Motivation bildete die Kritik an der in der Phonologie noch immer üblichen Vorgehensweise, Hypothesen lediglich anhand kleiner Sprachfragmente zu überprüfen. Diese Vorgehensweise kann dazu führen, daß wichtige Generalisierungen oder aber Regelinteraktionen übersehen werden. Gerade innerhalb von praktischen sprachtechnologischen Anwendungen, wie sie in den letzten Jahren immer wichtiger geworden sind, ist aber eine adäquate Modellierung des gesamten sprachlichen Phänomenbereichs notwendig. Diese Anforderungen erfordern außerdem eine formale Beschreibung der Regularitäten, da ansonsten keine größeren Datenmengen verarbeitet werden können und eine Integration der Ergebnisse in sprachtechnologische Applikationen ebenfalls unmöglich ist. Es wurden in der Motivation demnach zwei Problembereiche aufgespannt, die es zu lösen galt:

- Wie können prosodische Betonungsmuster objektiv erfaßt werden, ohne auf Introspektion zurückgreifen zu müssen?
- Wie können die Regularitäten oder phonologischen Hypothesen formal erfaßt und zu den objektiv erfaßten Daten in Beziehung gesetzt werden?

Die erste sich aus der Motivation ergebende Fragestellung bildete den Gegenstand von Kapitel 2, die zweite den Gegenstand von Kapitel 3.

7.2 Das Prominenzmodell

In Kapitel 2 wurde mit der *Prominenz* eine objektive phonetische Meßgröße eingeführt, welche introspektiv erhobene Daten zur Betonung ersetzen kann. Prominenz wurde als gradueller Parameter verstanden, welcher die wahrgenommene Betonungsstärke einer Silbe relativ zu ihrer Umgebung erfaßt. Die Prominenz eignet sich zur Erhebung von Betonungsmustern besonders gut, da gezeigt werden konnte, daß Hörer hinsichtlich der wahrgenommenen Prominenzmuster hohe Übereinstimmungen zeigen. Im Gegensatz zur Prominenz wird die linguistische Betonung jedoch als kategoriale Eigenschaft einer linguistischen Einheit auf unterschiedlichen Ebenen wie der Silbe, dem phonologischen Wort oder der Äußerung zugewiesen. Phonologische Modellierungen, wie sie aus der Generativen Phonologie, der Metrischen Phonologie und auch der Deklarativen Phonologie bekannt sind, gehen implizit davon aus, daß sich die relativen Betonungsstärken in unterschiedlichen Prominenzwerten niederschlagen. Diesen Schluß lassen Betonungsmodelle zu, die mit metrischen Gittern oder Zahlenwerten als Repräsentanten für Betonungsstärken operieren. Es wurde ebenfalls dargestellt, wie sich die phonologischen Modellierungen der Prominenz formal in Form endlicher Übergangsnetzwerke darstellen lassen. Die den phonologischen Modellen zugrundeliegende Intuition, daß sich linguistische Betonung als wahrgenommene Prominenz niederschlägt, konnte durch statistische Analysen und ein Perzeptionsexperiment bestätigt werden. Es zeigte sich, daß die linguistisch relevanten kategorialen Unterschiede sich auch in den wahrgenommenen Prominenzwerten widerspiegeln. Die Prominenzwerte enthalten aber darüber hinaus noch Informationen zu paralinguistischen Phänomenen wie z.B. der Emphase, mit der eine Sprecherin eine bestimmte Äußerung tätigt. Hier konnte anhand statistischer Analysen gezeigt werden, daß die paralinguistisch erklärbaren Prominenzphänomene offenbar nicht mit den linguistisch motivierten Prominenzmustern in Konflikt geraten. Die linguistisch relevanten äußerungsinternen Prominenzmuster bleiben auch in unterschiedlichen Sprechstilen weitgehend erhalten. Die Untersuchungen aus Kapitel 2 zeigen insgesamt, daß sich die Prominenz als phonetische Meßgröße eignet, um damit die Validität phonologischer Vorhersagen zur Betonung zu überprüfen.

7.3 Die Methode zum Vergleich phonetischer Messungen und phonologischer Vorhersagen

Neben dem Vorliegen einer phonetischen Meßgröße muß nunmehr eine Methode entwickelt werden, anhand deren mit Hilfe der gemessenen Daten eine phonologi-

sche Vorhersage zur Betonung tatsächlich überprüft werden kann. Hierzu wurden zunächst Gütekriterien aufgestellt, die für phonologische Grammatiken gelten sollten. Als grundlegende Anforderungen an eine Grammatik wurden in Anlehnung an (CHOMSKY 1965) die empirische Adäquatheit, die Minimierung des Regelapparats sowie die formale Darstellbarkeit formuliert. Der Anspruch der formalen Darstellbarkeit ist bzgl. einer Überprüfung der empirischen Adäquatheit schon deshalb notwendig, weil nur durch automatische Verfahren größere Datenmengen hinsichtlich ihrer Strukturen überprüft werden können. Weiterhin behandelt das Kapitel die Frage, ob und wie die phonetisch erhobenen Daten zu den phonologischen Vorhersagen überhaupt in Beziehung gesetzt werden können. Hierbei wurde auf die Ansätze der *Experimental Phonology* sowie der *Laboratory Phonology* zurückgegriffen. In beiden Forschungsprogrammen wird die introspektive Methode durch die Erhebung objektiven phonetischen Datenmaterials ersetzt. Die *Experimental Phonology* richtet ihr Augenmerk zusätzlich in erster Linie auf die phonologische Vorhersage mit Hilfe formaler Modelle, die es erlauben, automatische Vorhersagen über größere Datenmengen zu betrachten, aus denen Generalisierungen abgeleitet werden können. Die *Laboratory Phonology* widmet sich verstärkt dem Problem der Abbildung von diskreten phonologischen Einheiten auf graduelle phonetische Parameter. Die formale Darstellung phonologischer Betonungsmuster war bereits Gegenstand von Kapitel 2, das Abbildungsproblem wurde dort hinsichtlich der Prominenz ebenfalls diskutiert. Die Tatsache, daß sich aus phonetischen Prominenzmustern kategoriale Betonungsmuster ablesen lassen, spricht für die Möglichkeit eines Vergleichs phonologisch vorhergesagter und phonetisch gemessener Prominenzmuster. Als phonetisches Datenmaterial sollte für den Vergleich die Bonner Prosodische Datenbank verwendet werden, welche nach dem Verfahren von (FANT und KRUCKENBERG 1989) mit Silbeprominenzen annotiert wurde. Ein Vergleich phonologischer Betonung und phonetischer Prominenz kann auf der Ebene der Ordinalskala erfolgen, da sich für beide Modellierungen Ranglisten hinsichtlich relativer Prominenz aufstellen lassen. Bei einem Vergleich phonologischer Betonungsmuster und phonetisch erhobener Daten ist allerdings zu beachten, daß die phonologischen Beschreibungen von einer idealisierten Form ausgehen, die für alle Sprecher einer Sprache verbindlich zutreffen sollte. Ein phonetisches Datum hingegen bildet selbstverständlich genau das ab, was eine einzelne Sprecherin in einer bestimmten Situation geäußert hat. Um zu bestimmen, wie stark diese Variationen hinsichtlich der Prominenzmuster sind, wurde in allen auf Silbenebene vergleichbaren Äußerungen eine Korrelation zwischen den verschiedenen Datenbanksprecherinnen berechnet und als *Baseline* für die phonologischen Vorhersagen verwendet. Es wurde ferner ausgeschlossen, daß unterschiedliche Fokussierungen in erster Linie ursächlich für stark abweichende Prominenzmuster bei den Sprecherinnen waren, da ansonsten diese bei der phonologischen Vorhersage stark berücksichtigt werden müßten. Die Äußerungen sind sprecherinnenabhängig variabel. Auch eine gewisse Hörervariation hinsichtlich der Wahrnehmung von Prominenzen ist gegeben. Diese wird aber in der verwendeten Datenbank abgefangen, indem als relevanter Prominenzwert der Median über die

Annotationen verschiedener Hörer gebildet wird. Es steht somit also eine Methode zur Verfügung, mit deren Hilfe auf der Basis phonetischer Prominenzmessungen innerhalb einer Datenbank phonologische Vorhersagen zur Betonung überprüft und somit hinsichtlich ihrer empirischen Adäquatheit evaluiert werden können.

7.4 Anwendung der Methode auf die Äußerungsbetonung

Die Methode wurde in den folgenden Kapiteln 4 und 5 angewandt, indem verschiedene Vorhersagemodelle zunächst formalisiert und implementiert wurden. Dieses Vorgehen führte z.T. zu Vereinfachungen der Grammatiken, da die formale Herangehensweise Redundanzen aufdecken konnte. Anschließend wurden die Vorhersagen mit phonetisch erhobenen Prominenzmustern verglichen. Die Güte der Vorhersage wurde anschließend anhand der *Baseline* gemessen. Auf der Ebene der Äußerungsbetonung führte dies Vorgehen bei einem Vergleich einer syntaxbasierten Vorhersage der Wortbetonung sowie einer wortklassenspezifischen Vorhersage zu der Erkenntnis, daß letztere die bessere Modellierung liefert. Da die Vorhersageergebnisse aber noch weit entfernt von der angestrebten *Baseline* waren, wurden zusätzliche Erkenntnisse aus dem Bereich der Sprachtechnologie in die Vorhersage integriert. Eine Vorhersage, welche eine feinere Unterteilung in Wortklassen vornahm, konnte die Vorhersagen soweit verbessern, daß die angestrebte *Baseline* erreicht wurde. Die Berücksichtigung der Auftretenshäufigkeit eines Wortes in großen Sprachkorpora erzielte hingegen keine Verbesserung der Vorhersage. Regeln, welche auf eine stärkere Rhythmizität der Sprache abzielen, ergaben eine geringfügige Verbesserung. Allerdings konnte lediglich Evidenz dafür erbracht werden, daß Sprecher allzulange Folgen unbetonter Silben vermeiden. Folgen mehrerer betonter Silben hingegen führen nicht ersichtlich zu einer Verschiebung von Betonung oder einer Deakzentuierung. Diese Ergebnisse stellen Grundannahmen der phonologischen Literatur in Frage, die eine Vermeidung von sogenannten Akzentzusammenstößen als einen grundlegenden Prozeß betrachtet. Den Abschluß von Kapitel 4 bildet die Formalisierung der resultierenden Regelmenge mit Hilfe konkatenierter endlicher Automaten. Diese Art der formalen Darstellung generativer phonologischer Regelsysteme bietet — wie in Kapitel 2 gezeigt — eine gute Grundlage für eine Integration in aktuelle Grammatikformalismen.

7.5 Anwendung der Methode auf die Wortbetonung

Anschließend wurde die Methode auch auf die Wortbetonung angewandt. In Kapitel 4 wurde noch von der Wortbetonung abstrahiert, da sich hierbei die größten Schwächen einer erfolgreichen Vorhersage gezeigt hatten. Um phonologische Wortbetonungsregeln zu testen, mußte das Datenmaterial erweitert werden, da

das ursprüngliche Korpus keinen repräsentativen Querschnitt der deutschen Wortbetonung liefern konnte. Basierend auf der Zusammenfassung deutscher Wortbetonungsregeln von (JESSEN 1998), wurde ein Regelwerk erarbeitet, teilweise vereinfacht, implementiert und anhand der manuellen Annotationen der Verbmobil-Wortliste überprüft. Das resultierende Regelwerk stellte sich als recht erfolgreich heraus und kann mit den Ergebnissen von Vorhersagemodellen auf der Basis Neuroner Netze konkurrieren. Wird der Vorhersagealgorithmus mit den Regeln aus Kapitel 4 verknüpft und auf die Prominenzannotationen der BPD angewandt, sinkt die Korrelation zwischen vorhergesagten und wahrgenommenen Prominenzen allerdings nach wie vor auf einen Wert unterhalb der *Baseline*. Das Regelwerk basiert auf den folgenden wichtigen Grundannahmen, daß

- bestimmte native Flexionssuffixe als extrametrisch betrachtet und bei der Zuweisung der Wortbetonung entsprechend ignoriert werden müssen,
- finale schwere Silben die Betonung an sich ziehen,
- ansonsten die Betonung auf die vorletzte bzw. die dem rechten Wortrand am nächsten liegende betonbare Silbe fällt.

Eine Untersuchung der fehlerhaften Vorhersagen führte weiterhin zu dem Ergebnis, daß die Unterteilung in leichte und schwere Silben zu grob war. Daher wurde ein Produktionsexperiment durchgeführt, was zu einer Erweiterung der Silbengewichtshierarchie im Deutschen führte. Die resultierende Silbengewichtshierarchie beinhaltet zwei Kategorien mehr als die einfache Unterteilung in „leicht“ vs. „schwer“. Anschließend wurde die phonologische Regel der Akzentverschiebung in komplexen Komposita der Form (A(BC)) anhand perceptiver Studien untersucht. Wie bereits die Ergebnisse auf der Äußerungsebene vermuten ließen, ergaben auch diese Studien, daß die Akzentverschiebung im Deutschen keine große Relevanz hat. In viersilbigen Komposita konnte eine Akzentverschiebung allerdings tendenziell beobachtet werden.

Da bei der Analyse der fehlerhaften Vorhersage signifikant viele Eigennamen auftauchten, wurde ferner untersucht, ob es möglich ist, durch ein separates Regelwerk, welches speziell auf Eigennamen zugeschnitten wurde, bessere Vorhersagen zu erzielen. Untersuchungen an dem Eigennamenausprachelexikon ONOMASTICA sowie ein anschließendes Produktionsexperiment zeigten die starke Tendenz, daß bei Eigennamen schwere Endsilben nicht die Betonung an sich ziehen. Wurde diese Regel bei der Betonungsvorhersage weggelassen, konnte der Vorhersagefehler bei Eigennamen stark reduziert werden. Dennoch bleiben die Ergebnisse der Vorhersage aufgrund der starken individuellen Ausprägung von Eigennamen und vieler fremdsprachlicher Einträge noch weit hinter den Ergebnissen für die Verbmobil-Wortliste zurück. Die Tendenz, daß bei Eigennamen eher der „germanische Wurzelakzent“ zum Tragen kommt, der sich in einer Betonung des linken Wortrandes äußert, konnte nicht systematisch bestätigt, aber auch nicht widerlegt werden. Die Existenz des

„germanischen Wurzelakzents“ ist in neueren phonologischen Arbeiten meist abgestritten worden. Um überzeugende Belege für seine Relevanz zu erbringen, sind weitere Untersuchungen notwendig.

Das Kapitel endet mit einer Formalisierung der Regelwerke für Eigennamen und Nicht-Eigennamen auf der Basis endlicher Übergangnetzwerke, die sich in Grammatikformalismen und sprachtechnologischen Applikationen integrieren lassen.

7.6 Eine optimalitätstheoretische Modellierung der Ergebnisse

Die phonologische Theorienbildung hat sich in den letzten Jahren von generativen Ansätzen weg zu einem evaluationsbasierten Ansatz entwickelt. Bei diesem Ansatz geht es darum, aus einer Reihe von Kandidatenformen die optimale Form auszuwählen. Diese Form ist diejenige, welche die wenigsten *Constraints* verletzt, wobei die *Constraints* unterschiedlich gewichtet sind. Die Erkenntnisse in den hier vorgelegten Studien zur Äußerungs- und Wortbetonung wurden naheliegenderweise zunächst auf der Basis generativer Regeln überprüft. Zusätzlich wurde daher untersucht, ob sich die Ergebnisse auch in einen optimalitätstheoretischen Rahmen einbetten lassen. Es wurde ein Algorithmus vorgestellt, mit dessen Hilfe sich Kandidaten für mögliche Betonungsmuster als endliche Automaten darstellen lassen. Anschließend wurden ausgehend von den ermittelten Wortbetonungsregeln *Constraints* formuliert, auf deren Basis den Übergängen im Automaten Disharmoniewerte oder Kosten zugewiesen wurden. Durch Aufsummierung der Kosten beim Durchlaufen des Automaten kann so der optimale Kandidat ermittelt werden. Die Untersuchungen ergaben eine *Constraint*hierarchie, anhand derer unterschiedlich hohe Kosten vergeben werden. Diese Hierarchie reflektiert den Einfluß, welche die den *Constraints* entsprechenden Regeln bereits für die Vorhersagegüte im generativen Ansatz hatten. Es konnte daher gezeigt werden, daß der Erkenntnisgewinn generativer Methoden sich in einen optimalitätstheoretischen Ansatz integrieren läßt. Außerdem wurde eine Formalisierung optimalitätstheoretischer *Constraints* der Wortbetonung vorgestellt.

7.7 Fazit

Um den empirischen Defiziten phonologischer Modellbildungen im Bereich der Betonung entgegenzutreten, wurde eine Evaluationsmethode entwickelt und angewendet, die folgende Punkte berücksichtigt:

- Die phonologischen Vorhersagen zur Betonung wurden formal dargestellt und implementiert.

- *Die Vorhersagen zur Betonung wurden an größeren, repräsentativen phonetischen Datensätzen überprüft, die mit Hilfe eines reliablen perzeptiven Annotationsverfahrens etikettiert waren.*
- *Die phonetischen Daten spiegeln die phonologischen kategorialen Muster wider. Die darüber hinaus in den phonetischen Daten enthaltenen Informationen konnte von den phonologisch relevanten Mustern separiert werden, so daß ein Vergleich phonologischer und phonetischer Muster möglich war.*

Auf der Basis dieser Methode wurden verschiedene phonologische Modellierungen zur Wort- und Äußerungsbetonung überprüft. Die Regelsätze konnten dabei hinsichtlich ihrer empirischen Adäquatheit z.T. falsifiziert und verbessert werden. Teilweise konnten sie auch vereinfacht werden. Die wichtigsten Erkenntnisse dieser Evaluationen sind die folgenden:

- *Syntaktische Phrasierung spielt nur eine geringe Rolle für die Betonungszuweisung auf Äußerungsebene.*
- *Mit Hilfe der Wortklasseninformation, die über eine einfache Differenzierung in Inhalts- und Funktionswörter hinausgeht, läßt sich die Prominenz auf Äußerungsebene gut vorhersagen.*
- *Die Auftretenshäufigkeit eines Wortes in der deutschen Sprache steht in keinem direkten Zusammenhang zu seiner Betonungsstärke.*
- *Deakzentuierung spielt auf der Äußerungsebene, aber auch bei wortinternen provozierten Akzentzusammenstößen nur eine marginale Rolle im Deutschen.*
- *Lange Folgen unbetonter Silben, die insbesondere bei langen Wörtern auftauchen, werden im Deutschen vermieden.*
- *Das Silbengewicht spielt im Deutschen eine herausragende Rolle bei der Platzierung der Wortbetonung: Sofern die letzte Silbe signifikant schwerer ist als die vorletzte, fällt die Wortbetonung an den rechten Wortrand.*
- *Die Silbengewichtshierarchie muß allerdings erweitert werden, um alle Wortbetonungsphänomene erklären zu können.*
- *Ist die finale Silbe leicht, so fällt die Betonung i.d.R. auf die dem rechten Wortrand am nächsten stehende betonbare Silbe.*
- *Bei der Betonung von Eigennamen im Deutschen ist der Einfluß des Silbengewichts wesentlich weniger stark als bei der Betonung von Nicht-Eigennamen. Daher mußte hierfür ein separates Regelwerk entwickelt werden.*

Alle Erkenntnisse wurden formal notiert, so daß ihrer Integration in Grammatikformalismen sowie sprachtechnologischen Anwendungen nichts im Wege steht. Abschließend wurde gezeigt, daß die aus den Evaluationen gewonnenen Erkenntnisse sich auch in den nicht-generativ geprägten Formalismus der Optimalitätstheorie einbinden lassen. Auch für diese Modellierung wurde eine Formalisierung in Form von endlichen Übergangnetzwerken vorgeschlagen. Auf dem Eingabeband repräsentieren die Pfade durch den Automaten mögliche Kandidaten für eine Betonung. Jeder Übergang im Automaten stellt u.U. eine *Constraint*verletzung dar. Die Verletzung dieser *Constraints* schlägt sich in Form von Kosten auf dem Ausgabeband des Übergangnetzwerkes nieder.

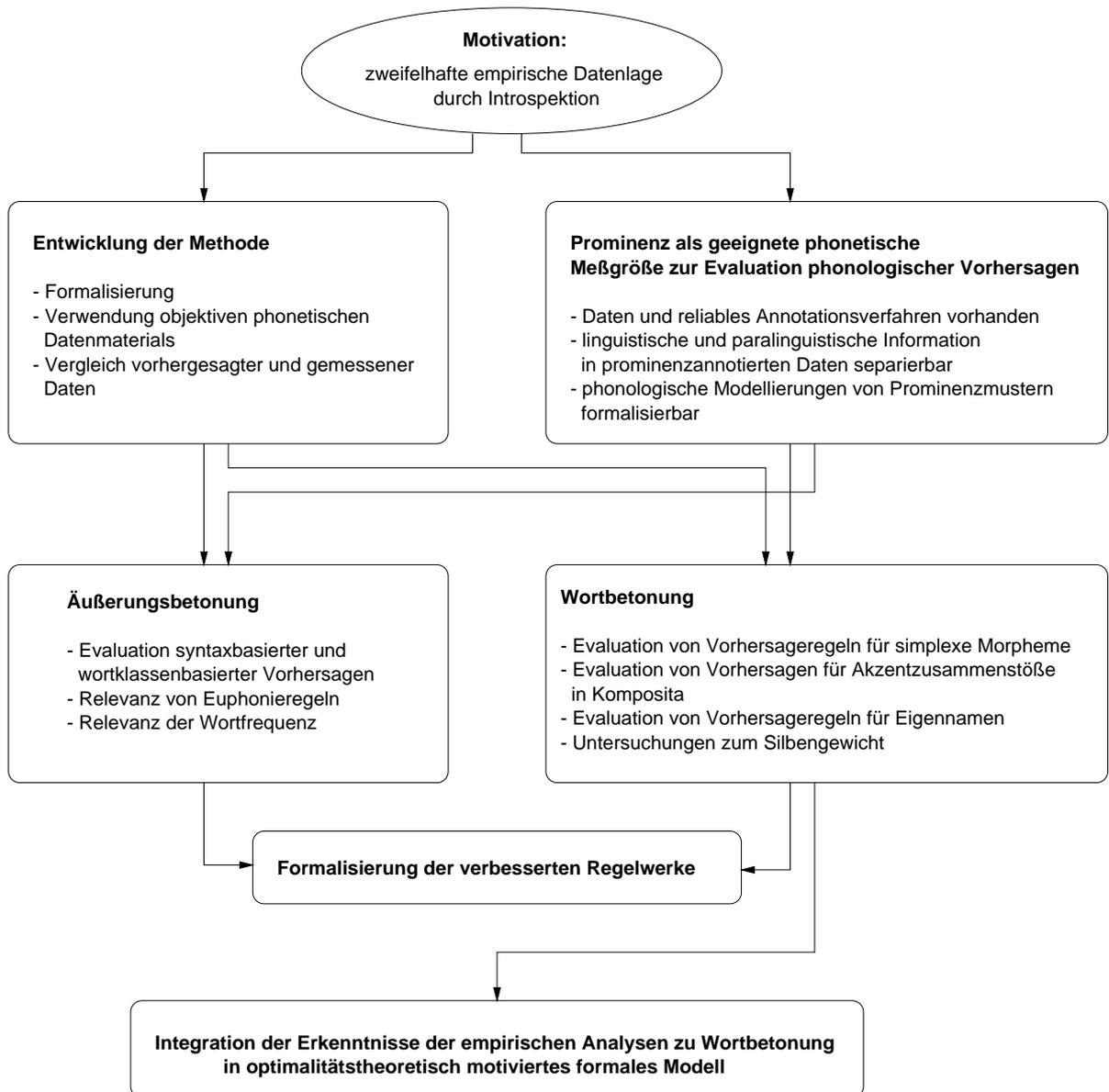


Abbildung 7.1: Gesamtüberblick über die zentralen Arbeitsschritte und Ergebnisse der Arbeit

Kapitel 8

Schlußbemerkungen und Ausblick

Ich beherrsche die deutsche Sprache, aber sie gehorcht nicht immer.
Alfred Polgar

Die Anwendung der in dieser Arbeit entwickelten Evaluationsmethode führte zu einer Präzisierung und Weiterentwicklung phonologischer Modellierungen der deutschen Wort- und Äußerungsbetonung. Die Erkenntnisse ließen sich elegant in einen nicht generativ geprägten Formalismus der Optimalitätstheorie integrieren, was den Schluß nahelegt, daß die hier vorgestellte Methode durchaus ähnlich genutzt werden kann, um auch optimalitätstheoretische *Constraints* zu evaluieren. Neben der Falsifikation einiger Regeln wie der sogenannten Euphonieregeln oder auch dem Infragestellen des Einflusses des Silbengewichts bei der Vorhersage von Eigennamen führten die Ergebnisse zu einigen Weiterentwicklungen bestehender Regelsysteme. So wurden Erkenntnisse sprachtechnologischer Forschung in die Vorhersage der Äußerungsbetonung integriert und Experimente zur Silbengewichtsproblematik durchgeführt. Die Formalisierung der Analyseergebnisse bietet neben ihrer Integrierbarkeit in moderne Grammatikformalismen die Möglichkeit einer Anwendung in sprachtechnologischen Applikationen wie der Sprachsynthese. Die Vorhersageergebnisse im Bereich der Wortbetonung sind für das resultierende Regelsystem ähnlich gut wie Vorhersagen, die speziell für Sprachsyntheseanwendungen auf der Basis Neuronaler Netze entwickelt wurden. So bietet die hier dargestellte Formalisierung neben dem Vorteil der weitgehenden empirischen Adäquatheit den Vorteil eines sprachwissenschaftlichen Erkenntnisgewinns.

Trotz der guten Vorhersageergebnisse ist zu berücksichtigen, daß sicherlich noch einige Verbesserungen der Vorhersageregeln notwendig sind. Insbesondere die spezielle Problematik der Betonung von Eigennamen ist in der hier vorliegenden Arbeit lediglich ansatzweise gelöst worden. Auch die Ergebnisse zu Akzentzusammenstößen in Komposita lassen noch einige Fragen offen, welche eventuell mit Experimenten auf der Basis delexikalisierter Sprache beantwortet werden können. Weitere ungelöste Fragestellungen ergeben sich im Hinblick auf unterschiedliche Sprechstile. Es ist unklar, ob sich die hier gewonnenen Erkenntnisse zur marginalen Rolle von

Euphonieregeln im Deutschen ohne weiteres aufrechterhalten lassen, sofern weitere Sprechstile wie schnelle Sprache oder Spontansprache berücksichtigt werden. Um diesen Problemen in Zukunft begegnen zu können, bedarf es allerdings der aufwendigen Erhebung geeignet annotierter Daten. Die Analyseergebnisse anhand gelesener Sprache sind sicherlich ein erster Schritt, aber noch nicht die Antwort aller offener Fragen. Die Möglichkeiten der hier entwickelten Methode sowie ihre erfolgreiche Anwendung sollten aber deutlich gemacht haben, daß diese weiteren Schritte sowohl durchführbar als auch lohnenswert sind.

Literaturverzeichnis

- ABB, B., C. GÜNTHER, M. HERWEG, K. LEBETH, C. MAIENBORN und A. SCHOPP (1993). *Incremental Syntactic and Phonological Encoding - An Outline of the SYNPHONICS Formulator*. In: SCHOPP, A. (ed.). *Hamburg Working Papers on Language Production - II*. S. 17–32. Universität Hamburg.
- ABB, B. und K. LEBETH (1992). *Der SYNPHONICS-Formulator*. In: HERWEG, M. (ed.). *Hamburg Working Papers on Language Production - I*. S. 53–81. Universität Hamburg.
- ABELN, S. und C. GÜNTHER (1993). *Event-Based Representation of Phonetic Knowledge in a Concept-To-Speech System*. In: SCHOPP, A. (ed.). *Hamburg Working Papers on Language Production - II*. S. 89–122. Universität Hamburg.
- BATLINER, A., R. HUBER, H. NIEMANN, E. NÖTH, J. SPILKER und K. FISCHER (2000). *The Recognition of Emotion*. In: *Verbmobil: Foundations of Speech-to-Speech Translation*. Springer: Berlin, Heidelberg.
- BAUMANN, S., M. GRICE und R. BENZMÜLLER (2000). *GToBI - a phonological system for the transcription of German intonation*. In: *Proceedings of Prosody 2000, Cracow*.
- BAUMANN, S. und J. TROUVAIN (2001). *The prosody of German telephone numbers*. In: *Proceedings of EUROSPEECH 2001*, Band 1, S. 557–560, Aalborg, Danmark.
- BECKMAN, M. (1986). *Stress and Non-Stress Accent*. Foris Publications: Dordrecht.
- BECKMAN, M. und J. EDWARDS (1990). *Lengthenings and shortenings and the nature of prosodic constituency*. In: KINGSTON, J. und M. BECKMAN, (eds.): *Between the Grammar and Physics of Speech*, Band I aus *Papers in Laboratory Phonology*, S. 152–178. Cambridge University Press: Cambridge.
- BECKMAN, M. und J. PIERREHUMBERT (1986). *Intonational structure in Japanese and English*. *Phonology Yearbook*, 3:15–70.
- BECKMAN, M. und J. KINGSTON (1990). *Introduction*. In: KINGSTON, J. und M. BECKMAN, (eds.): *Between the Grammar and Physics of Speech*, Band I aus *Papers in Laboratory Phonology*, S. 1–16. Cambridge University Press: Cambridge.

- BEHN, S. (1912). *Der deutsche Rhythmus und sein eigenes Gesetz. Eine experimentelle Untersuchung*. Karl J. Trübner: Straßburg.
- BENWARE, W. A. (1987). *Accent Variation in German Nominal Compounds of the Type (A(BC))*. *Linguistische Berichte*, 108:102–127.
- BERG, T. (1997). *Lexical Stress Differences in English and German: the special status of proper nouns*. *Linguistische Berichte*, 167:3–22.
- BERINSTEIN, A. (1979). *A Cross-Linguistic Study on the Perception and Production of Stress*. UCLA Working Papers in Phonetics, 47. Department of Linguistics, UCLA.
- BEUTNAGEL, M. und A. CONKIE (1999). *Interaction of Units in a Unit Selection Database*. In: *Proceedings of EUROSPEECH '99*, Band 3, S. 1063–1066.
- BIRD, S. (1992). *Finite-State Phonology in HPSG*. In: *Proceedings of the Fifteenth Conference on Computational Linguistics COLING '92*, S. 74–80. International Committee on Computational Linguistics.
- BIRD, S. und T. M. ELLISON (1992). *One Level Phonology*. Technischer Bericht, University of Edinburgh.
- BIRD, S. und T. M. ELLISON (1994). *One-level phonology: autosegmental representations and rules as finite automata*. *Computational Linguistics*, 20:55–90.
- BOERSMA, P. (1998). *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives.*, Band 11 aus *LOT International Series*. Holland Academic Graphics: The Hague. auch: Doctoral Thesis, University of Amsterdam.
- BOLINGER, D.L. (1972). *Accent is predictable (if you're a mind reader)*. *Language*, 48:633–644.
- BRAUN, W. (1993). *Etymologisches Wörterbuch des Deutschen*. Akademie Verlag: Berlin. Erarbeitet im Zentralinstitut für Sprachwissenschaft, Berlin, unter Leitung von Wolfgang Pfeifer.
- BROWMAN, C. und L. GOLDSTEIN (1986). *Towards an Articulatory Phonology*. *Phonology Yearbook*, 3:219–252.
- BROWMAN, C. und L. GOLDSTEIN (1989). *Articulatory gestures as Phonological Units*. *Phonology*, 6:201–251.
- BRUCE, G. und P. TOUATI (1990). *On the analysis of prosody in spontaneous dialogues*. In: HOUSE, D. und P. TOUATI (eds.): *Working Papers*, Band 34, S. 37–55. Lund University, Department of Linguistics, Lund.
- BUSSMANN, H. (1990). *Lexikon der Sprachwissenschaft*. 2. Auflage. Kröner: Stuttgart.

- CAMPBELL, N. (1996). *CHATR: A high definition Speech Re-Sequencing System*. In: *Proceedings of the 3rd ASA/ASJ Joint Meeting*, S. 1223–1228, Hawaii, USA. ASA/ASJ.
- CARSON-BERNDSEN, J. (1991). *Ereignisstrukturen für phonologisches Parsen*. Technischer Bericht ASL-TR-9-92/UBI, University of Bielefeld.
- CARSON-BERNDSEN, J. (1993). *Time Map Phonology and the Projection Problem in Spoken Language Recognition*. Dissertation, University of Bielefeld. Erschienen in überarbeiteter Version bei Kluwer (Dordrecht) unter dem Titel: *Time Map Phonology – Finite-State Models and Event Logics in Speech Recognition*.
- CARSON-BERNDSEN, J. (1998). *Time Map Phonology – Finite-State Models and Event Logics in Speech Recognition*. Kluwer: Dordrecht.
- CELEX (1995). *The CELEX Lexical Database, Release 2*.
<http://www.kun.nl/celex/>.
- CHAFE, W. (1976). *Givenness, contrastiveness, subjects, topics, and points of view*. In: LI, C. (ed.): *Subject and Topic*, S. 25–56. New York: Academic Press.
- CHOMSKY, N. (1957). *Syntactic Structures*. Band IV aus *Janua Linguarum*. Mouton & Co: The Hague.
- CHOMSKY, N. (1965). *Aspects of a Theory of Syntax*. MIT Press: Cambridge, Massachusetts.
- CHOMSKY, N. und M. HALLE (1968). *The Sound Pattern of English*. Harper & Row: New York.
- COLEMAN, J. (1992a). *The phonetic interpretation of headed phonological structures using overlapping constituents*. *Phonology*, 9:1–44.
- COLEMAN, J. (1992b). *'Synthesis-by-rule' without segments or rewrite rules*. In: C. BAILLY, C. BENOÎT, T.R. SAVALLIS (ed.): *Talking Machines: Theories, Models and Designs*, S. 43–60. Elsevier Science: Amsterdam.
- COLEMAN, J. (1998). *Phonological Representations. Their names, forms, and powers*. Cambridge University Press: Cambridge.
- COOPER, W.E, S. EADY und P. MUELLER (1985). *Acoustical Aspects of contrastive stress in question-answer contexts*. *Journal of the Acoustical Society of America*, 77(6):2142–2156.
- DAELEMANS, W., S. GILLIS und G. DURIEUX (1994). *The Acquisition of Stress: a Data-Oriented Approach*. *Computational Linguistics*, 20(3):421–451. Special issue on Computational Phonology (Steven Bird guest editor).

- DIRKSEN, A. und J. COLEMAN (1997). *All-prosodic Speech Synthesis*. In: SANTEN, J.P. VON, R. SPROAT, J. OLIVE und J. HIRSCHBERG (eds.): *Progress in Speech Synthesis*, S. 91–108. Springer: New York.
- DOGIL, G. (1979). *Autosegmental Account of Phonological Emphasis*, Band 7 aus *Slavic Linguistics*. Edmonton: Canada.
- DOLESCHAL, U. (1988). *Zum deutschen Kompositionsakzent, Tema con variazioni*. *Wiener Linguistische Gazette*, 40–41:3–28.
- DRESHER, B.E. und J. KAYE (1990). *A computational learning model for metrical phonology*. *Cognition*, 34(2):137–195.
- EISENBERG, P. (1991). *Syllabische Struktur und Wortakzent*. *Zeitschrift für Sprachwissenschaft*, 10(1):37–64.
- ELLISON, M. (1994a). *Phonological Derivation in Optimality Theory*. In: *Proceedings of the 15th International Conference on Computational Linguistics*, S. 1007–1013.
- ELLISON, M. (1994b). *Phonological Derivation in Optimality Theory*. In: *Proceedings of COLING '94*, Band II, S. 1007–1013, Kyoto, Japan.
- ELSNER, A. (2000). *Erkennung und Beschreibung des prosodischen Fokus*. Dissertation, Universität Bonn. <http://hss.ulb.uni-bonn.de:90>
- ELSNER, A., M. WOLTERS, T. PORTELE, M. RAUTH und G. SONNTAG (1998). *Designing and Labelling a prosodic database for American English*. In: *Proceedings on the Workshop for Language Resources, Granada*, S. 1079–1082.
- ENDRES, B. O. (2000). *Deutsches Referenzkorpus des IDS Mannheim DEREKO*. <http://www.ids-mannheim.de/dereko/>.
- EOM, K.-W., J.-Y. KIM und S.-M. KIM (1999). *A Primary Study on the Randomness Control of the Prosodic boundary index for Natural Synthetic Speech*. In: *Proceedings of EUROSPEECH '99*, Band 4, S. 1828–1830.
- FANT, G. und A. KRUCKENBERG (1989). *Preliminaries to the study of Swedish Prose Reading and reading style*. *STL-QPSR 2/1989*, S. 1–68.
- FANT, G. und A. KRUCKENBERG (1999). *Prominence Correlates in Swedish Prosody*. In: *Proceedings of ICPhS '99*, San Fransisco, USA.
- FANT, G., A. KRUCKENBERG und J. LILJENCRAKTS (1998b). *Acoustical Correlates of perceived prominence. A study of Swedish Prosody*. Unveröffentlichtes Manuskript.
- FÉRY, C. (1986). *Metrische Phonologie und Wortakzent im Deutschen*. *Studium Linguistik*, 20:16–43.

- FÉRY, C. (1988). *Rhythmische und tonale Struktur der Intonationsphrase*. In: ALTMANN, H. (ed.): *Intonationsforschung*, S. 41–64. Niemeyer: Tübingen.
- FÉRY, C. (1998). *German Word Stress in Optimality Theory*. Rutgers Optimality Archive ROA 301-0399: <http://ruccs.rutgers.edu/roa.html>.
- FISCHENBECK, E. (2001). *Vorhersagen der Metrischen Phonologie und deren perzeptive Überprüfung: Akzentzusammenstöße in deutschen Komposita*. Magisterarbeit, Universität Bonn.
- FRANK, R. und G. SATTA (1998). *Optimality Theory and the Generative Complexity of Constraint Violability*. Rutgers Optimality Archive ROA 228-1197: <http://ruccs.rutgers.edu/roa.html>.
- FRY, D. B. (1958). *Experiments in the perception of stress*. *Language and Speech*, 1:126–152.
- FRY, D. B. (1955). *Duration and Intensity as Physical Correlates of Linguistics Stress*. *Journal of the Acoustical Society of America*, 35:765–769.
- FUCHS, A. (1984). 'Deaccenting' and 'Default Accent'. In: GIBBON, D. und H. RICHTER (eds.): *Intonation, Accent and Rhythm. Studies in Discourse Phonology*, Band 8 aus *Research in Text Theory*, S. 134–164. de Gruyter: The Hague.
- FUJISAKI, H. und K. HIROSE (1984). *Analysis of voice fundamental frequency contours for declarative sentences of Japanese*. *Journal of the Acoustical Society of Japan*, 5(4):233–241.
- GAZDAR, G., E. KLEIN, G. PULLUM und I. SAG (1985). *Generalized Phrase Structure Grammar*. Blackwell: Cambridge.
- GIBBON, D. (1987). *Finite State Processing of Tone Systems*. In: *Proceedings of the Third Conference of the European Chapter of the Association for Computational Linguistics*, S. 291–297.
- GIBBON, D. (2001). *Finite state prosodic analysis of African corpus resources*. In: *Proceedings of EUROSPEECH 2001*, Band 1, S. 83–86, Aalborg, Danmark.
- GIEGERICH, H. (1984a). *Compound Stress in German*. *Edinburgh Work in Progress*, 14:83–97.
- GIEGERICH, H. (1984b). *Metrische Phonologie und Kompositionsakzent im Deutschen*. *Papiere zur Linguistik*, 28:3–25.
- GIEGERICH, H. (1985). *Metrical Phonology and phonological structure. English and German*. Cambridge University Press: Cambridge.
- GOLDSMITH, J. (1979). *Autosegmental Phonology*. Garland Press: New York.

- GOLDSMITH, J. (1990). *Autosegmental and Metrical Phonology*. Blackwell: Oxford.
- GRABE, E. und P. WARREN (1995). *Stress Shift: Do speakers do it or listeners hear it?* In: CORNELL, B. und A. ARVANITI, (eds.): *Phonology and Phonetic Evidence*, Band IV aus *Papers in Laboratory Phonology*, S. 95–109. Cambridge University Press: Cambridge.
- GREWENDORF, G., F. HAMM und W. STERNEFELD (1987). *Sprachliches Wissen. Eine Einführung in moderne Theorien der grammatischen Beschreibung*. Suhrkamp: Frankfurt am Main.
- GUNDEL, J.K. (1977). *The role of topic and comment in linguistic theory*. Indiana University Linguistics Club: Bloomington.
- GÜNTHER, C. (1999). *Prosodie und Sprachproduktion*, Band 401 aus *Linguistische Arbeiten*. Niemeyer: Tübingen.
- GUSSENHOVEN, C. (2000). *Vowel duration, syllable quantity and stress in Dutch*. Rutgers Optimality Archive ROA 2: <http://ruccs.rutgers.edu/roa.html>.
- GUSSENHOVEN, C. und A. RIETVELD (1988). *Fundamental Frequency declination in Dutch: testing three hypotheses*. *Journal of Phonetics*, 16:355–369.
- HAIN, H.-U. und H. G. ZIMMERMANN (2001). *Optimierung der Eingabe eines neuronalen Netzes zur Bestimmung der Wortbetonung mit Hilfe von Weight Decay*. In: HESS, W. und K. STÖBER (eds.): *Elektronische Sprachsignalverarbeitung*, Band 22 aus *Studentexte zur Sprachkommunikation*, S. 282–289, ITG, DEGA, GI, w.e.b.: Dresden.
- HAYES, BRUCE (1983). *A Grid-based Theory of English Meter*. *Linguistic Inquiry*, 14:357–397.
- HAYES, B. (1995). *Metrical Stress Theory. Principles and Case Studies*. The University of Chicago Press: Chicago and London.
- HENRICI, T. (2001). *Über den Zusammenhang von Syntax und prosodischer Phrasierung*. Magisterarbeit, Universität Bonn.
- HEUFT, B., T. PORTELE, F. HÖFER, J. KRÄMER, H. MEYER, M. RAUTH und G. SONNTAG (1995). *Parametric description of f₀-contours in a prosodic database*. In: *Proceedings of the International Conference on Spoken Language Processing, Stockholm*, Band 2, S. 378–381.
- HEUFT, B. (1999). *Eine prominenzbasierte Methode zur Prosodieanalyse und -synthese*, Band 2 aus *Computer Studies in Language and Speech*. Peter Lang: Frankfurt am Main.

- HEUFT, B., T. PORTELE, P. WAGNER, C. WIDERA und M. WOLTERS (2000). *Perceptual Prominence*. In: SENDLMEIER, W. F. (ed.): *Speech and Signals. Aspects of Speech Synthesis and Automatic Speech Recognition*, Band 69 aus *Forum Phonetikum*. Hector. Festschrift dedicated to Wolfgang Hess on his 60th birthday.
- HIRTUM, A. VAN und D. BERCKMANS (2001). *The fundamental frequency of cough by autocorrelation analysis*. In: *Proceedings of EUROSPEECH 2001*, S. 2435–2438, Aalborg, Danmark.
- HÖHLE, T. (1998). *An Architecture for Phonology*. to be published in CSLI series.
- HOPCROFT, J.E. und J. ULLMAN (1979). *Introduction to Automata Theory, Languages and Computation*. Addison Wesley Publishing Company: Reading, Mass.
- IPA (1996). *Chart of the International Phonetic Alphabet, revised 1993, corrected 1996*. Journal of the International Phonetic Association.
- JESPERSEN, O. (1904). *Lehrbuch der Phonetik*. B.G. Teubner: Leipzig.
- JESSEN, M. (1998). *German*. In: HULST, H. VAN DER (ed.): *Word Prosodic Systems in the Languages of Europe*. Mouton de Gruyter: Berlin.
- JOHNSON, C. . (1972). *Formal Aspects of Phonological Description*. Mouton: The Hague.
- JONES, D. (1940). *An Outline of English Phonetics*. Dutton: New York.
- JONES, D. (1950). *The Phoneme*. W. Heffer and Sons Ltd.: Cambridge, UK.
- KAPLAN, R. und M. KAY (1994). *Regular Models of Phonological Rule Systems*. *Computational Phonology*, 20:331–378.
- KARTTUNEN, L. (1993). *Finite-State Constraints*. In: GOLDSMITH, J. (ed.): *The Last Phonological Rule. Reflections on Constraints and Derivations*, *Studies in Contemporary Linguistics*. The University of Chicago Press: Chicago and London.
- KARTTUNEN, L. (1998). *On the proper treatment of Optimality in Computational Phonology*. Unveröffentlichtes Manuskript.
- KAUFMANN, H. (1959). *Bildungsweise und Betonung der deutschen Ortsnamen*. Carl Winter Universitätsverlag: Heidelberg.
- KAY, M. (1987). *Non-concatenative Finite State Morphology*. In: *Proceedings of the Third Meeting of the European Chapter of the Association for Computational Linguistics*, S. 2–10.
- KENSTOWICZ, M. (1994). *Sonority-Driven Stress*. Rutgers Optimality Archive ROA 33-1094: <http://ruccs.rutgers.edu/roa.html>.

- KIPARSKY, P. (1966). *Über den deutschen Akzent*. In: *Untersuchungen über Akzent und Intonation im Deutschen*, Band VII aus *Studia Grammatica*, S. 69–98. Akademie Verlag: Berlin.
- KLATT, D. H., J. ALLEN und M. S. HUNNICUTT (1987). *From text to speech*. Cambridge University Press: Cambridge.
- KLOSA, A., K. KUNZEL-RAZUM, W. STOLZE-STUBENRECHT und M. WERMKE (eds.) (2001). *Deutsches Universalwörterbuch*. 4. Auflage. Dudenverlag: Mannheim, Leipzig, Wien, Zürich.
- KLUGE, F. (1995). *Etymologisches Wörterbuch der deutschen Sprache*. 23. Auflage. De Gruyter: Berlin, New York.
- KOHLER, K.J. (1987). *Categorical pitch perception*. In: *Proceedings of the Eleventh International Conference on Phonetic Sciences*, Band 5, S. 149–152, Tallinn.
- KORNAL, A. (1991). *Formal Phonology*. Dissertation, Stanford University.
- KOSKENNIEMI, K. (1983). *Two level morphology: A general computational model for word-form recognition and production*. University of Helsinki. Department of General Linguistics. Publication No. 11.
- KRAHMER, E. und M. SWERTS (1998). *Reconciling two competing views on contrastiveness*. In: *Proceedings of ICSLP 98, Sydney*.
- LADD, D. R. (1992). *An introduction to intonational phonology*. In: DOCHERTY, G.D. und D. LADD, (eds.): *Gesture, Segment, Prosody*, Band II aus *Papers in Laboratory Phonology*, S. 321–334. Cambridge University Press: Cambridge.
- LADD, D.R. (1996). *Intonational Phonology*. Band 79 aus *Cambridge Studies in Linguistics*. Cambridge University Press: London.
- LADD, D.R., J. VERHOEVEN und K. JACOBS (1994). *Influence of adjacent pitch accents on each others perceived prominence: Two contradictory effects*. *Journal of Phonetics*, 22:87–99.
- LAHIRI, A. und J. KOREMAN (1988). *Syllable weight and quantity in Dutch*. In: *West Coast Conference of Formal Linguistics 7*, S. 217–228.
- LAVER, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge University Press: Cambridge.
- LEA, W. (1977). *Acoustic Correlates of Stress and Juncture*. In: HYMAN, L. (ed.): *Studies in Stress and Accent*, Band 4 aus *Southern California Occasional Papers in Linguistics*, S. 83–120. Department of Linguistics, University of Southern California, Los Angeles, CA.

- LEBEN, W. R. (1973). *Suprasegmental Phonology*. Dissertation, MIT. unpublished.
- LEHISTE, I. (1970). *Suprasegmentals*. MIT Press: Cambridge, MA and London, England.
- LEVELT, W. J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press: Cambridge, MA.
- LIBERMAN, M. und A. PRINCE (1977). *On stress and linguistic rhythm*. *Linguistic Inquiry*, 8(2):249–336.
- LIBERMAN, M. und J. PIERREHUMBERT (1984). *Intonational invariance under changes in pitch range and length*. In: ANONOF, M. und R. OEHRLE (eds.): *Language Sound Structure*, S. 157–233. MIT Press: Cambridge, MA.
- LINDBLOM, B. (1983). *Economy of speech gestures*. In: MACNEILAGE, P. F. (ed.): *The production of speech*, S. 217–245. Springer: Heidelberg.
- LOCAL, J. (1994). *Phonological Structure, Parametric Phonetic Interpretation and Natural Sounding Synthesis*. In: KELLER, E. (ed.): *Fundamentals of Speech Synthesis and Speech Recognition. Basic Concepts, State-of-the-Art and Future Challenges*, S. 253–270. John Wiley & Sons Ltd: Chichester.
- LÜNGEN, H., K. EHLEBRACHT, D. GIBBON und A. P. Q. SIMOES (1998). *Morphologie in VERBMOBIL Phase II*. VERBMOBIL Report 366, Universität Bielefeld.
- MANGOLD, M. (1975). *Phonetic Emphasis. A Study in Language Universals*, Band 10 aus *Forum Phonetikum*. Helmut Buske: Hamburg.
- MANGOLD, M. (1990). *Duden Aussprachewörterbuch. Wörterbuch der deutschen Standardaussprache*. Dudenverlag: Mannheim, Wien, Zürich.
- MAYER, J. (1997). *Intonation und Bedeutung. Aspekte der Prosodie-Semantik-Schnittstelle im Deutschen.*, Band 3(4) aus *Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung/Phonetik*. Universität Stuttgart.
- MENGEL, A. (1995). *A comparison of German Names and German Words*. In: *Proceedings of ICPHS '95*, Band 4, S. 554–556, Stockholm, Sweden.
- MENGEL, A. (1998). *Deutscher Wortakzent. Symbole, Signale*. Phorm Verlag: München.
- MENGEL, A. und A. WIRTH (1995). *Lexikon Deutscher Eigennamen*. CD-ROM. TU Berlin, Deutsche Telekom Technologiezentrum.
- MIXDORFF, H. und O. JOKISCH (2001). *Comparing a data-driven and a rule-based approach to predicting Prosodic Features of German*. In: HESS, W. und K. STÖBER (eds.): *Elektronische Sprachsignalverarbeitung*, Band 22 aus *Studentexte zur Sprachkommunikation*, S. 289–305, ITG, DEGA, GI, w.e.b.: Dresden.

- MIXDORFF, H. und C. WIDERA (2001). *Perceived Prominence in Terms of a Linguistically Motivated Quantitative Intonation Model*. In: *Proceedings of EUROSPEECH 2001*, Band 1, S. 403–406, Aalborg, Danmark.
- MÜLLER, A. und R. HOFFMANN (2001). *Accent Label Prediction by Time Delay Neural Networks Using Gating Clusters*. In: *Proceedings of EUROSPEECH 2001*, Band 1, S. 549–602, Aalborg, Danmark.
- MONAGHAN, A. (1993). *What determines accentuation? A reply to Cruttenden and Faber*. *Journal of Pragmatics*, 19:559–584.
- MÜLLER, S. (1996). *The Babel-System—An HPSG Prolog Implementation, Overview*. In: *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, S. 263–277, London.
- NESPOR, M. und I. VOGEL (1986). *Prosodic Phonology*. Foris Publications: Dordrecht.
- OGDEN, R., J. LOCAL und P. CARTER (1999). *Temporal Interpretation in PROSYNTH, a prosodic speech synthesis system*. In: *Proceedings of ICPhS '99*, Band 2, S. 1059–1062, San Francisco, USA.
- OHALA, J. J. und J. J. JAEGER (1986). *Introduction*. In: OHALA, J. J. und J. JAEGER (eds.): *Experimental Phonology*, S. 1–12. Academic Press: Orlando, Florida.
- PIERREHUMBERT, J. (1979). *The perception of fundamental frequency*. *Journal of the Acoustical Society of America*, 66:363–369.
- PIERREHUMBERT, J. (1980). *The Phonology and Phonetics of English Intonation*. Dissertation, M.I.T.
- PIERREHUMBERT, J. (1990). *Phonological and Phonetic Representation*. *Journal of Phonetics*, 17:231–278.
- PIERREHUMBERT, J. und J. HIRSCHBERG (1990). *The meaning of intonational contours in the interpretation of discourse*. In: P. COHEN, J. MORGAN und M. POLLACK (eds.): *Intentions on Communication*, S. 271–311. MIT Press: Cambridge, MA.
- PIKE, K. (1947). *Phonemics: A technique for reducing languages to writing*. University of Michigan Press: Ann Arbor.
- POLLARD, C. und I. SAG (1987). *Information-Based Syntax and Semantics*. Center for the Study of Language and Information, Stanford, CA.
- POLLARD, C. und I. A. SAG (1994). *Head Driven Phrase Structure Grammar*. *Studies in Contemporary Linguistics*. CSLI: Stanford.
- POPPER, K. (1934). *Logik der Forschung*. 1. Auflage. Mohr: Tübingen.

- PORTELE, T. (1998). *Perceived Prominence and Acoustic Parameters in American English*. In: *Proceeding of ICSLP 98, Sydney*.
- PORTELE, T. und B. HEUFT (1997). *Towards a Prominence-based synthesis system*. *Speech Communication*, S. 61–72.
- PRINCE, A. (1983). *Relating to the Grid*. *Linguistic Inquiry*, 14:19–100.
- PRINCE, A. (1990). *Quantitative Consequences of Rhythmic Organization*. In: *Parasession on the Syllable in Phonetics and Phonology*, S. 355–398. Chicago Linguistic Society.
- PRINCE, A. und P. SMOLENSKY (1993). *Optimality Theory*. Technischer Bericht RuC-CS TR-2, Center for Cognitive Science, Rutgers University.
- QUASTHOFF, U. (1998). *Deutscher Wortschatz im Internet*. LDV-Forum. Zeitschrift für Computerlinguistik und Sprachtechnologie, 15(2):4–23.
- REYELT, M., M. GRICE, R. BENZMÜLLER, J. MAYER und A. BATLINER (1996a). *Prosodische Ettikettierung des Deutschen mit ToBI*. In: GIBBON, D. (ed.): *Natural Language Processing and Speech Technology. Results of the 3rd KONVENS Conference, Bielefeld, October 1996*, S. 144–155. Mouton de Gruyter: Berlin, New York.
- REYELT, M., M. GRICE, R. BENZMÜLLER, J. MAYER und A. BATLINER (1996b). *Prosodische Ettikettierung des Deutschen mit ToBI*. In: GIBBON, D. (ed.): *Natural Language Processing and Speech Technology. Results of the 3rd KONVENS Conference, Bielefeld, October 1996*, S. 144–155. Mouton de Gruyter: Berlin, New York.
- RIETVELD, A.C.M und C. GUSSENHOVEN (1985). *On the relation between pitch excursion size and prominence*. *Journal of Phonetics*, 13:299–308.
- ROOTH, M. (1995). *Fokus*. In: LAPPIN, S. (ed.): *Handbook of Contemporary Semantic Theory*, S. 271–297. Blackwell: Oxford.
- ROSS, K. und M. OSTENDORF (1996). *Prediction of abstract prosodic labels for speech synthesis*. *Computer, Speech and Language*, 10:155–185.
- SAUSSURE, FERDINAND DE (1916). *Cours de linguistique générale*. Payot: Paris.
- SCHMITZ, H.-C., B. SCHRÖDER und P. WAGNER (2001). *Zur Akzentuierung semantischer und pragmatischer Fokusse*. In: HESS, W. und K. STÖBER (eds.): *Tagungsband der 12. Konferenz Elektronische Sprachsignalverarbeitung, Band 22 aus Studententexte zur Sprachkommunikation*, S. 151–158. ITG, DEGA, GI, w.e.b.: Dresden.
- SCHWEITZER, A. und M. HAASE (2000). *Zwei Ansätze zur syntaxgesteuerten Prosodiegenerierung*. In: *KONVENS 2000 - Sprachkommunikation*, Band 161 aus ITG Fachbericht, S. 197–202. ITG: Berlin, Offenbach.
- SCOBBI, J. (1991). *Attribute Value Phonology*. Dissertation, University of Edinburgh.

- SELKIRK, E.O. (1984). *Phonology and Syntax. The relation between sound and structure.* Cambridge University Press: Cambridge.
- SELKIRK, E.O. (1982). *The Syllable.* In: HULST, H. VAN DER & N. SMITH (ed.): *The Structure of Phonological Representations*, S. 337–383. Foris Publications: Dordrecht.
- SELTING, M. (1994). *Emphatic Speech Style – with special focus on the prosodic signalling of heightened emotional involvement in conversation.* *Journal of Pragmatics*, 22:375–408.
- SHATTUCK-HUFNAGEL, S. (1995). *The importance of phonological transcription in empirical approaches to “stress shift” vs. “early accent”: comments on Grabe and Warren, and Vogel, Bunnell and Hoskins.* In: CORNELL, B. und A. ARVANITI (eds.): *Phonology and Phonetic Evidence*, Band IV aus *Papers in Laboratory Phonology*, S. 128–140. Cambridge University Press: Cambridge.
- SLUIJTER, A. und V. VAN HEUVEN (1996). *Spectral Balance as an acoustic correlate of linguistic stress.* *Journal of the Acoustical Society of America*, 100(4):2471–2485.
- SONNTAG, G. P. und T. PORTELE (1998). *PURR - A method for prosody evaluation and investigation.* *Journal of Computer Speech and Language*, 12(4):437–451.
- SONNTAG, G. P. (1999). *Evaluation von Prosodie.* Dissertation, Universität Bonn. Shaker Verlag: Aachen.
- STÖBER, K., P. WAGNER, J. HELBIG, S. KÖSTER, D. STALL, M. THOMAE, J. BLAUERT, W. HESS und H. MANGOLD (2000). *Speech Synthesis Using Multilevel Selection and Concatenation of Units from Large Speech Corpora.* In: *VerbMobil: Foundations of Speech-to-Speech Translation.* Springer: Berlin, Heidelberg.
- STÖBER, K., T. PORTELE, P. WAGNER und W. HESS (1999). *Synthesis by Word Concatenation.* In: *Proceedings of EUROSPEECH'99*, Band 2, S. 619–622, Budapest, Hungary.
- STREEFKERK, B. M., L. C. POLS und L. F. TEN BOSCH (1998). *Automatic detection of prominence (as defined by listeners) in read aloud sentences.* In: *Proceedings of ICSLP '98*, Sydney.
- STREEFKERK, B. M., L. C. POLS und L. F. TEN BOSCH (2001). *Up to what level can acoustical and textual features predict prominence.* In: *Proceedings of EUROSPEECH 2001*, Band 1, S. 811–814, Aalborg, Danmark.
- TERKEN, JACQUES (1991). *Fundamental Frequency and perceived prominence of accented syllables.* *Journal of the Acoustical Society of America*, 87:1768–1776.

- T'HART, JOHAN (1981). *Differential Sensitivity to Pitch Distance, particularly in speech*. Journal of the Acoustical Society of America, 69:811–821.
- THOMPSON, H. (1980). *Stress and salience in English*. Dissertation, University of California, Berkeley, CA.
- TOLKIEN, J.R.R. (1996). *Der Herr der Ringe*. 24. Auflage. Klett Cotta: Stuttgart.
- TRASK, R.L. (1996). *A dictionary of phonetics and phonology*. Routledge: London and New York.
- UHMANN, S. (1991). *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Band 252 aus *Linguistische Arbeiten*. Niemeyer: Tübingen.
- VOGEL, I., H. BUNNELL und S. HOSKINS (1995). *The phonology and phonetics of the Rhythm Rule*. In: CORNELL, B. und A. ARVANITI (eds.): *Phonology and Phonetic Evidence*, Band IV aus *Papers in Laboratory Phonology*, S. 111–127. Cambridge University Press: Cambridge.
- WAGNER, P. (1997). *Phonologie und automatische Sprachsynthese*. Magisterarbeit, Universität Bielefeld.
- WAGNER, P. (1998). *Mutual Constraints at the Phonetics–Phonology Interface*. In: SCHRÖDER, B., W. LENDERS, W. HESS und T. PORTELE (eds.): *Proceedings of the 4th Conference on Natural Language Processing KONVENS-98*, Band 1 aus *Computer Studies in Language and Speech*, S. 207–212. Peter Lang: Frankfurt am Main.
- WAGNER, P. (1999). *The Synthesis of German Contrastive Focus*. In: *Proceedings of the Fourteenth International Congress of Phonetic Sciences*, Band 2, S. 1529.
- WAGNER, P. (2001). *Systematische Überprüfung deutscher Wortbetonungsregeln*. In: HESS, W. und K. STÖBER (eds.): *Elektronische Sprachsignalverarbeitung*, Band 22 aus *Studentexte zur Sprachkommunikation*, S. 329–338. ITG, DEGA, GI, w.e.b.: Dresden.
- WAGNER, P., S. BREUER und K. STÖBER (2000). *Automatische Prominenzetikettierung einer Datenbank für die korpusbasierte Sprachsynthese*. In: *Fortschritte der Akustik – DAGA 2000*. DEGA: Oldenburg.
- WAGNER, P. und P. KÜHNLEIN (1998). *A point-based event phonology for the phonetics-phonology interface*. In: *Proceedings of FCHG '98*.
- WAGNER, P. und T. PORTELE (1999). *Two Dimensions of Prominence*. In: *Proceedings of the ESCA Workshop on Dialogue and Prosody*, Eindhoven 1999.

- WAGNER, P. (2000). *Evaluating Metrical Phonology - A Computational-Empirical Approach*. In: *KONVENS 2000 - Sprachkommunikation*, Band 161 aus *ITG Fachbericht*, S. 243–248. ITG: Berlin, Offenbach.
- WAGNER, P. (erscheint). *Rhythmic Constraints in German*. In: *Proceedings of Prosody 2000—Special Session on Intonation Modelling*, Cracow, Poland.
- WALTHER, M. (1996). *OT SIMPLE - A construction-kit approach to Optimality Theory implementation*. Arbeiten des Sonderforschungsbereichs 282 'Theorie des Lexikons' Nr. 88, Seminar f. Allgemeine Sprachwissenschaft, University of Düsseldorf, Germany. (ROA-152): <http://ruccs.rutgers.edu/roa.html>.
- WALTHER, M. und B. KRÖGER (1994). *Phonetik-Phonologie-Kopplung in einem constraintbasierten gesturalen Modell*. In: *KONVENS 94. 2. Konferenz Verarbeitung natürlicher Sprache*.
- WIDERA, C. und T. PORTELE (1999). *Levels of Reduction in German Tense Vowels*. In: *Proceedings of EUROSPEECH '99*, S. 1695–1699, Budapest, Hungary.
- WIDERA, C., T. PORTELE und M. WOLTERS (1997). *Prediction of Word Prominence*. In: *Proceedings of EUROSPEECH '97*, S. 999–1003, Rhodes, Greece.
- WIEBE, B. (1992). *Modelling Autosegmental Phonology with Multiple-Tape Finite State Transducers*. Magisterarbeit, Simon Fraser University.
- WIESE, R. (1996). *The Phonology of German*. The Phonology of the World's Languages. Clarendon Press: Oxford.
- WOLTERS, M. und P. WAGNER (1998). *Focus Perception and Prominence*. In: B. SCHRÖDER, W. LENDERS, W. HESS und T. PORTELE (eds.): *Computers, Linguistics, and Phonetics between Language and Speech. Proceedings of the 4th Conference on Natural Language Processing KONVENS-98*, Band 1 aus *Computer Studies in Language and Speech*, S. 227–236. Peter Lang: Frankfurt am Main.
- WURZEL, W. U. (1970). *Zum Fremdwortakzent im Deutschen*. *Linguistics*, 56:87–108.
- WURZEL, W. U. (1980). *Der deutsche Wortakzent: Fakten—Regeln—Prinzipien. Ein Beitrag zu einer natürlichen Akzenttheorie*. *Zeitschrift für Germanistik*, 1:299–318.

Anhang A

Experimentelles Material

A.1 Zweisilbige Nonsense-Wörter

Der folgende Text wurde den Versuchspersonen zum Lesen und Beurteilen vorgelegt:

Die folgenden Sätze enthalten jeweils ein Wort, welches Ihnen unbekannt sein dürfte. Aus den Sätzen sollte aber ungefähr klar werden, um welchen Gegenstand es sich jeweils handelt. Bitte lesen Sie jeden Satz laut vor, damit sie einen Klangeindruck von dem Wort erhalten. Überlegen Sie nicht lang, bevor sie einen Satz lesen. Sollten Sie ein Wort wiedererkennen, so vermerken sie dies bitte informell auf dem Fragebogen. Bewerten sie anschließend, ob das jeweilige unbekannte Wort bei ihnen eine Assoziation weckt, zu der der Satz paßt.

Beispiel: Lesen Sie folgenden Satz laut vor!

“In dörflichen Regionen Ostwestfalens ist es ein noch immer beliebtes Brauchtum, vor einer Hochzeit bei den Verflommenen des Bräutigams Plömmen an die Tür zu nageln.”

Nun bewerten sie bitte auf der folgenden Skala, wie gut das Wort “Plömmen” (welches Sie vermutlich nicht kennen) in seiner lautlichen Struktur zu dem beschriebenen Kontext paßt:

sehr gut ordentlich mäßig schlecht

Bitte beantworten Sie vor Testbeginn noch einige Fragen:

1. Alter: _____
2. Geschlecht: weiblich männlich
3. Muttersprache deutsch? nein ja
4. Dialektaler Hintergrund (z.B. westfälisch, bayrisch):

5. Fremdsprachenkenntnisse: nein ja,
und zwar: _____

6. derz. Beruf/Ausbildungsrichtung/Schule/Studienfach:

Vielen Dank für Ihre Hilfe. :-)

1. Der Schlagzeuger beherrschte auch das Spiel auf der Sombu meisterhaft.
sehr gut ordentlich mäßig schlecht

2. Ein Sobrill ist der traditionelle Herd der Menschen in Papua-Neuguinea.
sehr gut ordentlich mäßig schlecht

3. Die Samivölker Nordfinnlands trugen auf ihren Wanderungen immer einen
Birkenholzbecher bei sich, den sie Saubokk nannten.
sehr gut ordentlich mäßig schlecht

4. Schon die alten Griechen verwendeten Kräuterliköre, sogenannte Obeis, um
Magenleiden zu kurieren.
sehr gut ordentlich mäßig schlecht

5. In Feuerland hüllen sich die Menschen in ihren Sobokk, einen farbenfrohen
Wollumhang.
sehr gut ordentlich mäßig schlecht

6. Wenn dir in Kuwait nahegelegt wird, dich mit an den Sobritt zu setzen, solltest
du besser darauf eingehen.
sehr gut ordentlich mäßig schlecht

7. In Laos gibt es in jedem Haushalt einen Altar für die Obinns, die freundlichen
Hausgötter.
sehr gut ordentlich mäßig schlecht

8. Ein Jemenite, der etwas auf sich hält, raucht täglich eine Obrill, eine Pfeife aus Wurzelholz.
sehr gut ordentlich mäßig schlecht
9. Neben dem Didgeridoo gibt es bei den Australiern auch das weitaus weniger bekannten Sombokk.
sehr gut ordentlich mäßig schlecht
10. Kleine Nebentäler in den Anden, die sogenannten Soburs, sind häufig von der Außenwelt völlig abgeschnitten.
sehr gut ordentlich mäßig schlecht
11. In der Mythologie der Eskimos wurde die Erde aus Packeis, Firn, Sauspritt, Pulverschnee und einer Tüte Kies erschaffen.
sehr gut ordentlich mäßig schlecht
12. Wenn dir im Ural neben Ziegenkäse und Roggenbrot noch Soba angeboten wird, so solltest du nicht nein sagen, denn davon wird man hundert Jahre alt.
sehr gut ordentlich mäßig schlecht
13. Wenn ein Zulukrieger das Mannesalter erreicht, so erhält er das traditionelle Zubehör für den Kampf: Pfeil und Bogen, Obat und Messer.
sehr gut ordentlich mäßig schlecht
14. Bei den Einwohnern der Osterinseln gehören zu einem Gesicht Mund, Nase und Augen genauso dazu wie der Saubei.
sehr gut ordentlich mäßig schlecht
15. Vor dem Import des Feuerwassers berauschten sich die amerikanischen Ureinwohner an Pilzen und Somburs sowie rauchbaren Pflanzen.
sehr gut ordentlich mäßig schlecht

16. Die Kelten fürchteten sich weder vor Blitz, Donner, Sturm noch den damals weitverbreiteten Ospritts.
sehr gut ordentlich mäßig schlecht
17. Wenn einem Sudanesen seine Saubills aus der Brieftasche fallen, dann ist der Tag gelaufen.
sehr gut ordentlich mäßig schlecht
18. Die Adelige wie die einfache Nomadin pflegt ihr Haar mit Kamm und Sobinn, so sie denn aus der Mongolei stammt.
sehr gut ordentlich mäßig schlecht
19. Hat der baskische Schmied keine Zange zur Hand, so tut es auch eine gutgepflegte Saubatt.
sehr gut ordentlich mäßig schlecht
20. Vergißt eine Frau in Sri Lanka, ihre Sombinn vernünftig zu verknoten, kann das in der Öffentlichkeit ganz schön peinlich werden.
sehr gut ordentlich mäßig schlecht
21. Um die dünne Humusschicht nicht unnötig zu strapazieren, verwendet man in Paraguay zum Pflügen gern einen Sobokk.
sehr gut ordentlich mäßig schlecht
22. Es ist unglaublich, wie die Völker Ozeaniens von Insel zu Insel gelangt sind und sich dabei nur an Sternen und Saubinn orientiert haben.
sehr gut ordentlich mäßig schlecht
23. In den Wäldern Sibiriens findet man noch den Tiger, den weißen Wolf, das Sombei und den Schwarzbären.
sehr gut ordentlich mäßig schlecht

24. Wenn eine chilenische Frau verliebt ist, legt sie ein mit Sobatts verziertes Taschentuch auf die Fensterbank ihres Angebeteten.
sehr gut ordentlich mäßig schlecht
25. Die Mexikaner sind berühmt für ihre rauschenden Feste, bei denen neben Fleisch auch viel gesundes Gemüse, Obst und Obtritt aufs Buffett kommt.
sehr gut ordentlich mäßig schlecht
26. Es ist viel komplizierter, eine Sonu zu verfassen als eine typisch europäische Gedichtform.
sehr gut ordentlich mäßig schlecht
27. Es ist erfreulich, daß neben Zelten und Tipis jetzt auch Saubritts wieder im Outdoorfachhandel erhältlich sind.
sehr gut ordentlich mäßig schlecht
28. In Timbuktu waren die öffentlichen sanitären Anlagen schon vor Jahrhunderten vorbildlich. Jeder Wohnblock verfügte über Badehaus, Sombatt und mehrere WCs.
sehr gut ordentlich mäßig schlecht
29. Beim tadschikischen Obertongesang ist es für den Sänger wichtig, sich neben dem Rachenraum auch auf das korrekte Oba zu konzentrieren.
sehr gut ordentlich mäßig schlecht
30. Nach Feierabend trifft man sich in Ungarn nicht nur in der Kneipe, sondern auch in den preiswerteren Sospritts.
sehr gut ordentlich mäßig schlecht
31. "Eine frühe Sauba bereitet noch keine gute Ernte" lautet eine Bauernregel aus Java.
sehr gut ordentlich mäßig schlecht

32. Viele Tierfreunde können nicht verstehen, daß manche Menschen niedliche Hermeline, Robbenbabies oder kleine Obeis wegen ihres Fells töten.
sehr gut ordentlich mäßig schlecht
33. "Wenn du dich auf dein Saubu konzentrierst, so brauchst du kein Schwert," lautet eine leider in Vergessenheit geratene Taliban-Weisheit.
sehr gut ordentlich mäßig schlecht
34. Auf dem Basar in Tunis kann man neben Melonen und Datteln auch Obur, Bonbons und Gewürze kaufen.
sehr gut ordentlich mäßig schlecht
35. Wenn man in Wladiwostock in der U-Bahn ausgeraubt wird, so ist eine bei den Behörden vorgelegte Sombritt von Vorteil.
sehr gut ordentlich mäßig schlecht
36. Wehe dem Mountainbiker, der vergißt, seine Bremsen zu überprüfen, bevor er einen äthiopischen Sobill hinunterrast.
sehr gut ordentlich mäßig schlecht
37. Die Georgier haben einen Hang zum Kitsch. Ihre Wohnzimmer sind überladen mit Plüschroddeln und goldenen Sobritts.
sehr gut ordentlich mäßig schlecht
38. Beim Klettern im Himalaya half mir ein hilfsbereiter Einheimischer mit seinem Obokk aus, nachdem ich einen Karabiner verloren hatte.
sehr gut ordentlich mäßig schlecht
39. "Vergiß nicht dein Heft, Schreiber, Saubur und Turnbeutel!" rufen ghanaische Eltern, wenn die Kinder morgens aus dem Haus gehen.
sehr gut ordentlich mäßig schlecht

40. Eine kleine grönländische Obei wärmt wesentlich effizienter als ein deutscher Kachelofen.
sehr gut ordentlich mäßig schlecht
41. "Wenn du deine Somba pflegst, werden die Götter die Glück und Zufriedenheit schenken," ist eine für uns nicht nachvollziehbare Weisheit aus Madagaskar.
sehr gut ordentlich mäßig schlecht
42. Tibetanische Forscher glauben, das Dalai-Lama-Gen für Göttlichkeit und immerwährendes Obrill jetzt identifiziert zu haben.
sehr gut ordentlich mäßig schlecht
43. Die Menschen am Kaspischen Meer fürchten, daß durch die Umweltverschmutzung nicht nur die Fische leiden, sondern auch die Insekten und So-beis.
sehr gut ordentlich mäßig schlecht
44. In den vedischen Schriften sind die Somspritts in ihrer zentralen Aussage vergleichbar mit der hebräischen Kabbala.
sehr gut ordentlich mittel schlecht

A.2 Unbekannte Namen

Der folgende Text wurde den Versuchspersonen zum Lesen vorgelegt:

Bitte lesen Sie folgenden Text laut vor. Machen Sie sich nicht allzu viele Gedanken über den Inhalt, sondern versuchen Sie, so flüssig wie möglich zu lesen. Wenn Sie sich versprechen sollten, so lesen Sie den Satz mit dem Versprecher noch einmal. Sollten Sie ein Wort nicht kennen, so vertrauen Sie auf ihre Intuition. Versprecher sind nicht schlimm! Vielen Dank für Ihre Hilfe. :-)

Bitte beantworten Sie vor Testbeginn noch einige Fragen. (s.u.)

1. Alter: _____
2. Geschlecht: weiblich männlich
3. Muttersprache deutsch? nein ja
4. Dialektaler Hintergrund (z.B. westfälisch, bayrisch):

5. derz. Beruf/Ausbildungsrichtung/Schule/Studienfach:

6. Wie gut kennen Sie das Buch "Der Herr der Ringe"?
bin Fan sehr gut nur gelesen gar nicht

Folgendes Gespräch konnte man neulich auf einer Tagung zum Thema "Sprachwandel und deutsche Dialekte" verfolgen:

I: Ach hallo, Herr Cirdan! Ich wollte sie nochmal wegen ihres Artikels über ost-westfälische Bezeichnungen für Küchengeräte ausfragen. Stimmt es wirklich, dass noch vor 150 Jahren die Menschen in der Region Bielefeld eine Schaumkelle als "Palantiri" bezeichnet haben?

II: Da haben sie mich mißverstanden, Frau Anduril. Das gilt mehr für Detmold, aber lustigerweise findet sich der Ausdruck auch in einigen Gebieten im südlichen Hinterpommern. In Bielefeld selbst holten die Bäuerinnen ihre Klösse nicht mit einer Schaumkelle, sondern mit einem Elrand aus dem Kochwasser.

I: Hat dazu die Kollegin Meier nicht im letzten Jahr auf der Tagung in Aragorn einen Vortrag gehalten?

II: Wo liegt denn Aragorn?

I: In der Nähe von Kansas City, soweit ich weiß. Das war die Konferenz, bei der das Thema "Küchengeräte" besonders kontrovers diskutiert wurde. Vielleicht

auch, weil es ja etwas heikel ist. Es ging da eben auch um geschlechtsspezifische Sprachverwendung. Im Fränkischen Raum sind nachgewiesenermaßen die Frauen eher als die Männer dazu übergegangen, ihren Eimer nicht mehr als Fingolfin zu bezeichnen. Vielleicht, weil es ja auch schneller geht, das Wort "Eimer" auszusprechen. Wenn die Frauen ständig schwere Eimer schleppen mußten, war das vielleicht angenehmer ...

II: Es kann aber doch auch sein, daß es den fränkischen Frauen zu umständlich war, ihren Männern immer zu sagen "Räum doch endlich mal den Fingolfin wieder in den Stall!"

I: Sie verstehen die Brisanz des Themas, wie ich sehe. Was machen sie denn so zur Zeit?

II: Ich schreibe mit Beutlin zusammen an einem Aufsatz über vergessene Bezeichnungen für Handwerksgeräte. Wußten sie schon, dass ein schlichter Hobel in der Grazer Gegend noch heute als Legolas bezeichnet werden?

I: Im Ernst? Naja, in der Steiermark treibt die Sprache interessante Blüten. Übrigens habe ich eben erst erfahren, daß Frau Bombadil jetzt ja doch die Professur in Clausthal-Zellerfeld bekommen hat.

II: Ach, das war aber auch verdient, nachdem sie diese exzellenten Schriften über die Friesen und ihre Bekleidung veröffentlichen konnte. Wie war das noch? Da gab es bei den Friesen doch dieses beliebte Damendessous im ausgehenden Mittelalter. Jetzt fällt mir der Name nicht mehr ein.

I: Sie meinen sicher diesen Strumpfhalter aus dünnen Hanfseilen. Den nannte man doch einen Gandalf, oder?

II: Ach ja, richtig! Gandalf!!!

Anhang B

Extrametrische und betonte Affixe

B.1 Suffixe

Dies ist die vollständige Liste der in der Vorverarbeitung zur Bestimmung der Wortbetonung abgetrennten Suffixe in IPA-Notation:

bə, əl, əlnd, əlst, əlt, əm, ən, ənd, əns, ənt, ənts, ɐ, əɣ, ɐl, əɣl, ɐn, əɣn, ɐnd, əɣnd, ɛst, əɣst, ɛt, əɣt, ət, əs, ɛst, haft, bæ, baɣ, baɣ, baɪr, ɕən, haɪt, ɪɕ, ɪɕt, ɣən, ɣə, ɪɕs, ɪk, ɪɕt, ɪɣ, nən, nər, ɪs, kaɪt, lam, lə, ləɣ, lər, lɪɕ, lɪŋ, man, nɐ, nəɣ, nis, sə, zaɪm, fə, stə, stəm, stən, tə, təl, təm, tən, təs, tɪɕ, tuɪm, tyɪm, ʊŋ, ɣʊŋ, vɛɣk, vɛrk, vɛɣts, vɛrts, tsɔɪk, faft, loɪs, səl, zəl, nis, tsɪɕ, ɪf, jə, i:ə, jən, ɪn, mɛɪsɪɕ

B.2 Unbetonte Präfixe

Dies ist die vollständige Liste der in der Vorverarbeitung zur Bestimmung der Wortbetonung abgetrennten Präfixe in IPA-Notation. Einige Präfixe, die auch betont auftreten können, sind in der Liste verblieben:

ɐ, ʔɛnt, bə, ʔyɪ.bə, ɣə, fɛɣ, fɛ, ʔɛɣ, tsɛɣ, ɣə, ʔʊm, duɣɕ, hɪn.tɐ, ʔʊn, tɐ, fəl, viɪ.də, ʔɪn.tɐ

B.3 Betonte Präfixe

Dies ist die Liste der in der Vorverarbeitung als die Wortbetonung an sich ziehend betrachteten Präfixe:

ʔap, ʔan, ʔaɪf, baɪ, ʔaɪn, ʔɛm.pɔɪɣ, fɔɣt, loɪs, mit, naɪx, niɪ.də, vɛk, vaɪ.tɐ, daɪɣ, hɛɣ, hɪn, fɔɣ, tsu:

B.4 (A(BC))-Komposita

Kunstdenkmal, Stadtpfarrkirche, Hauptfahrrinne, Hauptbahnhof, Ostdeutschland, Kreishauptstadt, Stabhochsprung, Kunstzeitschrift, Geldbriefträger, Sportflugzeug,

Fausthandschuh, Kriegsdienstzeit, Landgasthof, Sportfernsehen, Startzeitpunkt, Waldschwimmbad, Startbildschirm, Dienstfahrrad, Aidsfacharzt, Dampfschiffahrt, Stadtparkasse, Chefvolkswirt, Suchscheinwerfer, Eisschnellauf, Fachhochschule, Stadthaushalt, Textbaustein, Tatzeitraum, Hilfsbuchhalter, Windschutzscheibe, Heizkraftwerk, Feldhandball, Herzstillstand, Weltfußball

Anhang C

Lebenslauf

Name:	Petra Susanne Wagner
geb. am:	22.05.1971 in Minden/Westf.
Eltern:	Else Wagner geb. Dallmeier und Götz Wagner
Staatsangehörigkeit:	deutsch
Ausbildung:	
1977-1978	Grundschule Hellteich in Walsrode
1987-1981	Grundschule Hahlen in Minden/Westf.
1981-1987	Besselgymnasium Minden
1987-1988	Central Valley High School Spokane, WA (USA) als Stipendiatin des Parlamentarischen Patenschaftsprogramms des Deutschen Bundestages
1988-1990	Besselgymnasium Minden, Abitur
1990-1992	Lehramtsstudium Deutsch und Englisch (Sek I u. II) an der Universität Bielefeld
1992-1997	Magisterstudiengang Linguistik, Anglistik, Psychologie an der Universität Bielefeld
01/1998	Studienabschluß Magistra Artium in Linguistik, Anglistik, Psychologie
02/1998-10/2000	wissenschaftliche Mitarbeiterin in verschiedenen Drittmittelprojekten (BMBF, DFG) an der Rheinisch-Westfälischen Friedrich-Wilhelms-Universität Bonn am Institut für Kommunikationsforschung und Phonetik, Promotionsstudiengang Kommunikationsforschung und Phonetik, Anglistik, Psychologie
seit 02/2001	wissenschaftliche Mitarbeiterin am Lehrstuhl für Phonetik des Instituts für Kommunikationsforschung und Phonetik der Rheinisch-Westfälischen Friedrich-Wilhelms-Universität Bonn