Neuere Phonologie-Theorien und ihre Anwendung aufs Deutsche*

Non-linear Phonology, with its two main directions, Autosegmental Phonology and Metrical Phonology, has been applied to prosodic phenomena of German like accent and syllable structure. This is especially true of the autosegmental approach (cf. WIESE 1986, 1988 and 1996, VATER 1992 and 2004, NEEF 1996, BECKER 1998): Metrical Phonology was applied to German as well (although to a lesser degree), especially in the area of accent structure, e.g. by GIEGERICH (1985) and FÉRY (1996), the latter one combining it with the framework of Optimality Theory (OT): VENNEMANN (1982, 1988) and EISENBERG (1991) developed their own ways of analysing German syllable and accent structures. Non-linear approaches were also applied to the analysis of intonation (cf. UHMANN 1991 and FÉRY 1992) and of primarily segmental phenomena like final devoicing in German (cf. RUBACH 1990): All these analyses obtained new and interesting results thanks to the useful, flexible and inspiring framework of Non-linear Phonology.

0. Vorbemerkungen

Die wichtigsten und charakteristischsten neueren phonologischen Analysen des Deutschen basieren auf der Nicht-linearen Phonologie, die sich als Konsequenz aus den Beschränkungen des SPE-Modells in den achtziger Jahren entwickelt hat.[1] Dabei überwiegt der Ansatz der Autosegmentalen Phonologie als Bezugsrahmen. Die Metrische Phonologie spielt in der deutschen Phonologie eine mindere Rolle, wohl wegen der oft gewaltsamen, sehr abstrakten Lösungen, die die strikt binäre Strukturierung erzwingt.[2] Neben der Nicht-linearen Phonologie ist auch die Optimalitätstheorie (OT) aufs Deutsche angewandt worden, so auf Akzentstrukturen (vgl. FÉRY 1996). Eigene Wege in der Darstellung von Silben- und Akzentstrukturen gehen VENNEMANN (1982; 1988) und EISENBERG (1991).

---

* *Kwartalnik Neofilologiczny* (2005), LII/1, 23–50.

[1] „SPE“ = *Sound Pattern of English*, das Werk von CHOMSKY / HALLE (1968), das der Richtung den Namen gab.

[2] Eine gründliche und im Wesentlichen überzeugende metrische Analyse des Deutschen bietet GIEGERICH (1985). Doch auch hier findet sich Null-Onset bei Silben ohne Onset-Segment wie *in* und *aus*, da in der Metrischen Phonologie jede phonologische Einheit in ein stärkeres Element (s) und ein schwächeres (w) zerlegt wird.

1. Segmentale Strukturen

1.1 Konsonantismus

1.1.1 Die Stimmtonopposition

Die Stimmtonopposition betrifft alle Obstruenten des Deutschen und nur sie, denn Sonoranten sind spontan (und das heißt immer) stimmhaft. Zwei Sonderfälle sind zu beachten:

- Dem stimmhaften palatalen Frikativ /j/ steht der stimmlose palatale Frikativ [ç] gegenüber, der kein eigenes Phonem bildet, sondern ein Allophon des Phonems /x/.[3]
- Dem stimmlosen velaren Frikativ /x/ steht ein stimmhafter velarer Frikativ [ɣ] gegenüber, der nur als Allophon von /g/ in einigen Dialekten (z.B. im Berlinischen in [zɑ:ɣɳ] für *sagen*) vorkommt.

Insgesamt ergeben sich nach Ramers / Vater ([4]1995: 81) folgende Stimmtonpaare:

Tab. 1: Stimmtonpaare

p	f	t	s	ʃ	ç	k	x
b	v	d	z	ʒ	j	g	(ɣ)

Seit langem gibt es eine Diskussion darüber, welches Merkmal bei der Opposition zwischen diesen Paaren relevant ist. Während Wurzel (1970: 259 f.) die Stimmhaftigkeit vs. Stimmlosigkeit als zugrundeliegende Opposition annimmt, setzen Meinhold / Stock (1980) und Kloeke (1982) statt dessen eine Fortis-Lenis-Opposition an, die definiert werden kann als 'größere vs. geringere Artikulationsenergie' (vgl. Kohler / Künzel 1978: 118 und Ezawa 1972: 108–120).[4]

Im Standarddeutschen besteht, wie z.B. Untersuchugen von Meinhold / Stock (1963), Kohler / Künzel (1978) und Knetschke / Sperlbaum (1987) gezeigt haben, keine automatische Kopplung von 'Stimmlosigkeit + Fortisartikulation' und 'Stimmhaftigkeit + Lenisartikulation'. Lenisobstruenten werden häufig ohne Stimmton realisiert, im Süddeutschen fast ausschließlich, wie Bannert (1976) fürs Bairische nachweist; er nimmt an, dass nicht die Artikulationsspannung, sondern die Quantität der Sequenz 'Vokal + Konsonant' im Bairischen die distinktive Funktion in Paaren wie *wega* vs. *wecka* übernimmt. Entstimmlichung der Lenes ist vor allem im Anlaut zu beob-

[3] Die neue IPA-Version (1989) markiert den palatalen Frikativ als [ʝ] und unterscheidet ihn so vom Glide [j]. Im Deutschen kommt [ɣ] nur als dialektale Realisierung von [g] vor, z.B. [o:ɣɳ] *Augen* im Berlinischen.

[4] Die (variable) Artikulationsenergie kann einen größeren vs. geringeren intraoralen Luftdruck, eine längere vs. kürzere Obstruentendauer und (bei Plosiven) eine Aspiration vs. Nicht-Aspiration zur Folge haben

achten. Im Wortinlaut zwischen Vokalen werden andererseits manchmal – nach KNETSCHKE / SPERLBAUM (1987) in einem Viertel aller Fälle – sogar Fortes stimmhaft. EZAWA (1972: 109) nennt als Beispiele die inlautenden Frikative in Wörtern niederdeutscher Herkunft wie *quasseln*.[5] KLOEKE (1982) nimmt mit MEINHOLD / STOCK (1980: 123) zugrundeliegende stimmlose Lenes an und formuliert eine Regel, nach der stimmlose Lenes im Inlaut stimmhaft werden. In RAMERS / VATER ([4]1995: 82) werden Gründe angeführt, die gegen eine solche Regel sprechen:

- Auch zwischen Vokalen sind stl. Lenes möglich (als Aussprachevariante in *Ebbe*, *Widder*, *Roggen*);
- Im Silbenanlaut sind stimmhafte wie stimmlose Lenis möglich, in *Bar* und *Gasse* wie in *Tanzbar* und *Stadtgasse*. Wenn auch Stimmtonverlust auftritt (z.B. bei progressiver Assimilation über die Silbengrenze hinweg), so ist er doch keineswegs obligatorisch und bedeutet nicht gleichzeitige Fortisierung. Die meisten Sprecher des Deutschen halten die Opposition zwischen *Tanzbar* und *Tanzpaar* aufrecht.

Der phonologischen Opposition 'Fortis – Lenis' kann ein ganzes Bündel phonetischer Korrelate zugeordnet werden, wie Konsonantendauer, Artikulationsspannung, intraoraler Luftdruck, Aspiration und Stimmtonbeteiligung, die gemeinsam oder, in bestimmten Kontexten, bei verschiedenen Sprechern unterschiedlicher regionaler Herkunft oder in Abhängigkeit von Stilebene und Sprechtempo, auch einzeln zur Aufrechterhaltung der Opposition beitragen. Welche phonetischen Parameter jeweils vorhanden sind oder fehlen, kann in der Regel nicht vorhergesagt werden. Deshalb stehen stimmlose und stimmhafte Lenes in freier, regionaler oder sozialer Variation. Die phonologische Opposition zwischen Fortis- und Lenisobstruenten bleibt dagegen im Wort- und Silbenanlaut erhalten und wird nur im Auslaut zugunsten der Fortes neutralisiert.

Die **Auslautverhärtung** ist eine Entstimmlichung und Fortisierung, nach STAMPE (1973) ein natürlicher Prozess, der eine Erleichterung der Artikulation bewirkt.[6] Wie (1–01) zeigt, betreffen die durch „Auslautverhärtung" geschaffenen Alternationen nicht nur Flexionsformen des gleichen Worts, sondern auch verschiedene, durch Wortbildungsregeln aufeinander bezogene Wörter:[7]

---

[5] Stimmlose Lenes werden durch untergesetzten Kreis gekennzeichnet, stimmhafte Fortes durch untergesetztes Häkchen.

[6] Stimmlose Obstruenten sind im Auslaut einfacher zu produzieren als stimmhafte; englischsprachige Kinder auslautverhärten Obstruenten während des Spracherwerbs automatisch und lernen später, diesen Prozess zu unterdrücken (vgl. WURZEL 1981: 954). Zur Auslautverhärtung im Deutschen vgl. RUBACH (1990).

[7] Schreibweisen wie *Lo*[p] und *l*[o]*ben*, bei denen nur der jeweils relevante Bestandteil in phonetischer Transkription erscheint, wurden von KLOEKE (1982) übernommen. Die übliche Unterscheidung zwischen „/../" für phonologische und „[..]" für phonetische Repräsentation wird in diesem Beitrag zugunsten letzterer aufgehoben.

(1–01)	Gra[s]	–	Grä[z]er
	Land[t]	–	Län[d]er
	Lo[p]	–	lo[b]en
	San[t]	–	san[d]ig

Stimmhafte Obstruenten (bzw. Lenes) sind markierter als stimmlose (bzw. Fortes): Die Morphemgrenze ist als relevante Umgebung für die Auslautverhärtung inadäquat:

(1–02)	[zɛnd+ʊŋ]	*Sendung*
	[ve:b+ɐ]	*Weber*
	[lo:g+ɪʃ]	*logisch*

Die Wortgrenze erfasst wiederum nicht alle Fälle von Auslautverhärtung:

(1–03)	[frɑ:k+tə]	*fragte*
	[re:t+lɪç]	*redlich*
	[rø:s+çən]	*Röschen*

RAMERS / VATER ([4]1995: 84) notieren den Stimmtonverlust von Obstruenten in silbenfinaler Position und vor Obstruent als Fortisierung (s. obige Diskussion):

(1–04) $[-\text{son}] \rightarrow [+\text{fortis}] \,/\, \underline{\quad} \begin{pmatrix} [-\text{son}] \\ ]_\sigma \end{pmatrix}$

Will man Fälle wie *Magd* und *Obst* nicht erfassen (weil man nicht unbedingt zugrundeliegendes /g/ bzw. /b/ annehmen muss), dann kann man die Regel vereinfachen:

(1–05) $[-\text{son}] \rightarrow [+\text{fortis}] \,/\, \underline{\quad}\,]_\sigma$

Fälle wie *Jagd* (Pl. *Jagden*; Ableitung vom Verb *jagen*) sprechen m.E. jedoch für Version (1–04).

1.1.2 Affrikaten

Das SPE-Modell führte zu Problemen nicht nur im suprasegmentalen, sondern auch im segmentalen Bereich, besonders bei Affrikaten. Unter Affrikaten versteht man homorgane, d.h. an annähernd gleicher Artikulationsstelle gebildete, Verbindungen aus Plosiv und folgendem Frikativ. Im Standarddeutschen zählen zu dieser Klasse die Cluster [pf], [ts] und [tʃ]; in einigen Fremdwörtern findet sich außerdem [dʒ], z.B. in *Gin*, *Job* und

Budget.[8] Seit TRUBETZKOY (1939) gibt es eine Diskussion darüber, ob Affrikaten monosegmental oder bisegmental zu werten seien. Für beide Positionen wurden Argumente angeführt (vgl. WERNER 1972: 50–55 und MORCINIEC 2002). Phonotaktische Regularitäten des Deutschen sprechen für eine monophonematische Analyse, da sich Affrikaten in ihrer Verbindbarkeit wie einfache Konsonanten verhalten, wogegen ihre Merkmalzusammensetzung keine monophonematische Analyse erlaubt.[9] Da jedes Segment aus einem Bündel ungeordneter Merkmale besteht, gibt es einen Konflikt zwischen dem Merkmal [–dauernd] des plosiven Teils und dem Merkmal [+dauernd] des frikativen Teils; (1–06)a enthält zwei kontradiktorische Merkmale in einem einzigen Bündel neben weiteren nicht-kontradiktorischen Merkmalen (die hier durch „..." angedeutet sind): (1–06)b behandelt die Merkmalspezifizierungen, als ob sie zwei aufeinander folgende Bündel darstellten.

(1–06)a $\begin{bmatrix} \text{– dauernd} \\ \text{+ dauernd} \\ \text{...} \end{bmatrix}$ b $\begin{bmatrix} \text{– dauernd, + dauernd} \\ \text{...} \\ \text{...} \end{bmatrix}$

Eine Lösung bietet die Autosegmentale Phonologie mit ihren mehrschichtigen Repräsentationen: Affrikaten sind zweigliedrig in der Segmentschicht, wo einem Segment mit dem Merkmal [–dau(ernd)] ein Segment mit dem Merkmal [+dau(ernd)] folgt, verhalten sich aber wie eine Einheit in der „Skelettschicht", wo ihre „phonotaktische" Funktion, d.h. ihre Funktion im Silbenaufbau beschrieben wird. Da verhalten sie sich nämlich wie ein Einzelsegment: [pf] hat die gleichen Kombinationsmöglichkeiten wie [p] (vgl. *Plan* und *Pflanze, Prinz* und *Pfriem*); [ts] hat annähernd die gleichen Kombinationsmöglichkeiten wie [t] (vgl. *Tweed* und *zwischen*):[10] Auch der spiegelbildliche Silbenaufbau des Deutschen liefert ein Argument dafür, dass sich Affrikaten phonotaktisch wie eine Einheit verhalten. In der Autosegmentalen Phonologie stellt man das so dar, dass zwei Einheiten der Segmentebene mit einer Einheit der Skelettebene verbunden sind:[11]

---

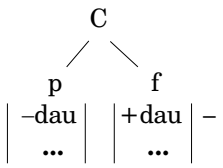
[8] [ps] in *Gips*, [pʃ] in *hübsch*, [ks] in *Keks* und [kʃ] in *mucksch* sind nicht homorgan, daher keine Affrikaten.

[9] Weitere externe Evidenz für bisegmentale Wertung liefern Fehlleistungen von Aphatikern, die z.B. *Bürtse* statt *Bürste* sagen. Auch Versprecher wie [nɛst] *Nest* für [nɛts] *Netz* (vgl. KLOEKE 1982: 43) sprechen dafür. Für eine monophonematische Wertung sprechen die Restriktionen für Konsonantencluster im Silbenanlaut: In dieser Position sind in der Regel nur zwei Konsonanten zugelassen (vgl. WIESE 1988: 59), z.B. [ps] in *Psalm* oder [ks] in *Xaver*, aber weder *[psl] noch *[ksl] oder andere. Allerdings bilden Kombinationen mit [ʃ] in Erstposition wie [ʃtr] in *Streit*, [ʃpl] in *Splitter* und [ʃpr] in *Sprache* eine Ausnahme zu dieser Silbenstrukturbedingung.

[10] Hier findet man viel seltener Korrespondenzen, da [t] vor [v] oft zu [k] geworden ist (vgl. *quer* < *twer* und *Quark* < *twarog*) oder aber in der hochdt. Lautverschiebung zu [ts] verschoben wurde: *twee* > *zwei*.

[11] Die Darstellung der Affrikaten [ts] und [tʃ] ist der Darstellung von [pf] analog.

(1–07)



Skelettschicht

Segmentschicht

Auf der Skelettschicht werden die phonotaktischen Verhältnisse mit Hilfe der Einheiten C und V markiert: „V“ (von ‘vowel’) bezeichnet den Silbengipfel, „C“ (von ‘consonant’) ein silbenmarginales Element. Die Skelettschicht wird daher auch „CV-Schicht“ genannt.

1.1.3 Nasalkonsonanten

Während /m/ und /n/ in allen Positionen in Opposition sind (vgl. *mein* : *nein* und *am* : *an*) scheint der velare Nasal [ŋ] zwar mit /m/ und /n/ im In- und Auslaut zu kontrastieren, nicht aber im Anlaut, wo er nie vorkommt:

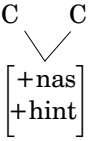
(1–08)	a	rammen	[ramən]	(1–09)	a	ramm [ram]
	b	rannen	[ranən]		b	rann [ran]
	c	Rangen	[raŋən		c	Rang [raŋ]

Bereits Isačenko (1963) legte dar, dass der klassische Phonembegriff, der auf Distribution und Opposition aufbaut, zur Erklärung der Eigenschaften von [ŋ] nicht ausreicht:

- [ŋ] unterliegt im Deutschen, anders als [m] und [n], starken Distributionsbeschränkungen
- [ŋ] ist im Deutschen nach Langvokalen und Diphthongen nicht zugelassen.
- Vor [ŋ] steht nie ein Konsonant in gleicher Silbe.

Der velare Nasal verhält sich offensichtlich phonotaktisch wie ein Konsonantencluster und nicht wie ein Einzelkonsonant. Man könnte ihn folglich innerhalb des CV-Modells so darstellen:

(1–10)



Ginge man von einem zugrundeliegenden Nasal /ŋ/ aus, so wäre dieser der einzige Langkonsonant im Deutschen und besäße somit einen völlig exzeptionellen Status. Man hat aber genügend Gründe dafür, von einem /n/ vor /g/ und /k/ auszugehen. In einheimischen („nativen“) Wörtern steht [n] morphemintern nur vor dem homorganen Plosiv [t], [ŋ] nur vor homorganem

[k] (vgl. WURZEL (1981: 908 f.). Diese Beschränkung gilt nicht für den labialen Nasal [m], wie die Beispiele *Imker*, *Lemgo*, *Amt*, *Hemd* und *Zimt* zeigen. Es liegt also nahe, den velaren Nasal vor [k] als Produkt einer regressiven Assimilation zu betrachten, zumal auch Fälle von progressiver Nasalassimilation an den vorausgehenden Plosiv vorkommen, wie z.B. in ['hɑ:kŋ] für *Haken*. Die Realisierung [nk] erscheint gewöhnlich nicht im gleichen Morphem, sondern nur über Morphemgrenzen hinweg, z.B. in *u*[nk]*lar* oder *a*[nk]*ommen*. Hier ist auch eine regressive Assimilation über Morphemgrenzen hinweg möglich (*a*[ŋk]*ommen* und *u*[ŋk]*lar*); aber dieser Prozess ist optional, während die Assimilation innerhalb von Morphemgrenzen obligatorisch ist (vgl. KLOEKE 1982: 122); es existiert beispielsweise im Standarddeutschen keine Aussprache *[tank].

Da die Sequenz [ng] morphemintern ausgeschlossen ist, ist auch in diesem Fall von einer regressiven Nasalassimilation auszugehen, die das Cluster /ng/ in /ŋg/ überführt. Jedoch erscheint /ŋg/ nur vor Vollvokalen in Fremdwörtern, vor Fremdsuffixen oder in Namen; sonst steht einfaches [ŋ] (vgl. KLOEKE 1982: 119). (1–11) veranschaulicht die komplementäre Verteilung von [ŋg] und [ŋ]:

(1–11)

[ŋg]	[ŋ]
Ungar	Hunger
Ingo	Inge
Anglist	englisch
fingieren	fingen

Vor [t], [s] und [st] im Silbenauslaut sind beide Formen möglich, allerdings wird /g/ aufgrund der Auslautverhärtung als [ŋk] realisiert. Für die Annahme eines zugrundeliegenden Clusters /ng/ auch in den Formen der rechten Spalte in (1–11) können, neben den oben erörterten distributionellen Kriterien, noch weitere Argumente angeführt werden:

– In einigen Fremdwörtern alternieren [ŋ] und [ŋg] (bzw. [ŋk]) in verwandten Wortformen, z.B. in '*Tria*[ŋ]*el* – *tria*[ŋg]*ulär*' (vgl. KLOEKE 1982: 119): Solche Alternationen existieren auch in dialektalen Varianten des Deutschen, z.B. in norddt. '*Di*[ŋk] –*Di*[ŋ]*e*' oder '*la*[ŋk] – *lä*[ŋ]*er*'.
– Die Tilgung von finalem Schwa in *Ge*-...-*e*-Kollektiva erfolgt regelmäßig nach einfachen Nasalen, nicht jedoch nach dem Cluster [nd] und nach [ŋ] (vgl. VENNEMANN 1970: 77 und KLOEKE 1982: 118):

(1–12)

Gelände	Gedärm
Gebinde	Gebein
Gewinde	Gestirn
Gehänge	Gestein
Gestänge	Gehörn

– Innerhalb der starken Verben mit [ɪ]-Stamm im Präsens zeigen auf [ŋ] auslautende Verbstämme im Partizip Perfekt die gleiche Form wie auf

‘Nasal + Plosiv’ endende Stämme, während Verben mit einfachem Nasal in finaler Position einem anderen Ablautmuster folgen (vgl. VENNEMANN 1970: 76):

(1–13)

I		II		III	
binden	gebunden	klingen	geklungen	rinnen	geronnen
sinken	gesunken	ringen	gerungen	spinnen	gesponnen
trinken	getrunken	singen	gesungen	sinnen	gesonnen

Zur Abbildung der phonologischen Repräsentationen /ng/ und /nk/ auf die phonetischen Repräsentationen [ŋ] (und vor Vollvokal, z.B. in *Ingo*, [ŋg]) und [ŋk] werden zwei phonologische Regeln benötigt, eine regressive Nasalassimilationsregel und eine g-Tilgungs-Regel. Beide Regeln sind intrinsisch geordnet, weil die Nasalassimilation den Input für die g-Tilgungs-Regel erst erzeugt: Die erste Regel füttert die zweite; es besteht also eine „feeding order“ zwischen beiden Regeln.

(1–14)

	/bank/	/bang/	**Phonologische Repräsentation**
1. regressive Nasalassimilation			
	/baŋk/	/baŋg/	
2. g-Tilgung		/baŋ/	
	[baŋk]	[baŋ]	**Phonetische Repräsentation**

Für die (innerhalb der Morphem-Domäne obligatorische) regressive Nasalassimilation wurde folgende Regel vorgeschlagen (vgl. RAMERS / VATER [4]1995: 91):

(1–15) $$\begin{bmatrix} +\text{nas} \\ -\text{lab} \end{bmatrix} \rightarrow [+\text{hint}] \,/\, __ \begin{bmatrix} -\text{son} \\ -\text{dau} \\ +\text{hint} \end{bmatrix}_{\mu} \Bigg]$$

Die Regel besagt, dass ein nicht-labialer Nasal hinten (also als [ŋ]) realisiert wird vor einem hinteren nicht-dauernden Konsonanten (d.h. vor hinterem Plosiv), der sich vor der Morphemgrenze befindet.

1.2 Vokalismus

1.2.1 Monophthonge

Im deutschen Lautsystem können insgesamt 16 Monophthonge unterschieden werden, ohne das vokalisierte [ɐ], das sich auf das konsonantische /r/ zurückführen lässt, und ohne nasalierte Vokale, die nur in Fremdwörtern

aus dem Französischen, z.B. *Balkon* und *Bassin*, vorkommen. /a/ und /ɑ:/ werden in der zugrundeliegenden Repräsentation als Hintervokale analysiert, obwohl sie, rein artikulatorisch betrachtet, als Zentralvokale (wie [ə] und [ɐ]) realisiert werden. Offensichtlich bilden die a-Laute mit den hinteren Vokalen eine natürliche Lautklasse. Das zeigt sich besonders daran, dass nach ihnen (wie nach [ɔ, o, ʊ, u]) das [x]-Allophon (der „Achlaut"), nicht das [ç]-Allophon (der „Ichlaut") steht (vgl. [ax], [dɔx], [bu:x] vs. [ɪç], [ɛçt], [mœçtə]).

Zur Differenzierung der Vokale von den übrigen Oberklassen, den Sonoranten, Obstruenten und Laryngalen, genügt die Spezifizierung [-kons, +son] (vgl. Tab. 2). Glides werden auf der segmentalen Ebene von „silbischen" Vokalen nicht unterschieden, weil Silbischkeit kein inhärentes Merkmal isolierter Segmente ist, sondern die relationale Eigenschaft kennzeichnet, innerhalb der prosodischen Domäne 'Silbe' das sonorste Element, d.h. den Silbengipfel zu bilden. Gleiches gilt für silbische Liquide und Nasale, die ebenfalls segmental von anderen Liquiden und Nasalen nicht unterscheidbar sind (zur Begründung dieser Analyse vgl. auch Kloeke 1982: 66 und Wiese 1988: 62 f.). Außer den Oberklassenmerkmalen werden zur Differenzierung der Vokale des Deutschen nur die Merkmale 'hinten', 'hoch', 'tief', 'labial'[12], 'gespannt' und 'lang' benötigt. Für das Vokalsystem des Standarddeutschen wird folgende vorläufige Merkmalsmatrix vorgeschlagen (ohne die zwei Oberklassenmerkmale):[13]

Tab. 2: Merkmalsmatrix für die Vokale des Standarddeutschen

	i	ɪ	e	ɛ	y	ʏ	ø	œ	ɑ	a	u	ʊ	o	ɔ	ə
hint	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+
hoch	+	+	–	–	+	+	–	–	–	–	+	+	–	–	–
tief	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–
lab	–	–	–	–	+	+	+	+	–	–	+	+	+	+	–
gesp	+	–	+	–	+	–	+	–	+	–	+	–	+	–	–

Die Vokale des Deutschen zerfallen in zwei Klassen, die traditionell als „Lang- und Kurzvokale" bzw. „geschlossene und offene Vokale" bezeichnet werden. Durch welches distinktive Merkmal die Vokale der beiden Gruppen jeweils differenziert werden, ist seit langem umstritten (vgl. Werner 1972: 24–30 und Ramers 1988). Auf die Frage nach dem distinktiven Merkmal

[12] Das Merkmal '**l**abial' wird anstelle des traditionellen '**r**und' verwendet, weil es auch für die bilabialen und labiodentalen Obstruenten des Deutschen benutzt werden kann, die nicht mit Lippenrundung, sondern nur mit aktiver Beteiligung der Lippen artikuliert werden. Weil im Deutschen keine nicht-runden Labialvokale neben den runden existieren, kann im Gesamtlautsystem auf das Merkmal '**r**und' verzichtet werden.

[13] Mit weniger Merkmalen für Vokale kommt Yu (1992a) aus, der hier Unterspezifikation (d.h. Verzicht auf Spezifizierung von Merkmalen, die auf Grund anderer Merkmale voraussagbar sind) annimmt.

innerhalb dieser Vokalopposition wurden in der Hauptsache drei Lösungen angeboten:
– das Quantitätsmerkmal,
– das Gespanntheitsmerkmal,
– die Anschlussart an den Folgekonsonanten (Silbenschnittkorrelation):

Die letzte, von TRUBETZKOY (1939) gewählte, Option wird hier nicht berücksichtigt, weil sich gezeigt hat, dass sie aufgrund der Vokal- oder Konsonantendauer voraussagbar ist (vgl. FISCHER-JØRGENSEN 1969, REIS 1974). Als phonetische Korrelate des Gespanntheitsmerkmals wurden in 1.1.1 Muskelspannung und Zentralisierung genannt. Für die Gespanntheit als distinktives Merkmal wird vor allem die Existenz gespannter und ungespannter Kurzvokale in Nebensilben und das Fehlen von Langvokalen in dieser Position als Argument angeführt, wie Tab. 3 (nach MOULTON 1962: 63) zeigt (zur Notierung vgl. Fn. 7):

Tab. 3

D[i]ner	d[ɪ]ffus
D[e]tail	D[ɛ]ssert
K[o]lumbus	K[ɔ]llege
k[u]rier	sk[ʊ]rril
am[y]sieren	m[ʏ]stifizieren
[ø]konom	[œ]strogen

Bei Aufrechterhaltung der Opposition „gespannt – ungespannt“ ist die Opposition „lang – kurz“ offenbar im Nebenton aufgehoben (vgl. REIS 1974). Vokallänge als von Gespanntheit und Akzent abhängige Eigenschaft wird mit einer phonologischen Regel (nach KLOEKE 1982: 9) abgeleitet:

(1–16) $\begin{bmatrix} \text{V} \\ +\text{gesp} \rightarrow \end{bmatrix}$ $\begin{bmatrix} \overline{} \\ +\text{akz} \end{bmatrix}$

Für das Gespanntheitsmerkmal lassen sich allerdings nur schwer phonetische Korrelate finden (vgl. RAMERS 1988, Kap. 2). WIESE (1988: 62–65) führt als weiteres Argument an, dass sich Langvokale phonotaktisch wie Diphthonge bzw. Kombinationen von Vokal und Konsonant verhalten. Dieser distributionelle Sachverhalt kann nur gedeutet werden, wenn man Kurzvokale als eine quantitative (metrische) Einheit wertet, Langvokale, Diphthonge und Vokal + Konsonant-Cluster dagegen als bimetrisch beurteilt (vgl. WIESE 1988: 64). Danach wäre Quantität im Deutschen nicht ein inhärentes segmentales, sondern ein suprasegmentales Merkmal. Externe Evidenz dafür liefern Versprecher, in denen Segmente ohne Änderung der Quantitätsverhältnisse vertauscht werden oder wo nur die Quantität ihren Wert ändert (vgl. STEMBERGER 1984 und BERG 1988); vgl. folgende Beispiele:

(1–17) a eine üppige W[y:]se
(statt: [i:]) (Segmentvertauschung ohne Quantitätsänderung)
b Du brauchst mich n[i:]cht zu fragen.
(statt: [ɪ]) (Änderung der Quantität)

RAMERS (1988) postuliert daher ein Schema zugrundeliegender Vokale, das von den ungespannten Vokalen nur [ə] enthält, das kurz und lang vorkommt (vgl. *kesse* [kɛsə] vs. *Käse* [kɛ:zə]):

Tab. 4: Zugrundeliegende Vokale im Deutschen

i	y	u
e	ø	o
ɛ		a

Für eine vollständige phonologische Repräsentation deutscher Wörter benötigt man dann noch suprasegmentale quantitative Strukturen der in ihnen enthaltenen Vokale wie in (1–18):

(1–18) V C V
a (in *Bahn*) a (in *Bann*)

Diese unterschiedliche quantitative Struktur hat auch Einfluss auf die qualitativen Eigenschaften der Vokalrealisierungen im phonetischen Output; hohe und mittlere Kurzvokale werden mit tieferer Zungenposition gebildet als die entsprechenden Langvokale (bei den a-Lauten ist es umgekehrt); sie nehmen auch eine zentralere Stellung in der horizontalen Zungenlage ein. Dadurch weicht die phonetische Repräsentation der Vokale des Deutschen erheblich von ihrer phonologischen ab (vgl. RAMERS / VATER [4]1995: 102):

Tab. 5: Phonetische Repräsentation deutscher Vokale

	i		y						u
	e	ɪ	ø	ʏ				ʊ	o
		ɛ	œ		ə			ɔ	
						a			
							ɑ		

Phonologische Regeln, die zwischen der phonetischen und der phonologischen Repräsentation vermitteln, schlägt RAMERS (1988: Kap. 4) vor. Diese an das System von H. P. JØRGENSEN (1969: 219) angelehnte Darstellung müsste durch weitere Untersuchungen bestätigt werden. Zudem müsste der (aus konsonantischem [r] abgeleitete) r-Vokal [ɐ] in Wörtern wie *Lehrer* [le:rɐ]

und *Meister* [majstɐ], dessen phonologische Merkmale schwer zu bestimmen sind, noch einbezogen werden.[14]

1.2.2 Diphthonge

Diphthonge können phonetisch als vokalische Laute charakterisiert werden, bei deren Artikulation eine Gleitbewegung der Zunge von einer Vokalposition in eine andere vollzogen wird; akustisch ergibt sich eine Änderung der Formantstruktur, die einen hörbaren Wechsel der Klangqualität erzeugt. Auch während der Produktion von Monophthongen ändert sich die akustische Struktur, nur sind diese Änderungen zu geringfügig, um wahrnehmbar zu sein. Phonologisch lässt sich ein Diphthong als (der gleichen Silbe angehörende) Folge aus Vokal + Glide charakterisieren.[15] Die Haupt-Diphthonge des Deutschen, [aj], [aw] und [ɔj][16] stehen zueinander in Opposition:

frei	Frau	freu'
freien	Frauen	freuen

[aj], [aw] und [ɔj] sind fallende Diphthonge, d.h. das erste Element bildet jeweils den Silbengipfel, es ist sonorer als das zweite (vgl. 2.1). Die Realisierung der Glides unterliegt starker Variation, die bei /aj/ von [i] über [ɪ] und [e] bis zu [ɛ] oder [ə] reicht. Die Repräsentationen /i/, /u/ und /y/ werden als zugrundeliegend angesehen, weil die tieferen und zentraleren Varianten als reduzierte Formen interpretierbar sind (vgl. Ramers / Vater [4]1995). Diphthonge nehmen, anders als Affrikaten, zwei Positionen auf der CV-Schicht ein, da sie sich phonotaktisch wie Langvokale verhalten:

- Sie stehen im Gegensatz zu Kurzvokalen im Wort- und Silbenauslaut: *Mai*, *Pleite*, *Bau*, *Heu, heute,* etc.
- Sie kommen im Standarddeutschen nie vor velarem Nasal vor, der auf das Cluster /ng/ zurückführbar ist (vgl. 1.1.2).
- Sie stehen nie vor der Affrikata [pf].[17]

In den Interjektionen *hui* und *pfui* tritt außerdem ein Diphthong [uj] auf. Darüber hinaus kommt im Deutschen jedoch noch eine Serie weiterer Diphthonge vor. Die von Moulton (1962) – vor der erstmaligen Beschreibung aller deutschen

---

[14] Man müsste für die Zungenhöhe offenbar einen zusätzlichen Grad (zwischen dem von [ə] und [a]) annehmen.

[15] Ein Glide hat (phonetisch) vokalische Eigenschaften, bildet aber (phonologisch) keinen Silbengipfel.

[16] Alternativ zu [ɔj] wird auch [ɔɥ] realisiert; [ɥ] ist das Transkriptionszeichen für einen gerundeten hohen vorderen Glide, wie er z.B. im französischen *nuit* [nɥi]) vorkommt.

[17] Ein weiteres Argument für die bisegmentale Lösung bildet die Umlautregel. Sie erfasst nur den zweiten Teil des Diphthongs [aw], der zu [ɥ] umgelautet wird. Das erste Element bleibt ein hinterer Vokal, der durch eine Rundungsassimilation an das vordere [œ] zum gerundeten [ɔ] wird.

/r/-Allophone bei ULBRICH (1972) – geschilderte r-Vokalisierung im Auslaut und vor Konsonanten bewirkt, dass Verbindungen aus Langvokalen und vokalischem /r/ (das als [ɐ] transkribiert wird) entstanden: [iːɐ, yːɐ, uːɐ, eːɐ, øːɐ, oːɐ, ɑːɐ].[18] So verdienstvoll Moultons Einsicht zu diesem frühen Zeitpunkt ist, so sind einige Mängel in seiner Darstellung nicht zu übersehen (vgl. VATER 2004):

- MOULTON (1962) zieht aus seinen Beobachtungen nicht die Konsequenz, dass die neu entstandenen Sequenzen [iːɐ, yːɐ, uːɐ, eːɐ, øːɐ, oːɐ, ɑːɐ] in *hier*, *für*, *Uhr*, *er*, *hör*, *Chor*, *Bar* als Sequenzen aus Vollvokal + Glide (oder wie er sagt „non-syllabic vowel") typische Kandidaten für Diphthonge sind.
- Moultons Spezifizierung der dem /r/ vorangehenden Vokale als „lang" stimmt nur bedingt: Gespannte Vokale sind nur lang unter Betonung, sonst kurz: In *Ozon* ist das erste /o/ kurz, das zweite lang. Vor [ɐ] verlieren m.E. die gespannten Vokale ihre Länge, so dass fürs Deutsche typische Diphthonge entstehen, also [iɐ], [yɐ], [eə] usw., nicht [iːɐ], [yːɐ], [eːɐ]. Moulton (1962: 24) nimmt im übrigen an, dass nach ungespannten Kurzvokalen, also in *irr*, *Herr* etc., /r/ oft vokalisiert wird.

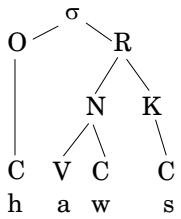
Die Einsichten Moultons wurden in RAMERS / VATER ([4]1995: 108–110) und VATER (2004) berücksichtigt, wobei auch die Darstellungen von KÖNIG (1989 und 1999) und BECKER (1995) kritisch gewürdigt wurden. Bei KÖNIG (1999) werden die Sequenzen [iːɐ, yːɐ, uːɐ, eːɐ, øːɐ] systematisch als Diphthonge angesehen. Allerdings stimme ich in einigen Punkten mit dieser Darstellung nicht überein, vor allem, weil ich im Gegensatz zu Königs experimental-phonetischen Studien theoretische Überlegungen im Rahmen der Autosegmentalen Phonologie angewandt habe. Ein Diphthong nimmt zwei Positionen in der Skelettschicht ein; er besteht aus einem Silbengipfel (V) und einem (silbenmarginalen) Glide, der als C eingestuft wird (vgl. *Haus* in (1–20)):

(1–20)

C V C C	**Skelettschicht**
| | | |	
h a w s	**Segmentschicht**

Die Verbindung der Skelettschicht mit Pike'schen Silbenkonstituenten (vgl. 2.1) ergibt (1–21):[19]

(1–21)



[18] MOULTON (1962: 24) wählt die amerikanischen Studenten bekannte Transkription [ʌ] (vgl. engl. *but* [bʌt]).

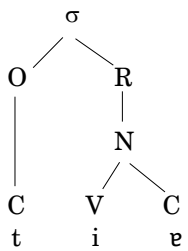
[19] O = Onset, R = Reim, N = Nukleus, K = Koda (vgl. die Beschreibung des PIKE-Schemas in 2.1):

Die erwähnten Untersuchungsergebnisse sprechen dafür, Realisierungen von [ɐ] nach Vokal wie in *ihr*, *Uhr*, *Ohr* als Diphthonge zu werten; vgl. BECKER (1995: 185):

> „Das vokalisierte /r/ bildet mit den Nukleusvokalen tautosyllabische Vokalverbindungen, die zu den Diphthongen gerechnet werden müßten, was auch einige Phonologen tun, z.B. Menzerath 1941: 63 und Vater 1992: 117.“[20]

(1–22) zeigt die Silbenstruktur von *Tier*:

(1–22)



Wie MOULTON (1962) nehme ich an, dass auch nach ungespannten Vokalen /r/ meistens (wenn auch nicht so kategorisch wie nach gespannten Vokalen) als Glide realisiert wird.[21] Auch BECKER (1995: 187) konstatiert (gestützt auf ULBRICH 1972:114), dass der „Gesamtstreubereich von [ɐ] auch [ɪ] und [ʊ] erfasst und damit „fast den gesamten Vokalraum“ abdeckt, und KÖNIG (1989: 78) setzt sechs Diphthonge mit ungespanntem erstem Bestandteil an. Durch „Glidisierung“ – wie ich diesen Vorgang bezeichne – sind also nicht nur die bereits erwähnten (und auch von KÖNIG 1999 angenommenen) 7 Diphthonge [iɐ, yɐ, uɐ, eɐ, øɐ, oɐ, ɑɐ] entstanden, sondern auch die Diphthonge [ɪɐ, ʏɐ, ʊɐ, ɛɐ, œɐ, ɔɐ, aɐ], wobei, da [ɑ] und [a] vermutlich vor r-Glide in [a] zusammenfallen, nur noch 6 Diphthonge dazukommen, in Wörtern wie *irr*, *dürr*, *Furt*, *Herr*, *örtlich*, *dort*. Die aus ungespanntem Vokal + [ɐ] gebildeten Diphthonge sind insofern sogar noch natürlicher, als das traditionelle Diphthongsystem nur Kombinationen aus ungespanntem Vokal und Glide kennt: [aj], [aw] und [ɔj]. (1–23) repräsentiert die Silbenstruktur des Worts *Stern* mit dem Diphthong [ɛɐ]:[22]

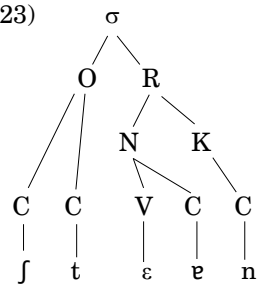
---

[20] Nach BECKER (1995: 185) weicht jedoch die Mehrheit der Phonologen der Frage aus bzw. zählt diese Verbindungen nicht zu den Diphthongen (wie z.B. MEINHOLD / STOCK 1982: 86 f.).

[21] Dabei fallen offenbar gespanntes [ɑ] und ungespanntes [a] zusammen: [haɐt] kann die Realisierung von *haart*, *hart* und *harrt* sein (vgl. ULBRICH 1972: 123, BECKER 1995:188): KÖNIG (1989 und 1999) nimmt an, dass nach [ɑ] immer und nach [o] meistens /r/ total wegfällt Ich habe jedoch oft genug den Diphthong [aɐ] gehört.

[22] Nicht-Silbischkeit wird nicht angezeigt, da [ɐ] durch die Assoziation mit C als nichtsilbisch zu werten ist.

(1–23)



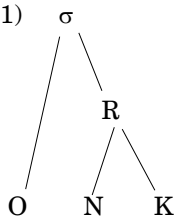
Auf Grund der vorangehenden Überlegungen sind fürs Deutsche insgesamt die 17 Diphthonge [aj, aw, ɔj, uj, iɐ, yɐ, uɐ, eɐ, øɐ, oɐ; ɪɐ, ʏɐ, ʊɐ, ɛɐ, œɐ, ɔɐ, aɐ] anzunehmen.[23] Die Erkenntnisse von Moulton (1962) und die darauf aufbauenden Studien zeigen, dass die gängigen phonologischen Beschreibungen des Deutschen (inkl. Aussprachewörterbücher) das Phonemsystem des Deutschen im vokalischen Bereich nur sehr unvollkommen darstellen.

2 Suprasegmentale Strukturen

2.1 Silbenstruktur

Pike / Pike (1947) postulieren eine hierarchische Silbenstruktur mit den Silbenkonstituenten O(nset), R(eim), N(ukleus) und K(oda):

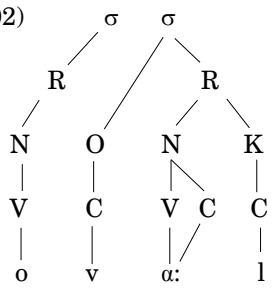
(2–01)



Selkirk (1982) greift diese Analyse auf und gibt eine Begründung für die Annahme von Silbenkonstituenten: Sie bilden Domänen für phonologische Regularitäten und Prozesse. So ist im Deutschen die Silbenkoda nach Vater (1992) die Domäne für die Auslautverhärtung und der Reim die Domäne für Nasalassimilationen. Die Silbenkonstituenten wurden von Selkirk (1982) mit der CV-Struktur kombiniert. Das Wort *oval* hat die Struktur (2–02):

[23] Die Frage, ob alle Diphthonge zugrunde liegend anzusetzen sind oder ob die „neuen" Diphthonge nicht eher als abgeleitet aus Vokal + r zu erklären sind, wird in Vater (2004) diskutiert. Becker (1995: 187 ff.) plädiert dafür, dass Diphthonge nicht unbedingt Phonemstatus haben müssen. Hierzu sind m.E. weitere Studien notwendig.

(2–02) (nach VATER 1992: 123)

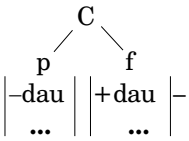


Die erste Silbe besteht aus einem N(ukleus) dominierenden R(eim), die zweite aus O(nset) und R, das in N und K(oda) verzweigt. N enthält V und C: Ein kurzer Vokal wird nur mit V assoziiert, ein langer mit V+C, wobei V den Silbengipfel, C die Länge (die wie ein konsonantisches Segment fungiert) repräsentiert.

VENNEMANN (1982: 268) analysiert die Silbe nach dem Pike-Schema, doch ohne Reim. Jede Silbe hat einen obligatorischen Nukleus und fakultative Ränder. Die Silbe ist nach dem Prinzip der Konsonantischen Stärke (CS) aufgebaut. CS ist die Umkehrung der Sonoritätshierarchie. Die CS nimmt nach den Rändern hin zu. Silben können überlappen in ambisyllabischen Elementen („Gelenken"): Später (vgl. VENNEMANN 1994: 13 f.) analysierte er Silben nach den von SIEVERS ([3]1885) erstmalig postulierten Silbenschnittkorrelationen. Danach unterscheiden sich Silben mit gespanntem Vokal von Silben mit ungespanntem dadurch, dass bei ersteren die Energie unmittelbar nach Artikulation des Langvokals abfällt („sanfter Schnitt"), bei letzteren mitten in der Koda, wobei ein „scharfer Schnitt" vorliegt und der Folgekonsonant zum „Gelenk", d.h. ambisilbisch wird.

Die Silbenstruktur des Deutschen lässt sich gut im Rahmen der Autosegmentalen Phonologie darstellen, wobei man auch die Pikeschen Silbenkonstituenten (nach SELKIRK 1982, VATER 1992 und WIESE 1996) berücksichtigen kann. In 1.1.2 wurde bereits dargestellt, dass die Annahme einer Skelettschicht, die mit der Segmentschicht verbunden ist, das Affrikatenproblem lösen kann. Affrikaten bilden in der Segmentschicht zwei Einheiten (die sich im Merkmal [±dauernd] unterscheiden), die mit einer C-Einheit der Skelettschicht verbunden sind; vgl. (2–03) (= (1–07)):

(2–03)

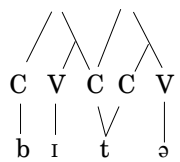


Skelettschicht
Segmentschicht

Ein „ambisilbischer" Konsonant wird nach WIESE (1988: 78 ff.; 1996: 35 f.) und RAMERS (1992) mit zwei C-Einheiten verbunden:[24]

[24] In VATER (1992, 1998) habe ich darauf hingewiesen, dass diese Darstellung problematisch ist, da sie normalerweise für lange Elemente (z.B. lange Vokale im Dt., lange Konsonanten im Italienischen oder Polnischen) reserviert ist. Ambisilbische Konsonanten im Deutschen sind normalerweise nicht lang. Ich habe gezeigt, dass das Deutsche durchaus Silben hat, die auf

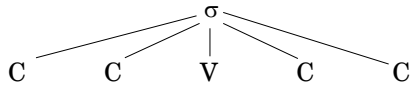
(2–04)



(Die Silbenkonstituenten O, R usw. wurden hier weggelassen)

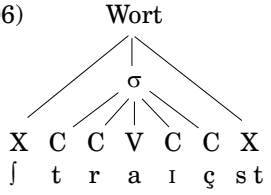
CLEMENTS / KEYZER (1983: 8) schlagen fürs Englische und andere Sprachen ein „flaches“ Silbenstruktur-Modell vor, wo die Silbe nicht in Konstituenten unterteilt ist. Ein solches Modell setzt WIESE (1986: 3) fürs Deutsche an, wobei er neben einer V-Einheit vier C-Einheiten annimmt:

(2–05) Das flache Silbenmodell von Wiese (1986)



Zusätzliche Elemente (wie den Endkonsonanten in *Mond, Obst* und *Herbst*) erklärt WIESE (1986, 1988) als extrasilbisch. Er weist darauf hin, dass alle extrasilbischen Konsonanten (/s/, /t/ oder /st/) koronal sind, wobei er /st/ als ein Element rechnet.

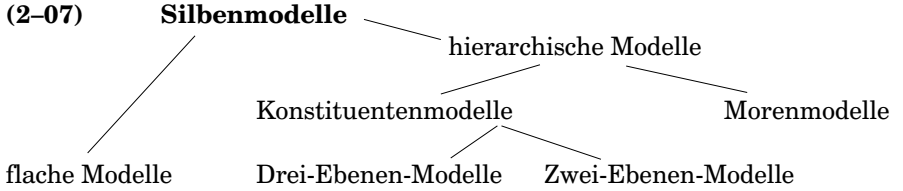
(2–06)



(nach WIESE 1988: 100)

Dem flachen Silbenmodell stehen mehrere hierarchischen Modelle gegenüber (vgl. VATER 1998):[25]

(2–07)



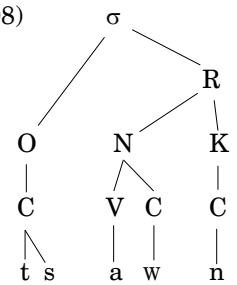
Ein zweischichtiges Konstituentenmodell – bei dem der Reimknoten fehlt – verwenden NOSKE (1992) und WIESE (1996): Ich befürworte das „klassische“

Kurzvokal enden (z.B. Schwa-Silben) und dass die Annahme ambisilbischer Konsonanten in Fällen wie *Ebbe*, *Kladde*, *Egge*, *meschugge, Rappen*, *Socke* und *bitte* nicht adäquat ist.

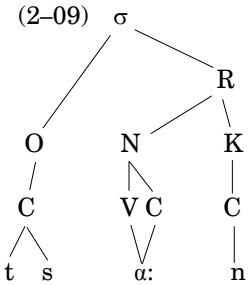
[25] Vgl. hierzu auch RAMERS (1998: 90–105).

dreischichtige – in (2–01) abgebildete – Konstituentenmodell von PIKE / PIKE (1947: 92), das von FUDGE (1969: 273) wiederentdeckt und von SELKIRK (1982: 341) ins Rahmenwerk der Autosegmentalen Phonologie integriert wurde. Dies Modell habe ich auf die Analyse der Silbe (speziell des Nukleus) im Deutschen angewandt (vgl. VATER 1992, 1998).[26] Durch Assoziation mit der Skelettschicht (C/V-Schicht) ergeben sich Strukturen wie (2–08) für *Zaun* und (2–09) für *Zahn*. (VATER 2002: 192):

(2–08)



(2–09)



Während die Konstituentenmodelle mit verschiedenen Konstituenten (O, R, N und K) operieren, ähnlich wie bei syntaktischen Konstituenten, teilen Morenmodelle die Silbe in gleichartige Teile, sogenannte Moren (Gewichtseinheiten): Die Konsonanten, die dem Silbengipfel vorangehen (im Konstituentenmodell: der Onset) werden als „gewichtslos“ angesehen (vgl. HAYES 1989, AUER 1991 und NOSKE 1992); sie tragen nicht zum Morengewicht bei. Jeder Kurzvokal und jeder Koda-Konsonant wird als eine More gezählt; ein Langvokal zählt als zwei Moren. Nach allgemeiner Übereinstimmung aller Phonologen spielt **Sonorität** (vgl. 1.2.1) eine zentrale Rolle in der Silbenstruktur Sonorität ist eine relationale Eigenschaft von Lautsegmenten; sie ist bei Vokalen größer als bei Konsonanten, doch gibt es innerhalb der Konsonanten noch große Unterschiede. Das sonorste Element in einer Silbe bildet den Silbengipfel.

Obwohl das phonetische Substrat der Sonorität noch nicht entdeckt wurde (vgl. HEIKE 1992), wissen wir, dass Sonorität vom Silbengipfel hin zu den Rändern kontinuierlich abnimmt. Die Sonoritätsskala in Silben gilt als universell. So sind Sonoranten (Liquide wie [l] und [r], Nasale wie [m], [n] und [ŋ]) in allen Sprachen sonorer als Obstruenten (Plosive und Frikative) und daher näher am Zentrum (d.h. am Silbengipfel); hier Beispiele aus dem Deutschen und Englischen:

(2–10)	a [tsɛlt]	b [grʊft]	c [pflɪçt]
	Zelt	*Gruft*	*Pflicht*
(2–11)	a [hænd]	b [klɛft]	c [rɪsk]
	hand	*cleft*	*risk*

[26] Vgl. auch GIEGERICH (1992) und YU (1992b) zum Onset sowie HALL (1992a,b) zur Koda.

Die Universalität der Sonoritätsskala wird jedoch natürlich beeinflusst durch sprachspezifische Besonderheiten (wie Fehlen von Glides). WIESE (1996: 260) schlägt die Sonoritätsskala (2–12) vor, in der Sonoranten in drei Gruppen differenziert sind, während Obstruenten undifferenziert bleiben:

(2–12) Sonoritätsskala nach WIESE (1996)

Obstruenten	Nasale	l	r	hohe Vokale	Vokale

Ich habe Kritik an dieser Skala aus folgenden Gründen (zu anderen Kritikpunkten vgl. NEEF 1996):

- Alle Vokale verhalten sich in Bezug auf Sonorität gleich; /a/ hat nicht mehr Sonorität als /e/ oder /i/; WIESE (1996) denkt wohl eher an die Glides /j/, /w/ (vgl. [bajn], [maws]) als an hohe Vokale.
- Glides sind offenbar sonorer als Sonoranten, da sie (zusammen mit Vokalen) im Nukleus vorkommen; vgl. /hajm/ *Heim*, /rawm/ *Raum*, /hajl/ *Heil*, /gawl/ *Gaul*. Das gilt auch für den r-Glide in Diphthongen (vgl. [uɐ] *Uhr*, [iɐ] *ihr* oder [Iɐ] *irr* und die anderen Beispiele in 1.2.2).
- Gewöhnlich wird angenommen, dass innerhalb der Obstruenten die Sonorität bei Frikativen stärker ist als bei Plosiven (in der Interjektion *pst* bildet der Frikativ sogar den Silbengipfel).

Ich schlug auf Grund dieser Überlegungen eine andere Sonoritätsskala vor (vgl. VATER 1998):

(2–13) Sonoritätsskala nach VATER (1998)

Plosive	Frikative	Nasale	silbische Nasale	Liquide	Glides	Vokale

Diese Skala basiert unter anderem auf meiner Beobachtung, dass silbische Nasale im Deutschen nach Nasalen im Onset auftreten; dies geschieht nach Elision von Schwa zwischen Nasalen:

(2–14) *nehmen* [ne:mən] → [ne:mn̩] → [ne:mm̩]
(2–15) *singen* [zɪŋən] → [zɪŋn̩] → [zɪŋŋ̩]

Nach Schwa-Elision zwischen Nasalen entstehen Sequenzen aus nichtsilbischen Nasalen im Onset und silbischen Nasalen im Nukleus (wo sie den Silbengipfel bilden): [rɛnn̩], [kɔmm̩]. Die Ansetzung zweier Sonoritätsstufen bei den Nasalen mag befremden; ich kann jedoch keine andere Erklärung finden für das Auftreten silbischer Nasale im Nukleus, wo sie offenbar höhere Sonorität erlangen als „normale“ Nasale im Onset. Niemand anders hat bis jetzt dies heiße Eisen angefasst.[27]

[27] Mein Vorschlag wurde kritisiert: Der hohe Sonoritätsgrad sollte allein auf Grund der Nukleusposition vorhersagbar sein. Diese Erklärung genügt, wenn ein Obstruent vorangeht wie in

Phonologen nehmen übereinstimmend an, dass der Nukleus des Deutschen niemals drei Positionen enthält: Es gibt im Deutschen weder Triphthonge noch Diphthonge, die aus Langvokal und Glide bestehen; Langvokale nehmen zwei Positionen in der Skelettschicht ein (s. oben): Ich nehme an, dass nicht nur die traditionellen Diphthonge des Deutschen, also /aj/, /aw/ und /ɔj/, deren vokalische Konstituenten ungespannt und kurz sind, bimetrisch sind (d.h. zwei Positionen einnehmen), sondern auch die durch r-Glidiserung entstandenen Diphthonge, deren erste Konstituente meist gespannt ist wie in *ihr* [iɐ], *er* [eɐ], *Uhr* [uɐ], *Ohr* [oɐ] (vgl. 1.2.2); kurze gespannte Vokale muss man ohnehin fürs Deutsche annehmen wie in der ersten Silbe von *oval* (vgl. (2–02)), *egal, Husar*. Vor [ɐ]-Glides hat der Vokal seine Länge. die er vor konsonantischem [r] hat, eingebüßt (vgl. *ihre*, *Uhren*, *Ehre*):

In Bezug auf das Minimum eines Silbennukleus besteht weniger Einigkeit. Wiese (1986: 5) nimmt an, dass im Deutschen der Silbennukleus immer aus zwei Elementen besteht. Abgesehen davon, dass die Annahme eines Nukleus (der aus dem Silbengipfel und einem zusätzlichen Element besteht) dem von Wiese (1986; 1988) postulierten flachen Silbenmodell widerspricht (vgl. (2–06)), ergibt sich noch die unangenehme Konsequenz, dass bei einem obligatorisch bimetrischen Nukleus die Unterscheidung leichter und schwerer Silben, die relevant ist für die Erklärung der Akzentverhältnisse im Deutschen (vgl. Ramers 1992: 256) unter den Tisch fällt, denn dann gibt es praktisch nur noch schwere Silben (mit zweielementigem Nukleus): Wiese (1986; 1996) vernachlässigt drei Silbentypen des Deutschen, die keineswegs marginal sind (vgl. Vater 1992):

- Silben, die auf Schwa enden; vgl. *Silbe* [zɪlbə] und *Vase* [vɑ:zə];
- Silben mit kurzem gespanntem Vokal wie in der ersten Silbe von *oval* und *zivil*;
- Silben mit silbischem Konsonant wie in *Himmel* [hɪml] und *Regen* [re:gɳ]. Sie erlauben kein weiteres Element im Nukleus.

Kurze gespannte Vokale kommen in zahlreichen Wörtern fremden Ursprungs vor (vgl. *Idee*, *Husar, Regal, oval, banal, Hydrant, Ödem*), oft sogar mehrmals in einem Wort wie in *Ökonomie* und *Banalität*, wo jeweils alle Silben gespannte Vokale enthalten, von denen nur der letzte lang ist, weil er den Hauptakzent trägt. Es gibt auch native Wörter mit gespanntem Kurzvokal wie *lebendig*, wo im Standarddeutschen der Vokal der ersten Silbe ein gespanntes /e/ ist (im Umgangsdeutschen wird dies [e] durch Schwa ersetzt): Aber auch in nativen einsilbigen Funktionswörtern wie *so*, *zu* und *wie*, die normalerweise unbetont sind (vgl. *das ist so/zu schwer*; *wie geht es?*), haben wir kurzen gespannten Vokal. Die Vokale in den zweiten Silben der nativen Wörter *Monat*, *Kleinod*, *Armut, Demut* und *Hochmut* – die Sekundärakzent

---

geben [ge:bm̩] und *Segen* [ze:gɳ], aber sie reicht nicht für Fälle wo ein Nasal vorangeht wie in (2–14) und (2–15). Offenbar erhöht sich die Sonorität eines Nasals, wenn er in die Nukleusposition gelangt.

haben – sind relativ kurz und sollten zu den kurzen Vokalen gerechnet werden, obwohl manche Grammatiken und phonologische Monographien sie als „halblang“ werten. Es ist fraglich, ob für deutsche Vokale drei Längengrade anzunehmen sind.

Es gibt Sprachen, in denen Konsonanten (gewöhnlich Sonoranten) ohne weiteres den Silbengipfel bilden können. Ein Beispiel ist das Tschechische, wie der berühmte Zungenbrecher zeigt:

(2–16) Strč prst skrs krk. (‘Steck den Finger durch den Hals’)

Im Deutschen können die Sonoranten /l/, /m/, /n/ und /ŋ/ ebenfalls den Silbengipfel bilden, alternativ zur Realisierung von Schwa wie in (2–17)).

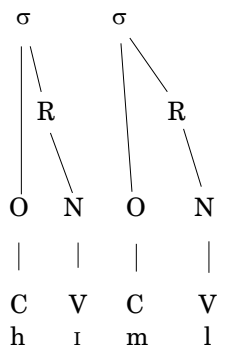
(2–17)	a [ge:bm]/[ge:bən]	*geben*		b [tre:tn] / [tre:tn]	*treten*
	c [zɑ:gŋ] / [zɑ:gən]	*sagen*		d [hɪml] / [hɪməl]	*Himmel*

Vibranten (/r/-Laute) werden in dieser Position vokalisch: [o:dɐ], [le:rɐ]. Evidenz für die komplette Vokalisierung von /r/ bieten unreine Reime, wo [ɐ] mit [a] reimt.

(2–18) a Und die <u>Tuba</u> bläst der <u>Huber</u>. (Titel eines Fernsehprogramms)
b Heut war ich bei der <u>Frieda</u>. (Schlager aus dem Jahr 1928,
Das tu ich morgen <u>wieder</u>. gesungen von Max Raabe in einem Kabarettprogramm, gesendet im WDR, 25.5.2001)

Dieser Reimtyp kommt auch im Wort *Pizzaflitzer* vor, das einen Boten bezeichnet, der Pizzas auf dem Motorrad transportiert. Im Berlinischen fällt [ɐ] total mit [a]zusammen. Hier bilden *Frieda* und *wieder* einen reinen Reim. Laterale verhalten sich wie Nasale. *Himmel* hat die Baumstruktur (2–19):

(2–19) (nach Vater 1998: 145)

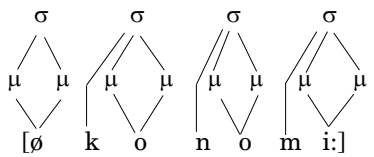


2.2 Akzentstrukturen

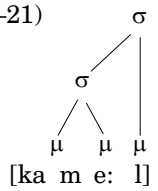
2.2.1 Akzentstrukturen in der Optimalitätstheorie

Das Deutsche ist nach FÉRY (1996) eine quantitäts-sensitive Sprache; das reguläre Fuß-Muster monomorphematischer Wörter ist ein moraischer Trochäus am rechten Wortende; davor kommt optional noch ein Trochäus vor. Bei komplexen Wörtern ist sind Flexive immer unbetont, während Derivative betont oder unbetont sind (vgl. *Schreib-eréi* vs. *Schréib-ung):* Komplexe Wörter verhalten sich mit einigen Ausnahmen wie einfache. In Komposita ist die erste Komponente betont und die zweite dann, wenn sie verzweigt (nach GIEGERICH 1985); hierzu gibt es einige Ausnahmen. FÉRY (1996: 64) zerlegt Silben in Moren (vgl. 2.1). Silben mit Schwa oder silbischem Sonoranten wie die Endsilben in *Ratte* [ratə], *Vogel* [fo:gl] und *schöner* [ʃø:nɐ]bewertet sie als nichtmoraisch.[28] Silben mit Vollvokal sind mindestens bimoraisch. Silben mit ungespanntem Vokal sind durch Konsonant geschlossen, damit bimoraisch. Offene Silben mit Kurzvokal – die monomoraisch sein müssten (vgl. die erste Silbe von *oval* in (2–02) oder die zweite Silbe in *Sofa*) – schließt sie aus. Alle Silben in (2–20) sind ihrer Meinung nach bimoraisch; d.h. sie behandelt alle vorkommenden Vokale als Langvokale, obwohl nur der letzte lang, da betont, ist. Eine Silbe, in der dem Vokal ein Koda-Konsonant folgt, ist trimoraisch (vgl. (2–21)):

(2–20) (FÉRY 1996: 66)



(2–21) (FÉRY 1996: 66)



Als Rahmen zur Beschreibung der Akzentverhältnisse wählt FÉRY (1996) die Optimalitätstheorie (OT): Außerdem sind die aus der Metrischen Phonologie bekannten „Gitterschläge“ relevant für die Akzentstruktur. Von den Constraints, die die Akzentstruktur aus der Silben- und Morenstruktur der Wörter ableiten (FÉRY 1996: 71), sind die in A „undominiert“, d.h. allen anderen übergeordnet.

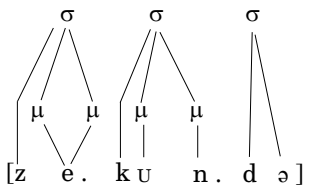
[28] NEEF (1996) fasst solche Silben als „Nebensilben“ zusammen.

Tab. 6: Constraints für Akzentmuster im Deutschen

A	GRID-CONSTRAINTS (GC)	a. Gitterschläge werden von minimal einer More oder Silbe projiziert
		b. Ein Gitterschlag wird von maximal zwei Moren projiziert (= B)
		c. Ein Gitterschlag wird von maximal einer Silbe projiziert.
		d. Der Kopf eines Fußes wird von genau zwei Moren projiziert.
B	THREEMORAS=TWOGRIDMARKS	Trimoraische Silben projizieren zwei Gitterschläge.
C	FOOT-BINARITY	Füße bestehen aus genau zwei Gitterpositionen
D	ALIGN-FOOT-RIGHT	Jedes Prosodische Wort endet auf einen Fuß
E	ALIGN-FOOT-LEFT	Jedes Prosodische Wort beginnt mit einem Fuß
F	FOOT-FORM(TROCHAIC)	Füße sind trochäisch
G	NOCLASH	Adjazente Köpfe sind verboten
H	ALIGN-HEAD	Der rechte Rand eines Prosod. Worts fällt mit seinem Kopf zusammen

Die Constraints sind verletzbar. Verletzungen höherrangiger Constraints sind schwerwiegender als Verletzungen niederrangiger Constraints. Optimal ist der Output, der die wenigsten bzw. nur die niederrangigsten Constraints verletzt (vgl. 3.2. in meinem Aufsatz im gleichen Band): Das Wort *Sekunde* hat die in (2–22) dargestellte prosodische Struktur.

(2–22)



Die Anwendung der Constraints auf die Struktur (2–22) ergibt Fuß- und Akzentstruktur in Tab.7:

Tab. 7 Constraint-Tableau für das Wort *Sekunde* (FÉRY 1996: 73)

μ μ /zekUndə/	NOCLASH	FOOT-FORM (TRO)	FOOT-BIN	ALIGN-FOOT RIGHT	ALIGN-FOOT LEFT
a. x (x .) ☞ Se kunde					*
b. (x) (x)(x .) Se kunde	*!		*		
c. x (x .) Sekun de				*!	
d. x (. x .) Sekunde		*!	*		

Der beste Kandidat ist der, der die wenigsten und niedrigsten Constraints verletzt. In diesem Fall ist das a., das nur den niedrigsten Constraint verletzt (die erste Silbe ist keinem Fuß assoziiert); b. (mit zwei Füßen) verletzt NOCLASH und FOOT-BIN; c. verletzt ALIGN-FOOT-RIGHT (das höheren Rang hat als ALIGN-FOOT-LEFT!); d. ist weder trochäisch noch binär. Beim Wort *Vitamin* [vita‘mi:n] nimmt FÉRY (1996: 73) in der schweren letzten Silbe drei Moren an, insgesamt fünf Moren im Wort. Die drei Silben sind zwei Füßen zugeordnet, von denen der erste aus zwei Silben, der zweite aus einer Silbe besteht: (x .)(x.), mit zusätzlichem Schlag auf *min*. Die richtig akzentuierte Variante a. *Vita‘min* verletzt keinen einzigen Constraint, während Variante b. *‘Vitamin* (mit gleicher Fußstruktur) ALIGN-HEAD verletzt und Variante c *Vi‘tamin* mit Fußstruktur (x)(x .) und Zusatzschlag auf *ta* sowohl den obersten Constraint, 3M = 2GM, verletzt sowie den zweiten, NOCLASH. Insgesamt handelt es sich um einen begrüßenswerten Versuch, die Akzentstruktur deutscher Wörter mit Hilfe der OT zu analysieren. Allerdings werden einige Annahmen der Verfasserin (wie z.B. die Außerachtlassung von Silben, die auf kurzen Vokal enden) nicht den Daten gerecht; andere Entscheidungen wie z.B. die Bewertung von Schwa-Silben als nichtmoraisch, sind m.E. fragwürdig. Der an sich interessante Ansatz von FÉRY (1996) ist daher revisionsbedürftig.

2.2.2 Eisenbergs Akzenttheorie

Peter Eisenberg hat (in EISENBERG 1991) eine alternative Akzenttheorie fürs Deutsche entwickelt, die sich auch aufs Englische und andere Sprachen

anwenden ließe. Diese Theorie ist nicht nur einfacher als die von LIBERMAN / PRINCE (1977) und GIEGERICH (1985), sondern m.E. auch insofern adäquater, als sie einige Probleme vermeidet, die im Zusammenhang mit der autosegmentalen und metrischen Theorie auftreten; sie reduziert auch die Zahl der Ausnahmen. EISENBERG (1991: 37 ff.) macht folgende Annahmen:

- Akzentmuster sind mit Füßen verbunden.
- „Wort“ als prosodische Einheit ist das flektierte Wort (nicht der Stamm, wie er z.B. im Nom. Sg. von Substantiven und im Imperativ von Verben realisiert wird).
- Zu unterscheiden sind „betonbare Silben“ und „betonte Silben“; die Relation zwischen ihnen ist relevant für die Bestimmung der prosodischen Struktur eines Worts.

Nach EISENBERG (1991: 44) ist es falsch, ein Wort wie *Haus* als einsilbig anzusehen. Man muss mit der flektierten (Plural-)Form *Häuser* arbeiten; dann zeigt sich, dass *Häuser* und *Meter* dem gleichen trochäischen Muster {–_} folgen, das das Hauptmuster einfacher (nicht-abgeleiteter) Wörter im Deutschen bildet. Alle Silben mit Vollvokal sind betonbar, alle Silben mit Schwa sind nicht betonbar. Innerhalb der Menge der betonbaren Silben gibt es eine Untermenge tatsächlich betonter Silben und innerhalb dieser Untermenge gibt es eine Untermenge von Silben, die den Hauptton tragen. „Betonungsstruktur“ und „Akzentstruktur“ sind zu unterscheiden. Die Akzentstruktur bestimmt die Silbe mit dem Hauptakzent innerhalb aller betonten Silben. In nicht-nativen Wörtern kann der Hauptakzent auf die letzte Silbe fallen, wie in [by’ro:] *Büro*, [kroko’di:l] *Krokodil*, [ɑnɑto’mi:] *Anatomie*. Um Wörter mit Hauptakzent auf der Penultima wie [‘aʊto] *Auto* zu analysieren oder auf der Antepenultima wie [‘ananas] *Ananas*, müssen wir wohl das Silbengewicht berücksichtigen; doch kann dieser Faktor nicht verantwortlich sein für Akzentunterschiede zwischen Wörtern wie ‘*Auto* and *Bü’ro* (wo das Silbengewicht gleich verteilt ist): EISENBERG (1991: 62) kann einen Teil dieser Betonungsunterschiede dadurch erklären, dass die Substantive verschiedenen Flexionsparadigmen angehören (vgl. *Fagótt* und *Márgot*).

Das Wort [‘ro:zə] *Rose* besteht aus einer betonbaren Silbe und einer unbetonbaren Schwa-Silbe; das Wort [‘ɛço] *Echo* „echo“ besteht aus zwei betonbaren Silben, die das gleiche Betonungs- und Akzentmuster zugeteilt bekommen:

(2–23)a	Silbenfolge:	{[ro:][zə]}	b	Silbenfolge:	{[ɛ][ço]}
	Betonbarkeitsstruktur:	{ +, _}		Betonbarkeitsstruktur:	{ +, + }
	Betonungsstruktur:	{ ¯, _}		Betonungsstruktur:	{ ¯, _}
	Akzentstruktur:	{ ˈ, _}		Akzentstruktur:	{ ˈ, _}!

Das Zeichen „ˈ“ zeigt Primärakzent an. Bei zweisilbigen Wörtern sind Betonungs- und Akzentstruktur identisch. *Rose* und *Echo* folgen beide dem trochäischen Muster. EISENBERG (1991:47) sagt: „Der Trochäus {ˈ,_} ist **das** Substantivmuster“. Auch FÉRY (1996) hält den Trochäus für das charakteristische

Betonungsmuster deutscher Substantive. Bei daktylischen Füßen wie ['rɛtətə] *rettete* ist die Struktur prädiktabel, da sie zwei unbetonbare Schwa-Silben nach der betonbaren Silbe enthält. Ein viersilbiger Fuß wie *goldenere* kann reduziert werden auf einen dreisilbigen wie *goldnere* (vgl. EISENBERG 1991: 62).

Bei abgeleiteten Wörtern besteht auch eine starke Tendenz zum trochäischen und daktylischen Muster. In mehr als dreisilbigen Formen finden wir gewöhnlich Kombinationen zwei- oder dreisilbiger Füße. Das Wort *Afri'kanerinnen* erhält folgende Betonungs- und Akzentstruktur:

(2–24) Silbenfolge:	{	[ɑ]	[fri]	[kɑ:]	[nə]	[rɪ]	[nən] }
Betonbarkeitsstruktur:	{	+	+	+	–	+	– }
Betonungsstruktur:	{	¯,	–	¯,	–,	¯	– }
Akzentstruktur:	{	¯,	–	ˈ,	–,	¯	– }

Komposita wie *Apfelbaum* und *Autoverkäufer* erhalten das selbe Betonungsmuster wie ihre Komponenten; der Hauptakzent fällt gewöhnlich auf die erste Komponente ('*Apfelbaum*, '*Autoverkäufer*): Eisenberg's Vorgehensweise hat mehrere positive Konsequenzen:

- Sie ist notationell einfacher als z.B. die Optimalitätstheorie.
- Sie kommt im Wesentlichen ohne das (problematische) Silbengewicht aus.
- Betonungsmuster nativer und nicht-nativer Wörter können nach gleichen Verfahren bestimmt werden.
- Das trochäische Muster erweist sich als das relevante Muster für einfache Wörter.

Es ist zu hoffen, dass diese interessante und einfache Akzenttheorie auf weitere Daten des Deutschen (vor allem im Bereich komplexer Wörter) angewandt wird.[29]

2.2.3 Intonationsstudien

Zum lange vernachlässigten prosodischen Bereich der Intonation im Deutschen gibt es einige interessante und ergebnisreiche neuere Untersuchungen.[30] Hier ist vor allem UHMANN (1991) zu nennen, ein Werk, in dem die nichtlineare Phonologie (im autosegmentalen und metrischen Ansatz) auf die Intonation des Deutschen angewandt wird. Auf der Grundlage des für englische Intonationsstrukturen entwickelten Modells von PIERREHUMBERT (1987) entwickelt sie ein eigenes Modell, zu dessen konstitutiven Bestandteilen Intonationsphrasen, metrische Gitter, tontragende Einheiten und Assoziationen von Tönen und Texteinheiten gehören. Die Verf. ordnet

[29] Zur hier vernachlässigten Frage des neutralen (bzw. „normalen") Akzents im Deutschen vgl. JACOBS (1982) und HÖHLE (1982).

[30] Als älteres (im Wesentlichen dem Strukturalismus verhaftetes) Standardwerk zur Intonation des Deutschen ist VON ESSEN (1956) zu nennen.

Intonationseigenschaften sogenannten Fokusstrukturen in Sätzen zu. „Fokus“ ist der Teil des Satzes, der inhaltlich im Vordergrund steht; nicht-fokussierte Teile bilden den Hintergrund. Die konstitutiven tonalen Eigenschaften deutscher Intonationsstrukturen werden durch phonetische Analysen bestimmt, in denen der Verlauf der Grundfrequenz-Maxima (F_0-Verlauf) ermittelt wird. In der Analyse der Phrasierungsgrenzen ergibt sich eine Häufung der Messwerte in einem bestimmten Frequenzbereich, wobei sich ein Grenztoninventar aus tiefem Grenzton *T%* und hohem Grenzton *H%* als relevant erweist. Den internen Verlauf einer Intonationsphrase bestimmen Akzenttöne, die den F_0-Verlauf auf metrisch prominenten Silben modifizieren. Fürs Deutsche nimmt die Verf. die bitonalen Konturtäne *H*+T* und *T*+H* sowie die monotonalen Töne *H** und *T** an.[31] Eine Intonationsphrase enthält obligatorisch einen Akzentton und einen Grenzton. Die Verf. konstruiert aus den möglichen Kombinationen von Grenz- und Akzenttönen eine *Intonationsgrammatik*.

3. Schlussbemerkungen

Der Überblick sollte die Anwendung neuerer phonologischen Theorien und Methoden aufs Deutsche deutlich machen. Es sind vor allem die Hauptansätze der Nicht-linearen Phonologie – Autosegmentale und Metrische Phonologie –, die erfolgreich bei der Analyse segmentaler und prosodischer Strukturen eingesetzt wurden. Beide Ansätze arbeiten im Gegensatz zum klassischen generativen Ansatz, dem SPE-Modell, nicht mit linearen, sondern mit hierarchischen Strukturen und sind deswegen vor allem geeigneter für die Analyse suprasegmentaler (prosodischer) Verhältnisse, die sich ja nicht an einzelnen Segmenten festmachen lassen, sondern größere phonologische Sequenzen betreffen. Aber auch im segmentalen Bereich bieten diese Ansätze Fortschritte, wie sich z.B. bei der Analyse von Affrikaten und Diphthongen (vgl. 1.1.2 und 1.2.2) zeigte.

Autosegmentale und Metrische Phonologie wurden vor allem im Bereich der Silbenstruktur – und darüber hinaus auch in den darüber liegenden prosodischen Bereichen (Fuß- und phonologische Wortstruktur) intensiv und, wie ich meine, effizient angewendet. Das Bild, das ich in 2.1 von der neueren Forschung über Silbenstrukturen des Deutschen entwarf, ist insofern einseitig, als ich nur über Analysen im Rahmen der Autosegmentalen Phonologie berichtet habe (die ich selbst in meinen Silben-Studien angewandt habe). Die Metrische Phonologie kam nicht zu Worte. Hier kann ich vor allem auf die ausgezeichnete und gründliche Studie von Giegerich (1985) verweisen. Interessanterweise zieht Wiese (1996) zur Analyse der Akzentstrukturen im Deutschen die Metrische Phonologie heran, obwohl er die Analyse der Silbenstruktur und segmentaler Strukturen im Rahmen der Autosegmentalen

[31] Bitonale Töne beziehen nachfolgende unakzentuierte Silben in den F_0-Verlauf ein.

Phonologie betreibt.[32] Autosegmentale und Metrische Phonologie wurden von UHMANN (1991) zur Untersuchung der Intonation im Deutschen herangezogen (vgl. 2.2.3).[33] Die international verbreitete Optimalitätstheorie (OT) wurde besonders auf Akzentstrukturen des Deutschen angewandt, so von FÉRY 1996, die die OT mit der Metrischen Phonologie und der Morenanalyse verbindet (vgl. 2.2.1).

Auf das Studium der Akzentverteilung im Deutschen erstreckt sich auch der eigenständige und m.E. vielversprechende theoretische Ansatz von EISENBERG (1991), den ich in 2.2.2 referiert habe. Von anderen Theorien ist vor allem noch die „Wortdesign-Theorie" (NEEF 1996, 1998) zu nennen, die ohne Annahme von zugrunde liegenden Strukturen auskommt und (phonologische und morphologische) Designbedingungen z.B. für Flexionsmuster des Deutschen formuliert.

Bibliographie

Abraham, Werner (Hg.) (1982): *Satzglieder im Deutschen*. Tübingen: Narr (= *SDG* 15).

Altmann, Hans / Batliner, Anton / Oppenrieder, Wilhelm (Hg.): Zur *Intonation von Modus und Fokus im Deutschen*. Tübingen: Niemeyer (= *LA* 234).

Archangeli, Diana (1997): Optimality Theory: An Introduction to Linguistics in the 1990s. In: Archangeli, D. / Langendoen, D.T. (Hg.), 1–32.

Archangeli, Diana / Langendoen, D. Terence (Hg.) (1997): *Optimality Theory. An Overview*. Oxford, Malden / MA: Blackwell.

Auer, Peter (1991): Zur More in der Phonologie. In: Ramers, K.H. / Wiese, R. (Hg.), 3–36.

Bannert, R. (1976): *Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage*. Lund: Gleerup.

Becker, Thomas (1995): *Das Vokalsystem der deutschen Standardsprache*. München: Universität (= Habilitationsschrift). Druckfassung 1998. Frankfurt am Main: Lang (= *Arbeiten zur Sprachanalyse* 32).

Berg, Thomas (1992): Umrisse einer psycholinguistischen Theorie der Silbe. In: Eisenberg, P. / Ramers, K.H. / Vater, H. (Hg.), 45–99.

Chomsky, Noam / Halle, Morris (1968): *The Sound Pattern of Engli*sh. New York: Harper & Row.

Clements, George N. / Keyzer, Samuel J. (1983): *A Generative Theory of the Syllable*. Cambridge/Mass.: MIT Press.

Durand, Jacques (Hg.) (1986): *Dependency and Non-Linear Phonology*. London: Croom Helm.

[32] Zur metrischen Akzentanalyse (inkl. Kritik an diesem Modell) vgl. auch RAMERS (1998), Kap.5 (speziell §5.3, 5.4), und das Kapitel „Phonologie" (speziell § 3.4.3) von RAMERS (2002) in MEIBAUER et al. (2002).

[33] Zu neueren Intonationsstudien vgl. auch ALTMANN / BATLINER / OPPENRIEDER (Hg.)(1989) und FÉRY (1992).

Eisenberg, Peter (1991): Syllabische Struktur und Wortakzent. In: Ramers K.H. / Wiese, R. (Hg.), 37–64.
Eisenberg, Peter / Ramers, Karl-Heinz / Vater, Heinz (Hg.) (1992): *Silbenphonologie des Deutschen*. Tübingen: Narr (= *SDG* 42).
Essen, Otto von (1956): *Grundzüge der hochdeutschen Satzintonation*. Ratingen / Düsseldorf: Henn.
Ezawa, Kennosuke (1972): *Die Opposition stimmhafter und stimmloser Verschlußlaute im Deutschen*. Tübingen: Narr (= *TBL* 29).
Féry, Caroline (1991): German schwa in prosodic morphology. In: Ramers, K.H. / Wiese, R. (Hg.), 65–85.
Féry, Caroline (1992): *Focus, topic, and intonation*. Stuttgart: Wiss. Zentrum der IBM GmbH.
Féry, Caroline (1996): German Foot and Word Stress in OT. *Nordlyd* 24, 63–96.
Giegerich, Heinz (1985): *Metrical phonology and phonological structure: German and English*. Cambridge: CUP.
Giegerich, Heinz (1992): Onset maximization in German. In: Eisenberg, P. / Ramers, K.H. / Vater, H. (Hg.), 134–171.
Goldsmith, John A. (Hg.) (1995): *The Handbook of Phonological Theory*. Oxford, Cambridge/MA: Blackwell.
Hall, Tracy A. (1992a): *Syllable Structure and Syllable-Related Processes*. Tübingen: Niemeyer (= *LA* 276).
Hall, Tracy A. (1992b): Syllable Final Clusters and Schwa Epenthesis in German. In: Eisenberg, P. / Ramers , K.H. / Vater, H. (Hg.), 208–245.
Halle, Morris (1992): Phonological Features. In: Bright, W. (Hg.): *International Encyclopedia of Linguistics*. Oxford: Oxf. Univ. Press, 207–212.
Hayes, Bruce (1989): Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.
Heike, Georg (1992): Zur Phonetik der Silbe. In: Eisenberg, P. / Ramers, K.H. / Vater, H. (Hg.), 1–44.
Höhle, Tilman (1978): Lexikalistische Syntax. *Die Aktiv-Passiv-Relation und andere Infinitkonstruktionen im Deutschen*. Tübingen: Niemeyer (= *LA* 67).
Höhle, Tilman (1982): Explikation für ‘normale Betonung’ und ‘normale Wortstellung’. In: Abraham, W. (Hg.), 75–340.
Isačenko, Alexander V. (1963): Der phonologische Status des velaren Nasals im Deutschen. *Zeitschrift für Phonetik* 16, 77–84. Wieder in: Steger, H. (ed.) (1970): *Vorschläge zu einer strukturalen Grammatik des Deutschen*. Darmstadt: Wiss. Buchgesellschaft (= *Wege der Forschung* 146), 468–479.
Jacobs, Joachim (1982): Neutraler und nicht-neutraler Satzakzent im Deutschen. In: Vennemann, T. (Hg.), 141–169.
Jørgensen, Hans Peter (1969): Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantfrequenzen. *Phonetica* 19, 217–245.

Kiparsky, Paul (1982): From cyclic phonology to lexical phonology. In: Van der Hulst, H. / Smith, N. (Hg.), 131–175.

Kloeke, Wus van Lessen (1982): *Deutsche Phonologie und Morphologie: Merkmale und Markiertheit*. Tübingen: Niemeyer (= *LA* 117).

Knetschke, E. / Sperlbaum, M. (1987): *Zur Orthoepie der Plosive in der deutschen Hochsprache. Eine auditiv-komparative Untersuchung*. Tübingen: Niemeyer (= *PHONAI. Lautbibliothek der deutschen Sprache* 33).

Kohler, Klaus / Künzel, H. (1978): The Temporal Organisation of Closing-Opening Movements for Sequences of Vowels and Plosives in German. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)* 10, 117–166.

König, Werner (1989): *Atlas zur Aussprache des Schriftdeutschen in der Bundesrepublik Deutschland*. Band 1: *Text*. Bd. 2*: Tabellen und Karten* (abgek.: *TK*): Ismaning: Hueber (= *Sprachen der Welt*).

König, Werner (1999): Das Diphthongsystem des Deutschen neu betrachtet. *Sprachwissenschaft* 24/1, 105–126.

Liberman, Mark / Prince, Alan (1977): On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249–336.

Meibauer, Jörg / Demske, Ulrike / Geilfuß-Wolfgang, Jochen / Pafel, Jürgen / Ramers, Karl Heinz / Rothweiler, Monika / Steinbach, Markus (2002): *Einführung in die germanistische Linguistik*. Stuttgart: Metzler.

Meinhold, Gottfried / Stock, Eberhard (1963): Stimmlosigkeit und Stimmhaftigkeit der Verschlußphase (Plosion) bei deutschen Medien im absoluten Anlaut und nach stimmlosen Lauten. *Zeitschrift für Phonetik* 16, 137–148.

Meinhold, Gottfried / Stock, Eberhard (1980, [2]1982): *Phonologie der deutschen Gegenwartssprache*. Leipzig: VEB Bibliographisches Institut.

Menzerath, Paul (1941): *Der Diphthong. Eine kritische und experimentelle Untersuchung*. Bonn: Dümmler (= *Phonetische Studien* 2).

Morciniec, Norbert (2002): Zur phonologischen Wertung der deutschen Affrikaten und Diphthonge. In: Cirko, L. / Kiedroń, St. (Hg.): *Norbert Morciniec. Studia Philologica. Ausgewählte Schriften zur Germanistik und Niederlandistik*. Wrocław: ATUT, 79–92.

Moulton, William (1962): *The Sounds of English and German*. Chicago: Chicago Univ. Press.

Neef, Martin (1996): *Wortdesign. Eine deklarative Analyse der deutschen Verbflexion*. Tübingen: Stauffenburg (= *SDG* 52).

Neef, Martin (1998): *Elemente einer deklarativen Wortgrammatik*. Hürth: Gabel (= *KLAGE* 32).

Noske, Roland (1992): Moraic vs. Constituent Syllables. In: Eisenberg, P. / Ramers, K.H./ Vater, H. (Hg.), 284–328.

Pierrehumbert, Jeannette (1987): *The phonology and phonetics of English intonation*. Bloomington, Ind.: Indiana University Linguistic Club.

Pike, Kenneth L. / Pike, Evelyn G. (1947): Immediate Constituents of Mazateco Syllables. *Internat. Journal of American Linguistics* 13, 78–91.

Pompino-Marschall, Bernd (1995): *Einführung in die Phonetik*. Berlin, New York: de Gruyter.
Ramers, Karl Heinz (1988): *Vokalquantität und -qualität im Deutschen*. Tübingen: Niemeyer (= *LA* 213).
Ramers, Karl Heinz (1992): Ambisilbische Konsonanten im Deutschen. In: Eisenberg, P. / Ramers, K.H. / Vater, H. (Hg.), 246–283.
Ramers, Karl Heinz (1998): *Einführung in die Phonologie*. München: Fink (= *UTB* 2008):
Ramers, Karl Heinz (2002): Phonologie. In: Meibauer, J. et al. (Hg.), 70–120.
Ramers, Karl Heinz / Vater, Heinz ([4]1995): *Einführung in die Phonologie*. Hürth: Gabel.
Ramers, Karl Heinz / Wiese, Richard (Hg.)(1991): *Prosodische Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
Reis, Marga (1974): *Lauttheorie und Lautgeschichte. Untersuchungen am Beispiel der Dehnungs- und Kürzungsvorgänge im Deutschen*. München: Fink.
Rubach, Jerzy (1990): Final Devoicing and Cyclic Syllabification in German. *Linguistic Inquiry* 21, 79–94.
Selkirk, Elizabeth O. (1982): The syllable. In: Van der Hulst, H. / Smith, N. (Hg.) II, 337–383.
Stampe, David (1973): *A dissertation on natural phonology*. (Diss. Chicago). Bloomington, Ind.: Indiana University Press / New York: Garland Press 1979.
Stark, J. (1974): Aphasiological evidence for the abstract analysis of the German velar nasal [N]. *Wiener Linguistische Gazette* 7, 21–37.
Stemberger, Joseph P. (1984): Length as a suprasegmental: Evidence from speech errors. *Language* 60, 895–913.
Trubetzkoy, Nikolai S. (1939): *Grundzüge der Phonologie*. Prag (= *TCLP* 7): (Wieder erschienen 1958 ([5]1971) bei Vandenhoeck & Ruprecht, Göttingen).
Uhmann, Susanne (1991): *Fokusphonologie. Eine Analyse deutscher Intonationsstrukturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer (= *L A* 252).
Ulbrich, Horst (1972): *Instrumentalphonetisch-auditive R-Untersuchungen im Deutschen*. Berlin: Akademie-Verlag (= *Schriften zur Phonetik, Sprachwissenschaft und Kommunikationsforschung* 13).
Van der Hulst, Harry / Smith, Norval (Hg.) (1982): *The structure of phonological representations* (Part 1 / Part II). Dordrecht: Foris.
Vater, Heinz (1992): Zum Silben-Nukleus im Deutschen. In: Eisenberg, P. / Ramers, K.H. / Vater, H. (Hg), 100–133.
Vater, Heinz (1994, [4]2002): *Einführung in die Sprachwissenschaft*. München: Fink (= *UTB* 1799).
Vater, Heinz (1998): Zur Silbenstruktur im Deutschen. In: Kröger, B. et al. (Hg.): *Festschrift für Georg Heike*. Frankfurt a. M. (= *Forum Phoneticum* 66), 137–149.

Vater, Heinz (2002): On the Syllable Structure of German. *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 7/2 (Fall 2002), 181–207.

Vater, Heinz (2004): Wie viele Diphthonge hat das Deutsche? *Estudios Filológicos Alemanes* 4, 7–21.

Vennemann, Theo (1970): The German Velar Nasal. A Case for Abstract Phonology. *Phonetica* 22, 65–81.

Vennemann, Theo (1972): On the theory of syllabic phonology. *LB* 18, 1–18.

Vennemann, Theo (1982): Zur Silbenstruktur der deutschen Standardsprache. In: Vennemann, Th. (Hg.), 261–305.

Vennemann, Theo (1988): *Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin.* Berlin: de Gruyter.

Vennemann, Theo (Hg.) (1982): *Silben, Segmente, Akzente*. Tübingen: Niemeyer (= *LA* 126).

Werner, Otmar (1972): *Phonemik des Deutschen*. Stuttgart: Metzler. (= *Sammlung Metzler* 108).

Wiese, Richard (1986): Zur Theorie der Silbe. *Studium Linguistik* 20, 1–15.

Wiese, Richard (1988): *Silbische und lexikalische Phonologie. Studien zum Chinesischen und Deutschen*. Tübingen: Niemeyer (= *LA* 211).

Wiese, Richard (1996): *The Phonology of German.* Oxford: Clarendon.

Wiese, Richard (Hg.) (1994): *Recent Developments in Lexical Phonology.* Düsseldorf: Heinrich-Heine-Universität (= *Theorie des Lexikons. Arbeiten des SFB 282.* Nr. 56).

Wurzel, Wolfgang U. (1970): *Studien zur deutschen Lautstruktur*. Berlin: Akademie-Verlag (= *SG* 8).

Wurzel, Wolfgang U. (1981): Phonologie: Segmentale Struktur. In. Heidolph, K. E. et al. (Hg.): *Grundzüge einer deutschen Grammatik*. Berlin: Akademie-Verlag, 898–990.

Yu, Si-Taek (1992a): *Unterspezifikation in der Phonologie des Deutschen*. Tübingen: Niemeyer (= *LA* 274).

Yu, Si-Taek (1992b): Silbeninitiale Cluster und Silbifizierung im Deutschen. In: Eisenberg, P. / Ramers, K.H. / Vater, H. (Hg.), 172–207.