

# Die Vokale des Gegenwartsdeutschen

*Eine generative Theorie des Silbenschnitts  
und der Reduktionssilbe im Deutschen*

Verfasser: Krisztián Tronka

Betreuer: Prof. Dr. Tamás Szende  
(Universitätsprofessor)

eingereicht

in der Werkstatt für Germanistik (Werkstattleiter: Dr. phil. Imre Szigeti)  
der Doktorandenschule für Linguistik und Literaturwissenschaft  
(Leiter der Schule: Dr. Miklós Maróth DSc.)  
der Katholischen Pázmány-Péter-Universität



Piliscsaba

2004

# A mai német nyelv magánhangzói

*A szótagmetszet és a redukált szótag  
generatív-fonológiai leírása*

Szerző: Tronka Krisztián

Témavezető: Prof. Dr. Szende Tamás  
(egyetemi tanár)

Pázmány Péter Katolikus Egyetem  
Nyelvészeti és Irodalomtudományi Doktori Iskola  
(vezető: Dr. Maróth Miklós DSc. egyetemi tanár)  
Germanisztikai Műhely  
(műhelyvezető: Dr. phil. Szigeti Imre, egyetemi docens)



Piliscsaba

2004

# Inhalt

<b>1. Einleitung</b> .....	<b>3</b>
<b>2. Die Vokale des Gegenwartsdeutschen – eine erste Annäherung</b> .....	<b>8</b>
<b>3. Der Silbenschnitt im Deutschen</b> .....	<b>14</b>
3.1 Vokalquantität und -qualität im Deutschen.....	15
3.1.1 <i>Quantität und Qualität aus phonetischer Sicht</i> .....	15
3.1.2 <i>Die zugrunde liegenden Vokalklassen</i> .....	26
3.1.3 <i>Die Frage nach der Distinktivität der Qualität und Quantität</i> .....	36
3.2 Zur Definition des Silbenschnitts.....	47
3.3 Phonetisches Korrelat des Silbenschnitts.....	51
3.4 Zur Repräsentation des Silbenschnitts .....	60
3.4.1 <i>Vennemann (1991b) und (1994)</i> .....	60
3.4.2 <i>Becker (1996b, 1998)</i> .....	63
3.4.3 <i>Maas (1999)</i> .....	66
3.4.4 <i>Lenerz (2000)</i> .....	67
3.4.5 <i>Kritik an den bestehenden Silbenschnittrepräsentationen</i> .....	69
3.4.6 <i>Silbenschnitt und die interne Struktur des Silbenkerns</i> .....	72
3.5 Silbenschnitt, Vokalquantität und Vokalqualität.....	77
3.6 Silbenschnitt und Nasalvokale.....	79
3.7 Silbenschnitt und Diphthonge.....	83
3.8 Zwischenbilanz.....	86
<b>4. Die Reduktionssilbe im Deutschen</b> .....	<b>87</b>
4.1 Das Schwa-Phänomen im Deutschen.....	87
4.1.1 <i>Phonetik des Schwa</i> .....	87
4.1.2 <i>Schwa und Akzent</i> .....	88
4.1.3 <i>Schwa und Silbenschnitt</i> .....	90
4.1.4 <i>Vorkommen des Schwa im Deutschen</i> .....	91
4.1.5 <i>Das Schwa-Phänomen: eine Zusammenfassung</i> .....	94
4.2 Die interne Struktur des Schwa .....	95
4.3 Frühere Beschreibungen zum Schwa im Deutschen.....	99
4.3.1 <i>Wiese (1988) und (1996)</i> .....	99

4.3.2	<i>Giegerich (1987)</i> .....	104
4.3.3	<i>Noske (1993)</i> .....	109
4.4	Zugrunde liegendes Schwa und Reduktionssilbe .....	113
4.5	Das lexikalische Schwa und die V-Epenthese .....	118
4.6	Postlexikalische Schwa-Prozesse .....	126
4.7	Schwa und <i>r</i> -Vokal .....	132
4.8	Zwischenbilanz .....	141
<b>5.</b>	<b>Zusammenfassung und Ausblick</b> .....	<b>144</b>
5.1	Das Vokalsystem des Gegenwartsdeutschen.....	144
5.2	Exkurs: Überlegungen zur Silbenstruktur im Deutschen.....	149
5.2.1	<i>Silbenstruktur und Silbifizierung im Deutschen</i> .....	150
5.2.2	<i>Silbenstruktur und Auslautverhärtung</i> .....	151
5.3	Ausblick: weiterführende Fragen.....	152
<b>Literatur</b>	.....	<b>153</b>

# 1. Einleitung

Die vorliegende Studie setzt sich zum Ziel, die Vokale des Gegenwartsdeutschen zu beschreiben. Für diese Beschreibung ist zweierlei charakteristisch. Erstens wird sie aus dem Blickwinkel zweier Konzepte durchgeführt, die zu den beiden meist diskutierten und daher aller Wahrscheinlichkeit nach größten Problemen des deutschen Vokalismus jeweils eine alternative Deskriptionsmöglichkeit bieten. So steht im Mittelpunkt des Silbenschnittkonzepts, das – wie Restles (1998) umfangreiche Untersuchung gezeigt hat – auf eine lange Geschichte zurückblickt, und zu dessen Verfechtern so große Persönlichkeiten wie Sievers, Jespersen, Trubetzkoy und neuerdings Vennemann zählen, die Frage nach dem primären phonologischen Unterscheidungsmerkmal zwischen den phonetisch als gespannten, unter Akzent langen sowie ungespannten kurzen Vokalen (vgl. etwa [i:] in *biete* vs. [ç] in *bitte*, [y:] in *fühlen* vs. [Y] in *füllen* oder [o:] in *Ofen* vs. [O] in *offen* usw.). Die Grundidee dieses Ansatzes besteht darin, dass für die Unterscheidung zwischen diesen beiden Vokalklassen entgegen den traditionellen Auffassungen weder die Quantität (markiert durch das Vorhandensein vs. Fehlen des Kolons in der obigen Auflistung) noch die Gespanntheit (markiert durch die unterschiedlichen IPA-Symbole in einem Paar in den oben gegebenen Beispielen) verantwortlich ist, vielmehr sind diese beiden deutlichen phonetischen Differenzen als bloße Folgeerscheinungen eines höheren prosodischen Kontrastes, nämlich des sog. Silbenschnittes zu betrachten. Da ungespannte Kurzvokale nur in geschlossenen Silben vorkommen, wobei die Schließung der Silbe sogar ‚virtuell‘, also durch Ambisilbizität erreicht wird, nimmt man an, dass die Vokalbildung hier durch eine vorgezogene Artikulation des postvokalischen Konsonanten ‚abgeschnitten‘ wird, während sie bei den gespannten Langvokalen von der Artikulation eines eventuellen postvokalischen Konsonanten nicht beeinflusst wird. So ist bei den gespannt-langen Vokalen über sanften Schnitt oder losen Anschluss, bei den ungespannt-kurzen dagegen über scharfen Schnitt oder festen Anschluss die Rede.

Den Kern des Konzeptes der Reduktionssilbe bildet der zweifelsohne eigenartigste Vokal des heutigen Deutsch: das Schwa, das so viele phonetische wie phonologische Besonderheiten (artikulatorische Minimalität, Vorkommen ausschließlich in unbetonter Position, Teilnahme an verschiedenen Alternationen: *Atem* vs. *atØmen*, *lecker* [ˈ] vs. *leckere* [WR], *bitten*: [Wn] vs. [n1] usw.) aufweist, dass man ihm mit Recht einen Sonderstatus im deutschen

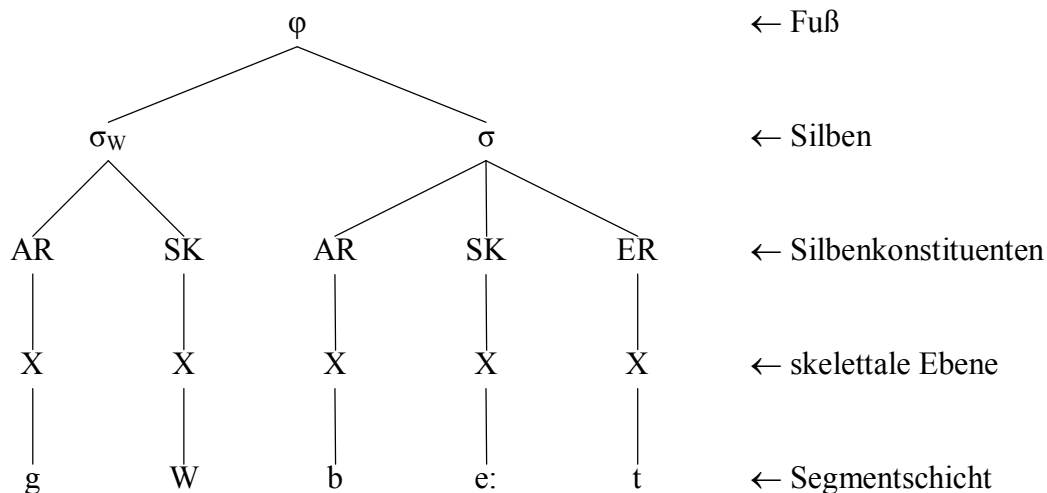
Vokalsystem zubilligen soll. Dieser besteht nach dem Konzept der Reduktionssilbe von Vennemann (1991a) darin, dass das Schwa nicht bloß eines der Vokalsegmente des Gegenwartsdeutschen darstellt, sondern vielmehr als Reflexion eines eigenen Silbentyps, nämlich der Reduktionssilbe zu verstehen ist, die sich gegenüber der Vollsilbe durch ihre generelle Unbetonbarkeit auszeichnet.

Die Verbindung dieser beiden Grundkonzepte ist durch zwei Punkte motiviert. Einerseits besteht guter Grund anzunehmen, dass das Schwa außerhalb des hier als Silbenschnittopposition analysierten Vokalgegensatzes des heutigen Deutsch steht, was sich – wie das gezeigt wird – durchaus mit der Annahme unterschiedlicher Strukturen für die Voll- und Reduktionssilben im Deutschen erfassen lässt. Somit stehen die beiden als Leitfäden der hier durchzuführenden Analyse gedachten Ansätze in einer ergänzenden Beziehung zueinander. Andererseits bieten die beiden Konzepte zusammen eine gute Grundlage für die Beschreibung des Gesamtsystems der deutschen Vokale, da sie zahlreiche als gewichtig erscheinende Fragen des deutschen Vokalismus direkt oder indirekt berühren – angefangen mit der inneren Beschaffenheit (Merkmalstruktur) der vokalischen Segmente des Gegenwartsdeutschen über die Nasalvokale und Diphthonge bis hin zum vokalischen *r*.

Das andere eingangs genannte Charakteristikum der vorliegenden Untersuchung betrifft die gewählte Rahmentheorie, als welche hier die generative Phonologie dienen wird. Da jedoch die generative Phonologie seit dem Erscheinen ihres Basiswerks von Chomsky & Halle (1968) zahlreiche Modifizierungen erfahren hat, und heute keine einheitliche Theorie darstellt, sondern vielmehr als Konglomerat einzelner partikulärer Ansätze zu verschiedenen Teilaspekten der Repräsentation und Derivation (oder neuerdings Evaluation im Sinne der Optimalitätstheorie) zu verstehen ist, scheint es relevant zu sein, schon vor der Durchführung der eigentlichen Untersuchung diejenigen Aspekte bzw. Ansätze anzusprechen, die in der vorliegenden Analyse berücksichtigt werden. So wird hier in Anlehnung an die repräsentationellen Erneuerungen des Grundmodells der generativen Phonologie von einer multilinearen phonologischen Repräsentation ausgegangen. Dies bedeutet einerseits die Annahme einer hierarchischen Merkmalstruktur (vgl. Clements 1985, 1987, Sagey 1986, McCarthy 1988, Clements & Hume 1995), in der die einzelnen phonologischen Merkmale zu größeren Klassen zusammengefasst werden, für welche Annahme im Vergleich zur herkömmlichen Auffassung von der Unstrukturiertheit der Merkmale zahlreiche Argumente (darunter der adäquate Ausdruck bestimmter phonologischer Prozesse – so der verschiedenen Arten der Assimilation) sprechen. Auf die hier vertretene merkmalgeometrische Repräsentation der deutschen Vokale werden wir bei der Behandlung der Minimalität des

Schwa eingehen. Andererseits wollen wir den gängigen Repräsentationsvorschlägen folgend auch oberhalb der segmentalen Ebene von mehreren Strukturen ausgehen, von denen für die vorliegende Untersuchung neben den Positionen auf der CV-Schicht (vgl. Clements & Keyser 1983), die sich bei der Repräsentation des Schwa als unentbehrlich erweisen wird, sowie dem metrischen Fuß, der bestimmte Schwa-Vorkommen steuert, besonders die Silbe als relevante Größe zu betrachten ist. Für das Deutsche wollen wir – ähnlich wie Hall (1992) – eine relativ einfache Silbenstruktur annehmen, in der der Silbenknoten drei Positionen dominiert: den Anfangsrand (AR), den Silbenkern (SK) und den Endrand (ER). Mit der Annahme dieser relativ flachen Silbenstruktur wird behauptet, dass der Reim als zusammenfassende Kategorie für den Silbenkern und den Endrand zumindest für die Phonologie der deutschen Vokale keine relevante Größe darstellt, m.a.W. für keines der von manchen Autoren mit ihm verbundenen Phänomene des deutschen Vokalismus, nämlich den Silbenschnitt bzw. die *r*-Vokalisierung als Domäne dient. Wichtig ist schon an dieser Stelle darauf hinzuweisen, dass diese relativ einfache Silbenstruktur hier für beide Silbentypen, also sowohl für die Voll- als auch für die Reduktionssilbe angenommen wird, und die oben angedeuteten strukturellen und auch repräsentationsmäßig erfassbaren Unterschiede innerhalb des Silbenkerns zu suchen sind. Somit ergibt sich für das Wort *Gebet* die prosodische Grobstruktur unter (1).

(1) Darstellung der prosodischen Grobstruktur des Wortes *Gebet*<sup>1</sup>



Schließlich sei noch die nicht mehr repräsentationelle, sondern die Derivation betreffende Erneuerung des Grundmodell der generativen Phonologie erwähnt, die besonders für die Beschreibung des Schwa als Rahmentheorie dienen wird, nämlich die von Kiparsky (1982)

<sup>1</sup> σ markiert eine Vollsilbe, σ<sub>w</sub> eine Reduktionssilbe.

und Mohanan (1982) entwickelte lexikalische Phonologie. Die Grundidee dieses Ansatzes besteht darin, dass gewisse phonologische Regeln nicht erst auf den Output der Syntax, sondern schon vorher, genau im Lexikon angewendet werden sollen, was eine Trennung zwischen einer lexikalischen und einer postsyntaktischen oder postlexikalischen Phonologie sinnvoll macht, wodurch die Phonologie nicht mehr mit zwei, sondern mit drei Repräsentationsebenen zu rechnen hat: einer zugrunde liegenden (markiert durch senkrechte Striche, vgl. |RA:t|), einer lexikalischen (gekennzeichnet mit Schrägstrichen, vgl. /RA:t/), sowie einer Oberflächenstruktur (markiert durch die eckigen Klammern, vgl. [RA:t]).

Zu den im Rahmen der vorliegenden phonologischen Untersuchung behandelten Themen, ja sogar zur Verbindung mancher dieser Themen und/oder Beschreibungsaspekte stehen in der phonetisch-phonologischen Literatur zum deutschen Vokalismus zahlreiche Aufsätze und Monographien zur Verfügung. Bei den beiden großen Fragestellungen außerhalb der Silbenschnitt- bzw. Reduktionssilbenphonologie sind u.a. besonders Moulton (1962), Reis (1974), Wurzel (1981), Kloeke (1982), Wiese (1988) und (1996), Giegerich (1987), Ramers (1988) und Hall (1992) zu nennen. Der Silbenschnitt wird speziell neben den mehr historischen Quellen von Sievers (1901), Jespersen (1932) und Trubetzkoy (1939) bzw. der großen Fülle phonetischer Arbeiten von Fischer-Jørgensen (1941) über von Essen (1962) bis hin zu Spiekermann (2000) in den neueren primär phonologisch konzipierten Werken Vennemann (1991ab, 1992 und 1994), Becker (1996ab, 1998 und 2002), Restle (1998), Maas (1999) und Lenerz (2000 und 2002) thematisiert. Von den zuletzt genannten Autoren wird von Lenerz der Versuch unternommen, das Konzept des Silbenschnitts in die generative Phonologie einzubetten, während Vennemann und Maas den Silbenschnitt mit der Reduktionssilbe verbinden, und Becker schließlich ausgehend von seinem in der strukturalistischen Phonologie wurzelnden Silbenschnittkonzept aus eine Beschreibung des Gesamtsystems der deutschen Vokale vornimmt. Das Novum der vorliegenden Arbeit im Vergleich zu diesen Quellen besteht

1. in der neuartigen Betrachtung a) des Silbenschnitts, b) der Beziehung zwischen Voll- und Reduktionssilbe und c) zwischen Silbenschnitt- und Quantitätensprachen, wobei im Mittelpunkt die Erkenntnis steht, dass der Unterschied sowohl zwischen den Voll- und Reduktionssilben als auch zwischen den Silbenschnitt- und Quantitätensprachen im Silbenkern zu lokalisieren ist;
2. in der Bestrebung, diese Erkenntnis auf eine konsequente Art und Weise in der generativen Phonologie zu implementieren, d.h. alle ihre relevanten repräsentationellen und derivationellen Bezüge voll auszuarbeiten; sowie



3. in der Bemühung, darauf beruhend und darüber hinausgehend eine Beschreibung des Gesamtsystems der Vokalsegmente des Gegenwartsdeutschen durchzuführen, d.h. auch solche Phänomene des deutschen Vokalismus zu beschreiben, die mit dem Silbenschnitt bzw. der Reduktionssilbe nicht unmittelbar zusammenhängen, so beim Silbenschnitt die Nasalvokale und Diphthonge, und bei der Reduktionssilbe die *r*-Vokalisierung(en) sowie die interne Struktur des Schwa und darüber hinausgehend aller anderen Vokale des heutigen Deutsch.

Die Arbeit ist wie folgt gegliedert. Kap. 2 bietet eine phonetisch orientierte Kurzbeschreibung der Vokalsegmente des Gegenwartsdeutschen. Kap. 3 ist dem Silbenschnitt gewidmet: Nachdem die fragliche Vokalopposition beschrieben wird, werden die beiden mehr traditionellen Konzepte zu ihrer phonologischen Erfassung dargestellt und mit Gegenargumenten widerlegt. Darauf folgt nun die eigentliche Beschreibung des Silbenschnittphänomens, in der die Definition, das phonetische Korrelat, die Repräsentation des Silbenschnitts thematisiert werden, bzw. auf die Beziehung des Silbenschnitts zu den Nasalvokalen und Diphthongen eingegangen wird. In Kap. 4 erfolgt die Besprechung des Konzepts der Reduktionssilbe. Nach der ausführlichen Beschreibung des Schwa-Phänomens werden die einzelnen repräsentationellen und derivationellen Aspekte des Schwa detailliert dargelegt. Kap. 5 dient nun als Zusammenfassung der Arbeit: Hier erfolgt eine Art Synthese der Ergebnisse der beiden vorausgehenden thematischen Kapitel, d.h. die Beschreibung des Gesamtsystems der deutschen Vokale. Aber auch an dieser Stelle wollen wir unsere weiterführenden Gedanken als Ausblick formulieren.

## 2. Die Vokale des Gegenwartsdeutschen – eine erste Annäherung

Im vorliegenden Kapitel wollen wir einen kurzen Überblick über die Vokale des Gegenwartsdeutschen bieten. Dabei wird es sich primär um eine phonetische Beschreibung handeln, d.h. wir werden versuchen, uns von jeder phonologischen Wertung zu enthalten. Auch werden längere phonetische Ausführungen bei dem einen oder anderen phonetischen Merkmal vermieden, um die Kürze der Beschreibung zu gewährleisten bzw. den Gegenstand der nachfolgenden thematischen Kapitel nicht vorwegzunehmen. Mit dieser Zusammenfassung der Vokalsegmente des Gegenwartsdeutschen verfolgen wir das Ziel, einen Einstieg in die thematischen Kapitel zu schaffen und zugleich die phonologische Beschreibung im letzten Kapitel vorzubereiten.

Traditionell wird behauptet, das Deutsche sei eine vokalreiche Sprache. Dass in dieser Aussage gewisse Wahrheit steckt, wird einem klar, wenn man einen Blick auf die Abbildung

rechts wirft, die die verschiedenen im Gegenwartsdeutschen anzutreffenden Vokalqualitäten enthält. Bei dieser Abbildung, dem sog. Vokalviereck oder Vokaltrapez handelt es sich um eine sowohl artikulatorisch als auch akustisch motivierte zweidimensionale Matrix. Die vertikale Achse entspricht

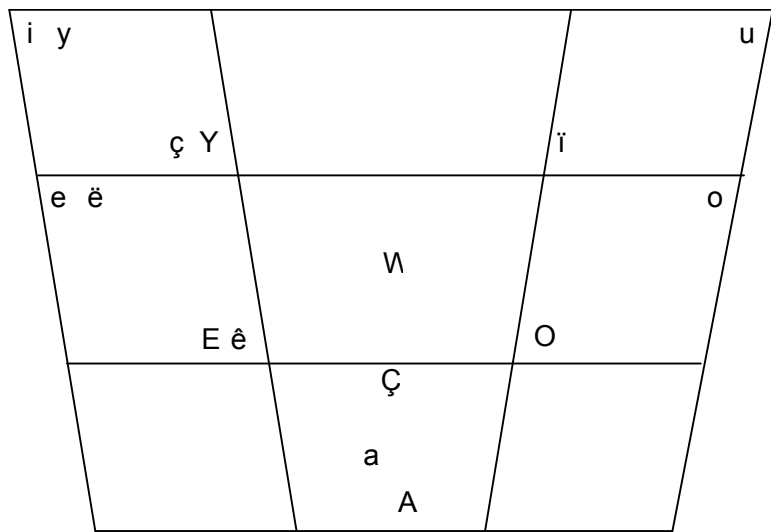


Abb.1 Die verschiedenen Vokalqualitäten des Deutschen

dabei der vertikalen, die horizontale der horizontalen Zungenbewegung: je höher ein Vokal im Artikulationsraum angeordnet ist, desto höher wird er im Mundraum artikuliert bzw. je weiter links sich ein Vokal im Vokalraum befindet, desto weiter vorne wird er gebildet. Die akustische Motivation des Vokalvierecks ergibt sich aus dem Zusammenspiel der für die Qualität der einzelnen Vokale verantwortlichen ersten beiden Formanten: je höher der Vokal im Vokalraum angeordnet ist, desto niedriger ist sein erster Formant ( $F_1$ ) bzw. je weiter links er platziert ist, desto höher ist sein zweiter Formant ( $F_2$ ).

Aus der schematischen Anordnung unter Abb.1 ist zu sehen, dass im Deutschen sowohl nach der vertikalen als auch nach der horizontalen Zungenbewegung jeweils drei Klassen angenommen werden: nach der horizontalen die hohen, mittleren und niedrigen, nach der vertikalen die vorderen, zentralen und hinteren Vokale. Zu diesen beiden Klassifizierungskriterien tritt noch die Lippenaktivität hinzu, nach der zwischen gerundeten und ungerundeten Vokalen unterschieden wird. Dabei sind alle zentralen Vokale des Deutschen ungerundet, und alle hinteren gerundet, während in der vorderen Vokalreihe von den paarweise angeordneten Vokalen der jeweils erste ein ungerundeter, der jeweils zweite dagegen ein gerundeter Vokal ist. Auf die phonologische Erfassung dieser qualitativen Unterschiede im deutschen Vokalsystem in Form von phonologischen Merkmalen wollen wir in Kap. 4 näher eingehen.

Im Vokalviereck in Abb.1 findet man zwei Vokale, die sich phonologisch von allen anderen abheben und zusammen oft als Reduktionsvokale bezeichnet werden: das Schwa, d.h. der mittlere ungerundete Zentralvokal [ʊ] sowie das vokalische *r* (auch *a*-Schwa genannt) [ɘ], das ebenfalls einen ungerundeten Zentralvokal darstellt, der sich jedoch nach der vertikalen Zungenbewegung zwischen [ʊ] und [a] befindet und in Abb.1 als niedriger Vokal eingestuft wird. Wichtig ist schon an dieser Stelle darauf hinzuweisen, dass das *a*-Schwa in Abhängigkeit davon, ob es als eigener Silbenträger fungiert oder nicht, in zwei Ausprägungen vorkommen kann: es kann silbisch (nuklear) oder unsilbisch (postnuklear) sein, vgl. (2). Die phonetischen und phonologischen Besonderheiten der Reduktionsvokale werden in Kap. 4 behandelt.

(2) Die Reduktionsvokale des Deutschen

- a. Schwa: [ʊ] *Stelle, (ich) beende, Gebirge*
- b. vokalisches *r*
  - silbisch: [ɘ] *Vater, Mutter, Schwester, Bruder*
  - unsilbisch: [ɘ&] *Tier, Tür, Teer, Tor, Gebühr, Bär, Bar*

Betrachtet man die übrigen Vokale in Abb.1, die man im Gegensatz zu den Reduktionsvokalen traditionell auch als Vollvokale bezeichnet, so ist zu sehen, dass sie die aufgrund der vertikalen und horizontalen Zungenbewegung sowie der Lippenaktivität ermittelten Vokalklassen paarweise belegen (vgl. Tab. 1), wobei die einzelnen Paare oft mit den für ihre Glieder charakteristisch(st)en Buchstaben bezeichnet werden (vgl. Tab. 2). Das vielleicht bekannteste Konzept zur Erfassung des Unterschieds zwischen den beiden Vokalen

dieser Paare geht von der Gespanntheit der Vokalartikulation aus und bezeichnet das jeweils erste Glied als gespannten, das jeweils zweite als ungespannten Vokal. Etwas komplizierter wird diese auf den ersten Blick als klarer segmentaler (qualitativer) Kontrast erscheinende Opposition dadurch, dass die Gespanntheitsdifferenzen (zumindest unter Akzent) mit Unterschieden in der relativen Vokaldauer korrelieren: gespannte Vokale sind lang – ungespannte kurz. Dieser Vokalgegensatz stellt den Gegenstand des Silbenschnittkonzepts dar, das – wie gesagt – von der Ablehnung der Annahme eines qualitativen (Gespanntheits-) bzw. quantitativen Kontrasts ausgeht.

	<i>vorn</i>		<i>zentral</i>	<i>hinten</i>
	<i>ungerundet</i>	<i>gerundet</i>	<i>ungerundet</i>	<i>gerundet</i>
<i>hoch</i>	<b>i</b> <b>ç</b>	<b>y</b> <b>Y</b>		<b>u</b> <b>ï</b>
<i>mittel</i>	<b>e</b> <b>E</b>	<b>ë</b> <b>ê</b>		<b>o</b> <b>O</b>
<i>niedrig</i>			<b>A</b> <b>a</b>	

Tab. 1 Klassen der deutschen Vollvokale nach den Zungenbewegungen und der Lippenaktivität I

	<i>vorn</i>		<i>zentral</i>	<i>hinten</i>
	<i>ungerundet</i>	<i>gerundet</i>	<i>ungerundet</i>	<i>gerundet</i>
<i>hoch</i>	<b>i-Vokale</b>	<b>ü-Vokale</b>		<b>u-Vokale</b>
<i>mittel</i>	<b>e-Vokale</b>	<b>ö-Vokale</b>		<b>o-Vokale</b>
<i>niedrig</i>			<b>a-Vokale</b>	

Tab. 2 Klassen der deutschen Vollvokale nach den Zungenbewegungen und der Lippenaktivität II

Doch scheint auch für diese – aus den eingangs genannten Gründen – etwas vereinfachte Beschreibung der Vokale des Gegenwartsdeutschen die Erwähnung gewisser (eher die Schnittstelle zwischen Phonetik und Phonologie betreffender) Probleme dieser Opposition schon an dieser Stelle angebracht, da sie darauf hinweisen, dass sich dieser Vokalgegensatz wegen seiner Komplexität von allen anderen Vokaloppositionen des Deutschen unterscheidet. So ist – auch wenn manche Autoren (s. weiter unten) das Gegenteil annehmen – davon auszugehen, dass Längenunterschiede nur auf betonte Silben beschränkt sind, während unter Unbetontheit nur Gespanntheitsdifferenzen bestehen. Dies macht den Fall insofern

komplexer, als phonetisch eigentlich nicht mit zwei, sondern mit drei Klassen zu rechnen ist, neben den oben genannten beiden nämlich auch mit der Klasse der gespannten Kurzvokale. Ob diese auch eine phonologisch relevante Vokalklasse darstellen, muss selbstredend untersucht werden. Noch komplizierter wird die Beschreibung des Phänomens dadurch, dass – wie traditionell angenommen wird – der ungespannte mittlere ungerundete Vordervokal [E] in bestimmten Wörtern wie generell die gespannten Vokale unter Akzent lang gesprochen wird, wodurch die Anzahl der phonetischen Vokalklassen auf 4 erhöht wird (vgl. Tab. 3 und die Auflistung mit Beispielen unter (3)), was bedeutet, dass – zumindest phonetisch und nach der ersten Annäherung – im Deutschen alle möglichen Kombinationen der Gespanntheits- und Quantitätswerte belegt sind. Dieser exzeptionelle Vokal soll daher ebenfalls Gegenstand einer gesonderten phonologischen und phonetischen Untersuchung sein, die – wie die Behandlung der fraglichen Opposition selbst – in Kap. 3 erfolgen wird.

	<i>vorn</i>				<i>zentral</i>				<i>hinten</i>					
	<i>ungerundet</i>		<i>gerundet</i>		<i>ungerundet</i>		<i>gerundet</i>		<i>ungerundet</i>		<i>gerundet</i>			
<i>hoch</i>	<b>i:</b>	<b>i</b>	<b>ç</b>		<b>y</b>	<b>y</b>	<b>Y</b>					<b>u</b>	<b>u</b>	<b>ï</b>
<i>mittel</i>	<b>e:</b>	<b>e</b>	<b>E</b>	<b>E</b>	<b>ë:</b>	<b>ë</b>	<b>ê</b>					<b>o</b>	<b>o</b>	<b>O</b>
<i>niedrig</i>								<b>A</b>	<b>A</b>	<b>a</b>				
	<i>lang</i>	<i>kurz</i>	<i>kurz</i>	<i>lang</i>	<i>lang</i>	<i>kurz</i>	<i>kurz</i>	<i>lang</i>	<i>lang</i>	<i>kurz</i>	<i>kurz</i>	<i>lang</i>	<i>lang</i>	<i>kurz</i>
	<i>gespannt</i>	<i>ungespannt</i>			<i>gespannt</i>	<i>ungespannt</i>			<i>gespannt</i>	<i>ungespannt</i>			<i>gespannt</i>	<i>ungespannt</i>

Tab. 3 Phonetische Klassifizierung der deutschen Vollvokale

(3) Die Oralvokale des Deutschen

- a. *i*-Vokale: [i:] *bi*ete, [ç] *bit*te, [i] *Bit*umen
- b. *ü*-Vokale: [y:] *Fü*ßen, [Y] *Fü*ssen, [y] *Phys*ik
- c. *e*-Vokale: [e:] *Be*et, [E] *Be*tt, [e] *Be*ton, [E:] *Dä*mon
- d. *ö*-Vokale: [ë:] *Hö*hle, [ê] *Hö*lle, [ë] *Ö*konom
- e. *a*-Vokale: [A:] *Ba*hn, [a] *Ba*nn, [A] *ba*nal
- f. *o*-Vokale: [o:] *Mo*l, [O] *Mo*ll, [o] *Mo*lekül
- g. *u*-Vokale: [u:] *Stu*dium, [ï] *Stu*ttgart, [u] *Stu*dent

Im Deutschen findet man in Wörtern französischer Herkunft außer den bisher ermittelten Vokalen einige weitere, die mit gesenktem Velum gebildet werden und daher als nasalierte oder kurz Nasalvokale bezeichnet werden. Diese Vokale stellen nach den beiden Zungenbewegungen sowie der Lippenaktivität im Vergleich zu den übrigen Vollvokalen keine eigenständigen Vokalqualitäten dar (aus diesem Grund werden sie in Abb.1 nicht verzeichnet). Aus der Auflistung unter (4) geht hervor, erstens dass im Deutschen alle nicht-hohen ungespannten Vollvokale auch nasaliert vorkommen können, und zweitens dass sie sich trotz ihrer ungespannten Artikulation ähnlich wie die gespannten Oralvokale verhalten, d.h. unter Akzent lang realisiert werden. Auf die Probleme der Nasalvokale wird in Kap. 3 näher eingegangen.

(4) Die Nasalvokale im Deutschen

a. nasaliertes *e*

lang: [E\$:] *Pointe, Timbre, Teint*

kurz: [E\$] *Impromptu, Saint-Simonist*

b. nasaliertes *ö*

lang: [ê\$:] *Parfum*

c. nasaliertes *o*

lang: [O\$:] *Bonbon, Kupon, Aplomb*

kurz: [O\$] *Bonbon, Ombré, Jongleur*

d. nasaliertes *a*

lang: [a\$:] *Cancan, Ensemble, Pendant*

kurz: [a\$] *Cancan, Ensemble, Pendant*

Schließlich kommen im Deutschen auch Diphthonge, d.h. zur selben Silbe gehörende Vokal-Vokal-Verbindungen vor. In Anlehnung an Becker (1998: 117ff) lassen sich Diphthonge nach der Kombination der Kriterien der Schallfülle und der Artikulationsbewegung in drei Klassen einteilen, wobei im Deutschen alle drei Klassen belegt sind. So gehören zu den schließenden Diphthongen (d.h. solchen mit einer Gleitbewegung des Zungenkörpers nach oben und einer Abnahme der Druckstärke vom ersten zum zweiten Teilvokal) [au&], [ai&] und [Oi&], zu den steigenden Diphthongen (d.h. solchen mit einer Abnahme der Druckstärke vom ersten zum zweiten Teilvokal und keiner Gleitbewegung des Zungenkörpers nach oben) die Diphthonge mit [i] bzw. [y] als erstem und einem beliebigen Vokal als zweitem Teilvokal und schließlich zu den öffnenden Diphthongen (also solchen mit

einer Gleitbewegung der Zunge nach unten und gleichzeitig keiner Zunahme der Druckstärke vom ersten zum zweiten Vokal) die Diphthonge mit einem beliebigen Vokal als erstem und dem (unsilbischen) vokalischen *r* [ʧ&] als zweitem Teilvokal, vgl. (5). Von diesen drei Diphthongtypen heben sich die schließenden Diphthonge insofern ab, als in ihnen die Vokalteile fest miteinander verbunden sind, weshalb sie durchaus als Diphthonge im engeren Sinn bezeichnet werden können.

(5) Die Diphthonge im Deutschen

a. schließende Diphthonge: ‚feste‘ Diphthonge, Diphthonge im engeren Sinn<sup>2</sup>

[au&]	<i>Auto, braun, kaum</i>
[ai&]	<i>nein, Schein, Mai, Mayer</i>
[Oy&]	<i>neun, Scheune, heute, Boiler</i>

b. steigende Diphthonge: *i-/u*-Diphthonge<sup>3</sup>

[i&+V]	<i>Folie, partiell, Meridian, Nation</i>
[u&+V]	<i>Linguist, virtuell, tendenziös, Guasch</i>

c. öffnende Diphthonge: *r*-Diphthonge

[V:+ʧ&]	<i>Bier, Tür, Heer, Öhr, stur, Tor</i>
---------	--

Diphthonge können jedoch auch nach dem Kriterium klassifiziert werden (vgl. etwa Pompino-Marschall 1995: 118f), ob sich die beiden Teilvokale auf zugrunde liegende Vokalsegmente (Vokalphoneme) oder auf die Verbindung eines Vokals mit einem Konsonanten zurückführen lassen: Erstere bezeichnet man primäre oder phonologische, Letztere sekundäre oder phonetische Diphthonge. Nach diesem Kriterium sind die schließenden und steigenden Diphthonge in (5)a–b phonologische, die öffnenden in (5)c jedoch phonetische Diphthonge.

---

<sup>2</sup> Außerdem können die nur in den Anglizismen *Lady*, *Baby* sowie *Show* vorkommenden beiden quasi-Fremddiphthonge [ei&] bzw. [ou&] hierher gerechnet werden.

<sup>3</sup> In manchen Wörtern kommen nach den Aussprachewörterbüchern auch steigende Diphthonge mit den Randvokalen [y&] und [o&] vor, vgl. *Etui* [y&i], *Libyen* [y&E], *Erinnyen* [y&E], *Toilette* [o&a], *Memoiren* [o&a] usw. Jedoch sind diese Diphthonge nach Becker (1998: 119) zum Teil unaussprechbar und werden von den Sprechern durch andere Formen ersetzt (z.B. *Etui* mit [u&i]/[vi], *Libyen* mit [i&E], *Memoiren* mit [u&a]), auf der anderen Seite ist [o&a] von [oʧ&] (etwa in *Tor*) kaum zu unterscheiden und sollte daher eher als öffnender Diphthong analysiert werden. Aus diesem Grund sind die Halbvokale [y&] und [o&] als erster Teil von steigenden Diphthongen aus dem System auszugrenzen.

### 3. Der Silbenschnitt im Deutschen

Im Mittelpunkt dieses Kapitels der Arbeit steht die folgende Vokalopposition:

(6)	a.	b.	Beispiele
	i:	ç	<i>Miete</i> vs. <i>Mitte</i>
	y:	Y	<i>fühlen</i> vs. <i>füllen</i>
	e:	E	<i>Weg</i> vs. <i>weg</i>
	ë:	ê	<i>Öfen</i> vs. <i>öffnen</i>
	A:	a	<i>Staat</i> vs. <i>Stadt</i>
	o:	O	<i>Ofen</i> vs. <i>offen</i>
	u:	ï	<i>Ruhm</i> vs. <i>Rum</i>

Wie die aufgelisteten Minimalpaare zeigen, ist die Wahl zwischen den gespannten Lang- und ungespannten Kurzvokalen im Deutschen phonologisch relevant: in einer funktionalistischen Sichtweise wirkt sie bedeutungsunterscheidend, generativ-phonologisch gesehen stellen die beiden Vokale einer Reihe jeweils eine idiosynkratische Information dar, d.h. sie können aus dem phonologischen Kontext nicht hergeleitet werden und müssen daher als Teil der zugrunde liegenden Repräsentation betrachtet werden. Aufgrund der Korrelation zwischen Gespanntheit und Länge stellt sich die Frage nach dem primären phonologischen Unterscheidungsmerkmal zwischen den Vokalklassen unter (6), auf welche Frage die beiden traditionellen Auffassungen zwei mögliche Antworten bieten: nach dem einen Konzept stellt die Gespanntheit, nach dem anderen die Quantität die zugrunde liegende phonologische Information dar. Der Silbeschnittansatz geht dagegen davon aus, dass in dieser Opposition weder die Gespanntheit noch die Quantität relevant sind, sondern beide aus den Silbeschnittarten hergeleitet werden können. Das vorliegende Kapitel setzt sich zum Ziel, diesen Ansatz eingehend vorzustellen.

Das Kapitel ist wie folgt gegliedert. Im Mittelpunkt von Abschn. 3.1 stehen die phonetischen und phonologischen Bezüge der Gespanntheit (Qualität) und Vokaldauer (Quantität). Dabei werden in einem ersten Schritt diese beiden Phänomene phonetisch beschrieben, um dann unter Berücksichtigung phonologischer Gegebenheiten die in Opposition stehenden beiden zugrunde liegenden Vokalklassen herzuleiten, sowie die oben genannten beiden traditionellen Konzepte vorzustellen und mit Gegenargumenten zu



widerlegen. In Abschn. 3.2 soll ausgehend von den in der einschlägigen Literatur auffindbaren Definitionen der Begriff des Silbenschnitts bestimmt werden. Abschn. 3.3 ist anschließend der zentralen Frage nach dem phonetischen Korrelat des Silbenschnitts im Deutschen gewidmet: es werden die Ergebnisse verschiedener experimentalphonetischer Untersuchungen zu diesem Thema kurz zusammengefasst, wobei einer dieser Untersuchungen, nämlich Spiekermann (2000) mehr Aufmerksamkeit geschenkt wird. In Abschn. 3.4 wollen wir dann ein anderes gewichtiges Problem in diesem Zusammenhang besprechen: die Frage nach einer angemessenen Repräsentation der beiden Silbenschnitte im Deutschen, die auch deswegen interessant ist, weil sie zugleich auch eine Reflexion auf die Struktur der Silbe im Deutschen erfordert. Auch hier wollen wir von den zahlreichen in der einschlägigen Literatur vorhandenen Darstellungsvorschlägen ausgehen, um durch Hinweise auf deren Schwachstellen dann unseren eigenen Repräsentationsvorschlag im Detail zu erörtern. Abschn. 3.5 soll das Verhältnis des Silbenschnitts zur Gespanntheit und Quantität behandeln, in Abschn. 3.6 und 3.7 kommen wir dann auf die Beziehung zwischen dem Silbenschnitt und den Nasalvokalen bzw. den schließenden und steigenden Diphthongen des Deutschen zu sprechen. In Abschn. 3.8 werden schließlich die Ergebnisse des Kapitels kurz zusammengefasst.

## **3.1 Vokalquantität und -qualität im Deutschen**

### **3.1.1 Quantität und Qualität aus phonetischer Sicht**

Dauermessungen zum deutschen Vokalismus zufolge unterscheiden sich die Vokale (6)a und (6)b in der relativen Vokaldauer. Dabei werden die ersteren in etwa doppelt so lang artikuliert wie die letzteren. So kann zwischen den beiden Vokaltypen bei den Versuchspersonen von Fischer-Jørgensen (1969) im Durchschnitt ein Verhältnis von 1:1,92, bei denen von Ramers (1988) eines von 1:2,08 und schließlich bei denen von Jessen et al. (1995) eines von 1:1,8 festgestellt werden. Dieser klare Dauerunterschied trifft auf Vokale in betonter Stellung zweifelsohne zu, was die traditionelle Bezeichnung ‚Langvokale‘ für (6)a und ‚Kurzvokale‘ für (6)b als durchaus gerechtfertigt erscheinen lässt. Doch herrscht in der Fachliteratur keine Übereinstimmung darüber, ob die Längenverhältnisse der Vokale auch unter Unbetontheit in irgendeiner Form weiter bestehen.

Während sich z.B. Moulton (1962) für eine Neutralisation der Dauerunterschiede in unbetonten Silben einsetzt (vgl. 63), vertreten bspw. Meinhold & Stock (1982) die entgegengesetzte Position. Ausgehend von den Ergebnissen Maacks (1951) experimentalphonetischer Untersuchung zu den Lang- und Kurzvokalen des Deutschen in Abhängigkeit vom Akzent stellen sie unter Akzentlosigkeit zwar in der Tat eine Kürzung der Langvokale fest, jedoch besteht auch in unbetonter Stellung eine deutliche Dauerdifferenz zwischen den beiden Vokaltypen, u.z. im Verhältnis von 1:1,43<sup>4</sup> (vgl. Meinhold & Stock 1982: 90) – ein Ergebnis, das ihrer Meinung nach gegen die Stichhaltigkeit des Moulton'schen Schlusses von der Aufhebung der Dauerunterschiede in unbetonter Position spricht:

Die Berücksichtigung dieser Sachverhalte erlaubt nicht den Schluß, den Moulton (1962) aus einer angeblichen Aufhebung des Gegensatzes kurz – lang unter Akzentlosigkeit zieht: Es bleibe lediglich der Gegensatz *gespannt – ungespannt*. Hier wird der Sachverhalt in einer Weise vereinfacht, daß die tatsächlichen Verhältnisse dadurch verdeckt werden: Die mögliche Kürzung von Langvokalen bei Akzentlosigkeit führt in sehr vielen Fällen zu einer völligen Gleichheit mit der Dauer der Kurzvokale. In anderen Fällen kommt es zu einer Verringerung der Dauer („halbe Länge“), oder es bleibt bei der vollen Länge. Diese Möglichkeiten müssen in Betracht gezogen werden, doch eben nur als Möglichkeiten. (ebda)<sup>5</sup>

Ramers (1988), der Maacks Messwerte<sup>6</sup> ebenfalls einer vergleichenden Analyse unterwirft, kommt zunächst auf ein noch deutlicheres Ergebnis: das Verhältnis der Kurz- und Langvokale unter Akzentlosigkeit beträgt 1:1,54<sup>7</sup> (vgl. 81). Durch Ausschluss der überdurchschnittlich langen Diphthonge kann er jedoch einen wesentlich niedrigeren Kurz-Lang-Quotienten in

---

<sup>4</sup> Die Autoren geben nicht die Langvokaldauer im Vergleich zur Kurzvokaldauer an, sondern umgekehrt: die Dauer der Kurzvokale im Vergleich zu der ihrer langen Pendanten: 1:0,7. Die Umrechnung wurde durchgeführt, um einen Vergleich der einzelnen Quellen zu ermöglichen.

<sup>5</sup> Vgl. auch Meyer (1906: 24): „Die gespannten Vokale kommen betont nur lang, die ungespannten Vokale betont nur kurz vor. Unbetont sind auch die gespannten Vokale halblang oder kurz.“

<sup>6</sup> Ramers (1988) ging dabei im Gegensatz zu Meinhold & Stock (1982) nicht von Maacks Aufsatz aus dem Jahre 1951, sondern von dem aus 1949 aus, vermerkt jedoch, dass die beiden Materialien bis auf einige Ausnahmen identisch sind (vgl. Ramers 1988: ebda).

<sup>7</sup> Auch Ramers gibt die Dauer der Kurzvokale im Vergleich zu der ihrer langen Pendanten an: 1:0,65.

unbetonter Stellung, nämlich ein Verhältnis von nur 1:1,28<sup>8</sup> feststellen, dessen perzeptive Relevanz jedoch in Frage zu stellen sei (vgl. ebda).

Die aufgrund von Maack (1949) bzw. (1951) durch Meinhold & Stock (1982) und Ramers (1988) errechneten Werte sowie die darauf beruhende Schlussfolgerung von Meinhold & Stock (1982) scheinen in den Ergebnissen der Untersuchung von Delattre & Hohenberg (1981) auf den ersten Blick eine Bestätigung zu finden. Die beiden Autoren gehen in ihrem Aufsatz der Frage nach, ob die Dauer zur Unterscheidung zwischen den gespannten und ungespannten Vokalen des Deutschen in unbetonten Silben beiträgt. Ihre Untersuchung liefert eine positive Antwort auf diese Frage: zwischen den unbetonten Kurz- und Langvokalen können sie im Durchschnitt ein Verhältnis von 1:1,54 feststellen. Nach Becker (1998) besitzt dieses Ergebnis jedoch keine Aussagekraft dafür, dass die lang-kurz-Unterscheidung im Deutschen auch in unbetonter Stellung aufrechterhalten wird. Aus Fußnote 6 von Delattre & Hohenberg (1981) geht nämlich zweierlei hervor: einerseits unterscheiden die beiden Autoren zwischen 4 Akzentstufen: dem Hauptakzent (Akzentstufe 4), dem Nebenantwort (Akzentstufe 3), der unbetonten Position (Akzentstufe 2) und den Schwa-Silben (Akzentstufe 1); andererseits verstehen sie unter ‚betont‘ Akzentstufe 4, also ‚hauptbetont‘, unter ‚unbetont‘ dagegen Akzentstufe 3 und 2, also ‚nicht hauptbetont‘.<sup>9</sup> Becker kommt bei einer näheren Betrachtung ihres Untersuchungsmaterials zu der Erkenntnis, dass die Autoren die Dauerverhältnisse in drei Kontrasten untersuchen: erstens im Kontrast zwischen morphologischem Nebenantwort und unbetonter Silbe (Stufe 3 vs. Stufe 2), zweitens im Kontrast zwischen gespanntem und ungespanntem Vokal unter morphologischem Nebenantwort (Stufe 3 vs. Stufe 3) sowie drittens im Kontrast zwischen gespanntem Vokal in offener unbetonter Silbe und ungespanntem in geschlossener unbetonter Silbe (Stufe 2 vs. Stufe 2) (vgl. Becker 1998: 82). Nach ihm könnte das Ergebnis der Untersuchung Delattre & Hohenbergs (1981) von der Tatsache stark beeinflusst worden sein, dass die Autoren keine Unterscheidung zwischen ‚nebantwort‘ und ‚unbetont‘ vornahmen. So hätten sie höchstens die Relevanz des morphologischen Nebenantworts für die lang-kurz-Unterscheidung nachgewiesen, dass also diese Vokalopposition auch unter morphologischem Nebenantwort erhalten bleibt (vgl. Becker 1998: 83).

---

<sup>8</sup> Auch Ramers gibt die Dauer der Kurzvokale im Vergleich zu der ihrer langen Pendanten an: 1:0,78.

<sup>9</sup> Sie schließen aus ihrer Untersuchung die Schwa-Silben aus.

Angesichts dieser Kritik Beckers sind die Ergebnisse der Untersuchung Delattre & Hohenbergs zwar nicht in der Lage, die z.B. von Meinhold & Stock (1982) vertretene Position zu bestätigen. Sie machen jedoch darauf aufmerksam, dass zur Klärung dieser Frage eine Untersuchung notwendig ist, in der die Dauer der fraglichen Vokalpaare in unbetonter Stellung (d.h. auf Akzentstufe 2 auf Delattre & Hohenbergs Skala) gemessen wird. Jessen et al. (1995) bieten eine solche Messung. Die Autoren gehen bei der Suche nach den akustischen Korrelaten des Wortakzents und der Gespanntheitsopposition im Vokalsystem des Deutschen wie folgt vor: sie nehmen quasi-Minimalpaare mit gespannten und ungespannten Vokalen im Lautkontext [t<sup>h</sup>\_l]: *Ventil* vs. *Tormentill*, *Klientel* vs. *Kartell* usw., fügen diesen die Ableitungssuffixe *-isch* und *-ist* hinzu, wodurch sich die fraglichen Vokale beim ersteren Suffix in einer hauptbetonten (Stufe 4), beim letzteren dagegen in einer unbetonten Silbe (Stufe 2) befinden. In den auf diese Weise gewonnenen und von 10 Versuchspersonen vorgelesenen Wörtern werden dann zahlreiche akustische Parameter, darunter auch die Vokaldauer gemessen. Während die Autoren für die betonten Vokalpaare (6)a–(6)b ein Verhältnis von 1:1,8 feststellen konnten, ergab sich für dieselben Vokale unter Akzentlosigkeit lediglich ein Verhältnis von 1:1,07, das Moultons Annahme bestätigt: Vokallänge ist im Deutschen in der Tat akzentbedingt.

Im vorausgehenden Kapitel wurde zur Erfassung der qualitativen Differenzen zwischen den Vokalen in (6)a vs. (6)b als das vielleicht bekannteste Konzept die Gespanntheit genannt, das auf Sievers zurückgeht, der für den qualitativen Unterschied zwischen den beiden Vokalklassen den unterschiedlichen Grad der Muskelanspannung aller an der Artikulation beteiligten Organe als Grund angibt:

Beim langen *i\_*, *e\_* fühlt man bei einiger Aufmerksamkeit leicht, wie die Zunge zumal in dem articulirenden Vordertheil straff angespannt ist; geht man dann zu *i-*, *e-* über, so wird sie schlaffer und sinkt gewissermassen in sich zusammen. [...] Bei genauerem Aufmerken findet man übrigens leicht, dass der hier geschilderte Spannungsunterschied sich nicht auf die Zunge allein beschränkt, sondern in analoger Weise bei allen an der Lautbildung beteiligten Organen wiederkehrt, welche überhaupt verschiedene Spannungsgrade gestatten. Dies gilt bei der Vocalbildung einmal von den Lippen, bei deutlicher Rundung [...] oder spaltförmiger Erweiterung [...], sodann aber namentlich auch von den Stimmbändern. (Sievers 1901: 98)

Dementsprechend unterscheidet der Autor zwischen *gespannten* und *ungespannten* Vokalen (vgl. 99).

Sievers' Gespanntheitsmerkmal hat sich in der phonetischen und phonologischen Theorie weitgehend durchgesetzt. Viele Autoren (vgl. Moulton 1962 oder Jørgensen 1969a) erblicken

im Spannungsgradunterschied ein primäres Unterscheidungsmerkmal und betrachten andere qualitative Unterschiede (z.B. die Zentralisiertheit/Nicht-Zentralisiertheit – s. unten) als bloße Folgeerscheinungen der Gespanntheit. Jedoch erscheint das Gespanntheitsmerkmal nicht als unproblematisch, da zur Zeit Sievers' an entsprechender technischer Ausrüstung zur objektiven Messung der Artikulationsintensität fehlte, wodurch die von ihm angenommenen Gespanntheitsunterschiede höchstens als Ergebnis indirekter und/oder introspektiver Beobachtungen angesehen werden können.<sup>10</sup> Über die Frage, ob zwischen den ‚gespannten‘ und ‚ungespannten‘ Vokalen in der Tat solche Spannungsunterschiede bestehen, könnte nur eine elektromyographische Untersuchung Aufschluss geben, die eine Messung des elektrischen Potentials der Muskeln ermöglicht. Solche Untersuchungen gibt es aber leider nur zum Englischen und „die Übertragung auf die deutsche Opposition ist natürlich problematisch“ (Becker 1998: 44f.).<sup>11</sup> Somit kann das Sieverssche Gespanntheitskonzept im Deutschen – auch wenn es m.E. intuitiv einleuchtend erscheint – als experimentell bis heute unbestätigt erachtet werden: die Kategorien ‚gespannt‘–‚ungespannt‘ sind als „rein deskriptiv“ zu verstehen (Pompino-Marschall 1995: 217).

Ein viel zuverlässigeres, relativ leicht messbares phonetisches Korrelat zum qualitativen Unterschied zwischen den beiden Vokalklassen unter (6) bietet das Konzept der Zentralisierung, nach dem die ‚ungespannten‘ Vokale im Vergleich zu den ‚gespannten‘ nahe

---

<sup>10</sup> Davon zeugt auch die von ihm vorgeschlagene ‘Technik’ zur Feststellung des Spannungsgrades durch Betasten der „vorn zwischen den beiden Unterkieferknochen eingebetteten Weichteile“ mit zwei Fingern (98).

<sup>11</sup> Außerdem muss man vor Augen halten, dass diese Untersuchungen nicht einmal für das Englische eine eindeutige Korrelation zwischen Artikulationsintensität und ‚Gespanntheit‘ nachweisen konnten. So werden zwar die Ergebnisse MacNeilage & Sholes' (1964) von einigen Autoren (vgl. MacKay 1977: 329 oder Jørgensen 1969a: 243f.) als die Bestätigung der These vom größeren Spannungsgrad bei den gespannten Vokalen als bei den ungespannten interpretiert; Ramers wendet jedoch gegen diese Untersuchung ein, dass die Autoren „überhaupt keine Zweiteilung der Vokale in die Gruppen ‘tense’ und ‘lax’ vor[nehmen]“ (vgl. Ramers 1988: 129f.), wodurch ihre Analyse nicht zur Rechtfertigung der These des Gespanntheitsunterschieds im englischen Vokalismus verwendet werden kann. Die Untersuchungen von Smith & Hirano (1968), Raphael & Bell-Berti (1975) sowie Alfonso & Baer (1982) haben eine größere EMG-Aktivität des *Genioglossus posterior* (des v.a. für die Vorwärtsbewegung der Zungenmasse verantwortlichen Muskels) bei den gespannten Vokalen gefunden als bei den ungespannten, aber die umfangreichste und ausschließlich den Gespanntheitsverhältnissen im Englischen gewidmete Untersuchung von Raphael & Bell-Berti (1975) konnte bei 4 weiteren Muskeln keinen oder manchmal einen negativen Zusammenhang feststellen.

der Indifferenzlage des Artikulationsraumes gebildet werden. Auch dieses Konzept blickt auf eine lange Geschichte zurück: Die zentralisierte Bildung der ‚ungespannten‘ Vokale wurde – wie darauf Jørgensen (1969a) hinweist – bereits von Techmer (1884) bemerkt. Seine Erkenntnis lässt sich auf den Röntgenfilmbildzeichnungen von Wängler (1964), Valaczkai (1981) bzw. auf den Nachzeichnungen der kineradiographischen Aufnahmen von Bolla & Valaczkai (1986) bestätigen: In der Tat ist der höchste Punkt des Zungenkörpers bei den vorderen nichtniedrigen Vokalen mehr hinten und unten, bei den hinteren mehr vorne und unten, bzw. auch bei den niedrigen *a*-Vokalen ist – neben einer mehr vorderen Zungenaufwölbung – eine (wenn auch schwache) höhere Zungenposition des ‚ungespannten‘ [a] im Vergleich zum ‚gespannten‘ [A] zu beobachten. Um diesen subjektiven Eindruck zu erhärten, wurden auf den Nachzeichnungen eigene Messungen durchgeführt. Nach einer Digitalisierung der Abbildungen wurden sie auf ein einheitliches Format gebracht: durch Verkleinerung oder Vergrößerung wurde erreicht, dass der Abstand zwischen der Spitze der oberen Schneidezähne und der ihm gegenüber liegenden Rachenwand (als zwei mehr oder weniger unbeweglichen ‚konstanten‘ Größen) auf allen Abbildungen der gleiche (nämlich 80 mm) wird. Anschließend wurden 1. die Entfernung der Stelle der größten Zungenaufwölbung von der Spitze der oberen Schneidezähne (für die horizontale Zentralisierung), 2. die Entfernung des höchsten Punktes der Zungenmasse von dem genau ihm gegenüber liegenden Punkt des Gaumens (für die vertikale Zentralisierung) gemessen, und 3. die Verschiebung von der ‚gespannten‘ zur ‚ungespannten‘ Vokalartikulation, d.h. die Differenz zwischen dem Abstand der höchsten Zungenstelle der ‚ungespannten‘ und dem der höchsten Zungenstelle der ‚gespannten‘ Vokale sowie 4. die Durchschnittswerte dieser horizontalen bzw. vertikalen Verschiebung für die einzelnen Vokalqualitäten errechnet. Dabei ließen sich folgende Durchschnittswerte errechnen:

	<i>i-Vokale</i>	<i>ü-Vokale</i>	<i>e-Vokale</i>	<i>ö-Vokale</i>	<i>u-Vokale</i>	<i>o-Vokale</i>	<i>a-Vokale</i>
<i>Vertikal</i>	3,33	5	6,66	2	8,66	5,66	-3,5
<i>Horizontal</i>	7,33	3	12,66	2	-4,33	-7,33	-3,66

Tab. 4 Durchschnittswerte der Vokalzentralisierung

Die Messwerte bestätigen bis auf einige Fälle<sup>12</sup> den obigen subjektiven Eindruck: Im Bereich der nichtniedrigen vorderen Vokale ließ sich beim ‚ungespannten‘ Laut eine jeweils positive Verschiebung sowohl in vertikaler als auch in horizontaler Richtung feststellen. Beim ‚ungespannten‘ Segment der nichtniedrigen hinteren Vokalreihe war in der vertikalen Dimension eine positive, in der horizontalen dagegen eine negative Verschiebung zu beobachten. Und schließlich konnte zwischen dem ‚ungespannten‘ und ‚gespannten‘ *a* eine jeweils negative horizontale und vertikale Verschiebung ermittelt werden. Die ‚ungespannten‘ Vokale sind also in der Tat zentralisiert, d.h. sie werden näher der Indifferenzlage des Artikulationsraumes gebildet. Das genaue Ausmaß der Zentralisierung ist anhand der gemessenen Daten jedoch nicht präzise festzustellen. Dies ist einerseits wegen der Ungenauigkeit der Nachzeichnungen so: selbst die ‚konstanten‘, d.h. unbeweglichen Sprechwerkzeuge im Mundraum, so die oberen Schneidezähne oder der Hartgaumen, nehmen auf den verschiedenen Nachzeichnungen desselben Sprechers unterschiedliche Gestalten auf, was die Feststellung der Abweichungen von nur einigen Millimetern ziemlich erschwert. Andererseits liefert auch die hier angewendete Messmethode keine zuverlässigen Werte: so ist beispielsweise die Festlegung des höchsten Punktes der Zungenmasse oft (besonders bei einem relativ flachen Zungenrücken) der Willkür der analysierenden Person ausgesetzt.

Zur Bestimmung der vertikalen Zungenposition bei der Artikulation deutscher Vokale kann die glossometrische Untersuchung von Bohn et al. (1992) aufschlussreich sein. Hier wurde mit Hilfe einer opto-elektrischen Vorrichtung, des Glossometers der Abstand zwischen Zunge und Hartgaumen gemessen.<sup>13</sup> Die Mittelwerte der einzelnen Messwerte sowie die

---

<sup>12</sup> Während die Messwerte bei Wänglers (1964) *ö*-Nachzeichnungen (d.h. eine 2 mm lange negative – anstatt einer positiven – Verschiebung in der vertikalen Dimension) den subjektiven optischen Eindruck (höhere Zungenstelle beim ‚ungespannten‘ [ê] als beim ‚gespannten‘ [ë]) und dadurch den wirklichen exzeptionellen Status dieses Vokalpaars in Wänglers Nachzeichnungen bestätigt, widersprechen die errechneten Daten bei Valaczkais (1981) *a*-Nachzeichnungen (d.h. eine 2 mm lange positive – anstatt einer negativen – vertikalen Verschiebung) der sichtbaren größeren Zungenwölbung beim ‚gespannten‘ [A] als beim ‚ungespannten‘ [a]. Diese Diskrepanz ist jedoch eine Folge der unzuverlässigen Nachzeichnungs- und Messtechnik (beim [A] zeigt sich die der Zungenwölbung entgegenliegende Gaumenstelle konvex, beim [a] ist sie dagegen konkav, wodurch die Messwerte verzerrt werden).

<sup>13</sup> Dabei wurde der Versuchsperson, einem Sprecher des Nordddeutschen (möglicherweise dem deutschen Koautor Ocke-Schwen Bohn), ein 0,3 mm dünner künstlicher Gaumen mit vier Leuchtdioden und Fotosensoren in der Mittellinie zwischen Zahndamm und Weichgaumen aufgesetzt. Die Leuchtdioden strahlten infrarotes Licht aus, das von der Zungenoberfläche reflektiert wurde, dieser reflektierte Strahl

daraus errechneten Unterschiede zwischen den gespannten und ungespannten Vokalen (vgl. Tab. 5) sowie die Durchschnittswerte für die einzelnen Vokalklassen (vgl. Tab. 6) bestätigen das Konzept der vertikalen Zentralisierung: Für die nichtniedrigen ungespannten Vokale konnte bei den einzelnen Sensoren<sup>14</sup> im Vergleich zu ihren gespannten Pendanten ein höherer, für die niedrigen *a*-Vokale dagegen ein niedrigerer Wert gemessen werden, was im Bereich der nichtniedrigen Vokale auf eine mehr gesenkte, im Bereich der niedrigen *a*-Vokale dagegen eine mehr gehobene Zungenmasse bei der Artikulation der ungespannten Glieder der Vokalpaare hinweist. Da alle diese Differenzen im Mittelwert zumeist über 1,0 mm liegen (vgl. Tab. 5), welcher Unterschied nach den Autoren für die Distinktion zwischen zwei Vokalqualitäten ausreicht (vgl. 8)<sup>15</sup>, sind sie als signifikant, d.h. auch als perzeptiv relevant zu betrachten. Anzumerken ist jedoch in Bezug auf diese Untersuchung, dass sie das Konzept der vertikalen Zentralisierung im Großen zwar bestätigt, jedoch kaum zu einer Verallgemeinerung geeignet ist, da ihr die Aussprache eines einzigen Sprechers zugrunde liegt (vgl. 24).

<i>i</i> -Vokale	<i>ü</i> -Vokale	<i>e</i> -Vokale	<i>ö</i> -Vokale	<i>u</i> -Vokale	<i>o</i> -Vokale	<i>a</i> -Vokale
------------------	------------------	------------------	------------------	------------------	------------------	------------------

---

wurde dann von den Fotosensoren aufgefangen. Gemessen wurde die Intensität des von der Zungenoberfläche reflektierten und von den Fotosensoren aufgefangenen infraroten Lichtstrahls, woraus das System den Abstand zwischen dem Fotosensor und der reflektierenden Zungenoberfläche darunter errechnete. Für jeden Vokal wurden jeweils 10 Realisationen aufgenommen, aus den einzeln gemessenen Abstandswerten wurden dann für jeden Vokal über die vier Sensoren der Mittelwert und die Standardabweichung errechnet.

<sup>14</sup> Dies gilt für alle Vokalklassen bis auf die *ö*-Vokale, zwischen denen nur minimale Differenzen und bei den hinteren beiden Sensoren ein negativer Zusammenhang festzustellen sind, was die aufgrund der Nachzeichnungen oben errechneten Messwerte zwar bestätigt, den unten noch zu besprechenden  $F_1$ -Werten jedoch widerspricht. Hier handelt es sich möglicherweise entweder um eine sprecherspezifische Artikulation oder um dialektale Einflüsse.

<sup>15</sup> Diese Annahme der Autoren wird einerseits durch die Ergebnisse der Untersuchung von Flege (1988) begründet, nach denen Zungenpositionsunterschiede von etwa 1,0 mm die Vokalidentifizierung beeinflussen können. Andererseits lagen die von ihnen errechneten Standardabweichungen der Zungenabstandsmittelwerte zumeist auch unter diesem Wert, bzw. haben ihre vorläufigen Untersuchungen ergeben, dass Sprecher durchaus in der Lage sind, Zielpositionen mit Abweichungen zwischen 0,5 und 1,0 mm nachzuahmen (vgl. ebda).



S <sub>1</sub>	4	2,6	3,2	0,7	3	2	-3,2
S <sub>2</sub>	11,3	8,9	8,2	0,9	3,3	1,4	-2,2
S <sub>3</sub>	10,1	8,2	7,6	-0,5	5,9	2,7	-2,5
S <sub>4</sub>	3,9	2,9	3,8	-2,8	9,3	6,4	-2,4

Tab. 5 Zungenhöhendifferenzen der ungespannt-gespannten Glieder bei den einzelnen Vokalpaaren

	<i>nichtniedrige Vordervokale</i>	<i>nichtniedrige Hintervokale</i>	<i>a-Vokale</i>
S <sub>1</sub>	2,625	2,5	-3,2
S <sub>2</sub>	7,325	2,35	-2,2
S <sub>3</sub>	6,35	4,3	-2,5
S <sub>4</sub>	1,95	7,85	-2,4

Tab. 6 Durchschnittliche Zungenhöhendifferenzen der ungespannt-gespannten Glieder bei den Vokalklassen

Eine viel zuverlässigere Bestätigung für das Zentralisierungskonzept bieten jedoch die Ergebnisse akustischer Untersuchungen zum deutschen Vokalismus. Aufgrund der eingangs genannten Korrelation zwischen Akustik und Artikulation sollten die nichtniedrigen ungespannten Vordervokale höhere F<sub>1</sub>- und niedrigere F<sub>2</sub>-, die nichtniedrigen ungespannten Hintervokale höhere F<sub>1</sub>- und F<sub>2</sub>-, die niedrigen ungespannten Vokale dagegen zumindest niedrigere F<sub>1</sub>-Werte haben. Und genau davon zeugen die akustischen Analysen der deutschen Vokale von Rausch (1972), Jørgensen (1969a), Narahara & Shimoda (1977) und Ramers (1988) – um nur einige zu nennen. Anhand der von diesen Autoren festgestellten Mittelwerte der Formantfrequenzen der untersuchten Sprecher wurden nach der Formel  $(F_u - F_g) : F_g * 100$  (wo **F<sub>u</sub>** und **F<sub>g</sub>** für die Formantenfrequenz der ungespannten bzw. gespannten Vokale steht) die Werte der prozentualen Verschiebung, d.h. der prozentualen Verhältnisse der Unterschiede zwischen den Formantfrequenzmittelwerten der beiden Glieder der einzelnen Vokalpaare errechnet, die daraus ermittelten Durchschnittswerte der prozentualen

Verschiebung für die einzelnen Vokalpaare sind in Tab. 7, für die nichtniedrigen Vordervokale, die nichtniedrigen Hintervokale sowie die *a*-Vokale in Tab. 8 enthalten.

	<i>i</i> -Vokale	<i>ü</i> -Vokale	<i>e</i> -Vokale	<i>ö</i> -Vokale	<i>u</i> -Vokale	<i>o</i> -Vokale	<i>a</i> -Vokale
F <sub>1</sub>	32,99%	30,05%	50,60%	40,93%	33,68%	45,54%	-6,89%
F <sub>2</sub>	-10,92%	-18,11%	-12,19%	-5,72%	39,06%	37,45%	4,69%

Tab. 7 Durchschnittswerte der prozentualen Verschiebung (Vokalpaare)<sup>16</sup>

	<i>nichtniedrige Vordervokale</i>	<i>nichtniedrige Hintervokale</i>	<i>a</i> -Vokale
F <sub>1</sub>	38,64%	39,61%	-6,89%
F <sub>2</sub>	-11,73%	38,25%	4,69%

Tab. 8 Durchschnittswerte der prozentualen Verschiebung (Vokalklassen)

Die tabellarischen Übersichten legen folgende Interpretation nahe: Die nichtniedrigen vorderen Vokale weisen in der F<sub>1</sub>-Dimension eine prägnante positive, in der F<sub>2</sub>-Dimension eine weniger prägnante negative prozentuale Verschiebung auf. Bei den nichtniedrigen hinteren Vokalen ist dagegen eine prägnante positive prozentuale Verschiebung von den gespannten zu den ungespannten Vokalen sowohl in der F<sub>1</sub>- als auch in der F<sub>2</sub>-Relation zu beobachten. Schließlich ist bei den niedrigen *a*-Vokalen lediglich eine – im Vergleich zu den anderen beiden großen Vokalklassen – sehr geringe prozentuale Verschiebung in der F<sub>1</sub>-Dimension in negative, in der F<sub>2</sub>-Dimension in positive Richtung.<sup>17</sup> Diese Daten bestätigen in der Tat die Auffassung, dass es sich bei den ungespannten Vokalen im Vergleich zu ihren gespannten Pendanten auch im akustischen Sinne um „eine Annäherung an die ‘Indifferenzlage’“ (Jørgensen 1969a: 241) handelt.<sup>18</sup>

<sup>16</sup> Bei der Errechnung der Durchschnittswerte wurden die von der Zentralisierungstendenz abweichenden Einzelwerte nicht berücksichtigt, da sie das Gesamtbild verzerrt hätten.

<sup>17</sup> Zu den phonologischen Wertungsmöglichkeiten dieser Abweichung der Intensität der prozentualen Verschiebung zwischen den nichtniedrigen und niedrigen Vokalen kommen wir weiter unten noch zu sprechen zurück.

<sup>18</sup> Eine ähnliche Konklusion findet sich auch z.B. bei Fliflet (1962a) oder Wodarcz & Wodarcz-Magdics (1971).

Viele Autoren (u.a. Jakobson, Fant & Halle 1951, Moulton 1962, Jørgensen 1969a) führen die artikulatorische (und dadurch auch die akustische) Zentralisierung der ungespannten Vokale auf den Gespanntheitsunterschied zwischen den beiden Vokalklassen zurück. Der Zusammenhang beruht auf einer relativ einfachen Prämisse: Je kleiner die Artikulationsenergie – desto geringer kann sich die Zungenmasse von ihrer Ruhelage heraus entfernen. Bei den gespannten, also intensiver artikulierten Vokalen steht relativ viel Energie zur Verfügung, was eine relativ große Entfernung des Zungenkörpers von der neutralen Position ermöglicht. Ungespannte Vokale werden dagegen weniger intensiv, quasi ‚energiearm‘ artikuliert, was also eine geringere Entfernung der Zungenmasse von ihrer Ruheposition zur Folge hat. Auch wenn dieser Zusammenhang jedoch als intuitiv einleuchtend erscheint, kann er als reine Spekulation angesehen werden, solange – wie bereits oben gezeigt – der Unterschied im Spannungsgrad zwischen den beiden Vokalklassen des Deutschen experimentell nicht nachgewiesen wird. Aus diesem Grund werden die beiden Termini ‚gespannt‘ und ‚ungespannt‘ in der vorliegenden Arbeit als deskriptive Kategorien, nämlich als Synonyme zu den phonetisch untermauerten Begriffen ‚zentralisiert‘ und ‚nicht zentralisiert‘ verwendet.<sup>19</sup>

Schließlich sei angemerkt, dass neben diesen beiden Konzepten in der Fachliteratur auch weitere Vorschläge zur Erfassung des qualitativen Unterschieds zwischen den beiden Vokalklassen in (6) existieren, die sich jedoch – wie aus dem umfangreichen kritischen Überblick über die einschlägige phonetische und phonologische Literatur von Ramers (1988, Abschn. 2.4) hervorgeht – kaum als mögliche phonetische Korrelate des fraglichen qualitativen Unterschieds eignen. So erweist sich die Stimmlippenspannung oder die daraus resultierende Stärke des Luftdrucks, wie sie etwa von Meyer (1913) und Schuhmacher (1972)

---

<sup>19</sup> Wir gehen hier anders vor als Wurzel (1981: 911ff), der Zentralisiertheit und Gespanntheit als zwei verschiedene (oder z.T. verschiedene) Phänomene betrachtet, von denen er Ersterem den Status eines binären phonologischen Merkmals im nichtnativen Teil des deutschen Vokalsystems, Letzterem dagegen den einer bloßen phonetischen Folgeerscheinung der Quantität und Zentralisiertheit zubilligt. Diese Trennung der Gespanntheit und Zentralisiertheit beruht – wie Ramers (1988: 136f) anmerkt – auf einem auf Wängler (1960) zurückgehenden Missverständnis, nämlich auf der Gleichsetzung der Gespanntheit mit der Geschlossenheit der Artikulation (d.h. der vertikalen Zungenstellung): die Umkehrung der Geschlossenheitswerte bei den *a*-Vokalen wird automatisch als Umkehrung der Gespanntheitswerte betrachtet, [a] wird dadurch gespannt, [A:] ungespannt. Gegen diese Wertung spricht nach Ramers (1988: 137), dass bei [A:] die Zungenmasse weit mehr von der Indifferenzlage des Artikulationsraumes entfernt ist, was eine größere Deformation des Ansatzrohrs und folgerichtig einen größeren Spannungsgrad bedeutet.

vorgeschlagen wird, deswegen als ungeeignet, weil sie möglicherweise einen von der Silbendauer abhängigen Faktor darstellt (vgl. Fischer-Jørgensen 1969: 149). Von dem Merkmal ATR (*Advanced Tongue Root*, d.h. ‚vorgeschobene Zungenwurzel‘), das die gespannt-ungespannt-Unterscheidung auf die unterschiedliche Positionierung der Zungenwurzel zurückführt (bei Gespanntheit vorgeschoben, bei Ungespanntheit zurückgezogen), stellt Ramers (1988: 147) in Anlehnung u.a. an Lindau (1978) dagegen fest, es könne deswegen nicht auf die europäischen Sprachen, spezifisch auf das Englische oder Deutsche angewendet werden, weil es im Englischen oder Deutschen im Gegensatz zu den westafrikanischen Sprachen mit Vokalharmonie, für deren Beschreibung es zuerst angewendet wurde, keinen von der Zungenhöhe unabhängigen Parameter darstelle. Das Merkmal ATR ist auch deswegen sehr interessant, da es oft in merkmalsgeometrischen Beschreibungen vorkommt, so z.B. in Hall (1992) oder Wiese (1996). Doch hängt die Wahl in diesen Fällen weniger mit phonetischen Befunden bezüglich dieses Merkmals im Deutschen zusammen, sondern vielmehr mit theorieinternen Gründen. Die Integration von [±gespannt] und [±zentralisiert] in den Strukturbaum wäre kaum möglich, beim ersteren Merkmal wegen seines kaum definierbaren phonetischen Inhalts, beim letzteren dagegen der Komplexität seines phonetischen Korrelats. Bei [±ATR] ergeben sich dagegen keine Schwierigkeiten: es kann unter dem Knoten für den Artikulationsort mühelos untergebracht werden, da es an einen bestimmten Artikulator (Zungenwurzel) gebunden ist. Diese vermutliche Zwiespalt um die Einordnung von [±ATR] sehe ich spezifisch bei Wiese (1996) bestätigt. So begründet der Autor die Wahl genau dieses Merkmals damit, dass es den fraglichen Kontrast wie die anderen Merkmale durch eine spezifische artikulatorische Geste ausdrücke (vgl. 20). An einer späteren Stelle gibt er zu, dass sein Vorschlag zur Anbringung des ATR-Merkmals im Strukturbaum auch von der Korrektheit der Annahme abhängig sei, dass in der Tat dieses Merkmal für den Kontrast zwischen den gespannten und ungespannten Vokalen verantwortlich ist (vgl. 32). All diese Probleme erübrigen sich jedoch, wenn man den qualitativen Unterschieden zwischen den beiden Vokalreihen in (6) jede phonologische Relevanz abstreitet. In 3.5 wird eine Möglichkeit skizziert, wie das erreicht werden könnte.

### 3.1.2 Die zugrunde liegenden Vokalklassen

Im vorausgehenden Kapitel haben wir gesehen, dass im Deutschen nach den beiden Gespanntheits- und Vokaldauerwerten phonetisch sogar vier Vokalklassen angenommen werden könnten (s. die Tabelle rechts), von denen phonologisch jedoch lediglich zwei als relevant zu betrachten sind. Wie diese Reduktion der Anzahl der zugrunde liegenden Vokalklassen zu erfolgen

	<i>gespannt</i>	<i>ungespannt</i>	<i>gespannt</i>	<i>ungespannt</i>
	<i>lang</i>	<i>kurz</i>	<i>kurz</i>	<i>lang</i>
<i>i</i> -Vokale	ï:	ç	i	
<i>ü</i> -Vokale	y:	Y	y	
<i>e</i> -Vokale	e:	E	e	E:
<i>ö</i> -Vokale	ë:	ê	ë	
<i>a</i> -Vokale	Ä:	a	Ä	
<i>o</i> -Vokale	o:	O	o	
<i>u</i> -Vokale	u:	ö	u	

Tab. 9 Maximal belegte phonetische Vokalklassen im Deutschen

hat, ist Gegenstand des vorliegenden Abschnitts.

Wie aus Tab. 9 hervorgeht und wie bereits oben angesprochen erscheint die Klasse der ungespannten Langvokale am problematischsten, da sie lediglich einen einzigen Vokal, nämlich den mittleren vorderen ungerundeten ungespannten Langvokal [E:] (wie in *Dämon*, *Käse* usw.) enthält und dadurch die Ökonomie des Systems stört. Daher wäre eine Beschreibung, die mit weniger Klassen und daher Klassifizierungskriterien auskommt, jedoch phonetisch motiviert und auch phonologisch begründbar ist, jedenfalls vorzuziehen.

Neben diesem Intergrationsproblem weist das lange offene [E:] jedoch auch andere Probleme auf: So wird ihm oft vorgeworfen, es habe eine ‚illegitime Herkunft‘ es komme außerdem nicht auf dem ganzen deutschen Sprachgebiet vor bzw. es werde von vielen Sprechern nur in einem informellen Register verwendet. Ob dieser Vokal wirklich einen ‚Fremdkörper‘ im deutschen Vokalsystem darstellt, ist unklar: Während viele Autoren die Auffassung vertreten, es sei ein Vokal ohne phonologische Geschichte im gewöhnlichen Sinne (vgl. Moulton 1961: 35 bzw. 1962: 69), und seine Existenz könne nur auf den etymologisierenden Orthographiereform im 17. Jahrhundert sowie die schriftnachahmende Aussprache zurückgeführt werden (vgl. Reis 1974: 178), sind andere der Ansicht, bei der Entstehung dieses Segments hätten nachweislich auch dialektale Einflüsse eine Rolle gespielt,

was gegen seine illegitime Herkunft spreche (vgl. Sanders 1972: 58).<sup>20</sup> Was seine Verbreitung bzw. seine Bindung an ein bestimmtes Register betrifft, so wird oft davon ausgegangen, dass es eher auf den südlichen Teil des deutschen Sprachraums beschränkt ist und im nördlichen durch [e:] ersetzt wird, sowie dass es auf der anderen Seite im formalen häufiger als im informalen Register der Sprecher vorkommt (vgl. Moulton 1962: 69). Diese Auffassung wird mehr oder weniger von Königs (1989) umfangreicher Untersuchung bestätigt: [E:] wird zwar nicht generell in Norddeutschland, doch auf einem großen Gebiet des nördlichen Teils der ehemaligen Bundesrepublik durch [e:] ersetzt, aber selbst da kommt es im formalen Register größtenteils vor, das Gegenteil gilt nur für eine nicht gerade bedeutsame Gegend innerhalb dieses großen Gebietes. Die Abhängigkeit der E:-Verwendung von dem Grad der Formalität könnte nahelegen, dass es sich bei dieser Aussprache um Bildungslautung handelt, d.h. um eine Aussprache die von sprachexternen Faktoren, wie Bildungsgrad, Situation usw. abhängt. Während jedoch die Bildungslaute (-allophone) in einer strukturalistischen Annäherung funktionslos sind, wirkt die Wahl zwischen e: und E: bedeutungsunterscheidend (vgl. Becker 1998: 20), wie folgende Minimalpaare zeigen:

- (7) *Beeren* vs. *Bären*  
*dehnen* vs. *Dänen*  
*gebe* vs. *gübe*  
*lese* vs. *läse*  
*sehe* vs. *sähe*

Während die genannten Probleme um diesen Vokal manche Autoren (so z.B. Szulc 1966) dazu veranlassen, ihn aus dem Vokalsystem des Gegenwartsdeutschen auszugrenzen, gewähren ihm viele einen Sonderstatus im deutschen Vokalsystem, der größtenteils mit der oben genannten Einordnung dieses Segments als ‚ungespannter mittlerer Langvokal‘ identisch ist. Doch zieht diese Systemintegration von [E:] schwerwiegende Konsequenzen nach sich. Unabhängig davon, ob im Deutschen ein Quantitäts-, Qualitäts- oder Silbenschnittkontrast angenommen wird, muss nämlich zur Unterscheidung dieses zugrunde liegenden Segments ein eigenes Merkmal eingeführt werden. So werden in Meinhold & Stock (1982: 82) [ $\pm$ lang] und [ $\pm$ gespannt], in Wiese (1996: 21) ein prosodischer Längenkontrast und

---

<sup>20</sup> Zur E:-Diachronie vgl. auch Hinderling (1978).

[±ATR] und in Kloeke (1982: 11ff) [±tense] und ein abstraktes diakritisches Längenmerkmal [±L] angenommen.

Neben dieser Ansicht der exzeptionellen Einordnung des langen [E:] wird von manchen Autoren (so z.B. von Wurzel 1981: 911 oder Becker 1998: 15ff) ein Harmonisierungsversuch unternommen, bei dem das betreffende Segment nicht als mittlerer, sondern als niedriger Vokal aufgefasst wird. Geht man nämlich von einer gesonderten Betrachtung der Lang- und Kurzvokalreihen in (6) aus, ergibt sich, dass zwischen den beiden Langvokalen [E:] und [e:] dasselbe Verhältnis besteht wie zwischen [e:] und [i:]: [E:] wird im Vergleich zu [e:] genauso um eine ganze Stufe niedriger artikuliert wie [e:] im Vergleich zu [i:]. Dass dabei der mittlere ungespannte Kurzvokal [E] und das lange niedrige [E:] qualitativ identisch sind, soll diese Systematisierung nicht stören. Zentralisierung der kurzen ungespannten Vokale bedeutet nämlich, dass sie bis zu einer ganzen Stufe niedriger (oder ggf. sogar noch niedriger) artikuliert werden als ihre langen gespannten Pendants, sodass der hohe ungespannte Kurzvokal [ç] dem mittleren gespannten Langvokal [e:] qualitativ näher steht als seinem eigenen hohen gespannt-langen Gegenstück [i:]. Angesichts dieser Gegebenheit erscheint also die phonetische Ähnlichkeit zwischen [E] und [E:] nicht mehr so außerordentlich. Die phonologische Motivation für diese Einstufung des langen offenen [E:] erblickt Wurzel (1981: 933f) in der Umlautbildung. Dabei werden die phonetisch hinteren Monophthonge, d.h. die *u*- und *o*-Vokale in ihre in der vertikalen Zungenposition, der Lippenrundung und der Länge übereinstimmenden vorderen Gegenstücke übergeführt. Dass das an der Umlautung teilnehmende lange [A:] in das lange vordere [E:] verwandelt wird, spricht für den phonologischen Status von [E:] als niedrigen Vorderzungenvokal. Die Umlautbildung, genauer das analoge Verhalten der *a*-Laute im Vergleich zu den übrigen nicht-vorderen Monophthongen des Deutschen erlaubt nach Wurzel (ebda) eine weitere Harmonisierung des deutschen Vokalsystems. Nach der Auffassung des langen offenen [E:] als mittlerer ungespannter Langvokal weist dieses Segment gleich zwei Besonderheiten auf: es ist nämlich nicht nur der einzige ungespannte Langvokal im Vokalsystem des Gegenwartsdeutschen, sondern auch der einzige Langvokal ohne kurzes Pendant. Aufgrund der Einordnung des langen offenen [E:] als niedrigen Vokal sowie des analogen Verhaltens der *a*- und der restlichen nicht-vorderen Vokale (vgl. (8)) betrachtet der Autor das Ergebnis der Umlautbildung des kurzen [a], nämlich [E] als niedrigen vorderen Kurzvokal, d.h. in der mittleren und niedrigen vorderen nicht-runden Vokalreihe handelt es sich um einen Zusammenfall der Kurzvokale.

- (8) a. *Fuß* → *Füße* /u:/ → /y:/, *Fluss* → *Flüsse* /ɨ/ → /Y/  
b. *Floß* → *Flöße* /o:/ → /ë:/, *floss* → *flösse* /O/ → /ê/  
c. *fraß* → *fräße* /A:/ → /E:/, *Fass* → *Fässer* /a/ → /E/

Mit der Annahme Wurzels entsteht jedoch eine merkwürdige Situation im Deutschen: der eingangs genannte qualitative Zusammenfall bezieht sich demnach nicht nur auf zwei, sondern auf drei zugrunde liegende Segmente, nämlich auf den niedrigen langen ([E:] in *Dämon*), den niedrigen kurzen ([E] in *Fässer*) sowie den mittleren kurzen ungerundeten Vorderzungenvokal ([E] in *Fessel*). Somit erscheint es sinnvoll, die traditionelle Auffassung von dieser phonetischen Nivellierung der Unterschiede zu überprüfen. Zwar steht eine vergleichende phonetische Untersuchung der genannten beiden Kurzvokale m.W. aus, sodass hier ihr phonetischer Zusammenfall unter Verweis auf die Notwendigkeit der Durchführung eines solchen phonetischen Vergleichs angenommen wird. Auf der anderen Seite liefern die oben ausführlich besprochenen phonetischen artikulatorischen, akustischen und auditiven Untersuchungen zum deutschen Vokalismus aufschlussreiche Ergebnisse über das Verhältnis von [E:] und [E]. So konnten Bohn et al. (1992) in ihrer glossometrischen Untersuchung der Vokalartikulation ihrer Versuchsperson eine niedrigere Zungenposition bei [E:] als bei [E] feststellen, was für die Betrachtung des ersteren Segments als niedrigen Vokal spricht (vgl. 10). Die akustischen Messungen von Jørgensen (1969), Narahara & Shimoda (1977) und Ramers (1988) liefern jedoch ein etwas differenzierteres Bild über die phonetischen Verhältnisse der beiden Vokale. Während bei [E] im Vergleich zu [E:] überall ein höherer F<sub>2</sub>-Wert festzustellen ist, wobei diese horizontale Zentralisierung des ersteren Segments im Vergleich zum letzteren bezüglich deren Ausmaßes jedoch eine große Variation aufweist (die Werte schwanken zwischen 0,2% und 26,2%), sind die F<sub>1</sub>-Daten nicht mehr so einheitlich. Von den 11 Sprechern hat [E:] bei 5 höhere, bei wiederum 5 niedrigere F<sub>1</sub>-Werte im Vergleich zu [E], während sich bei einem Sprecher die beiden Segmente bezüglich des ersten Formanten nicht unterscheiden. Dies bedeutet bei der ersten Sprechergruppe eine offenere, bei der zweiten eine geschlosseneren [E:]-Artikulation, beim zuletzt genannten Sprecher ist dagegen kein Unterschied in der vertikalen Zungendimension anzunehmen. Anzumerken ist jedoch dabei, dass innerhalb der beiden entgegengesetzten Datengruppen bei jeweils 4 Sprechern der F<sub>1</sub>-Unterschied sehr gering war (bei der ersten Gruppe beträgt er im Durchschnitt 2%, bei der zweiten 1,25%), sowie dass in beiden Gruppen jeweils ein Sprecher eine besonders hohe F<sub>1</sub>-Differenz (in Gruppe 1: 18,8%, in Gruppe 2: 25,1%) produziert hat. Diese Daten belegen entgegen der traditionellen Auffassung, dass zwischen [E:] und [E]



qualitative Unterschiede bestehen können, u.U. sogar relativ bedeutende. Infolge der enormen Variation bezüglich der beiden Zungenparameter bzw. Formantfrequenzwerte erscheint jedoch nicht wahrscheinlich, dass diese qualitative Differenz neben den viel stabileren Quantitätsunterschieden als primäres Unterscheidungsmerkmal zwischen den beiden fraglichen Segmenten anzusehen ist. Dass aber die Qualität doch als eine Art sekundäres Unterscheidungsmerkmal betrachtet werden soll, zeigen die Ergebnisse des Perzeptionstests von Sendlmeier (1981), nach denen sie irgendeine Rolle bei der Identifikation der beiden Vokale spielt: ein künstlich gedehntes [E] wurde nämlich nur von etwas weniger als der Hälfte (47,2%) der Versuchspersonen als [E:] wahrgenommen, während etwas mehr als die Hälfte (51,6%) es immer noch als [E] (die restlichen 1,2% dagegen als [e:]) empfunden haben (vgl. 297).<sup>21</sup>

Angesichts dieser phonetischen Untersuchungen sowie der auf dem Hörerlebnis und der obigen phonologischen Argumentation (des Analogieschlusses Wurzels) beruhenden Annahme vom phonetischen Zusammenfalls der mittleren und niedrigen nichttrunden kurzen Vorderzungenvokale erscheint angebracht, auch die Frage zu untersuchen, ob das dadurch entstandene ‚neue‘ niedrige Vokalpaar [E:]–[E] eine Sonderstellung unter den deutschen Vokalen einnimmt. Ein Blick auf die im oben besprochenen artikulatorischen und akustischen Messwerte der deutschen Vokale in Tab. 4 bis Tab. 8 zeigt, dass auch zwischen den beiden *a*-Vokalen ähnliche Verhältnisse vorliegen, d.h. im Durchschnitt nicht große vertikale und horizontale bzw. F<sub>1</sub>- und F<sub>2</sub>-Unterschiede und große Variation mit u.U. nur ganz minimalen Abweichungen sowie manchmal mit der Umkehrung der Anordnung der beiden Vokale im artikulatorisch-akustischen Vokalraum. Ob diese qualitativen Unterschiede ähnlich wie bei [E:] und [E] auch bei den beiden *a*-Vokalen zumindest als sekundäres Unterscheidungsmerkmal bei der Vokalperzeption anzusehen sind, wird von Sendlmeiers (1981) Experiment zunächst nicht bestätigt: das künstlich gedehnte [a] wurde bei ihm von 98% als [A:], ein gekürztes [A:] dagegen von allen Versuchspersonen als [a] wahrgenommen. Allerdings soll dabei auch Heikes (1972) spezifisches Hörexperiment erwähnt werden, bei dem die Versuchspersonen die Dauer synthetisch erzeugter *a*-Vokale mit einem Regler selber einstellen konnten. Dieser Versuch hat ergeben, dass ein *a*-Vokal mit der Formantstruktur von [A] bei einer Durchschnittsdauer von 363 ms als [A:], bei einer von 147 ms als [a] empfunden wurde, während zur Wahrnehmung eines *a*-Vokals mit der Formantstruktur von [a] als [A:]

---

<sup>21</sup> Zum Vergleich zwei ähnliche Vokalpaare: [ç:] wurde von 64,4% als [e:] und [i:] von 68,3% als [ë:] perzipiert.

bzw. [a] der synthetische Vokal 463 bzw. 168 ms lang sein musste (vgl. 728), woraus der Autor der Schluss zieht, „daß ‚helles‘ /a/ [d. h. /a/ – Anm. v. Verf.] mit perceptiver Kürze korreliert und deshalb durch größere Dauerwerte kompensiert werden muß“ (ebda). Für das ‚dunkle‘ /A/ dagegen gelte der umgekehrte Sachverhalt (ebda). Folgerichtig ist nicht auszuschließen, dass die u.U. minimalen qualitativen Unterschiede zwischen den beiden *a*-Lauten doch irgendeine Rolle bei ihrer Perzeption spielen. Alles in allem ist die phonetische Ähnlichkeit der fraglichen beiden Vokalpaare nicht zu übersehen. Nicht zufällig sieht Becker (1998: 21) in diesen ähnlichen phonetischen Verhältnissen die ähnliche phonologische Klassifizierung der beiden Vokalpaare, d.h. ihre Bestimmung als niedrige Vokale, bestätigt.

Schließlich scheint es im Interesse einer Vermeidung umständlicher und dadurch das Verständnis des Textes beeinträchtigender Umschreibungen nicht abwägig zu sein, für den zugrunde liegenden niedrigen nichtrunden vorderen Langvokal [E:], der nach den Ergebnissen der zitierten phonetischen Untersuchungen eine andere Vokalqualität haben kann als das traditionell mit demselben Transkriptionszeichen wiedergegebene Segment [E], ein eigenes Symbol einzuführen.<sup>22</sup> So wollen wir in der vorliegenden Untersuchung das Symbol [E:] ähnlich wie Pilch (1966) durch [é:] ablösen. Das Symbol [a#], wie es etwa in Vennemann (1991a) und Becker (1998) zur Unterscheidung des [E:] von [E] verwendet wird, scheint uns weniger geeignet, da es der deutschen Orthographie entnommen ist, während alle anderen Symbole aus dem IPA-Alphabet stammen – allerdings ist das nur ein ‚Schönheitsfehler‘. Nach Wiese (1996: 21) spreche gegen die Verwendung des Symbols [é:], dass [E] und [é] in Sprachen, in denen sie vorkommen, zwei unterschiedliche Vokalqualitäten darstellten, dass sich ersteres im Deutschen jedoch qualitativ in keinerlei Weise von [E] unterscheidet. Die Ergebnisse der oben erwähnten phonetischen Untersuchungen entkräften jedoch sein Argument. Ein anderer potentieller Einwand gegen die Verwendung von [é:] könnte auf die

---

<sup>22</sup> Für den niedrigen nichtrunden vorderen Kurzvokal wollen wir hier kein eigenes Symbol einführen. Damit soll sein angenommener phonetischer Zusammenfall mit dem mittleren nichtrunden vorderen Kurzvokal zum Ausdruck gebracht werden. Wie weiter unten noch gezeigt wird, spielt ein eventueller qualitativer Unterschied zwischen [E:] und [E] in der Phonologie des Deutschen sowieso keine Rolle.

IPA-Vokaltabelle Bezug nehmen (vgl. die Abbildung rechts) [é] soll demnach für einen halbtiefen nichtrunden Vorderzungenvokal stehen, der also mit einer niedrigeren vertikalen Zungenposition gebildet wird als der untermittelhohe Vokal [E]. Die angeführten phonetischen Experimente haben jedoch gerade in dieser Hinsicht kein eindeutiges Ergebnis geliefert: bei einem Teil der Sprecher

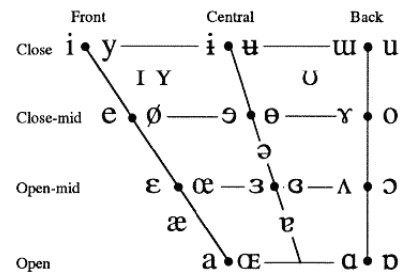


Abb.2 IPA-Vokaltabelle

war genau dieses, bei dem anderen dagegen das entgegengesetzte Verhältnis zu beobachten. Dass dieser Einwand doch nicht ganz akzeptabel ist, zeigen jedoch die deutschen Vokale [ç], [e] und [E]: [ç] ist laut IPA-Tabelle als halbhoch, [e] als obermittelhoch und [E] als untermittelhoch einzustufen. Die akustischen Messwerte zeugen jedoch davon, dass [ç] nicht selten im ‚Zwischenraum‘ von [e] und [E] gebildet wird. Wenn also [é] deswegen nicht akzeptiert werden soll, weil es einen Vokal bezeichne, für deren Bildung eine konsequent tiefere Zungenstellung charakteristisch sei als für [E], dann sollte mit demselben Argument auch [ç] aus der transkriptionsmäßigen Wiedergabe der deutschen Vokale eliminiert werden.

Wenn also im Standarddeutschen in der Tat einen im Vergleich zum mittleren ungerundeten langen |e:| offeneren Vokal gibt, dann kann er als niedriger vorderer ungerundeter gespannter Langvokal aufgefasst werden, dessen kurzes Pendant mit dem des mittleren |e:| zusammenfällt, vgl. Tab. 10. Somit sollten die bisher ermittelten 7 qualitativen Vokalklassen durch eine achte ergänzt werden, die – wie bei allen diesen Vokalklassen – in Anlehnung an die Orthographie als *ä*-Vokale bezeichnet werden könnten. Mit dieser Klassifizierung spricht nichts für die Annahme der Klasse der ungespannten Langvokale im Deutschen.

	<i>vorn</i>		<i>hinten</i>	
	<i>nichtrund</i>	<i>rund</i>	<i>nichtrund</i>	<i>rund</i>
<i>hoch</i>	i(:)    ç	y(:)    Y )		u(:)    ü )
<i>mittel</i>	e(:)    E	ë(:)    ê		o(:)    O )

<i>niedrig</i>	é(:) E	A(:) ) a
----------------	--------	-------------

Tab. 10 Klassen der Vollvokale des Deutschen

Durch Ausschuss der ungespannten Langvokale im Deutschen ergeben sich rein kombinatorisch drei phonetische Vokalklassen, vgl. Tab. 11. Es stellt sich nun die Frage, ob alle diese Vokalklassen phonologisch relevant sind.

Betrachtet man die Beispiele unter (9), so sieht man, dass zwischen den gespannten Lang- und Kurzvokalen eine komplementäre Verteilung besteht: gespannte Langvokale erscheinen – wie besprochen – nur unter Akzent, gespannte Kurzvokale dagegen ausschließlich unter Akzentlosigkeit. Andererseits zeugen diese Daten auch von einem systematischen Wechsel zwischen den beiden Vokalklassen: ein

	<i>gespannt</i>	<i>ungespannt</i>	<i>gespannt</i>
	<i>lang</i>	<i>kurz</i>	<i>kurz</i>
<i>i</i> -Vokale	i:	ɨ	i
<i>ü</i> -Vokale	y:	Y	y
<i>e</i> -Vokale	e:	E	e
<i>ö</i> -Vokale	ë:	ê	ë
<i>a</i> -Vokale	A:	a	A
<i>o</i> -Vokale	o:	O	o
<i>u</i> -Vokale	u:	ö	u

gespannter Vokal erscheint dabei immer nur unter Betontheit lang, unter Unbetontheit wird er kurz realisiert. Dies legt nahe, die beiden phonetischen Klassen der gespannten Lang- und der gespannten Kurzvokale phonologisch (zugrunde liegend) als eine einzige Vokalklasse aufzufassen und die quantitativen Verhältnisse aus der Position des Akzents herzuleiten.

(9) *Musík* [muzi:k], *MúsiKER* [mu:zikÇ], *musikÁlisch* [muzikA:lçS], *MusikalitÁt* [muzikAlité:t]

Es gibt jedoch ein potentiell Argument, das für die zugrunde liegende Trennung der gespannten-langen und gespannt-kurzen Vokale spricht, und das genau die Akzentverhältnisse in den nichtnativen Wörtern betrifft. Fasst man nämlich das Deutsche als eine quantitativsensitive Sprache an, so soll sich die Stelle des Wortakzents aus der Stelle der Langvokale ergeben, während sich die Kurzvokale – unabhängig davon, ob sie gespannt oder ungespannt sind – für die Akzentzuweisung als irrelevant erweisen. Wurzel (1981) stellt bei

der Untersuchung des Verhältnisses zwischen Wortakzent, Vokallänge und Zentralisiertheit<sup>23</sup> fest, dass die Betonung in den nichtnativen Wörtern aufgrund der Vokallänge, der Silbenstruktur sowie bestimmter morphologischer Gegebenheiten voraussagbar ist (vgl. 917), und formuliert eine Akzentzuweisungsregel, die der letzten schweren Silbe (d.h. der letzten Silbe mit einem Langvokal, Diphthong oder einer Kurzvokal-Konsonant-Sequenz) den Wortakzent zuweist, wenn diese nicht morphologisch unbetonbar ist (vgl. 918). So erhält in *Ökonom* die finale Silbe den Wortakzent, da sie zugrunde liegend einen langen nichtzentralisierten Vokal enthält und nicht unbetonbar ist, aber auch in *ökonomisch* und *ökonomischer* wird der Wortakzent der Silbe mit diesem zugrunde liegenden Vokal /o:/ zugewiesen, da das Suffix *-isch* bzw. die Flexionsendung *-er* zwar schwere Silben darstellen, aber als native Flexive bzw. Wortbildungsformative nicht betonbar sind (vgl. ebda). Dadurch ergibt sich für *ökonomisch* Pänultima-, in *ökonomischer* Antepänultimabetonung.

Diese Auffassung ist jedoch nicht ohne Probleme. Wiese (1988) (vgl. auch Wiese 1996) weist auf einige Schwachstellen einer quantitätssensitiven Wortakzenttheorie des Deutschen hin. Erstens erscheint der Zusammenhang zwischen Silbenschwere und Akzentzuweisung infolge der morphologischen Unbetonbarkeit mancher Suffixe als nicht eindeutig (vgl. Wiese 1988: 74).<sup>24</sup> Zweitens macht die obige quantitätssensitive Akzentregel des Deutschen manchmal falsche Voraussagen. So ist die jeweils letzte Silbe der Wörter unter (10) schwer, sodass die Regel allen diesen Wörtern finalen Akzent zuweist. Dies ist jedoch nur für die Fälle in (10)a. das richtige Akzentmuster, die Wörter in (10)b. sind anfangsbetont (vgl. ebda).

(10) a.	b.
<i>Graphít</i>	<i>Fázit</i>
<i>Dekán</i>	<i>Jápan</i>
<i>Motór</i>	<i>Mótor</i>
<i>Modúl</i>	<i>Kónsul</i>
<i>Baróck</i>	<i>Ámok</i>

---

<sup>23</sup> Zu der von ihm vorgenommenen Trennung zwischen Gespanntheit und Zentralisiertheit s. Fn.19.

<sup>24</sup> Dazu kommt noch, dass die von Wurzel (1981) als unbetonbar, d.h. betonungsunfähig erklärten Suffixe u.U. doch einen bestimmten Grad an Akzentuierung erhalten können. So hat die vorletzte Silbe in *malerische* oder *kinderliche* eine stärkere Betonung als die umgebenden Schwa-Silben. Damit ist das Konzept der Unbetonbarkeit (Akzentunfähigkeit) mancher Suffixe gewissermaßen entkräftet, gänzlich unbetont sind im Deutschen lediglich die Schwa-Silben (vgl. Wiese 1988: 103).

<i>Damást</i>	<i>Bállast</i>
<i>Papíer</i>	<i>Fákir</i>
<i>apárt</i>	<i>Léopard</i>
<i>urbán</i>	<i>Túrban</i>
<i>Hermelín</i>	<i>Pínguin</i>

Und schließlich gibt es im Deutschen eine Fülle von Wörtern, in denen zwar Vokallänge und Wortakzent miteinander zusammenhängen, jedoch nicht im Sinne der obigen quantitätssensitiven Akzentregel, sondern vielmehr umgekehrt: Akzentuiertheit führt zur Längung, Akzentlosigkeit zur Kürzung eines zugrunde liegend gespannten Vokals (vgl. Wiese 1996: 279), vgl. *Chemie* [Åe.}mi:] vs. *Chemiker* [}Åe:.mi.kÇ], bzw. *Philosoph* [filo}zo:f] vs. *Philosophie* [filozo}fi:]. Zur Herleitung der Akzentverhältnisse sowie des Quantitätswechsels in solchen Wörtern führt Wurzel (1981) eine Art *default*-Akzentzuweisungsregel, die mangels einer schweren Silbe im Wort, der dritt- oder zweitletzen leichten Silbe den Wortakzent zuweist (vgl. 918); sowie eine Quantitätsänderungsregel mit der Form in (11) ein (vgl. 930).

(11) Quantitätsänderungsregel nach Wurzel (1981: 930)

$$[+silbisch] \rightarrow [\alpha \text{ lang}] / \left[ \begin{array}{c} \text{_____} \\ [\alpha \text{ hauptbetont}] \end{array} \right] \&$$

Der Autor nimmt dann im Stamm /Åem/ einen zugrunde liegend gespannten Kurzvokal an, der sowohl in *Chemie* als auch in *Chemiker* in einer leichten Silbe steht. In *Chemie* ist der finale Vokal zugrunde liegend lang, sodass hier die quantitätssensitive Akzentregel der letzten Silbe den Wortakzent zuweist. *Chemiker* besteht aus zwei leichten und einer morphologisch unbetonbaren Silbe, hier wird *per default* die drittletzte (d.h. die initiale) Silbe betont, was die Anwendung der Regel in (11) auslöst: der gespannte Kurzvokal wird unter Akzent gelängt. In der letzten Silbe von *Philosoph* wird dagegen ein zugrunde liegend gespannter Langvokal angenommen, der dadurch Akzent auf sich zieht. In *Philosophie* ist aber auch die finale Silbe schwer, sodass hier diese betont wird, was zur Folge hat, dass die Regel in (11) einen Input findet: der zugrunde liegend gespannte Langvokal wird unter Akzentlosigkeit gekürzt.

Daraus, dass also ein auf der Annahme der Quantitätssensitivität basierendes Akzentmodell des Deutschen nur mit ganz starken und u.U. nicht unproblematischen Restriktionen (Betonungsunfähigkeit mancher Suffixe sowie eine Quantitätsänderungsregel), die die Grundannahme der Quantitätssensitivität teilweise außer Kraft setzen, funktionieren

kann, jedoch selbst dadurch falsche Akzentmuster herleiten kann, folgt, dass Vokallänge und Wortakzent zwar miteinander eng zusammenhängen, nicht aber, dass Letzterer von Ersterer herleitbar ist: das Deutsche ist möglicherweise keine quantitätssensitive Sprache (vgl. Wiese 1988: 74, bzw. 1996: 280).<sup>25</sup> Dadurch wird jedoch auch das eingangs genannte Argument für die Aufrechterhaltung aller drei auf der Oberfläche erscheinenden Vokalklassen in Tab. 11 entkräftet: im Deutschen sollen zugrunde liegend nur die beiden Vokalklassen in (6) angenommen werden. Wenden wir uns jetzt der Beschreibung der beiden Konzepte der zugrunde liegenden Qualität und Quantität, bzw. der Formulierung der Argumente *pro* und *kontra* zu.

### 3.1.3 Die Frage nach der Distinktivität der Qualität und Quantität

Betrachtet man in der fraglichen Vokalopposition – mit Moulton (1962), Reis (1974), Kloeke (1982) oder Jessen (1998) – die Qualität als primäres Unterscheidungsmerkmal, so ergeben sich folgende zugrunde liegenden Vokalpaare:

(12)	a.	b.
	i	ç
	y	Y
	e	E
	ë	ê
	A	a
	u	ū

---

<sup>25</sup> Diese Schlussfolgerung von Wiese (1988) bzw. (1996) ist in einem Punkt zu relativieren. Mit seinen obigen Ausführungen ist nämlich nur das in Wurzel (1981) vertretene quantitätssensitive Konzept widerlegt. Versucht man jedoch, mit Vennemann (1994) die Quantitätssensitivität auf eine andere Art und Weise zu erfassen, die der spezifischen Quantität des Deutschen besser Rechnung tragen kann, so kann das Deutsche schon als quantitätssensitive Sprache aufgefasst werden, in der jedoch nicht so sehr die Längen-, sondern die Silbenschnittverhältnisse der Vokale zur Akzentzuweisung verwendet werden. Unser Ziel bestand jedoch darin zu zeigen, dass Wurzels quantitätssensitive Akzentmodell für das Deutsche nicht akzeptabel ist, und nicht etwa, welche Theorie sich zur Beschreibung der Wortakzentverhältnisse im Deutschen besser eignet. Da zu diesem Zweck vollkommen ausreicht, Wurzels Theorie mit Wieses Gegenargumenten zu konfrontieren, haben wir bewusst auf eine Darstellung des Vennemannschen Akzentmodells verzichtet. Dazu wäre mindestens ein Vergleich von Wiese (1996: Kap.8) und Vennemann (1994) notwendig.

|o|      |O|

Mit der Annahme eines Qualitätskontrastes im Deutschen wird der Quantität jedwede phonologische Relevanz abgesprochen und bloß die Rolle einer phonetischen Begleiterscheinung unter Akzent zugebilligt, was in (12) durch die Verwendung unterschiedlicher Transkriptionszeichen durch gleichzeitiges Weglassen des von der IPA vorgeschlagenen Längezeichens ‚:‘ ausgedrückt wird. Dies macht die Eliminierung der Dauerunterschiede aus der Derivation und der Oberflächenrepräsentation sinnvoll: ob ein Vokal lang oder kurz zu sprechen ist, wird von den phonetischen Realisierungsregeln bestimmt. Diese berechnen dann aufgrund der vorhandenen phonologischen Informationen die aktuelle Länge eines Vokals in einer Segmentkette. Zu diesen phonologischen (zugrunde liegenden und abgeleiteten) Informationen zählen zunächst einmal der Gespanntheitswert des fraglichen Vokals und der Akzent. Die phonetischen Realisierungsregeln können bei der Berechnung der aktuellen Länge eines Vokals jedoch auch weitere phonologische Informationen berücksichtigen, so z.B. den unmittelbaren Lautkontext des Vokals, seine Stellung in den höheren prosodischen Konstituenten wie dem phonologischen Wort oder der phonologischen Phrase, oder eben seine Position im Vergleich zur Hauptakzentstelle. Bekannt ist z.B. der dehnende Charakter eines postvokalischen *r*-Konsonanten (vgl. Becker 1998), oder dass ein unbetonter gespannter Vokal im absoluten Wortauslaut trotz der Unbetontheit länger artikuliert wird (vgl. Neppert & Pétursson 1992: 161), oder aber dass ein gespannter Vokal vor der Hauptakzentstelle kürzer realisiert wird als danach (vgl. Ramers 1988). Daraus ist zu sehen, dass die phonetischen Realisierungsregeln nicht nur zwei mit Quantitätswerten operieren, sondern Vokaldauer als eine skalare Größe behandeln. Dies ist deswegen so, weil die Quantität nach diesem Ansatz kein abstraktes phonologisches Phänomen darstellt, das zur Distinktion zwischen zwei Vokalen dient, sondern rein phonetischer Natur ist. So soll man nach dem Qualitätsansatz in *Muss* zugrunde liegend einen als ungespannt spezifizierten Vokal annehmen, dem die phonetischen Realisierungsregeln phonetische Kürze zuweisen. In *Mus*, *Musik*, *musisch*, *Musiker*, *musikalisch*, *Musikalität* ist der erste Vokal dagegen zugrunde liegend gespannt, sodass er unter Akzent (also in *Mus*, *musisch* und *Musiker*) von den phonetischen Realisierungsregeln gedehnt wird, sonst (unter Akzentlosigkeit) wird er kurz gesprochen.

Ein großer Vorteil der Auffassung der Vokalopposition in (6) als Qualitätskontrast besteht zweifellos in der einheitlichen Behandlung der Quantitätsphänomene des Deutschen: die Länge eines Vokals wird von phonetischen Realisierungsregeln berechnet, die dabei



zahlreiche phonologische Informationen, u.a. den Akzent berücksichtigen, der auf die gespannte Vokale dehnend auswirkt. Auch erübrigt sich in diesem Fall die Debatte um das genaue phonetische Verhältnis zwischen den gespannten Kurz- und Langvokalen. Untersucht werden sollte in diesem Fall vielmehr die Frage, welche phonologischen Faktoren die Dehnung eines nichtzentralisierten Vokals bewirken, welche von diesen universell und welche sprachspezifisch sind, und eventuell worauf der vokaldehnende Charakter dieser phonologischen Faktoren zurückzuführen ist.

Trotz dieses klaren Vorteils der qualitativen Analyse gibt es ganz starke Argumente, die gegen die Annahme eines Qualitätskontrastes im deutschen Vokalsystem sprechen. Das erste davon ist phonetischer Natur und betrifft die niedrigen Vokale. Wie im vorausgehenden Abschnitt gezeigt weisen die beiden *a*-Vokale des Deutschen in Bezug auf das qualitative Merkmal eine wichtige Besonderheit auf: zwischen [A] und [a] besteht im Durchschnitt nur ein minimaler Unterschied in der Gespanntheit, der von Sprecher zu Sprecher eine große Variation zeigen kann. Aus diesem Grund wird zwischen den beiden *a*-Vokalen von vielen Autoren primär eine Quantitätsdifferenz angenommen, was sich auch in der Verwendung der gleichen Transkription für die *a*-Vokale (/a:–/a/) widerspiegelt (vgl. u.a. Meinhold & Stock 1982 oder Wiese 1988, 1996). Auch das DUDEN Aussprachwörterbuch (vgl. Mangold 1990) schließt sich dieser Tradition an, während das *Große Wörterbuch der deutschen Aussprache* (vgl. Krech et al. 1982) – ähnlich wie etwa Wurzel (1981) – konsequent zwei IPA-Symbole für die beiden *a*-Vokale (/A:–/a/) verwendet. Diese nur minimalen qualitativen Unterschiede im Bereich der niedrigen Vokale machen eher unwahrscheinlich, dass die Qualität (Gespanntheit/Zentralisiertheit) als primäres Unterscheidungsmerkmal im deutschen Vokalsystem dienen könnte. Scheinbar bietet sich die Annahme von zwei Unterscheidungsmerkmalen: einem Qualitätsmerkmal für die nichtniedrigen, und einem Quantitätsmerkmal für die niedrigen Vokale, an (eine ähnliche Position wird z.B. in Meinhold & Stock 1982 vertreten). Eine solche Analyse ist m.E. deswegen nicht akzeptabel, weil dadurch der Schein erweckt würde, im Bereich der niedrigen und nichtniedrigen Vokale handelte es sich nicht um dieselbe Opposition. Zwischen /A:/ und /a/, /i:/ und /ç/, /y:/ und /Y/ usw. besteht jedoch dieselbe Opposition, diese kann aber kein Qualitätskontrast sein.

Ebenfalls gegen eine Qualitätopposition spricht der Sonderstatus, den das Gespanntheitsmerkmal unter den anderen segmentalen (inhärenten) Merkmalen einnehmen würde: es wäre das einzige Merkmal, das in der Lage ist, den Vokal unter Akzent zu

dehnen.<sup>26</sup> Fasst man den fraglichen Vokalgegensatz jedoch als eine prosodische (z.B. als Quantitäts-) Opposition auf, so erübrigt sich dieses Problem. Für eine prosodische Beschreibung dieses Vokalkontrastes sprechen auch gewisse phonotaktische Unterschiede zwischen zentralisierten und nichtzentralisierten Vokalen, denen ein rein segmentaler Ansatz nicht Rechnung tragen kann (vgl. Vennemann 1991a: 214ff und Becker 1998: 49ff). Erstens kommen betonte ungespannte Vokale nur in gedeckter Stellung vor (und nicht am Wortende und im Hiat), die im Falle eines einzigen postvokalischen Konsonanten sogar durch Ambisilbizität (Gelenkbildung), d.h. durch ‚virtuelle‘ Schließung der Silbe erreicht wird. Zweitens kann nach einem zentralisierten Vokal um ein Segment mehr in der Silbe folgen als einem nichtzentralisierten Vokal oder eben einem Diphthong, m.a.W. sind gespannte Vokale, Diphthonge bzw. die Sequenz eines ungespannten Vokals und eines Konsonanten phonotaktisch äquivalent, vgl. *viel – feil – Film, doof – drauf – Dorf, Dienst – raubst – Herbst, Obst – läufst – wirfst*. Und schließlich zeigt sich im Deutschen tendenziell eine ‚Gespantheitspolarität‘, nach der Kurzvokale vorzugsweise vor Fortisobstruenten, Langvokale vor Lenisobstruenten stehen.<sup>27</sup>

Die genannten Argumente sprechen also gegen eine qualitative und für eine prosodische Beschreibung der fraglichen Vokalopposition. Da Quantität traditionell zu den prosodischen Phänomenen gerechnet wird, und zwischen den beiden Vokalklassen des Deutschen in (6) (zumindest unter Akzent) ein deutlicher Dauerunterschied besteht, liegt es auf der Hand, in dem Vokalgegensatz die Quantität als primäres Unterscheidungsmerkmal anzusehen. Im Folgenden soll diese Möglichkeit näher charakterisiert werden.

Mit der Annahme der zugrunde liegende Quantität ergeben sich für den Vokalgegensatz in (6) die zugrunde liegenden Vokalpaare in (13). Die Verwendung identischer Transkriptionssymbole sowie des Kolons deuten darauf hin, dass es sich hier zugrunde liegend nicht um einen segmentalen, sondern um einen prosodischen Kontrast, genauer um die Opposition zweier Quantitätsstufen handelt.<sup>28</sup> Da es sich jedoch dabei lediglich um eine

---

<sup>26</sup> Dazu kommt noch – wie oben besprochen – seine problematische Unterbringungsmöglichkeit in der Merkmalgeometrie.

<sup>27</sup> Zu der segmentalen Beschreibungsmöglichkeit dieser letzten phonotaktischen Regularität sowie deren Handicaps vgl. Becker (1998: 50f).

<sup>28</sup> Bei der Wahl zwischen den beiden möglichen IPA-Symbolen, d.h. denen für ‚gespannte‘ und denen für ‚ungespannte‘ Vokale (etwa /i/ vs. /ç/ oder /o/ vs. /O/ usw.) haben wir uns hier für erstere entschieden. Dass hier also die *a*-Vokale zugrunde liegend mit /A/, und nicht mit /a/ wie das z.B. in Krech et al. (1982) oder

IPA-mäßige *ad hoc* Darstellungsmethode handelt, stellt sich nun die Frage, wie Vokalquantität im Deutschen zu repräsentieren ist.

(13)	a.	b.
	/i:/	/i/
	/y:/	/y/
	/e:/	/e/
	/ë:/	/ë/
	/A:/	/A/
	/u:/	/u/
	/o:/	/o/

Lehiste (1970) nennt drei Möglichkeiten zur phonologischen Beschreibung eines Quantitätskontrastes: als segmentale Länge, als suprasegmentale Länge oder als Geminatio (vgl. 43). Die Beschreibung der Vokallänge als Geminatio geht auf Trubetzkoy (1939) zurück, der die Sprachen in zwei Klassen teilt, je nachdem, ob der lange Silbenträger in ihnen weiter zerlegbar ist oder nicht. Weiter zerlegbar ist ein langer Silbenträger nach ihm z.B. im Finnischen, in dem „zwischen Anfang und Ende eines solchen Silbensträgers eine morphologische Grenze liegen kann“ (170). Als Beispiel gibt er die finnische Partitivendung *-a/-ä* an, die an einen auf *-a* bzw. *-ä* auslautenden Stamm angeschlossen die Längung des auslautenden Vokals zur Folge hat, so wird *kukka* ‘Blume’ im Partitiv zu *kukkaa*, *leipä* ‘Brot’ zu *leipää* (ebda). Solche Sprachen nennt Trubetzkoy nach den Zeiteinheiten, in die ein langer Silbenträger zerlegt werden kann, „moren-zählend“, diejenigen dagegen, die keine Zerlegung ihres langen Silbenträgers zulassen, „silbenzählend“ (174). Er rechnet zwar das Deutsche zu den silbenzählenden Sprachen, Moulton (1956)<sup>29</sup> entdeckt jedoch eine Regularität in dieser Sprache, nach der sich ein Langvokal wie ein polyphonematischer Diphthong bzw. die Sequenz Kurzvokal + Konsonant verhält, z.B. was die Anzahl der Konsonanten betrifft, die diese in einer Silbe folgen können. Dies macht die Einordnung des Deutschen zu den morenzählenden Sprachen sinnvoll. Dementsprechend analysiert er die deutschen Langvokale als Cluster identischer Kurzvokale, so z.B. langes /i:/ als /ii/ (vgl. 374). Nach Kloeke (1982:

---

Wiese (1988) und (1996) der Fall ist, wiedergegeben werden, ist eine Konsequenz dieser unserer rein technischen Entscheidung und soll nicht dermaßen interpretiert werden, dass hier zugrunde liegend eine einzige *a*-Qualität angenommen wird, die der des gespannten Oberflächenvokals [A] gleichkommt.

<sup>29</sup> Zitiert nach Ramers (1988: 69).

7) ermöglicht die Geminationsanalyse Moultons jedoch keine adäquate Formulierung der deutschen Umlautregel, die zwar den gesamten Langvokal (also beide ‚Teilsegmente‘ dessen), aber nur das zweite Segment der Diphthonge betrifft. Gegen Moulton (1956) spricht nach Wiese (1988: 67) außerdem das Prinzip der Obligatorischen Kontur, das identische benachbarte Segmente verbietet.

Die Auffassung der Länge als segmentales Merkmal, etwa in Form  $[\pm\text{lang}]$ , ist nach Lehiste die einfachste Beschreibung eines Quantitätskontrasts. Sie fügt außerdem hinzu, dass diese Beschreibungsmöglichkeit für die generative Phonologie charakteristisch ist. Sie ist jedoch der Meinung, dass sich das segmentale Längenmerkmal am besten für Sprachen eignet, in denen die Quantitätsopposition nur auf eine kleine Anzahl von Phonemen beschränkt ist; als Beispiel nennt sie das Spanische, in dem ein distinktiver Längenunterschied nur bei den *r*-Lauten besteht, vgl.  $/r/$  vs.  $/rr/$ : *pero* ‘but’ vs. *perro* ‘Hund’ (vgl. 43). Eine prosodische Analyse der Quantitätsopposition schlägt sie dagegen für Sprachen vor, in denen dieser Kontrast für das gesamte Vokal- und/oder Konsonantensystem charakteristisch ist, und zwar aus ökonomischen Gründen, da dadurch das Phoneminventar der betreffenden Sprache reduzierbar sei (ebda).

Entsprechend Lehistes Bemerkung führt Wurzel (1981) die deutsche Vokalopposition in (13) auf ein segmentales Längenmerkmal zurück, das er in die Merkmalmatrix der deutschen Vokalphoneme integriert. Ähnlich wie Wurzel (1981) gehen auch Meinhold & Stock (1982) vor.

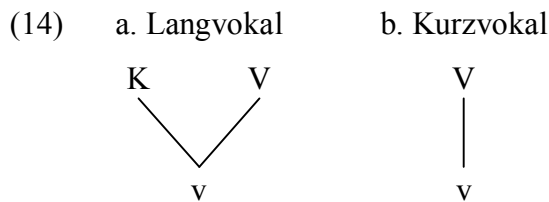
Wiese (1988) argumentiert gegen eine Beschreibung der Vokallänge als segmentales Merkmal, die nach ihm einen Rückschritt im Vergleich zu Jakobson & Halles Erkenntnis bedeutet, nach der Länge ein prosodisches Phänomen ist (vgl. 63). Ein prosodisches Merkmal nach diesen Autoren „is displayed only by those phonemes which form the crest of the syllable, and it may be defined only with reference to the relief of the syllable or of the syllable chain” (Jakobson & Halle 1956: 33). Es kann somit den inhärenten Merkmalen gegenübergestellt werden, denn ein inhärentes Merkmal „is displayed by phonemes irrespective of their role in the relief of the syllable, and the definition of such a feature does not refer to the relief of the syllable or of the syllable chain” (ebda)<sup>30</sup>. Wiese ist der Ansicht, dass diese Erkenntnis Jakobson & Halles nur in den nichtlinearen phonologischen Theorien ausgeführt werde (vgl. Wiese 1988: 63), genauer in der in Kap. 1 angesprochenen CV-Phonologie, die eine Trennung zwischen einer melodischen (segmentalen) und einer

---

<sup>30</sup> Vgl. auch Mayerthaler (1974: 17).

skeletalen (rhythmischen) Schicht vorsieht, wodurch z.B. Affrikaten mit zwei Segmenten, aber nur einer skeletalen Position verbunden sind.

Da sich Langvokale – wie oben besprochen – phonotaktisch wie Diphthonge oder Kurzvokal+Konsonant-Sequenzen verhalten, aber ‚melodisch‘ einheitlich sind, liegt es auf der Hand, für sie die spiegelbildliche Struktur als für die Affrikaten anzunehmen, d.h. ein Segment auf der melodischen Schicht, das mit zwei Positionen auf der CV-Schicht verbunden ist. Bei den Kurzvokalen gilt dagegen eine eins-zu-eins-Beziehung zwischen der melodischen und der rhythmischen Ebene, vgl. (14). Mit dieser Annahme wird die Anzahl der Segmente des deutschen Vokalsystems auf 8 reduziert, die sich voneinander in ihrer Merkmalstruktur unterscheiden, und für die es jeweils zwei distinkte zugrunde liegende Quantitäten gibt.<sup>31</sup>



Mit der Auffassung der Quantität als zugrunde liegendes Unterscheidungsmerkmal für die fragliche Vokalopposition wird Gespanntheit entweder als eine untergeordnete phonologische oder bloß als eine phonetische Größe angesehen. Im ersteren Fall, der z.B. in Wiese (1988) oder in Hall (1992) vertreten ist, wird den Vokalen bis auf das lange |E:| im Laufe der Derivation ein Gespanntheitswert zugewiesen. Dies ist bei Wiese (1988) wegen des Sonderstatus des |E:| als mittleren ungespannten Langvokals so, der schon zugrunde liegend für Gespanntheit spezifiziert ist, um die Unterscheidung zwischen ihm und |e:| sichern zu können. Bei Hall (1992) kommt noch ein weiteres Argument dazu, auf das wir weiter unten eingehen werden. Wenn man |E:| jedoch als regulären Vokal der niedrigen vorderen Reihe betrachtet – wie das hier der Fall ist, kann Gespanntheit durchaus als rein phonetisches Phänomen aufgefasst werden, die den zugrunde liegend langen bzw. kurzen Vokalen – ähnlich wie die aktuelle Vokaldauer im Qualitätskonzept – von phonetischen Realisierungsregeln zugewiesen wird. Dabei wird ein Vokal, der in der zugrunde liegenden Repräsentation mit zwei Positionen auf der CV-Ebene assoziiert ist, als ‚gespannt‘, einer dagegen, der in der zugrunde liegenden Repräsentation mit nur einer einzigen Position auf der skeletalen Schicht verbunden ist, als ‚ungespannt‘ interpretiert. Die Verwendung der Labels

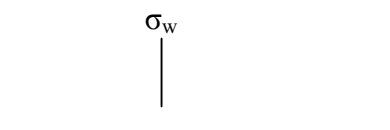
---

<sup>31</sup> Klein geschriebenes ‚v‘ soll dabei einen Vokal auf der segmentalen Ebene bezeichnen.

„ungespannt“–„gespannt“ anstatt von binären phonologischen Merkmalen wie etwa [±gespannt] zeigt, dass es sich hier – ähnlich wie bei der Vokaldauer im Qualitätskonzept – nicht um abstrakte binäre, sondern um weniger abstrakte – mehr konkrete – phonetische Merkmale handelt. Es wäre auch nicht unangebracht, nicht nur zwischen zwei phonetischen Qualitätswerten zu unterscheiden, sondern Gespanntheit als skalare Größe aufzufassen. Das bedeutete, dass bei der Berechnung der Qualität eines Vokals neben der Anzahl der CV-Positionen auch andere Kriterien berücksichtigt werden sollten. So könnte eine einzige Position auf der CV-Schicht in Verbindung mit dem Merkmal [niedrig] einen weniger zentralisierten Vokal ergeben als etwa eine CV-Position mit dem Merkmalswert [hoch].

Eine andere mit dem Quantitätskonzept verbundene Frage bezieht sich auf die Erfassung des oben beschriebenen Quantitätenwechsels in Abhängigkeit von der Akzentstelle und somit auf die Unterscheidung zwischen gespannten Lang- und Kurzvokalen. Wiese (1988) schlägt hierzu eine strikte Trennung „zwischen der **Quantität** in der Silbe und **Länge** auf der phonetischen Oberfläche“ (74) vor: die oben beschriebene klare phonotaktische Regularität des Deutschen, d.h. die phonotaktische Äquivalenz zwischen Langvokalen, Diphthongen und Kurzvokal-Konsonant-Sequenzen legt auf der einen Seite zwei Quantitätswerte nahe, auf der anderen existieren auf der Oberfläche – wie oben bereits ebenfalls besprochen – möglicherweise mehr als nur zwei Längenwerte. Während die beiden Quantitätswerte in der phonologischen Repräsentation, genauer schon zugrunde liegend vorhanden sind, werden die aktuellen Längenwerte den Vokalen von den phonetischen Realisationsregeln zugewiesen. Hall (1992) bietet eine alternative Analyse zu Wieses (1988) Beschreibung. Während Wiese (1988) den Quantitätenwechsel als Teil der phonetischen Realisation betrachtet, geht Hall (1992) von einer phonologisch relevanten Alternation aus, die er mit der Regel in (15) beschreibt. Nach dieser Regel wird ein zugrunde liegend langer Vokal unter Akzentlosigkeit gekürzt. Damit als Ergebnis der Anwendung dieser Regel keine ungespannten Kurzvokale entstehen, muss der Autor seine Gespanntheitszuweisungsregel vor (15) anwenden. Dies ist nun der andere Grund, warum der Autor Gespanntheit als phonologische Größe zu betrachten hat.

(15) Vokalkürzung nach Hall (1992: 32)<sup>32</sup>



<sup>32</sup> Tiefgestelltes ‚w‘ neben dem Silbenknoten soll auf die metrische Schwäche der Silbe hinweisen.

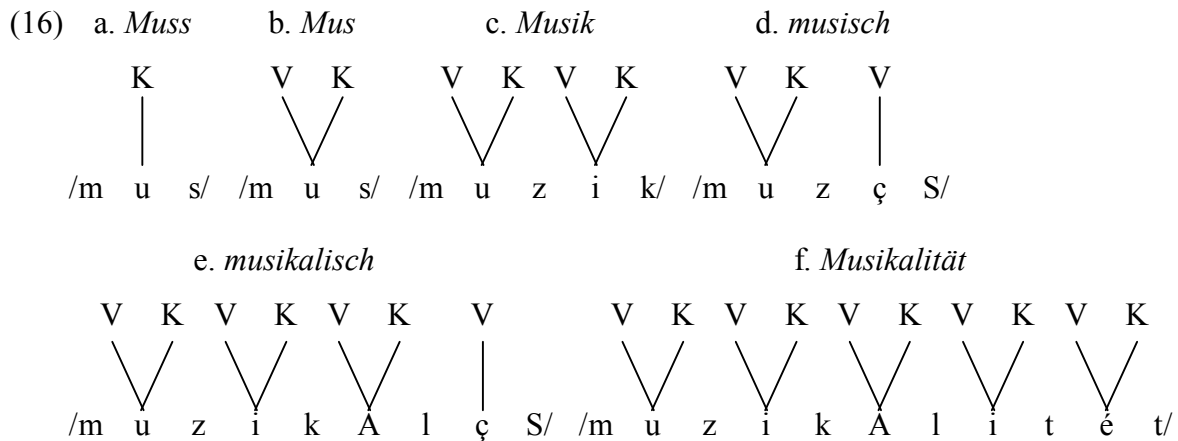


N

X X

[-kons]

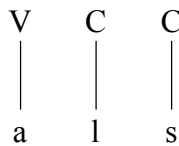
In *Muss* ist demnach ein zugrunde liegend mit einer skeletalen Position assoziierter Vokal anzunehmen (vgl. (16)a), der von den phonetischen Realisierungsregeln bzw. im Falle der Gespantheit eventuell von einer phonologischen *default*-Regel als der kurze ungespannte Vokal [ɔ̃] interpretiert wird. In *Mus*, *Musik*, *musisch*, *Musiker*, *musikalisch*, *Musikalität* ist der erste Vokal dagegen zugrunde liegend mit zwei skeletalen Positionen verbunden (vgl. (16)b-f), er wird unter Akzentuierung (also in *Mus*, *musisch* und *Musiker*) von den phonetischen Realisierungsregeln bzw. bei der Gespantheit eventuell im Laufe der Derivation als langes gespanntes [u:], sonst (unter Akzentlosigkeit) dagegen als kurzes gespanntes [u] erklärt.



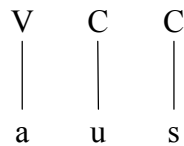
Eine solche quantitative Beschreibung der Vokalopposition in (6) hat zahlreiche Vorteile: die oben genannten mit dem qualitativen Ansatz verbundenen Schwierigkeiten werden in ihr substanzlos. Einerseits erübrigt sich hier das phonetische Problem der niedrigen Vokale: zwischen /A:/ und /a/ wird primär genauso ein Quantitätsunterschied angenommen wie zwischen /i:/ und /ç/, /e:/ und /E/ oder /u:/ und /ü/ usw. Dass dabei die Quantität im Bereich der niedrigen Vokale mit weniger ausgeprägten (bei manchen Sprechern eben keinen) qualitativen Differenzen einhergeht als im Bereich der nicht-niedrigen Vokale, ist eine für die phonologische Beschreibung nur weniger interessante Frage. Andererseits lassen sich in diesem Modell die oben genannten phonotaktischen Regularitäten der betreffenden beiden Vokalklassen des Deutschen adäquat zum Ausdruck bringen. Die phonotaktische Äquivalenz der Langvokale, Diphthonge und Kurzvokal+Konsonant-Sequenzen kann mit der

Quantitätsanalyse und der CV-Phonologie adäquat beschrieben werden (vgl. Wiese 1988: 62): nach der vokalischen Position (V) stehen in der Silbe des Deutschen genau zwei konsonantische Positionen (K) auf der skeletalen Schicht. Bei den Kurzvokal+Konsonant-Sequenzen nimmt der Kurzvokal die vokalische, der Konsonant die erste konsonantische Position ein (vgl. (17)a). Der silbische Diphthongteil besetzt die vokalische und der unsilbische die erste konsonantische Position (vgl. (17)b). Langvokale werden schließlich mit der V- und der ersten K-Position assoziiert (vgl. (17)c). In allen drei Fällen bleibt in der Silbenstruktur genau eine postvokalische K-Position frei, die in den Beispielen von (17) mit dem Konsonanten /s/ besetzt wird.<sup>33</sup>

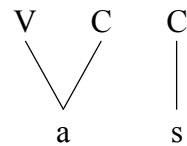
(17) a. Kurzvokal+Konsonant



b. Diphthong



c. Langvokal



Die phonotaktische Regularität der Kurzvokale, dass sie nur in geschlossenen Silben vorkommen kann, wird mit der oben bereits erwähnten Minimalitätsbedingung für die Silbenstruktur ausgedrückt, die jedoch auf der zugrunde liegenden Quantität sowie auf deren adäquater Repräsentation in der CV-Phonologie beruht. So besitzt nach Wiese (1988: 67f) der Silbenkern im Deutschen immer zwei Positionen: eine V- und eine C-Position. Diese Bedingung macht sicher, dass ein einziger Kurzvokal (der wie oben in (14)b gezeigt zugrunde liegend mit einer V-Position assoziiert ist) alleine keine Silbe bilden kann: die K-Position muss dann mit dem postvokalischen Konsonanten verbunden werden, auch wenn dieser dadurch ambisilbisch (zu zwei Silben gehörig) wird.

Und schließlich kann die tendenzielle ‚Gespanntheitspolarität‘, nach der Kurzvokale vorzugsweise vor Fortisobstruenten, Langvokale vor Lenisobstruenten stehen, in der CV-Phonologie plausibel erfasst werden, indem die erste postvokalische K-Position als ‚Stärkeposition‘ erklärt wird, in der Lenisobstruenten nur ausnahmsweise vorkommen (vgl. Becker 1998: 53).

Doch scheint eine quantitative Beschreibung ebenfalls nicht ganz unproblematisch. Sie ist nämlich bezüglich der Behandlung der Vokaldauer dem Qualitätskonzept weit unterlegen.

<sup>33</sup> Die letzten beiden (koronalen) Konsonanten in *Herbst*, *Obst* und *raupst* werden als extrasilbisch (d.h. außerhalb der Silbenstruktur liegend) betrachtet (vgl. Wiese 1988: 94).



Vokallänge ist im Deutschen ein Akzentphänomen – wie das zahlreiche Autoren von Moulton (1962) bis Vennemann (1991) behaupten. Durch die Annahme der Quantität als einzig oder primär relevanten phonologischen Größe bleibt diese m.E. sehr wichtige Generalisierung verdeckt. Auch mögliche Abhilfen ändern daran nicht viel: Durch Wieses (1988) Trennung zwischen phonologischer Quantität und phonetischer Dauer wird erlaubt, dass die phonetische Realisation eine zugrunde liegende phonologische Eigenschaft überschreibt, woraus „kurze Langvokale“ (Becker 1998: 52) resultieren. Halls (1992) Kürzungsregel (vgl. (15)) bedeutet dabei nur eine minimale Änderung: Zwar wird dadurch der phonetischen Realisation keine ‚Phonologie überschreibende‘ Funktion zugesprochen, aber im Endeffekt entsteht dasselbe wie bei Wiese (1988): sehr viele der zugrunde liegend langen Vokale (vgl. diejenigen in den Beispielen unter (16)) erscheinen auf der Oberfläche kurz. Dazu kommt noch, dass Halls (1992) Beschreibung im Gegensatz zu Wieses (1988) Analyse nicht einmal die Möglichkeit bietet, Gespanntheit als phonetische Größe aufzufassen, da sie diese Eigenschaft noch vor der Anwendung der Kürzungsregel benötigt.

Andererseits lässt sich mit Vennemann (1991ab, 1992, 1994) und Becker (1996ab, 1998) annehmen, dass das Deutsche eine besondere Art von Quantität hat, die nicht mit der in den ‚echten‘ Quantitätensprachen wie etwa dem klassischen Latein, dem Tschechischen, dem Finnischen oder eben dem Ungarischen gleichzusetzen ist. Betrachtet man die Interaktion von Akzent, Vokal- und Konsonantenlänge (wobei letzterer im Deutschen die Ambisilbizität gleichkommt), so ergibt sich eine typologische Verwandtschaft des Deutschen mit den nordwest-germanischen Sprachen, etwa dem Norwegischen oder Schwedischen, in denen Vokallänge unter Akzent nur unter ganz spezifischen Bedingungen möglich ist. Daher wäre eine Beschreibung der fraglichen Vokalopposition des Deutschen, die a. diese typologischen Aspekte berücksichtigt, sowie b. die Vorteile (nicht jedoch die Nachteile) der Qualitäts- und Quantitätsanalyse verkörpert, den besprochenen Analysemöglichkeiten jedenfalls vorzuziehen. Im Folgenden wollen wir das Silbenschnittkonzept vorstellen und zeigen, dass es diesen Anforderungen durchaus gerecht werden kann.

### **3.2 Zur Definition des Silbenschnitts**

Gewöhnlich wird als erste Quelle des Silbenschnittkonzepts Sievers (1901) betrachtet. Jedoch gab es auch vor ihm zumindest silbenschnittähnliche Konzepte – wie das neulich Restles

(1998) umfangreiche Forschung ergeben hat.<sup>34</sup> Sievers gilt somit nicht als der ‚Erfinder‘, sondern vielmehr als ‚Verbreiter‘ des Silbenschnittskonzepts.

Er definiert den Silbenschnitt („Silbenaccent“) als „die Art wie oder unter welchen Druckverhältnissen die Silbe vom Silbengipfel ab ihr Ende erreicht“ (Sievers 1901: 222) und unterscheidet zwischen stark und sanft geschnittenem Silbenaccent. Der Vokal im stark geschnittenem Silbenaccent wird „durch den folgenden Consonanten in einem Moment abgelöst, wo er noch voll und kräftig ertönt (unmittelbar hinter dem Silbengipfel), der jähe Absturz der Expiration fällt in den oder die silbenschiessenden Consonanten, die daher kräftig beginnen, aber mehr oder weniger abrupt endigen“ (222f.). Beim schwach geschnittenen Silbenaccent beginnt die Artikulation des Konsonanten erst, wenn der Vokal „bereits deutlich geschwächt ist (also eine merkbare Zeit nachdem der Silbengipfel passiert ist); der Consonant setzt daher auch mit nur mässiger Stärke ein, kann aber bei dem langsamern Decrescendo der Silbe deutlich und bequem ausklingen“ (223).

Aus den obigen Zitaten geht hervor, dass Sievers die Intensität der Artikulation des Nachfolgekonsonanten (Fortischarakter beim stark und Lenischarakter beim schwach geschnittenen Silbenaccent) als eine Begleiterscheinung des Silbenschnitts betrachtet. Er sieht zwar keinen systematischen Zusammenhang zwischen Vokaldauer und Silbenschnitt, gibt aber als Normalkonstellation den sanften Schnitt mit Langvokal und den starken Schnitt mit Kurzvokal an, wobei er Letzteres damit begründet, dass „es nicht üblich ist, den Vocal in voller Stärke längere Zeit auszuhalten“ (223).<sup>35</sup>

Jespersen (1932) übernimmt im Großen und Ganzen das Sieversche Konzept, gibt jedoch als definierendes Kriterium des Silbenschnitts im Gegensatz zu Sievers nicht „die Art wie [...] die Silbe vom Silbengipfel ab ihr Ende erreicht“ (Sievers 1901: 222), sondern betont die Interaktion zwischen Vokal und Folgekonsonanz und führt im Zusammenhang damit eigene Termini ein. Den scharfen Schnitt bezeichnet er als ‚festen Anschluss‘ zwischen Vokal und Folgekonsonant, der eintritt, wenn der Folgekonsonant „schnell kommt“ und „den Vokal in einem Augenblick ab[bricht], wo dieser am kräftigsten gesprochen wird“ (202). Der sanfte

---

<sup>34</sup> So kann als das wahrscheinlich älteste silbenschnittähnliche Konzept Ickelsamers ‘Teutsche Grammatica’ (vgl. Ickelsamer 1534<sup>3</sup>) erachtet werden.

<sup>35</sup> Restle (1998) bemerkt, dass Sievers Erläuterungen zur gegenläufigen Konstellation (d.h. starker Schnitt mit Langvokal bzw. sanfter Schnitt mit Kurzvokal) gewissermaßen unverständlich sind, sowie dass ein Vergleich dieser Textteile mit anderen Stellen zu einem völlig anderen Bild von der Auffassung des Sieverschen Silbenschnittkonzepts führt (vgl. Restle 1998: 8ff.).

Schnitt erscheint bei ihm als ‚loser Anschluss‘ zwischen Vokal und Folgekonsonant, u.z. wenn der Folgekonsonant „erst einige Zeit nach der kräftigsten Aussprache des Vokals kommt, wenn der Vokalklang also schon vor Eintritt des Konsonanten etwas geschwächt ist“ (ebda).

Trubetzkoy (1939) akzeptiert Jespersens Definition des Silbenschnitts, er geht also von der Interaktion zwischen Vokal und Folgekonsonant aus. Er übernimmt auch Jespersens Bezeichnungen ‚fester‘ bzw. ‚loser Anschluss‘, entwickelt dabei aber auch eigene Termini: an manchen Stellen ist von ‚scharfem‘ vs. ‚sanftem Silbenschnitt‘<sup>36</sup> bzw. vom ‚geschnittenem‘ und ‚ungeschnittenem Vokal‘ die Rede. Nach Trubetzkoy kommt ‚fester Anschluss‘ vor, wenn „der Konsonant in einem Augenblicke ein[setzt], wo der Vokal noch nicht den Höhepunkt seines normalerweise steigendfallenden Ablaufes überschritten hat“ (Trubetzkoy 1939: 196). Im Falle des ‚losen Anschlusses‘ läuft der Vokal dagegen „noch vor dem Einsatz des Konsonanten zur Gänze“ ab (ebda).

Aus der oben skizzierten Eigenart der Trubetzkoy'schen Phonologie folgt, dass er anders als Sievers die Vokaldauer als eine phonetische Begleiterscheinung der Silbenschnittkorrelation betrachtet:

Wenn dabei der Vokal mit festem Anschluß kürzer als der Vokal mit losem Anschluß ist, so ist dies nur eine phonetische Folgeerscheinung. (ebda)

Er gibt auch den Grund für diese Korrelation zwischen Silbenschnitt und (phonetischer) Vokaldauer an:

Der feste Anschluß ‚schneidet‘ sozusagen das Ende des Vokals ab und daher muß der so ‚geschnittene‘ Vokal kürzer als der normale, ungeschnittene Vokal sein. (ebda)

Die phonologische Irrelevanz der Vokaldauer zeigt sich nach Trubetzkoy deutlich in offenen unbetonten Silben, in denen sanft geschnittene Kurzvokale vorkommen, als Beispiel gibt er *le-béndig*, *Ho-lúnder*, *spa-zíeren* und *Ka-pi-tán* an. Dies legt nun nahe, dass Vokale mit festem Anschluss immer kurz sind, während solche mit losem Anschluss in Abhängigkeit von der Betonung lang (betont) oder kurz (unbetont).

Schon aus der eben zitierten Formulierung („der normale, ungeschnittene Vokal“) geht hervor, worauf der Autor auch explizit eingeht, dass der sanfte Schnitt das merkmallose Glied

---

<sup>36</sup> Dieses Terminuspaar geht – wie Restle (1998: 30) bemerkt – offensichtlich auf Sievers' „stark“ und „schwach geschnittenen Silbenaccent“ zurück.

in der Silbenschnittkorrelation darstellt. Dies zeigt sich auch bei der Aufhebung der Korrelation im Auslaut und vor Vokalen, wo – in Abwesenheit eines Folgekonsonanten – der merkmallose, sanft geschnittene Vokal vorkommt.

Vennemann (1991a), der das Konzept der Silbenschnitte nach langer Pause wieder aufgreift, stellt in den obigen Definitionsvorschlägen erhebliche Defizite fest. Aus den Erörterungen Sievers' schließt er darauf, dass dieser beim scharfem Schnitt eine steigende Intensität (Crescendo) vom Silbengipfel bis zum Silbenende und beim sanftem Schnitt einen Wechsel von steigender zu fallender Intensität (also von Crescendo zu Decrescendo) noch vor dem Silbenende annimmt. Er wendet jedoch gegen Sievers ein, dass es sogar unter seinen Beispielen Silben gibt, die ins Decrescendo übergehen, jedoch nicht scharf geschnitten sind (z.B. *voll*). Seiner Ansicht nach ist einzig und allein wichtig, wie der postvokalische Konsonant vom Silbengipfel erreicht wird (ob also im Crescendo oder Decrescendo), und nicht wie das Silbenende, wie Sievers behauptet (218). Jespersen und Trubetzkoy, die – wie gezeigt – bei der Silbenschnittkorrelation die Rolle des Folgekonsonanten betonen und also unter diesem Kontrast eine Opposition zwischen dem festen und losen Anschluss eines vokalischen Silbenträgers an einen folgenden Konsonanten verstehen, wirft Vennemann vor, dass eine Silbe auch dann Schnitteigenschaften aufweisen kann, wenn kein Nachfolgekonsonant vorhanden ist (ebda). Aus diesen beiden Kritikpunkten ergibt sich für ihn als Definitionskriterium des Silbenschnitts die Weise, wie das Ende des Nukleus erreicht wird. Endet der Nukleus der Silbe auf ein Crescendo und fällt das Decrescendo auf die Silbenkoda oder die nachfolgende Silbe, liegt scharfer Silbenschnitt vor. Beginnt dagegen das Decrescendo bereits im Nukleus, handelt es sich um sanften Silbenschnitt (219).

Becker (1996a, b und 1998) geht zu einer Definition seines Silbenschnittbegriffes von einer besonderen Ausprägung der Vokalquantität in den Silbenschnittsprachen im Gegensatz zu den Quantitätssprachen aus, wobei er darunter folgendes versteht:

Die Besonderheit besteht darin, daß der Kurzvokal durch den folgenden Konsonanten 'abgeschnitten' wird. Die Artikulationsbewegung ist beim Kurzvokal eine ballistische Bewegung, die durch die überlagerte Artikulation des folgenden Konsonanten abgefangen wird, während die Artikulation in einem Langvokal ausläuft, wobei ein möglicher folgender Konsonant lose angeschlossen wird – im Gegensatz zu dem festen Anschluß des Konsonanten nach Kurzvokal. (Becker 1996a: 4)

Indem er also die Rolle des bei den Kurzvokalen obligatorischen, bei den Langvokalen dagegen bloß möglichen Folgekonsonanten für die Vokalquantität und -qualität in den

Silbenschnittsprachen betont, vertritt er eine angesichts der Kritik von Vennemann revidierte Version des Jespersen-Trubetzkoy'schen Silbenschnittkonzepts.

Da – wie das im nächsten Abschnitt gezeigt wird – die von Sievers und Vennemann angenommenen Beziehungen zwischen den Energiekonturen den beiden Silbenschnittarten in Spiekermanns (2000) experimentalphonetischer Untersuchung auf keinerlei Weise nachgewiesen werden konnten, in Beckers Ansatz die von Vennemann genannten Mangelhaftigkeiten sowohl der Bestimmung von Sievers als auch der von Jespersen und Trubetzkoy beseitigt werden, wollen wir uns hier diesen Ansatz anschließen. Silbenschnitt soll daher als vorhandene oder fehlende Interaktion zwischen Vokal und Folgekonsonanz, eine Art Anschlussfähigkeit des Vokals, oder besser die Notwendigkeit des festen Anschlusses bei der Vokalklasse in (6)b im Gegensatz zu der in (6)a, definiert werden. Aus diesen Erörterungen und unserer Definition des Silbenschnitts (oder Anschlusses) geht deutlich hervor, dass in diesem Konzept genau die Beobachtung als phonologisch einzig relevant erklärt wird, dass ungespannte Kurzvokale im Deutschen nur in geschlossenen Silben vorkommen können, wobei die Schließung der Silbe mit einem solchen Vokal sogar ‚virtuell‘, d.h. durch Gelenkbildung (Ambisilbizität) gewährleistet wird. Wie im vorausgehenden Abschnitt gezeigt, will die quantitative Analyse diesem Phänomen durch eine der Silbenstruktur gestellte Minimalitätsbedingung Rechnung tragen, nach der das Deutsche einen verzweigenden Silbenkern hat, sodass auf den ungespannten Kurzvokal obligatorisch ein weiteres Segment folgen muss. Der Grundgedanke des Silbenschnittkonzeptes ist dagegen der, dass dieses Phänomen keine silbenstrukturelle Bedingung darstellt, sondern primär für die Vokalopposition in (6) verantwortlich ist. Die Opposition zwischen sanftem und scharfem Silbenschnitt (losem und festem Anschluss) lässt sich – analog zu (12) und (13) – in einer ersten Annäherung wie in (18) darstellen.

(18)	a. <sup>37</sup>	b.
	/i/	/i-/
	/y/	/y-/
	/e/	/e-/
	/ë/	/ë-/
	/A/	/A-/
	/u/	/u-/

---

<sup>37</sup> Zur Wahl der hier verwendeten Transkriptionssymbole vgl. Fn. 28.

/o/      /o-/

(18) zeigt, dass sich die beiden Vokalreihen (a und b) ähnlich wie bei der Annahme eines Quantitätskontrastes in (13) segmental-phonologisch voneinander in keinerlei Weise unterscheiden. Dass der einzige Unterschied zwischen ihnen darin besteht, dass die Vokale der Spalte b nie ‚selbstständig‘ vorkommen, d.h. dass ihnen in der Silbe mindestens ein weiteres Segment folgen muss, wird durch den waagerechten Strich hinter den Vokalsymbolen der Spalte b. zum Ausdruck gebracht. Auf eine phonologisch mehr motivierte autosegmentale Repräsentation des Silbenschnitts wird weiter unten in 3.4 noch detailliert eingegangen. Vorher muss aber die Frage nach dem phonetischen Korrelat des Silbenschnitts geklärt werden.

### 3.3    **Phonetisches Korrelat des Silbenschnitts**

Aus dem vorausgehenden Abschnitt geht hervor, dass die Definitionsversuche des Silbenschnitts auf zahlreichen phonetischen Annahmen stützen. Somit stellt sich die Frage, ob sich diese Annahmen experimentell-phonetisch bestätigen lassen. Seit den 40-er Jahren des 20. Jahrhunderts wurden zahlreiche phonetische Untersuchungen durchgeführt, die sich auf den Nachweis dieser phonetischen Annahmen gerichtet haben. Im vorliegenden Abschnitt wollen wir einen Überblick über diese Experimente bieten. Dabei wollen wir unsere Aufmerksamkeit besonders auf die neueste dieser Untersuchungen richten, die im Gegensatz zu allen früheren Experimenten ein eindeutiges Korrelat des Silbenschnitts ergeben hat.

Das erste Experiment zum Silbenschnitt führt Eli Fischer-Jørgensen (vgl. Fischer-Jørgensen 1941)<sup>38</sup> durch, die den Intensitätsverlauf der betonten Vokale bei zwei Versuchspersonen untersucht, da der Unterschied zwischen den beiden Silbenschnittarten des Deutschen nach Sievers in der unterschiedlichen Lage des Silbengipfels (im Vokal beim sanften bzw. im Folgekonsonanten beim scharfen Schnitt) liegen soll. Jedoch kann sie keine signifikanten Unterschiede in der Lage des Intensitätsmaximums bei Lang- und Kurzvokalen feststellen. Zu ähnlichem Ergebnis kommen auch von Essen (1962) und Jørgensen (1969b), von Essen findet z.B., dass der Konsonant „in jedem Falle erst nach der Überschreitung des letzten Intensitätsmaximums ein[setzt]“ (592).

---

<sup>38</sup> Zitiert nach Ramers (1988: 108).

Jørgensen (1969) setzt sich außerdem auch „mit der anderen Hälfte der Sieversschen Theorie“ auseinander, „daß der folgende Konsonant beim festen Anschluß kräftiger anfängt als beim losen Anschluß“ (181). Dabei findet er zwar keine eindeutigen Differenzen, wohl aber zwei, wenn auch schwache Tendenzen: erstens ist die Intensität des Konsonanten nach Kurzvokal oft größer als nach Langvokal, zweitens sind Langvokale oft kräftiger und Kurzvokale schwächer als die Folgekonsonanten. Die erste Tendenz zeigt sich bei einem paarweisen Vergleich von Wörtern (wie z.B. *Höhle* vs. *Hölle*) bei drei Versuchspersonen relativ deutlich. Jedoch ist vor einer Überinterpretation dieser Ergebnisse zu warnen, da diese einerseits äußerst instabil sind und da andererseits die Konsonantenintensität nur bei den drei Lauten *l*, *m*, *n* gemessen werden konnte, bei „folgendem *s* z.B. gibt es keine Unterschiede, und bei den Verschlusslauten ist die Implosion wahrscheinlich (sic) so schwach, daß sie von keiner perzeptorischen Bedeutung sein kann“ (183f). Diese von Jørgensen als tendenziell festgestellten Unterschiede in der Konsonantenintensität bedürfen nach Fischer-Jørgensen jedoch weiterer Untersuchungen (vgl. Fischer-Jørgensen 1969: 163).

Von Essen entdeckt während seines oben zitierten Experiments einen interessanten Zusammenhang zwischen der Anschlussart und dem „Gefälle des Abklingvorgangs“ des Vokals zum Folgekonsonanten: er konstatiert ein deutlich geringeres Abklinggefälle bei Langvokalen als bei Kurzvokalen (von Essen 1962: 592). Zur numerischen Erfassung dieses Sachverhalts nimmt er zunächst ein durchschnittliches, also geradlinig verlaufendes Gefälle an und errechnet den Neigungswinkel dieses Gefälles und der Nullinie. Da aber die Hüllkurve oberhalb und unterhalb der Nullinie meistens asymmetrisch verläuft, errechnet er beide Neigungswinkel. Die so gewonnene Summe nennt er *Konvergenz*. Der Autor fasst seine Ergebnisse folgendermaßen zusammen:

Der Unterschied zwischen dem scharfen und sanften Schnitt besteht „in den (sic) *Grade der Konvergenz*: der Abklingvorgang erfolgt bei Kurzvokalen rascher, der Abfall zum folgenden Konsonanten ist steiler, der Übergang abrupter als bei Langvokalen. (595)

Auch wenn von Essens Ergebnisse auf den ersten Blick als plausibel erscheinen, sind sie nicht unproblematisch. So kritisiert Eli Fischer-Jørgensen in einem dem Aufsatz beigelegten Diskussionsbeitrag (von Essen 1962: 595-597) das vom Autor angewendete Messgerät, das wegen seines äußerst begrenzten Messintervalls meistens nur eine Messung des Intensitätsverlaufs des ersten Formanten ermöglicht.<sup>39</sup> Jørgensen (1969a: 167) bemängelt

---

<sup>39</sup> Vgl. auch Fischer-Jørgensen (1969: 140) sowie Jørgensen (1969b: 167).

außerdem von Essens relativ kleines Korpus (insgesamt 32 Wörter), die dazu von nur einem Sprecher gesprochen wurden. Er nennt ferner auch Fälle, bei denen mit der Messmethode von Essens gar kein Intensitätsgefälle errechnet werden kann.<sup>40</sup> Diese Kritikpunkte berücksichtigend führt er eigene Untersuchungen durch, bei denen er jedoch keine nennenswerten Differenzen im Intensitätsfall zwischen scharfem und sanftem Schnitt finden kann, woraus er folgenden Schluss zieht:

In dem hier vorgelegten Material wurde in den Diagrammen nie eine Streuung gefunden, die mit von Essens Theorie übereinstimmt. (174)

Schließlich sei hier von den früheren experimentalphoneischen Untersuchungen zum Silbenschnitt Eli Fischer-Jørgensens Experiment aus dem Jahre 1969 zu nennen, denen angesichts des Umfangs des analysierten Korpus und der großen Anzahl der untersuchten Parameter (Luftstrom, Luftdruck, Lippendruck sowie Dauer) besonders große Bedeutung beigemessen werden kann.

Ihre Messungen bezüglich des Luftstroms ergeben, dass dieser bei der Sequenz Kurzvokal+Konsonant wesentlich stärker ist als bei der Verbindung Langvokal+Konsonant. Da aber in der überwiegenden Mehrheit der untersuchten Kurzvokal+Konsonant-Sequenzen bereits der prävokalische Konsonant einen stärkeren Luftstrom aufweist, kommt Fischer-Jørgensen zu dem Schluss, „daß der Luftstrom eher mit der Silbendauer als mit dem Anschluß zusammenhängt“ (149). In Bezug auf den intraoralen Luftdruck stellt sie ein etwas höheres Luftdruckmaximum des folgenden Konsonanten nach Kurzvokal als nach Langvokal fest, die Unterschiede sind jedoch im Großen und Ganzen relativ klein. Ihre Lippendruckmessungen zeigen einen größeren Organdruck nach kurzem als nach langem Vokal, welcher Unterschied statistisch signifikant ist, jedoch hält sie den Parameter Organdruck für kein geeignetes Korrelat der Silbenschnittopposition, da dessen akustisches Gegenstück nur schwer zu finden sei (153). Bei der Analyse der Dauerverhältnisse stellt sie einen deutlichen Unterschied fest: die Langvokale waren nach ihren Messungen etwa doppelt so lang wie die Kurzvokale. Im Bereich der postvokalischen Konsonanten findet sie die umgekehrte Tendenz: Konsonanten sind nach Kurzvokalen etwas länger als nach Langvokalen. Da diese quasi-komplementäre Verteilung von Vokal- und Konsonantenlänge als potentielles phonetisches Korrelat des

---

<sup>40</sup> Es handelt sich dabei um Fälle, bei denen „der Intensitätsgipfel mit dem Vokalschluß zusammenfällt, so daß es kein Intensitätsgefälle des Vokals vor dem folgenden Konsonanten gibt“. Als Beispiel gibt er ein Ozillogramm des Wortes *Hölle* an, „wo die Intensität durch den ganzen Vokal steigt“ (ebda).



Silbenschnitts gelten könnte, berechnete Fischer-Jørgensen den Quotienten Konsonantenlänge/Vokallänge für alle Einzelwortdurchschnitte. Dabei stellte sich jedoch heraus, dass dieser Quotient äußerst variabel ist, „offenbar weil verschiedene zum Teil unabhängige Faktoren diesen Quotienten beeinflussen“ (155). So kommt sie zu dem Schluss:

Da der Unterschied der Vokaldauer wie genannt viel stabiler ist, ist es wahrscheinlich, daß die Vokaldauer wichtiger ist. (158)

Die Ergebnisse der angeführten experimentalphonetischen Untersuchungen zum phonetischen Korrelat des Silbenschnitts lassen sich nun folgendermaßen zusammenfassen. Das von Sievers angenommene phonetische Korrelat des Silbenschnitts (unterschiedliche Lage des Intensitätsmaximums) sowie von Essens Konvergenz-Theorie können eindeutig als falsifiziert betrachtet werden. Die Parameter Luftstrom, Luftdruck sowie Organdruck beim postvokalischen Konsonanten können ebenfalls nicht als phonetische Korrelate gelten. Die Intensität des postvokalischen Konsonanten kann so lange nicht diese Rolle übernehmen, bis sie von weiteren Untersuchungen bestätigt wird. Als der einzig stabile phonetische Unterschied zwischen den beiden Silbenschnittarten scheinen die Dauerverhältnisse zu sein.<sup>41</sup>

Diese negativen experimentalphonetischen Ergebnisse bezüglich des phonetischen Korrelats der Silbenschnittopposition führten in der phonologischen Beschreibung zu einem skeptischen Standpunkt gegenüber dem Silbenschnittkonzept, was beispielsweise auch aus Ramers' (1988) folgender Formulierung ersichtlich ist:

Die Anschlußart ist als distinktives Merkmal der Vokalopposition im Deutschen solange unbrauchbar, wie kein akustisches Korrelat dieses Gegensatzes gefunden werden kann. (122)

Gewissermaßen differenzierter äußert sich bezüglich des Problems des fehlenden phonetischen Korrelats des Silbenschnitts jedoch Becker, wenn er Folgendes schreibt:

---

<sup>41</sup> Die Relevanz der Vokaldauer bei der Wahrnehmung des Anschlusses bestätigen auch Fliflets (1962b) Experimente, bei denen der Autor Vokale am Anfang und postvokalische Konsonanten im Innern verkürzte bzw. verlängerte, und die so gewonnenen Stimuli einem umfangreichen Perzeptionstest unterwarf. Becker (1998: 60f.) zweifelt jedoch an der Aussagekraft der Flifletschen Untersuchungen, er setzt sich generell gegen Schneideexperimente, „da bei der Gestaltwahrnehmung fehlende Aspekte des Signals ergänzt werden“ (61). Restle (1998: 64) hält Beckers Kritik für gewissermaßen übertrieben, da Fliflets Ergebnisse eindeutig zu konsistent sind: „man beachte die spiegelbildlichen Auswirkungen von Verkürzung und Verlängerung sowohl von Vokal als auch Konsonant auf die Perzeption“ (ebda). Jedoch meint er ähnlich wie Becker, dass Fliflets Ergebnisse mit modernen Methoden überprüft werden sollten.

Man kann fast sagen, der Stand der Forschung ist, daß die Silbenschnittopposition kein akustisches Korrelat hat und daher trotz ihrer intuitiven Plausibilität unbrauchbar ist.

Zunächst ist eine Präzisierung vorzunehmen: Das akustische Korrelat der Silbenschnittopposition im Deutschen ist natürlich kürzere Dauer und Zentralisierung der Kurzvokale. Das ist aber nicht das Problem; gesucht wird nach dem akustischen Korrelat des Unterschieds von Silbenschnittsprachen und Quantitätssprachen. Kürzere Dauer und Zentralisierung der Kurzvokale sind auch für die Vokalopposition des Tschechischen charakteristisch [...], das nicht zu den Silbenschnittsprachen zählt, daher scheidet diese Merkmale aus. (58f)

Um das phonetische Korrelat des Silbenschnitts zu finden, sollte also nach Becker nicht das Deutsche für sich untersucht werden, viel mehr sollten das Deutsche und eine Quantitätssprache einer kontrastiven experimentellen Untersuchung unterzogen werden.

Spiekermann (2000) bietet genau eine solche Untersuchung. Der Autor untersucht dabei „die in der Forschungsliteratur häufig als mögliche Korrelate des Silbenschnittes genannten akustischen Parameter Formantfrequenz, Vokaldauer, Grundfrequenz und Energieverlauf“ (39). Als Untersuchungsmaterial dienen dabei zwei Korpora: das erste davon enthält 225 von einem männlichen Sprecher geäußerte *nonsense*-Wörter der phonologischen Struktur [gWK<sub>1</sub>VK<sub>2</sub>W], wobei V einen Vokal, K<sub>1</sub> und K<sub>2</sub> identische Konsonanten bezeichnen, und die fraglichen Wörter in den Satz *Ich habe \_\_\_\_\_ gesagt* eingebettet sind, wodurch sich aus dem Satzkontext ergebende potentielle Störungsfaktoren ausgeschlossen wurden. Das zweite Untersuchungskorpus bestand dagegen aus 8 natürlichen Gesprächsausschnitten aus Nachrichtensendungen, Talkshows und Fernsehreportagen. Mit der Untersuchung dieses Standardkorpus hatte der Autor die Absicht, die anhand des *nonsense*-Korpus gemachten Aussagen zu verifizieren bzw. zu falsifizieren.

Von den genannten akustischen Parametern wurden im Falle des Energieverlaufs drei Merkmale untersucht: die Anzahl der Energiemaxima (kurz: *E-Zahl*), die Position eines Energiemaximums im Vokal (kurz: *E-Pos*) sowie die Gesamtkontur, genauer die Frage, wie stark ein hohes Energiemaximum ausgehalten wird (kurz: *E-Halt*). Die Untersuchung der *E-Zahl* lässt sich damit begründen, dass Sievers (1901) und Vennemann (1991a) „von einem Fehlen eines Energieabfalls auf den betonten Vokal im Falle eines scharfen Silbenschnittes ausgehen“ (ebda), was eine geringere Anzahl der Energiemaxima bei scharfem als bei sanftem Schnitt erwarten lässt. Die Motivation für die Analyse der *E-Pos* betrifft alle Silbenschnittdefinitionen: wie im vorausgehenden Abschnitt gezeigt, wird oft über ein Abschneiden des Vokals durch den Folgekonsonten unter scharfem Schnitt im Vergleich zum sanftem Schnitt die Rede. Dies legt die Erwartung nahe, „daß bei scharf geschnittenen

Vokalen die Position eines Energiemaximums relativ weit hinten, bei sanft geschnittenen Vokalen jedoch in jedem Falle weiter vorne zu finden sein müßte“ (40). Zu einer numerischen Erfassung der Position des Energiemaximums werden die untersuchten Vokale in neu gleichmäßige Positionen eingeteilt. Schließlich lässt sich die Analyse des *E-Halt* mit der gängigen Annahme begründen, dass sanft geschnittene Vokale voll und kräftig ausgehalten werden (dass sie „austrudeln“ Maas: 1999: 176), im Gegensatz zu den scharf geschnittenen Vokalen, bei denen dies infolge des ‚Abschneidens‘ durch die Folgekonsonanz nicht möglich ist. Zu erwarten wäre daher ein starkes Halten eines hohen Energiemaximums unter sanftem und ein schwaches unter scharfem Schnitt. Zur Erfassung dieses Merkmals wurde der höchste Energiewert auf dem Vokal als Referenzwert genommen. Wenn die Energie während des Gesamtverlaufs des Vokals weniger als 5% von diesem Referenzwert sinkt, wurde ein starkes Halten eines hohen Energiemaximums festgestellt, bei einem Wert zwischen 5 und 10% spricht der Autor von einem leichten, bei einem über 10% dagegen um kein Halten.

Bei den anderen untersuchten Parametern wurden die Dauer und die Formantenstruktur zueinander in Beziehung gesetzt: zunächst wurde aus den ersten beiden Formanten eines Vokals ein Differenzwert gebildet, der sich dann auf einem zweidimensionalen Diagramm direkt mit der Dauer vergleichen lässt. Schließlich wurde die Grundfrequenz auf vier gleichmäßig auf dem Vokal verteilten Punkten gemessen. Die Untersuchung der Grundfrequenz lässt sich mit Maas & Tophinkes (1993) Annahme begründen, nach der unter scharfem Schnitt eine steigende, unter sanftem eine fallende  $F_0$ -Kontur vorliegt.

Die Ergebnisse der Untersuchung der Energieverläufe von Spiekermanns (2000) lässt sich wie in Tab. 12 zusammenfassen.

<i>Korpus</i>	<i>sanfter Schnitt</i>			<i>scharfer Schnitt</i>			<i>Differenz</i>		
	<i>E-Zahl</i>	<i>E-Halt</i>	<i>E-Pos</i>	<i>E-Zahl</i>	<i>E-Halt</i>	<i>E-Pos</i>	<i>E-Zahl</i>	<i>E-Halt</i>	<i>E-Pos</i>
<i>Nonsense</i>	1,88	2,59	4,76	1,14	2,17	2,67	0,73	0,41	2,09
<i>Standard</i>	1,19	2,46	5,02	0,93	1,95	3,39	0,26	0,51	1,63

Tab. 12 Ergebnisse der Energiemessungen von Spiekermann (2000: 47)

Diese Daten zeigen in der Tat deutliche Unterschiede zwischen sanft und scharf geschnittenen Vokalen, die den oben als Ausgangspunkt genommenen phonetischen Annahmen mehr oder

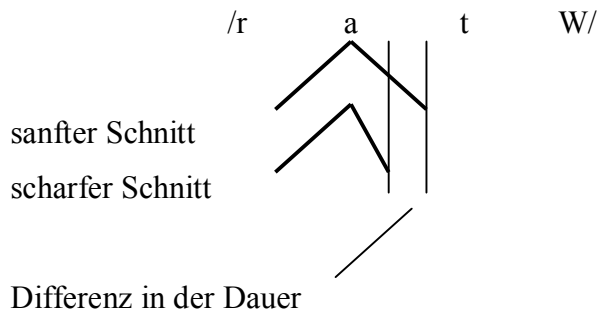
weniger entsprechen (vgl. 48). Erstens sind beim sanften Schnitt mehr Energiemaxima vorhanden als beim scharfen, was also auf ein Fehlen des Energieabfalls auf den betonten Vokal beim letzteren hinweisen könnte. Zweitens ist bei sanftem Schnitt ein starkes Halten eines hohen Energieniveaus festzustellen, was für das ‚Austrudeln‘ der sanft geschnittenen Vokale im Vergleich zum ‚Abschneiden‘ des scharf geschnittenen durch die Folgekonsonanz spricht. Drittens liegt beim sanften Schnitt das Energiemaximum weiter hinten im Vokal als beim scharfen, was mit der oben bereits zitierten gängigen Annahme von der entgegengesetzten Tendenz verträglich ist. Dies legt nach dem Verfasser folgende Modifizierung der phonetischen Beschreibung des Silbenschnitts nahe:

Während beim sanften Silbenschnitt der Vokal sein Energiemaximum langsam erreicht und sich anschließend langsam abschwächt, erreicht er beim scharfem (sic) Silbenschnitt sein Maximum sehr schnell und klingt dann aus. (49)

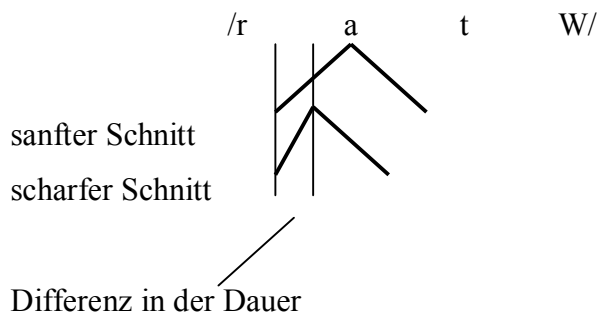
Diese Änderung des phonetischen Konzeptes lässt sich nach dem Autor graphisch wie in (20) verdeutlichen, zum Vergleich ist die gängige auf Sievers zurückgehende Annahme in (19) gegeben (vgl. 48).

Spiekermann (2000) schließt seine Darstellung der Ergebnisse der Untersuchung bezüglich der Energieverläufe damit ab, dass die Energiekurven die Silbenschnittunterschiede zu bestätigen scheinen: besonders das Merkmal *E-Halt* sei als stabiles phonetisches Korrelat der beiden Silbenschnitte zu betrachten, das durch die anderen beiden Energiemerkmale unterstützt werde (vgl. 58).

(19) Position des Energiemaximums nach Sievers



(20) Position des Energiemaximums nach der Untersuchung in Spiekermann (2000)



Spiekermanns (2000) Untersuchung ergab für die anderen beiden mit untersuchten akustischen Parameter Folgendes: Seine Dauer- und Formantmessungen bestätigte das weiter oben Gesagte: er stellt fest, „daß im Standarddeutschen sowohl eine qualitative als auch eine quantitative Unterscheidung der Varianten des Vokales gegeben ist“ (62). Seine Messungen bezüglich der Grundfrequenz liefern dagegen keine eindeutigen Ergebnisse, sodass er die von Maas & Tophinke (1993) vertretene Meinung über das Verhältnis zwischen Silbenschnittart und Grundfrequenz als falsifiziert betrachtet (vgl. 66).

Spiekermanns (2000) experimentalphonetische Untersuchung scheint für das Silbenschnittkonzept nicht nur deswegen äußerst relevant zu sein, weil sie in den Energiekurven eine phonetische Bestätigung der Silbenschnittopposition entdecken konnte, sondern auch, weil sie den in Becker (1998) formulierten Anforderungen Rechnung tragend seine Ergebnisse auch an nicht-Silbenschnittsprachen testet. Seine Untersuchung der Vokale des Finnischen, einer der prototypischen Quantitätensprachen (vgl. Becker 1998) liefert Folgendes. Die Anzahl der Energiemaxima zwischen den Lang- und Kurzvokalen des Finnischen zeigt eine mit dem Deutschen vergleichbare Variation: in Langvokalen sind mehr Energiemaxima vorhanden als in Kurzvokalen. Dies lässt sich nach dem Autor dermaßen interpretieren, „daß die Anzahl der Energiemaxima in erster Linie von der Dauer des Vokals abhängt, bzw. daß eine starke Korrelation zwischen der Dauer des betrachteten Vokales und

dem Merkmal *E-Zahl* besteht“ (72). In Bezug auf das Merkmal *E-Halt* stellt er bei allen finnischen Vokalen ein starkes Halten eines hohen Energiemaximums fest, was also einen eindeutigen Unterschied im Vergleich zu den Ergebnissen bezüglich des Standarddeutschen darstellt (vgl. 73). Bei der Untersuchung der *E-Pos* ergab sich eine interessante Gegebenheit: die Vokale des Finnischen zeigen ein Maximum relativ weit hinten im Vokal, was der Autor dahingehend interpretiert, dass dadurch *E-Pos* als potentielles Korrelat des Silbenschnitts im Deutschen ausscheidet (ebda). Bezüglich der Dauer stellt er erwartungsgemäß große Differenzen fest, die anders als im Deutschen mit keiner ausgeprägten qualitativen Unterschiede einhergehen (vgl. 74).

Von den mit untersuchten zwei Sprachen Tschechisch und Französisch, konnte der Autor bei Ersterer ähnliche Verhältnisse wie im Finnischen feststellen, was die phonetischen Unterschiede zwischen Silbenschnitt- und Quantitätensprachen weiter untermauert, da das Tschechische ebenfalls zum zuletzt genannten Sprachtyp gehört. Die Untersuchung der französischen Vokale erweist sich deswegen als interessant und zugleich wichtig, da in dieser Sprache Vokalquantität eher eine untergeordnete Rolle spielt:

Grundsätzlich ist eine größere Vokaldauer nur in Verbindung mit längenden Konsonanten beobachtbar und tritt auch dann nur in „korrekter Sprechweise“ oder bei gewollter phonologischer Opposition [...] auf. (77)

Die Verhältnisse im Französischen bestätigen die bisherigen Ergebnisse. So hängt die Anzahl der Energiemaxima auch in dieser Sprache stark von der Vokaldauer ab, bzw. zeigen sich keine nennenswerten Unterschiede beim Merkmal *E-Halt*, was also eine Silbenschnittopposition eher ausschließt.

Zusammenfassend kann man zu Spiekermanns (2000) Ergebnissen Folgendes festhalten. Zwischen den gespannten langen und den ungespannten kurzen Vokalen des Deutschen besteht ein Unterschied im akustischen Merkmal *E-Halt*, d.h. bei Ersteren ist ein starkes Anhalten eines hohen Energiemaximums zu beobachten, während bei Letzteren das Energiemaximum schnell sinkt. Dass dies nicht mit der Vokaldauer abhängt, sondern ein Spezifikum des Gegenwartsdeutschen als eine nicht-Quantitätensprache darstellt, zeigt sich darin, dass sich die Vokale in den echten Quantitätensprachen durch ein gleichmäßig starkes Anhalten eines hohen Energiemaximums auszeichnen. Das scheint zwei Hypothesen bezüglich des Silbenschnitts zu bestätigen. Dass zwischen dem Deutschen als Silbenschnittsprache und den Quantitätensprachen der wichtigste phonotaktische Unterschied die Unmöglichkeit des Vorkommens eines kurzen ungespannten Vokals in offenen Silben im

Deutschen betrifft, legt einerseits nahe, dass für das fehlende starke Anhalten des hohen Energiemaximums bei diesen Vokalen in der Tat die Folgekonsonanz, besser der feste Anschluss verantwortlich ist, diese Folgekonsonanz ‚schneidet‘ zwar den Vokal nicht im ‚herkömmlichen‘ Sinne ‚ab‘, sondern wirkt auf ihn gewissermaßen als ‚energiewegziehend‘ aus. Dadurch wird das Silbenschnittkonzept, und genauer die Becker’sche Silbenschnittsdefinition bestätigt. Auf der anderen Seite findet in den Ergebnissen von Spiekermann (2000) auch die Erkenntnis Vennemanns (1991a) Bestätigung, dass in Sprachen, die keine Silbenschnittopposition kennen, alle Vokale sanft geschnitten sind: diese haben nämlich in Bezug auf das akustische Korrelat des Silbenschnitts *E-Halt* genau denjenigen Wert, der in Silbenschnittsprachen für den sanften Schnitt charakteristisch ist. Im nächsten Abschnitt soll nun die Frage nach einer adäquaten Darstellung des Silbenschnitts geklärt werden.

### **3.4 Zur Repräsentation des Silbenschnitts**

Gegenstand des vorliegenden Abschnitts ist die adäquate Darstellung der beiden Silbenschnitte im Deutschen. Als Ausgangspunkt dienen dabei die von den oben bereits zitierten Autoren entwickelten Repräsentationsvorschläge. Diese sollen zuerst kurz vorgestellt und kommentiert werden, um ausgehend von den dabei formulierten Kritikpunkten unseren eigenen Repräsentationsvorschlag darstellen zu können.

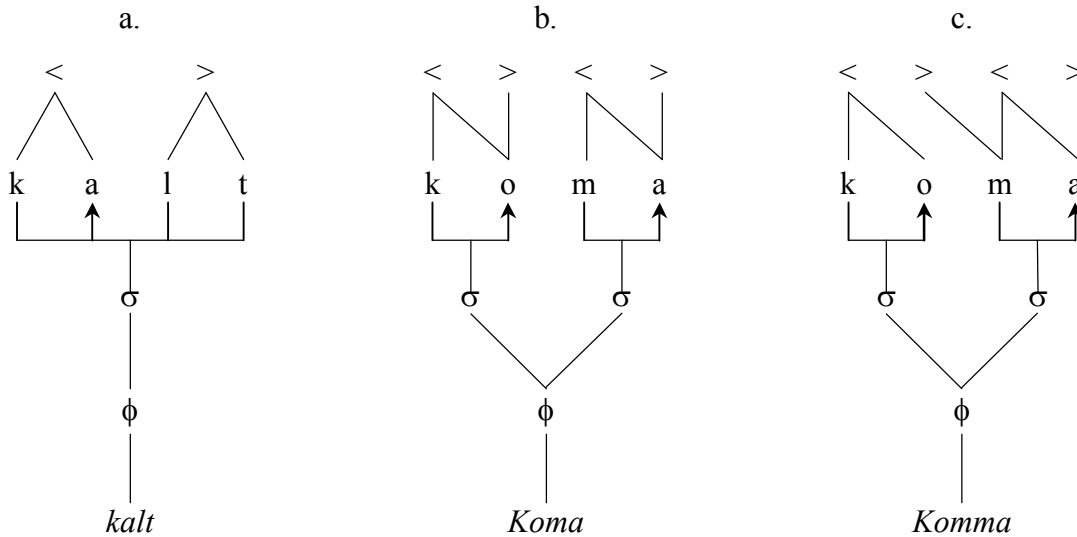
#### **3.4.1 Vennemann (1991b) und (1994)**

Vennemann (1991b) schlägt eine auf den Grundprinzipien der autosegmentalen Phonologie basierende Repräsentation vor. Die wesentlichen Züge dieser Repräsentation werden in Vennemann (1994) wie folgt zusammengefasst:

Dem jeweiligen Wort ist seine wortphonologische Form  $\phi$  zugeordnet, dieser die Folge der Silben  $\sigma$  [...], jeder von diesen ihr Morenraster, auf dem die Anzahl der zu der Silbe gehörigen Zeiteinheiten (Moren) und – durch den Pfeil – die Stelle des Silbengipfels ausgedrückt sind. Den Zeiteinheiten sind Sprachlaute zugeordnet. [...] Diese platzierten Sprachlaute wiederum sind durch Verbindungslinien auf eine bestimmte Weise dem Crescendo (<) und dem Decrescendo (>) der betreffenden Silbe zugeordnet. (Vennemann 1994: 8)

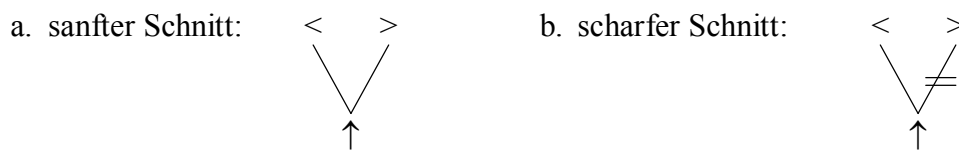
In (21) wird zeigt, wie die Silbenstruktur der Wörter *kalt*, *Koma* und *Komma* nach dieser Darstellungskonvention repräsentiert wird.

(21) Silbenstruktur der Wörter *kalt*, *Koma* und *Komma* (Vennemann 1991b: 91, 94)



Vennemanns Repräsentationsverfahren ermöglicht auch eine schematische Darstellung der beiden Silbenschnitte. Beim sanften Schnitt beginnt das Decrescendo bereits im Silbennukleus. Das macht eine Darstellung sinnvoll, in der der Gipfel ( $\uparrow$ ) sowohl mit dem Crescendo ( $<$ ) als auch mit dem Decrescendo ( $>$ ) verbunden ist (vgl. (22)a). Beim scharfen Schnitt endet der Nukleus auf Crescendo, was am besten durch die Verbindung des Gipfels ( $\uparrow$ ) mit dem Crescendo ( $<$ ) und durch das Verbot der Verbindung des Gipfels mit dem Decrescendo ( $>$ ) ausgedrückt wird (vgl. (22)b).

(22) Schematische Darstellung der beiden Silbenschnitte (Vennemann 1991b: 90)



Die oben skizzierte Darstellung der Silbenstruktur weist jedoch gewisse Redundanzen auf, deren Beseitigung der Autor in Vennemann (1994) unternimmt. Erstens erscheint die Anzeigung des Silbengipfels durch den Pfeil als redundant, denn der Silbengipfel ist „diejenige Stelle, auf die die letzte Verbindungslinie des Crescendos zuläuft“ (10). Eine zweite Redundanz ergibt sich aus der Anzeigung „des in phonologischen Kontexten selbstverständlichen Wortformknotens  $\phi$ „ (15). Die dritte und zugleich am meisten störende Redundanz ist jedoch die Markierung der Silbe einerseits durch den Silbenknoten  $\sigma$  und

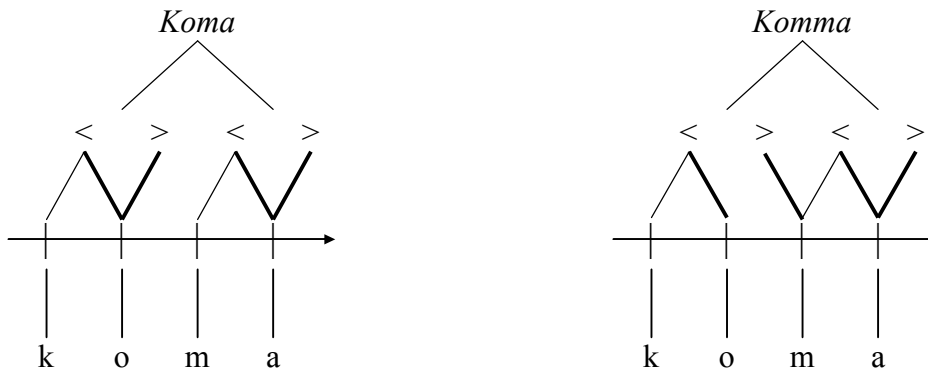


andererseits durch die Crescendo-Decrescendo-Paare. Da diese Paare aus Crescendo und Decrescendo für die Darstellung der Silbenschnitte unentbehrlich sind, kommt der Autor zum Schluss, den Silbenknoten und dadurch die ganze Silbenschicht aus der Darstellung zu eliminieren und also die Wortform direkt mit den Crescendo-Decrescendo-Konturen zu verknüpfen. Neben der Eliminierung der drei Elemente Silbengipfelmarker ( $\hat{\uparrow}$ ), Silben- und Wortformknoten ( $\sigma$  bzw.  $\phi$ ) wird der Darstellung auch ein Element hinzugefügt, nämlich der Zeitfeil zum Ausdruck der zeitlichen Linearität. Vennemann begründet die Erweiterung der Repräsentation durch den Zeitfeil mit folgenden Worten:

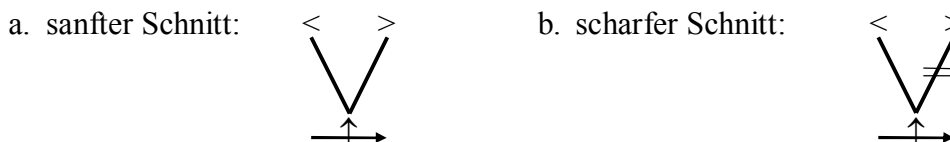
Die Konturen drücken von links nach rechts den (intuitiv aufgefaßten) Intensitätsverlauf in der Zeit aus. Auch die Sprachlaute schreiben wir konventionell von links nach rechts zum Ausdruck des Ablaufs in der Zeit. Man wird schon deshalb vermuten, daß es richtig ist, den Zeitfeil zwischen den Konturen und den Sprachlautfolgen anzuordnen. (12)

Somit lassen sich die obigen Beispielswörter *Koma* und *Komma* nach Vennemann (1994) wie in (23) darstellen. (24) enthält die schematische Darstellung der beiden Silbenschnitte nach der veränderten Konvention.

(23) Silbenstruktur der Wörter *Koma* und *Komma* (Vennemann 1994: 13)<sup>42</sup>



(24) Schematische Darstellung der beiden Silbenschnitte (Vennemann 1994: 13)

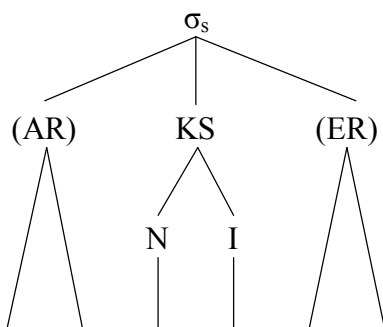


<sup>42</sup> Fettgedruckte Verbindungslinien sollen lediglich bei der Identifizierung der Silbengipfel helfen.

### 3.4.2 Becker (1996b, 1998)

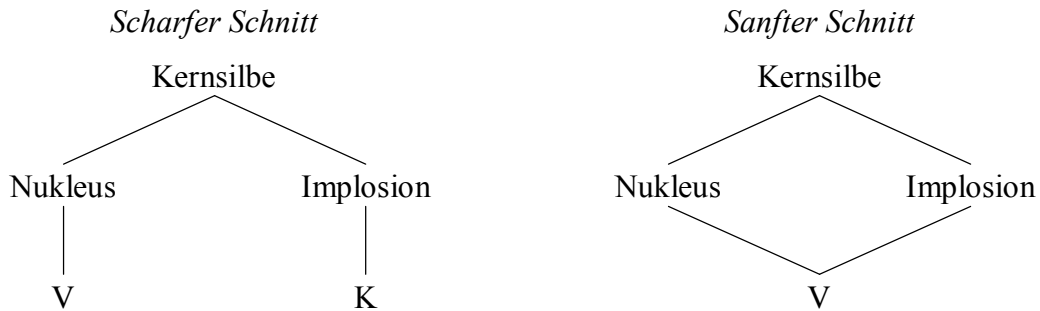
Beckers (1996b, 1998) Repräsentationsvorschlag geht von der Minimalitätsbedingung aus, nach der eine Tonsilbe im Deutschen obligatorisch einen Langvokal, einen Diphthong oder die Sequenz Kurzvokal+Konsonant enthält. Um diesem Sachverhalt gerecht zu werden, nimmt er eine Silbenkonstituente an, die er als Kernsilbe bezeichnet. Vor und nach der Kernsilbe können je zwei Konsonanten in der Tonsilbe vorkommen, die die beiden fakultativen Silbenränder Anfangsrand (AR) und Endrand (ER) bilden. Um die oben besprochene Äquivalenz zwischen Langvokal, Diphthong und Kurzvokal+Konsonant-Sequenz adäquat darzustellen, hat seine Kernsilbe ähnlich wie der Silbenkern des CV-phonologischen Silbenmodells von Wiese (1988: 67) eine verzweigende Struktur: die V-Position nennt Becker Nukleus, die C-Position Implosion. Beckers Implosionsposition kommt eine Sonderrolle zu, die durch mehrere Eigenschaften motiviert ist (vgl. Becker 1998: 75): a) sie ist obligatorisch, b) nur sie kann ambisilbisch sein, c) der Velarnasal [N] kann nur in dieser Position vorkommen, es sei denn, er ist ambisilbisch und besetzt daher sowohl die Implosionsposition der ersten als auch die Anfangsrandposition der zweiten Silbe, d) der Hauchlaut [h] kann in dieser Position nicht vorkommen und e) diese Strukturposition ist eine ‘Stärkeposition’, „d.h. in dieser Position werden die Sprachlaute allophonisch stärker artikuliert, Fortisobstruenten werden gegenüber Lenisobstruenten bevorzugt, *r* wird nicht vokalisiert oder verdrängt, die Laute sind beliebig dehnbar (*Wa::hn*, *Wan::d*, *Wat::te*) und Vokale tendieren zur Diphthongierung” (ebda). Die Implosionsposition spielt auch in der Repräsentation der beiden kontrastierenden Silbenschnitte eine große Rolle: beim sanften Schnitt wird der Vokal mit dem Nukleus und der Implosion assoziiert, beim scharfen Schnitt besetzt der Folgekonsonant die Implosionsposition. Somit hat die deutsche Tonsilbe die Struktur in (25) aus. (26) enthält die schematische Repräsentation der Silbenschnitte, (27) die Silbenstruktur der Wörter *kalt*, *Koma* und *Komma*.

(25) Struktur der Tonsilben (Becker 1996a: 12; 1998: 77, 80)

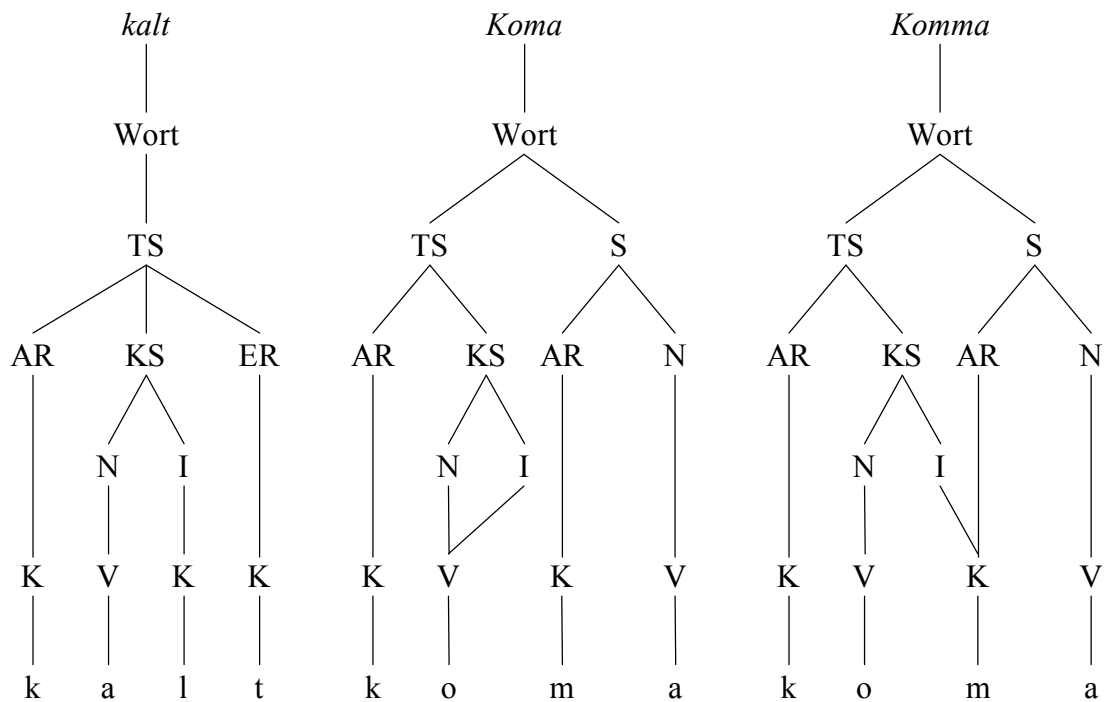


(K K) V K (K K)

(26) Repräsentation der Silbenschnitte (Becker 1996a: 11, 1998: 78)



(27) Silbenstruktur der Wörter *kalt*, *Koma* und *Komma* (Becker 1998: 79, 81)



Das bisher über Beckers Vorschlag zur Repräsentation des Silbenschnitts Gesagte gilt jedoch nur für die betonten Silben. Der Autor geht nämlich davon aus, dass die Opposition in (6) nur auf Tonsilben beschränkt ist. Dass diese Annahme nicht ganz falsch ist, zeigt die Tatsache, dass sich für die beiden Silbenschnitte zwar in betonter Position viele Minimalpaare finden lassen (vgl. (28)a), für die durch Ausscheidung der Vokaldauer in unbetonter Stellung entstandenen gespannt-ungespannten Vokalpaare jedoch keine.<sup>43</sup> Die bei Unbetontheit

<sup>43</sup> Die aufgrund der Aussprachewörterbücher aufstellbaren Minimalpaare und quasi-Minimalpaare, vgl. etwa *H[e].léne* vs. *H[E]l.léne*, *p[i].lieren* vs. *p[ç]l.lieren*, *K[o].latür* vs. *K[O]l.latür*; *K[o].lónne* vs. *K[O]l.lége*, *M[i].liéu* vs. *M[ç]l.lión*, *kons[u].mieren* vs. *s[i]m.mieren*, weist Becker (1998: 88) mit dem Argument

auffindbaren Gespanntheitsunterschiede weisen vielmehr eine komplementäre Verteilung auf: gespannte Kurzvokale kommen nur in offenen, ungespannte nur in geschlossenen Silben vor (vgl. (28)b).<sup>44</sup> Dies führt den Autor zur Annahme eines Allophonieverhältnisses zwischen gespannten und ungespannten Kurzvokalen unter Unbetontheit (vgl. Becker 1998: 85).<sup>45</sup> Zwar gibt es zu dieser Allophonieregel einige Ausnahmen, diese sind nach dem Autor jedoch systematisch (vgl. 82ff). So kommen in den auf einen Sonoranten auslautenden wortfinalen Silben zwar gespannte Langvokale vor (vgl. *Kámerun* [{}kamWru:n], *Pinguin* [{}pçNgui:n]), aber diese Silben erhalten einen rythmischen Nebenakzent bzw. können auch hauptbetont werden. Die übrigen Abweichungen sind morphologisch, unter Berufung auf die sog. Morphemkonstanz zu erklären, nach der in Komposita, sowie in abgeleiteten und flektierten Wörtern die morphologischen Bestandteile ihre phonologische Struktur (hier: Gespanntheit und Länge) beibehalten, vgl. *Jahrhundert* [{}jA:Ç&{}hündÇt], *Réchtsstaat* [{}rE'tËsStA:t]; *Fahreréi* [{}fA:rW{}raËi], *Ziereréi* [{}tSi:rW{}raËi]; *Hindus* [{}hçndus], *beáutot* [{}bW{}aËütot] usw.<sup>46</sup>

(28) a. betonte Silben	b. unbetonte Silben	
<i>Miete</i> vs. <i>Mitte</i>	<i>Ph[i].losophie</i> <sup>47</sup>	<i>Ph[ç]l.harmonie</i>
<i>Hüte</i> vs. <i>Hütte</i>	<i>St[y].ropór</i>	<i>St[Y]m.phalide</i>
<i>Beet</i> vs. <i>Bett</i>	<i>R[e].kórd</i>	<i>R[E]k.tión</i>
<i>Höhle</i> vs. <i>Hölle</i>	<i>B[ë].ótien</i>	<i>B[ê]r.síáner</i>
<i>Ofen</i> vs. <i>offen</i>	<i>T[u].rín</i>	<i>T[i]r.bíne</i>

---

zurück, dass es sich bei ihnen um Überlautung, d.h. die Interpretation des Schriftbildes handelt. Die Verdoppelung eines konsonantischen Buchstabens habe nämlich nur nach betontem Vokal eine Funktion (Markierung der Länge), nach unbetonten Vokalen sei sie funktionslos, was auch daran zu erkennen sei, dass sie bei der Eindeutschung der Schrift oft weggelassen (vereinfacht) werde, vgl. *Bajonett* (< fr. *baïonnette*), *Perücke* (< fr. *perruque*), *Pionier* (< fr. *pionnier*) usw.

<sup>44</sup> Vgl. auch Moulton (1962: 63f): „[...] with few exceptions, a tense vowel is to be pronounced if it is followed in the spelling by a single consonant letter, but a lax vowel if it is followed in the spelling by a double consonant letter or a consonant cluster.“

<sup>45</sup> Vgl. auch Ramers (1988).

<sup>46</sup> Einen Sonderfall stellen Wörter wie *Demut* [{}de:mu:t] oder *Beelzebub* [{}bW{}EltsWbu:p]/ [{}be:ltS Wbu:p] dar, die die Volksetymologie nach Becker als Komposita deutet: *Demut* ist eine Art *Mut*, *Beelzebub* eine Art *Bub*.

<sup>47</sup> Punkte markieren Silbengrenzen, der Akut steht für Betonung.

*Ruhm* vs. *Rum*

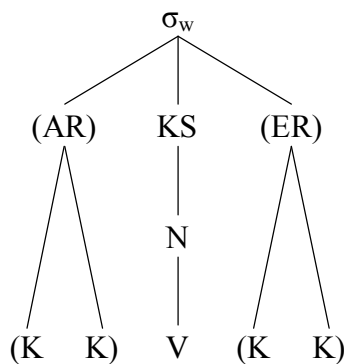
*Pr[o].letariát*    *S[O]l.dát*

*Staat* vs. *Stadt*

*m[A].liziós*    *M[a]l.tose*

Dass die gesamte Silbenschnittopposition nur auf Tonsilben beschränkt ist, will der Autor durch die Annahme einer unterschiedlichen Strukturierung für Tonsilben und für unbetonte Silben zum Ausdruck bringen. Während Tonsilben im Deutschen die Struktur in (25) haben, schlägt Becker für die unbetonten Silben eine Struktur vor, in der die für den Silbenschnitt verantwortliche Silbenkonstituente, nämlich die Implosionsposition fehlt, somit besteht die Kernsilbe einer unbetonten Silbe lediglich aus dem Nukleus (29).

(29) Struktur der unbetonten Silbe (Becker 1996a: 13; 1998: 80)

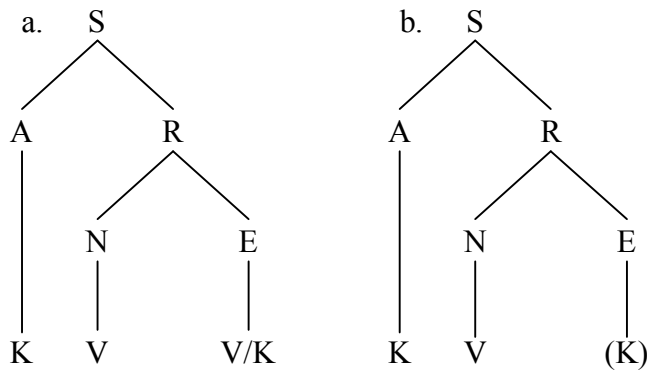


### 3.4.3 Maas (1999)

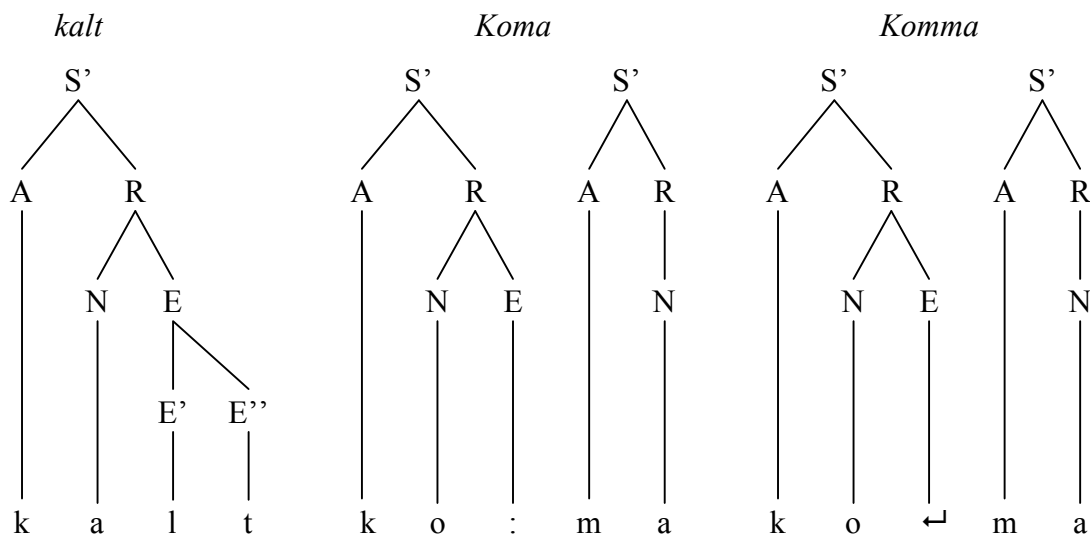
Maas (1999) schlägt für das Deutsche ähnlich wie Becker (1998) eine Silbenstruktur mit drei Silbenkonstituenten vor: Anfangsrand (A), Nukleus (N) und Endrand (E). Doch anders als dieser fasst er den Nukleus und den Endrand unter dem Reim (R) zusammen. Der Reim stellt im Konstituentenmodell von Maas (1999) denjenigen Teil der Silbe dar, in der der Silbenschnitt zum Ausdruck gebracht wird. Nach seinem Vorschlag ist der Reim einer prominenten (betonten) Silbe obligatorisch verzweigend. Dabei wird ein sanft geschnittener Vokal sowohl mit dem Nukleus als auch mit dem Endrand assoziiert, während ein scharf geschnittener Vokal nur mit dem Nukleus verbunden ist. Da Maas (1999) keine ambisilbischen Konsonanten annimmt (diese kommen nach ihm nur in Dialekten des Deutschen vor, in denen sie auch lang also als echte Geminaten gesprochen werden), muss er in einer offenen Silbe mit scharf geschnittenem Vokal den scharfen Schnitt mit einem extra Zeichen ‚↵‘ markieren. Somit ergibt sich die Struktur der prominenten Silbe im Deutschen unter (30)a. Da der Autor ähnlich wie Becker (1998) der Meinung ist, die Silbenschnittopposition besteht nur in betonten (prominenten) Silben, schlägt er ebenfalls

unterschiedliche Strukturen für die prominenten und die nicht-prominenten Silben vor. Die nicht-prominente Silbe unterscheidet sich von der prominenten dadurch, dass der Reim (R) in Ersterer nicht obligatorisch verzweigend ist, sowie dass ihr Endrand keinen Vokal enthalten darf (vgl. (30)b). Aufgrund des bisher über den Repräsentationsvorschlag von Maas (1999) Gesagte lässt sich die Silbenstruktur der drei Beispielswörter *kalt*, *Koma* und *Komma* in diesem Modell wie in (31) angeben.

(30) Struktur der prominenten (a) und nicht-prominenten (b) Silbe nach Maas (1999: 136)



(31) Silbenstruktur der Wörter *kalt*, *Koma* und *Komma* aufgrund von Maas (1999)

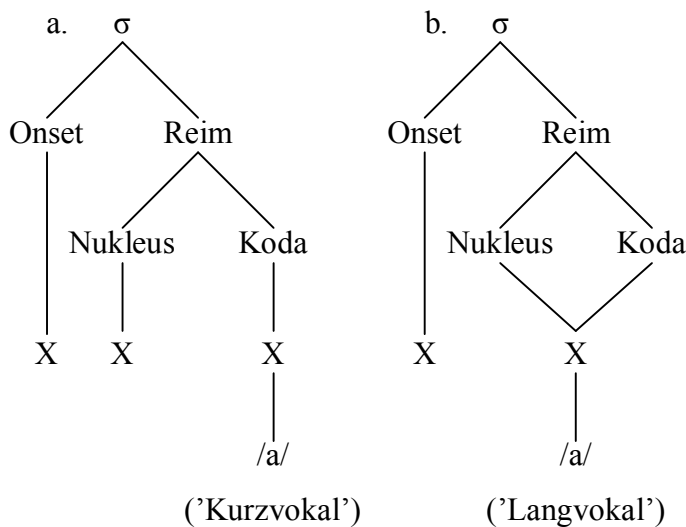


### 3.4.4 Lenerz (2000)

Im Gegensatz zu Becker (1996ab, 1998) und Maas (1999) geht Lenerz (2000) davon aus, dass Silbenschnittunterschiede sowohl in betonten als auch in unbetonten Silben bestehen. Eine Ähnlichkeit des Lenerz'schen und des Maas'schen Ansatzes besteht jedoch darin, dass beide eine mehrmals verzweigende binäre Silbenstruktur im Deutschen annehmen, wo der

Silbenknoten in die Konstituenten Onset (Anfangsrand) und Reim, ferner Letzterer dann in den Nukleus (Silbenkern) und die Koda (Endrand) zerfällt. Auch Lenerz plädiert dafür, den Silbenschritt in der Reinkonstituente unterzubringen: Aus (32) geht hervor, dass er für den Unterschied zwischen den beiden Silbenschritten ebenso die unterschiedliche Assoziation des Vokals mit den Reimkonstituenten verantwortlich macht: ein scharf geschnittener Vokal ist nur mit dem Nukleus (vgl. (32)a), ein sanft geschnittener dagegen sowohl mit dem Nukleus als auch mit der Koda verbunden (vgl. (32)b).

(32) Struktur der Silben mit scharfem (a) und sanftem Schnitt (b) (Lenerz 2000: 190)



Für die zugrunde liegende Repräsentation der beiden Silbenschritte muss Lenerz (2000) folgerichtig annehmen, dass ein Teil der Silbentruktur schon im Lexikon festgelegt ist. Dieser Teil ist nach Ansicht des Autors die Koda, die bei sanftem Schnitt zugrunde liegend mit dem Merkmal [=N] (vgl. (33)a), und bei scharfem Schnitt mit dem Merkmal [ $\neq$ N] (vgl. (33)b) assoziiert ist. Im ersteren Fall wird der Nukleusvokal auch mit der Koda verbunden, im letzteren unterbleibt diese Assoziation und die Koda wird mit dem postvokalischen Segment assoziiert, das entweder ein eigenes Segment sein oder auch mit dem Onset der Folgesilbe verbunden werden kann, vgl. (34). Daraus folgt, dass Lenerz (2000) im Gegensatz zu Maas (2000) Ambisilbilität im Standarddeutschen annimmt.

(33) Repräsentation der beiden Silbenschritte (Lenerz 2000: 193)

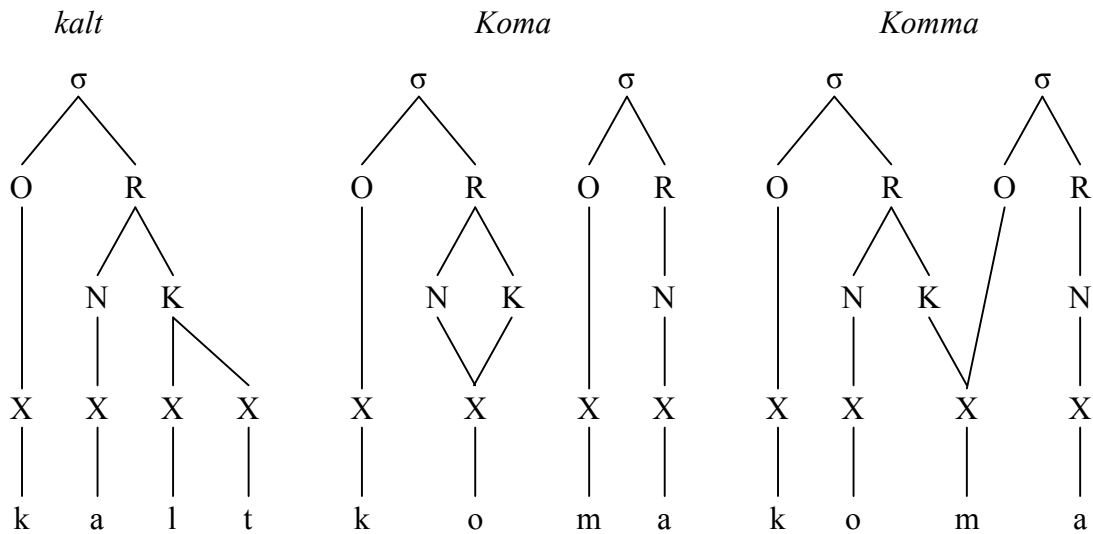


[=N]

[≠N]



(34) Struktur der Wörter *kalt*, *Koma* und *Komma* aufgrund von Lenerz (2000)



### 3.4.5 Kritik an den bestehenden Silbenschnittrepräsentationen

Im vorliegenden Teil der Arbeit wollen wir dafür argumentieren, dass keine der oben dargelegten Repräsentationsvorschläge in der Lage sind, den Silbenschnitt im Deutschen unter Berücksichtigung seiner phonetischen, phonologischen und typologischen Aspekte adäquat zu erfassen.

Gegen Vennemanns (1994) Darstellungsweise spricht dasselbe, was überhaupt im Lichte der Befunde Spiekermanns (2000) gegen seinen Ansatz spricht: er geht von einer phonetischen Annahme aus, nach der das Energiemaximum unter sanftem Schnitt im Vokal selbst, unter scharfem dagegen erst in der Folgekonsonanz liege, die die zitierte experimentalphonetische Untersuchung auf keinerlei Weise nachweisen konnte. Dadurch erscheint seine Repräsentation des Silbenschnitts im Deutschen als weniger akzeptabel.

Auch der Ansatz von Lenerz (2000) weist m.E. Defizite auf, wenn man von den Markiertheitsverhältnissen und den Ähnlichkeiten und Unterschieden zwischen den Silbenschnitt- und Quantitätssprachen ausgeht. Man kann nämlich mit Trubetzkoy (1939: 196) – wie oben besprochen – davon ausgehen, dass der sanfte Schnitt das merkmallose, der scharfe dagegen das merkmalhafte Glied in der Silbenschnittopposition darstellt. Dafür sprechen einerseits, dass bei der Aufhebung der Opposition (im Auslaut und vor Vokalen) im Deutschen nur sanft geschnittene Vokale vorkommen können, und andererseits, dass – wie Vennemann (1991a: 217) formuliert – genau die scharf geschnittenen, d.h. phonetisch als kurz und ungespannt realisierten Vokale die mehr eingeschränkte Klasse in dieser Opposition

darstellen. Die Lenerz'sche Darstellung kann dieser Beobachtung keineswegs Rechnung tragen: in ihr werden genau die scharf geschnittenen Vokale mit einer unmarkierten Struktur (Assoziation des Vokals nur mit dem Nukleus) erfasst, während die sanft geschnittenen strukturell als markiert (Assoziation des Vokals mit dem Nukleus und der Koda) zu betrachten sind. Vennemann (1991a: 220) sieht die Unmarkiertheit des sanften Schnittes auch in der Vermutung bestätigt, dass Sprachen ohne Silbenschnittopposition nur sanft geschnittene Vokale kennen. Diese Vermutung wird weiter untermauert, wenn man die Ergebnisse der oben zitierten experimentalphonetischen Untersuchung Spiekermanns (2000) betrachtet: Alle Vokale des Finnischen und Tschechischen als Quantitätssprachen zeigen ein gleichmäßig starkes Anhalten eines hohen Energiemaximums, das im Deutschen als Silbenschnittsprache nur für die sanft geschnittenen Vokale charakteristisch ist. Wollte man nun typologisch korrekt vorgehen, so könnte man diesen Sachverhalt mit der Darstellungsweise von Lenerz (2000) nur dadurch zum Ausdruck bringen, dass man alle Vokale der Quantitätssprachen, u.z. unabhängig von ihrer Quantität, wie die sanft geschnittenen Vokale des Deutschen repräsentiert, d.h. gleichzeitig mit dem Nukleus und der Koda verbindet, was ja naturgemäß nur eine inadäquate Darstellung der phonologischen Struktur im Finnischen oder Tschechischen sein kann. Somit kann die von Lenerz (2000) vorgeschlagene Darstellungsweise die typologischen Unterschiede zwischen Silbenschnitt und Quantität kaum vollständig erfassen, auch wenn der Autor genau den typologischen Aspekt als eines der wichtigsten Argumente für seine Repräsentation betrachtet: die Unterbringung der Quantität und des Silbenschnitts auf zwei unterschiedlichen Ebenen (Erstere auf der skelettalen Schicht – Letzterer auf der Ebene der Reimkonstituenten) kann die Beziehungen der beiden Sprachtypen nur teilweise ausdrücken.

Da die von Maas (1999) entwickelte Darstellungsweise des Silbenschnitts im Deutschen dem Lenerz'schen Vorschlag ähnlich ist, kann sie mit demselben Argument verworfen werden. Ein anderes Problem bereitet bei ihm m.E. das diakritische Symbol ◀ (vgl. die Repräsentation von *Komma* in (31)), mit dem der Autor die ambisilbischen Segmente abzulösen versucht: es ist phonetisch wie phonologisch kaum interpretierbar, auf der anderen Seite verschleiern die unterschiedlichen Strukturen in *Wand* und *Wanne*, dass es sich im Prinzip um dasselbe Phänomen (scharfen Schnitt) handelt. Meiner Ansicht nach ist das Konzept des Silbenschnitts von dem der Ambisilbizität im Deutschen kaum trennbar.

Auch Beckers (1998) Vorschlag ist nicht ganz ohne Probleme. Erstens halte ich seine Einschränkung des Silbenschnitts nur auf betonte Silben für grundsätzlich falsch. Der Autor begründet seine Position damit, dass die Silbenschnittopposition nur in betonten Silben

vorliegt, d.h. die Wahl zwischen den beiden Silbenschnitten nur unter Akzent bedeutungsunterscheidend wirkt. Die gespannten und ungespannten Kurzvokale in unbetonten Silben unterliegen somit lediglich einer allophonischen Verteilung: gespannte Kurzvokale kommen dabei in offenen – ungespannte in geschlossenen Silben vor. Was die Einschränkung der Distinktivität der Silbenschnitte nur auf betonte Silben betrifft, hat Becker offensichtlich recht, doch muss das m.E. nicht unbedingt bedeuten, dass in unbetonten Silben keine Silbenschnittunterschiede angenommen werden könnten. Was die gespannten und ungespannten Kurzvokale angeht, muss eine Präzision vorgenommen werden: Gespanntheit ist sowohl in betonten als auch in unbetonten Silben allophonisch verteilt – das ist einer der Grundgedanken der Silbenschnitttheorie. Wenn Becker die Allophonie der Gespanntheit in betonten und unbetonten Silben jeweils anders formuliert, so verschleiert er nach Lenerz (2000) geradezu das verbindende Moment. Nach Vennemann (1991a) haben nämlich die gespannten und ungespannten Kurzvokalen in unbetonten Silben nicht zufällig die oben genannte Verteilung, dies entspricht genau der Normalverteilung der gespannten und ungespannten Vokale unter Akzent. Zur expliziten Erfassung dieser auch wichtigen Generalisierung muss man davon ausgehen, dass Silbenschnitt sowohl in betonten als auch in unbetonten Silben besteht, wobei er in Ersteren bedeutungsunterscheidend ist, in Letzteren nicht und dort daher auch der Aufhebung ausgesetzt ist – wie das weiter unten noch gezeigt wird. Auch erübrigt sich mit diesem Argument die strukturelle Trennung zwischen betonten und unbetonten Silben, wie sie von Becker vorgenommen wird. Ob dabei die Silbe generell die von ihm für die Tonsilbe vorgeschlagene Struktur erhalten soll, erscheint jedoch im Lichte der Kritik von Lenerz (2000) als fraglich, in deren Mittelpunkt die Implosionsposition als einziger Unterschied zwischen betonten und unbetonten Silben steht. Sie stellt nach Lenerz (2000) eine zusätzliche, eingeschobene Silbenkonstituente dar (nach ihr können im Endrand genauso viele (nämlich zwei) weitere Segmente stehen wie in den unbetonten Silben), die nur für Silbenschnittsprachen charakteristisch ist und in Quantitätssprachen gänzlich fehlt, wodurch weder eine adäquate Typologisierung gegeben ist, noch der typologische Wandel von der Quantität zum Silbenschnitt angemessen erfasst werden kann.

Ein weiteres Problem der Becker'schen Repräsentation des Silbenschnitts im Deutschen betrifft m.E. seine ‚Kernsilbe‘. Diese stellt nämlich eine silbentypologisch sonderbare Konstituente dar, die auf den ersten Blick genau das leistet, was Wieses (1996) Silbenkern. In der Tat handelt es sich dabei jedoch um eine ‚Mischkategorie‘, die den traditionellen Nukleus und einen Teil der traditionellen Koda umfasst, die zusammen die Domäne für die Identifizierung des Silbenschnitts darstellen. So gesehen ist die Becker'sche

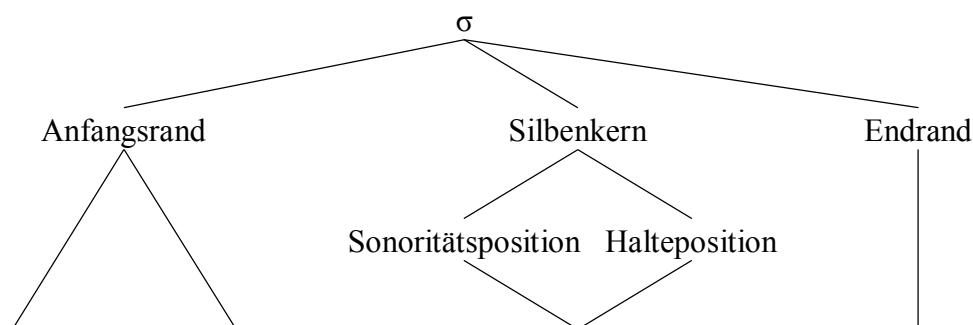
Darstellungsweise mit dem Repräsentationsvorschlag von Maas (1999) und Lenerz (2000) verwandt mit dem nicht unwesentlichen Unterschied, dass Letztere als Silbenschnittdomäne einen Teil einer silbentypologisch etablierten Konstituente, nämlich des Reims ansetzen, während sich Erstere einer silbentypologisch nicht belegten Kategorie bedient. Wenn diese Ähnlichkeit zwischen der Repräsentation von Becker (1998) und Maas (1999) bzw. Lenerz (2000) in der Tat vorliegt, dann lässt sich gegen Beckers Ansatz genau dasselbe Argument wie bei Maas (1999) und Lenerz (2000) nennen: Aus dem sprachtypologischen Aspekt erscheint als einzig adäquate Darstellungsweise die, in der der Silbenschnitt nicht im Reim lokalisiert ist. Wie das erreicht werden kann, wird weiter unten gezeigt.

### 3.4.6 Silbenschnitt und die interne Struktur des Silbenkerns

Alle bisher genannten Probleme werden substanzlos, wenn wir annehmen, dass die Silbenschnitte im Silbenkern unterzubringen sind. Da der scharfe und sanfte Schnitt wie in Becker (1998), Maas (1999) und Lenerz (2000) am besten durch unterschiedliche Assoziation der Vokalsegmente oder besser der skeletalen Positionen mit Silbenkonstituenten aufgefasst werden, wollen wir im Deutschen als Silbenschnittsprache einen binär verzweigenden Silbenkern annehmen, die beiden Silbenkernkonstituenten sollen als Sonoritäts- und Halteposition (SP bzw. HP) bezeichnet werden. Während mit der Sonoritätsposition stets das sonorste Segment der Silbe verbunden ist, ist die Halteposition für das starke vs. schwache Anhalten eines hohen Energiemaximums verantwortlich: Ist diese mit dem Vokal assoziiert, so ist für den Vokal ein starkes Anhalten eines hohen Energiemaximums, also sanfter Schnitt charakteristisch. Ist die Halteposition nicht mit dem Vokal, sondern mit dem postvokalischen Segment verbunden, unterbleibt im Vokal das starke Anhalten eines hohen Energiemaximums, das Vokalsegment ist folgerichtig scharf geschnitten, vgl. (35) und (36).

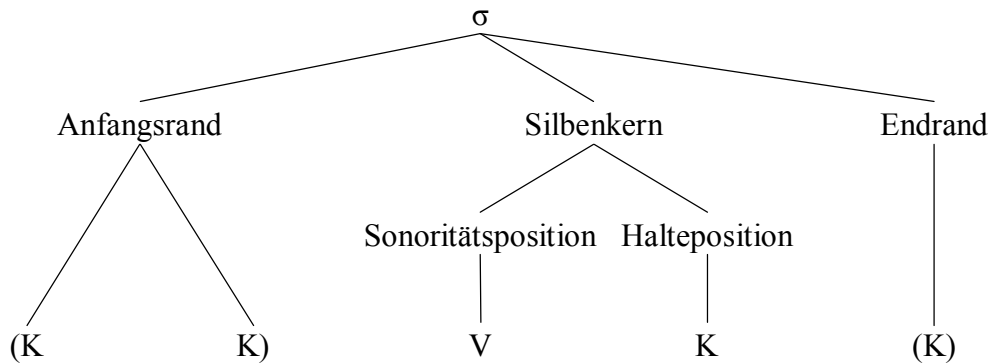
(35) Struktur der Silbe im Deutschen

a. unter sanftem Schnitt

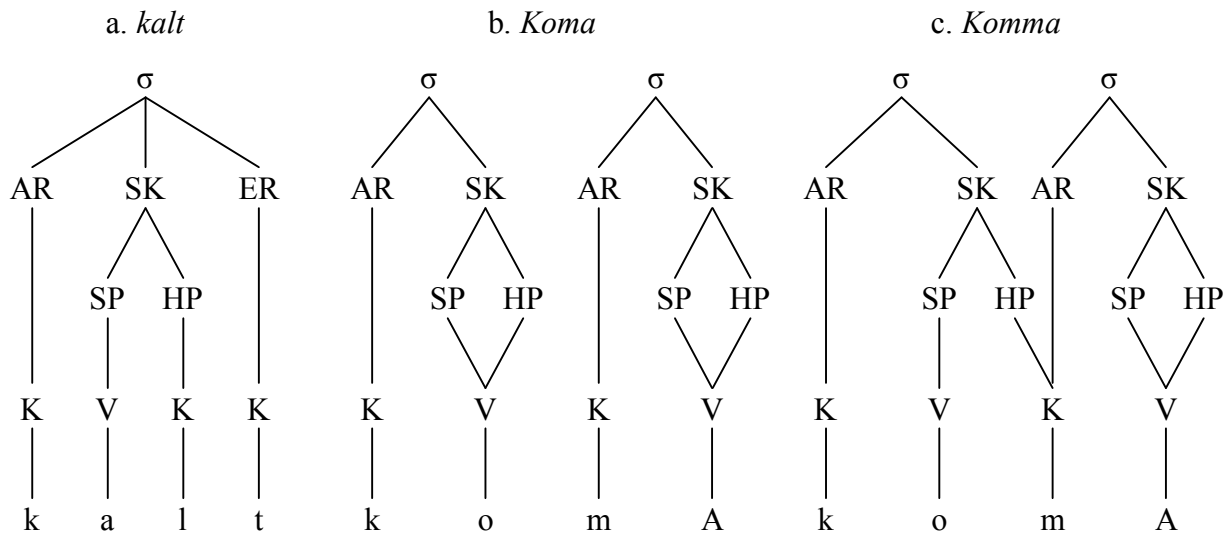


(K) (K) V (K)

b. unter scharfem Schnitt

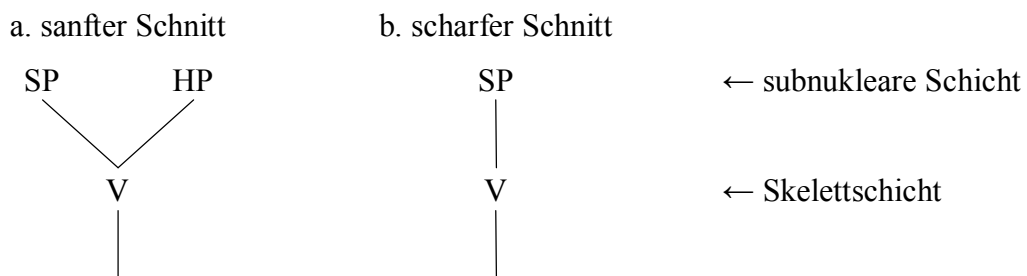


(36) Silbenstruktur der Wörter *kalt*, *Koma* und *Komma*



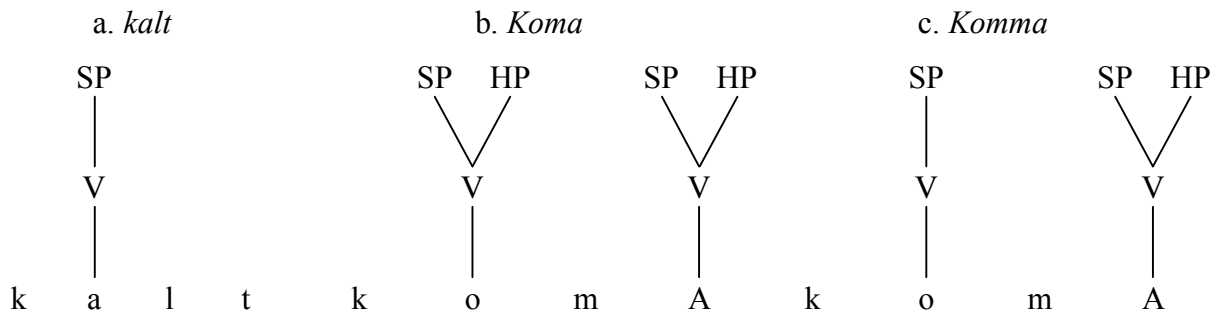
Die Unterscheidung zwischen zugrunde liegenden sanft und scharf geschnittenen Vokalen wollen wir im Grunde genommen ähnlich wie Lenerz (2000) dadurch erreichen, dass ein Teil der Silbenstruktur, genau die subnuklearen Konstituenten Sonoritätsposition und Halteposition, im Lexikon festgehalten sind, vgl. (37). Somit sollten die Wörter *kalt*, *Koma* und *Komma* zugrunde liegend die unter (38) abgebildete prosodische Struktur haben.

(37) Repräsentation der Silbenschnitte



v

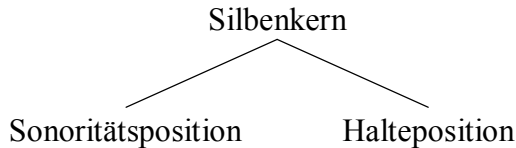
v

(38) Zugrunde liegende Repräsentation der Wörter *kalt*, *Koma* und *Komma*

Die für die Unterscheidung der beiden Silbenschnitte verantwortlichen zugrunde liegenden subnuklearen Konstituenten werden dann im Laufe der Silbifizierung mit dem Silbenkern assoziiert. Die wichtigste Voraussetzung für diese Assoziation beruht auf der eingangs besprochenen Annahme von dem binär verzweigenden Silbenkern und besagt, dass der Silbenkern obligatorisch eine Sonoritäts- und eine Halteposition enthalten soll. Nun ist diese Silbenkernbedingung dafür verantwortlich, dass die zugrunde liegenden subnuklearen Positionen SP und HP der beiden Vokale in *Koma* (vgl. (38)b) mit dem Knoten für den Silbenkern verbunden werden (vgl. (36)b). Und ähnlich erfolgt die Verbindung der zugrunde liegenden Sonoritätsposition des ersten scharf geschnittenen Vokals in *Komma* (vgl. (38)c) mit dem Silbenkernknoten, die durch die Kernbedingung vorgeschriebene Halteposition wird dann mit dem postvokalisches Nasal (besser mit der damit verbundenen skeletalen Position) assoziiert, was die Ambisilbizität dieses Segmentes ergibt (vgl. (36)c). Jedoch weist diese Beschreibung einen Nachteil auf. Nimmt man nämlich im Deutschen eine sprachspezifische Silbenkernbedingung an, so erhält der Silbenschnitt eine doppelte Markierung: Silbenschnitt wird einerseits in der zugrunde liegenden Repräsentation und andererseits auf der subnuklearen Ebene in Form einer Silbenstrukturbedingung gekennzeichnet. Darüber hinaus stellt sich die Frage nach der Erfassung der bei Lenerz (2000) kritisierten Markiertheitsverhältnisse und typologischen Adäquatheit: inwiefern stellt der sanfte Schnitt das unmarkierte Glied in der Opposition dar bzw. wie lässt sich neben dem Unterschied zwischen den Silbenschnitt- und Quantitätssprachen, der ja wie bei Lenerz (2000) durch die Unterbringung der beiden Phänomene auf jeweils einer anderen Repräsentationsebene erfasst wird, auch das oben formulierte verbindende Moment zwischen ihnen zum Ausdruck bringen? Um diese Probleme loszuwerden, schlagen wir vor, die fragliche

Silbenkernbedingung nicht als sprachspezifische, sondern als universelle Silbenstrukturbedingung aufzufassen und wie unter (39) zu formulieren.

(39) **Silbenkernbedingung:** Der Silbenkern enthält universell eine Sonoritäts- und eine Halteposition:



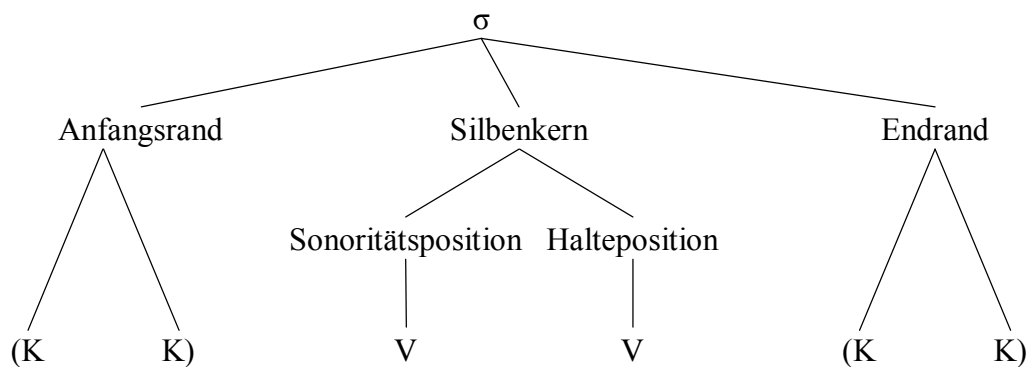
Mit der Universalität der Kernbedingung wird die doppelte Markierung des Silbenschnitt im Deutschen beseitigt: er ist lediglich zugrunde liegend markiert, die obligatorische Verzweigung des Silbenkerns kann aus dem universalen Prinzip in (39) hergeleitet werden. Auch die Unmarkiertheit des sanften Schnittes wird damit adäquat erfasst: das merkmallose Glied der Opposition wird mit einer unmarkierten Repräsentation verbunden, die Besetzung des Silbenkerns nur durch ein Vokalsegment beim sanften Schnitt ist ja durchaus als der Normalfall zu betrachten, während die Erscheinung eines weiteren Segmentes im Silbenkern, das in der Mehrheit der Fälle sogar ein Konsonant ist, sicherlich als markiert anzusehen ist. Und schließlich kann auch der verbindende Punkt zwischen Silbenschnitt- und Quantitätssprachen ausgedrückt werden: in den Quantitätssprachen kann man ein eins-zu-eins-Verhältnis zwischen Silbenkern und Vokal annehmen, alle Vokale sind also in diesen Sprachen sanft geschnitten und zeigen dadurch ein starkes Anhalten eines hohen Energiemaximums. Für das Ungarische, das keine Silbenschnittsprache darstellt, lässt sich somit aufgrund des bisher Gesagten sowie Törkenczy (1994) die Silbenstruktur unter (40) (nächste Seite) annehmen, die eine silbenstrukturelle Repräsentation der ungarischen Beispielswörter *dalt* (,Lied' Akk.), *kóma* (,Koma') und *koma* (,Gevatter') unter (41) nahelegen. Hinzuzufügen ist dieser Darstellung, dass sich in ihr die Aufspaltung des Silbenkerns eigentlich als redundant erweist, da die dadurch entstehenden beiden subnuklearen Positionen SP und HP stets mit demselben Vokal assoziiert sind, daher kann im Ungarischen auf die explizite Verzeichnung der Verzweigung des Silbenkerns in eine Sonoritäts- und Halteposition in einer vereinfachten Repräsentation verzichtet werden.

Anhand der oben dargelegten Auffassung über den Silbenschnitt sowie die silbenstrukturellen Unterschiede zwischen Quantität und Silbenschnitt kann man nun auch den typologischen Wandel von einer Quantitätssprache zu einer Silbenschnittsprache auf eine

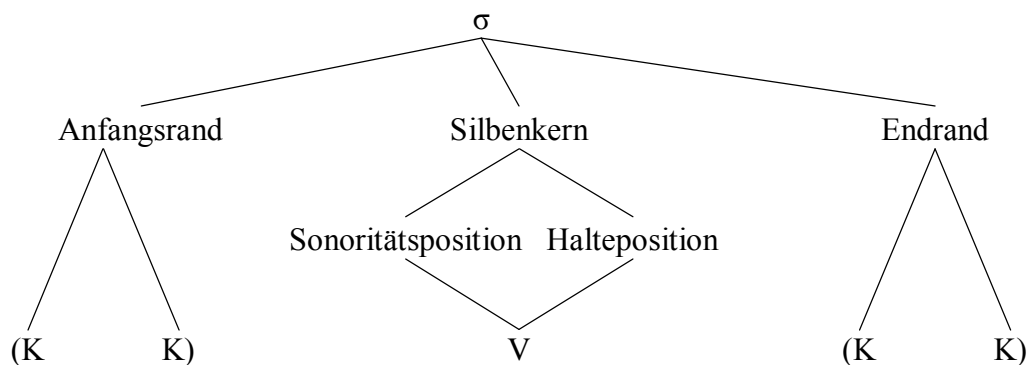
einfache Weise erfassen. Er erfolgt nicht wie bei Becker (1998) durch die Einfügung einer ansonsten nicht motivierten subsilbischen Konstituente, der Implosion als Stärkeposition, sondern wie bei Lenerz (2000) um die Verschiebung des Phänomens von der skelettalen Schicht auf die Ebene der Silbenkonstituenten. Anders als bei ihm findet hier dabei parallel keine ‚Domänenerweiterung‘ statt, d.h. die Domäne des Phänomens wird nicht vom Silbenkern (Quantität) zum Reim (Silbenschnitt). Vielmehr bleibt der Silbenschnitt innerhalb des Silbenkerns: beim Übergang von der Quantität zum Silbenschnitt findet eine phonologisch relevante ‚Kernspaltung‘ statt, d.h. die beiden subnuklearen Silbenpositionen übernehmen von den skelettalen Position die Rolle der Distinktivität.

(40) Struktur der Silbe im Ungarischen

a. mit Langvokal

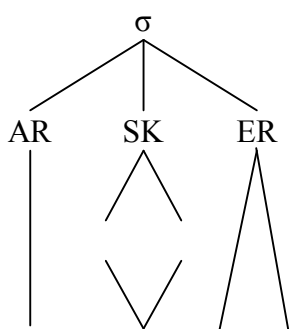


b. mit Kurzvokal

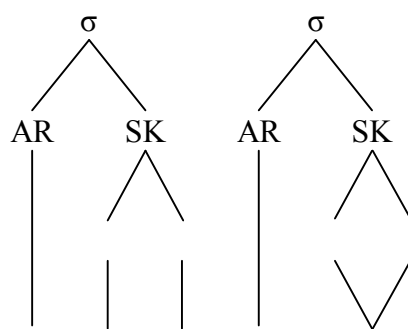


(41) Silbenstruktur der ung. Wörter *dalt* („Lied“ Akk.), *kóma* („Koma“) und *koma* („Gevatter“)

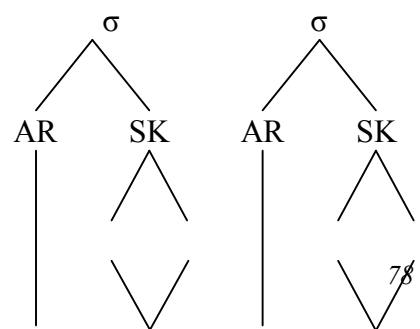
a. *dalt*



b. *kóma*



c. *koma*



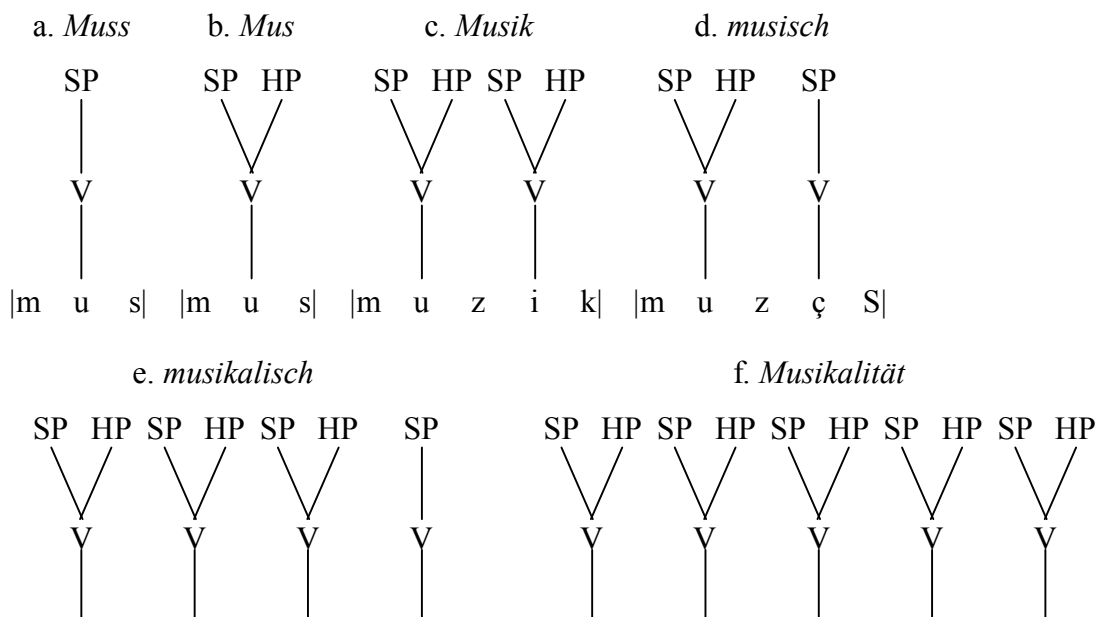


SP	HP		SP	HP		SP	HP		SP	HP		SP	HP	
K	V	K	K	K	V	V	K	V	K	V	K	V	K	V
d	o	l	t	k	o	m	o	k	o	m	o	m	o	

### 3.5 Silbenschnitt, Vokalquantität und Vokalqualität

Ähnlich wie Trubetzkoy (1939), Vennemann (1991a), Becker (1998) und Spiekermann (2000) wollen wir hier davon ausgehen, dass Gespanntheit und Vokaldauer bloße Folgeerscheinungen des Silbenschnitts darstellen. Da sie rein phonetische Größen im Deutschen sind, sollen sie der hier vertretenen Auffassung nach Gegenstand der phonetischen Realisation sein. So hat man in den Wörtern *Mus*, *Musik*, *musisch*, *musikalisch* und *Musikalität* sowohl zugrunde liegend als auch im Laufe der gesamten Derivation sanften Schnitt, in *Muss* dagegen scharfen Schnitt. Das von Hall (1992) als Quantitätsalternation aufgefasste Phänomen wird hier als reine phonetische Angelegenheit betrachtet, was damit erklärt werden kann, dass in dem hier vertretenen Modell die Quantität keine phonologische Eigenschaft darstellt, und in den betreffenden Wörtern in Abhängigkeit von der Akzentstelle kein systematischer Wechsel zwischen scharfem und sanftem Schnitt beobachtet werden kann.

(42) Zugrunde liegende Repräsentation der Vokale in *Muss*, *Mus*, *Musik*, *musisch*, *musikalisch* und *Musikalität*



|m u z i k A l ç S| |m u z i k A l i t é t|

Im vorausgehenden Abschnitt wurde bei der Besprechung des Ansatzes von Becker (1996ab, 1998) erwähnt, dass aus einem strukturalistischen Gesichtspunkt gesehen zwischen dem scharfen und sanften Schnitt in unbetonter Stellung ein Allophonieverhältnis angenommen werden soll. Eine diesbezügliche Allophonieregel weist im Deutschen jedoch eine wichtige Besonderheit auf: sie wird von den Sprechern des Gegenwartsdeutschen häufig nicht eingehalten. Moulton (1962) schiebt diese Abweichung noch auf das informelle Register, und gibt als Ergebnis dieser vollständigen Nivellierung der Opposition dort einen „mehr oder weniger ungespannten“ Vokal an, wenn er folgendermaßen formuliert:

In the normal relaxed speech of many persons, the suspension of the tense–lax opposition applies [...] to all [...] tense–lax pairs as well. For example, many people who in formal speech distinguish /koʃlu\*mbu\*s/ Kolumbus (with tense /oʃ/) and /ko\*le5gW/ Kollege (with lax /o\*/) often use in informal speech the same vowel in both words: /ko}lu\*mbu\*s/ like /ko}le5gW/, both with a more or less lax /o/. Similarly, formal /fi5lo5zo5}fi5/ Philosophie (with tense unstressed /i5 o5 o5/) may alternate with informal /filozo}fi5/ (with more or less lax /i o o/), etc. (64)

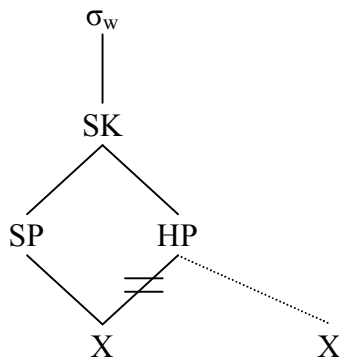
Dagegen rechnet Wurzel (1981) die gespannten Kurzvokale der „allerhöchsten Norm“ zu, um die sich die Sprecher bemühen müssen, die aber „in der normalen gesprochenen Sprache“ durch ungespannte Kurzvokale abgelöst werden:

In der normalen gesprochenen Sprache des wohl überwiegenden Teils der Sprecher des Deutschen erscheinen in Wörtern wie *direkt*, *Minister*, *Physik*, *Pyramide*, *aktuell*, *Musik*, *Metall*, *Methoden*, *Ökonom*, *möblieren*, *Kolonne*, *Hotel*, *Atom* und *Kalender* nicht die kurzen nichtzentralisierten Vokale, sondern statt dessen die kurzen zentralisierten Vokale, ebenso übrigens in nativen Wörtern wie *vielleicht*, *lebendig*, *zumal* und *sowohl*. Die kurzen nichtzentralisierten Vokale treten fast bei solchen Sprechern auf, die sich ganz bewußt an der höchsten Aussprachenorm des Deutschen orientieren, z.B. bei Berufssprechern. (919)

Im Hintergrund der Ersetzung der kurzen gespannten Vokale durch kurze ungespannte steht der „Systemzwang“ (Meinhold & Stock 1982: 90), die „phonologische Anpassung der nichtnativen Wörter an den Vokalismus der nativen Wörter“ (Wurzel 1981: 919): die unter Akzent beobachtbare Korrelation zwischen Quantität und Gespanntheit (lang: gespannt, kurz: ungespannt) wird auch auf die unbetonte Position nichtnativer Wörter ausgedehnt, es handelt sich also um einen „phonetischen Zusammenfall von sowohl Gespanntheit als auch Quantität“ (Hinderling 1982: 291). Infolge dieser freien Variation kann das Wort *Metal* hochsprachlich

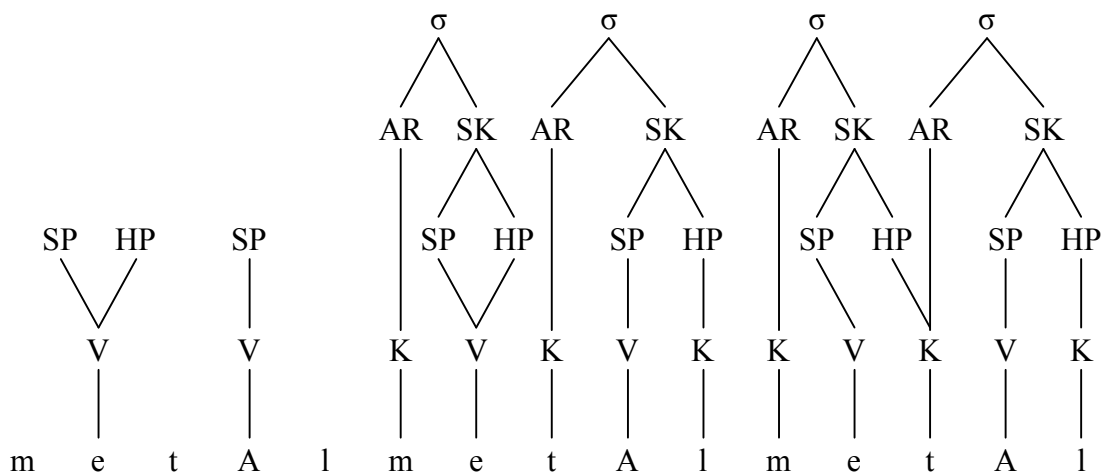
als [metal] umgangssprachlich dagegen als [mEtal] realisiert werden. Es stellt sich die Frage, wie dieser freien Variation in dem hier vertretenen Ansatz Rechnung zu tragen ist. Diesen Fall wollen wir als ‚Schärfung‘, d.h. als Wechsel vom sanften zum scharfen Silbenschnitt bezeichnen, die mit der Regel in (43) erfasst werden kann. Diese Regel umfasst einerseits die Tilgung der Assoziationslinie zwischen der Halteposition und einem sanft geschnittenen Vokal und gleichzeitig die Assoziation der Halteposition mit dem nachfolgenden Segment, auch wenn dadurch Ambisilbizität entsteht. Da diese Regel – wenn sie überhaupt zur Anwendung kommt – ohne Ausnahme arbeitet, soll sie in der postlexikalischen Komponente untergebracht werden. In (44) wird die Funktion dieser Regel am Beispiel *Metal* veranschaulicht.

(43) Schärfungsregel (fakultativ)



(44) Zugrunde liegende, Lexikalische und fakultative Oberflächenform von *Metal*

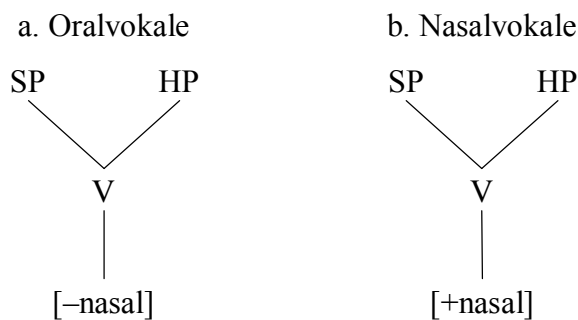
- a. Zugrunde liegende Form    b. Lexikalische Form    c. fakultative O-Form



### 3.6 Silbenschnitt und Nasalvokale

Aus der in Kap. 2 gegebenen Beschreibung der vier Nasalvokale unter (4) geht hervor, dass diese Vokale auf den ersten Blick kein eindeutiges Verhältnis zum Silbenschnitt zeigen. Sie kommen wie die sanft geschnittenen in offenen Silben vor, wo sie in Abhängigkeit vom Akzent phonetisch lang (wenn betont) und kurz (wenn unbetont) realisiert werden – doch werden sie wie die scharf geschnittenen Vokale ungespannt gesprochen. Da für ihre Unterbringung bei den sanft geschnittenen Vokalen mehr (und gewichtigere) Argumente sprechen, wollen wir sie hier als Vokale unter sanftem Schnitt analysieren, und ihre ungespannte Artikulation mit Becker (1998: 28) als Folge der Nasalität, artikulatorisch gesehen also des gesenkten Velums betrachten. Die Unterscheidung zwischen Oral- und Nasalvokalen sollte dann mit Hilfe eines phonologischen Merkmals (etwa [nasal]) zum Ausdruck gebracht werden, vgl. (45). Die aktuelle Länge und die Ungespanntheit der nasalierten Vokale werden dann im Laufe der phonetischen Realisierung errechnet – Erstere ähnlich wie bei den entsprechenden sanft geschnittenen Oralvokalen in erster Linie aufgrund der Akzentstelle, Letztere dagegen aus dem Zusammenspiel des Silbenschnitts und des Merkmalwertes [+nasal].

(45) Repräsentationelle Unterscheidung zwischen den Oral- und Nasalvokalen



Doch erscheint eine solche Integration der Nasalvokale ins Vokalsystem des Deutschen nicht unproblematisch. Wie die Aussprachewörterbücher Krech et al. (1982) und Mangold et al. (1990) zeigen, existieren zu den Nasalvokalen auch Ersatzformen zumeist der Form entsprechender Oralvokal plus Velarnasal:

(46) *Pointe* [E\$:] → [EN], *Bonbon* [O\$:]/[O\$] → [ON], *Cancan* [a\$:]/[a\$] → [aN]...

Angesichts dieser Substitution stellt sich die Frage, ob es nötig ist, nasalierte Vokale im Deutschen anzunehmen, oder ob man vielmehr in den fraglichen Wörtern die genannte

Sequenz annehmen soll. In dieser Frage gehen die Meinungen auseinander. So berichtet Kohler (1995) darüber, dass in der Standardlautung die Tendenz bestehe, „das deutsche Phonemsystem durch vier nasalierte Vokale zu ergänzen“ (173), während Wurzel (1981) eine geradezu feindselige Haltung gegenüber diesen vier Vokalen einnimmt und für deren Ausgrenzung aus dem deutschen Vokalsystem plädiert, wenn er folgendermaßen formuliert:

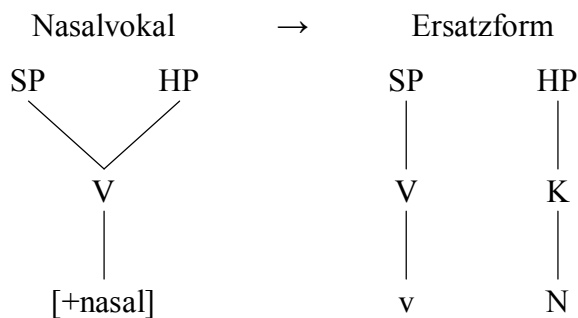
Die normale, den phonologischen Regularitäten des Deutschen folgende und für den Sprecher ohne einschlägige Französischkenntnisse einzig mögliche Aussprache ist die mit nichtnasalem Vokal plus velarem Nasalkonsonanten. Die Variante mit Nasalvokal setzt dagegen zumindest partielle Französischkenntnisse voraus und hat faktisch den Charakter eines fremdsprachigen Zitats. Wir wollen deshalb annehmen, daß es im Deutschen nur nichtnasale Vokale gibt. (907)

Nun erscheint einmal als fraglich, ob eine generelle Ersetzbarkeit der Nasalvokale durch die Sequenz entsprechender Oralvokal + Velarnasal im Deutschen in der Tat anzunehmen ist. Einerseits sei die Ersetzung dieser Art bei /a\$/ und /ê\$/ selten bzw. unmöglich. Doch könne es nach Meinhold & Stock (1982: 108f) zu Ersterem eine andersartige Ersatzform geben, bei der das Schriftbild interpretiert wird: *Emblem* wird daher entweder als [a\$ble:m] oder als [Emble:m] gesprochen. Etwas Ähnliches ist auch beim letzteren Vokal zu beobachten: zu *Parfum*, dem möglicherweise einzigen Wort mit diesem Nasalvokal, gibt es auch die Schreibweise *Parfüm*, die dem deutschen Vokalismus entsprechend als [paRfy:m] gesprochen wird. Andererseits scheint die Ersetzbarkeit (durch welche Form auch immer) laut Becker (1998: 27) auch damit zusammenzuhängen, ob die Wörter, in denen sie vorkommen, zum Grundwortschatz gehören oder eher als Bildungswörter zu betrachten sind. Während bei Ersteren die Ersetzung am ehesten akzeptiert werden kann, wird ein Nasalvokal in Letzteren kaum substituiert. Doch stehen Beckers Worte nicht unbedingt in einem unüberwindbaren Widerspruch zu Wurzels Behauptung. Die nicht substituierten Nasalvokale kommen in Bildungswörtern vor, von denen die Sprecher, die sie verwenden, und die sich sozial bestimmen lassen, wissen, „daß es sich bei diesen Wörtern um besondere Wörter handelt“ (ebda). Dieser besondere Status ist m.E. durchaus mit dem „Charakter eines fremdsprachigen Zitats“ zu vergleichen. Angesichtes dieser beiden Quellen soll die oben zitierte Argumentation Wurzels nicht völlig verworfen, sondern vielmehr in einer modifizierten Form angenommen werden. Demnach können in den Wörtern des (nativen) Standardwortschatzes zwar Nasalvokale gesprochen werden, doch werden solche Wörter zumeist nicht mit diesen, sondern mit verschiedenen Ersatzformen gesprochen. Die Ersetzung eines Nasalvokals kann im Deutschen auf zweierlei Art und Weise erfolgen. Im Falle der phonetisch-phonologisch

bedingten Substitution wird das für die Nasalität verantwortliche Merkmal [+nasal] quasi zu einem eigenständigen Nasalsegment, u.z. zum Velarnasal ‚verstärkt‘, der die Halteposition besetzt, in der Sonoritätsposition erscheint dann ein Oralvokal, der über dieselbe segmentale Ausstattung verfügt wie der entsprechende Nasalvokal, hier findet also ein systematischer ‚Silbenschnittumschlag‘ statt, vgl. (47). Aus dem obigen Zitat Wurzels geht hervor, dass der Autor diese Art der Ersetzung als die „normale, den phonologischen Regularitäten des Deutschen folgende und für den Sprecher ohne einschlägige Französischkenntnisse einzig mögliche Aussprache“ betrachtet. Doch bedarf dieser Behauptung m.E. einer etwas ausführlicheren Erklärung. Zwar kann nämlich das Erscheinen des Velarnasals nach einem in der Wurzel’schen Beschreibung kurzen, nach der hier vertretenen Auffassung scharf geschnittenen Vokal als durchaus normal betrachtet werden: es ist genau die Halteposition (ähnlich wie bei Becker die Implosion), auf die das Vorkommen dieses Sonoranten beschränkt ist. Andererseits wird er in der generativen Phonologie des Deutschen zumeist – so auch in Wurzel (1981) – nicht als Teil der zugrunde liegenden Repräsentation betrachtet, sondern aus der zugrunde liegenden Sequenz [ng] durch Anwendung zweier phonologischer Regeln (regressive Nasalassimilation und g-Tilgung) hergeleitet, was u.a. durch das oben erwähnte stark eingeschränkte Vorkommen des Velarnasals im Deutschen motiviert ist. Will man mangels entsprechender phonologischer Motivation nun die Möglichkeit der Annahme der zugrunde liegenden Sequenz [ng] für das Oberflächen-[N] als Ersatz der Nasalität der Nasalvokale im Deutschen von vornherein ausschließen, stellt sich die Frage, inwiefern eine Ersetzung als systemkonform aufgefasst werden kann, die auf ein zugrunde liegend nicht vorhandenes Segment zurückgreift. Nun wollen wir an dieser Stelle auf eine ausführliche Behandlung der Phonologie der Sonoranten des Gegenwartsdeutschen selbstverständlich nicht eingehen, das wäre ja mit dem Anliegen der vorliegenden Studie kaum zu vereinbaren; jedoch zeigen, dass diese Art der Substitution der Nasalvokale aus phonetischer wie phonologischer Sicht als begründet angesehen werden kann. Jedenfalls normal und systemkonform ist die Ersetzung eines Nasalvokals im Standarddeutschen durch die Folge Oralvokal plus Nasalkonsonant, da im Deutschen nasalierte Vokale ansonsten nicht vorkommen, die Ersatzsequenz dagegen reichlich vertreten ist, und da Nasalvokale eine komplexere Artikulation erfordern, durch ihre ‚Aufspaltung‘ in einen Oralvokal und einen Nasalkonsonanten jedoch ihre beiden wichtigen artikulatorischen Momente (Vokalartikulation und Nasalität) aufrecht erhalten bleiben. Die Wahl für den Velarnasal ist

m.E. phonetisch begründet und phonologisch begünstigt. Einerseits besteht zwischen einem Nasalvokal und der Folge Oralvokal + [N] kein bedeutender auditiver Unterschied<sup>48</sup>, was mit den zwischen ihnen bestehenden nicht großen artikulatorischen Differenzen zusammenhängt: neben der unterschiedlichen zeitlichen Koordination der Hebung des Velums (am Anfang der Vokalartikulation vs. erst später) unterscheiden sie sich durch eine artikulatorische Geste: die Hebung des an der Vokalartikulation sowieso teilnehmenden Zungenkörpers im Bereich des gesenkten Velums bis zum totalen Verschluss. Das Erscheinen des Velarnasals als Ersatz der Nasalität der Nasalvokale wird phonologisch durch die oben bereits genannte distributionelle Einschränkung dieses Sonoranten im Deutschen begünstigt: das Vorkommen des [N] ist stets an die Halteposition gebunden.

(47) Phonetisch-phonologisch bedingte Substitution der Nasalvokale im Deutschen



Die andere Art der Nasalvokalersetzung könnte als orthographisch bedingte Substitution bezeichnet werden, da es sich dabei – wie gezeigt – um die Interpretation des Schriftbildes handelt, bei der die Analogie keine unwesentliche Rolle spielt. So wird in Anlehnung an die Schreibweise in der ersten Silbe von *Emblem* ein scharf geschnittenes (phonetisch kurz-ungespanntes) [E], in der eingedeutschten Schreibweise von *Parfüm* dagegen ein sanft geschnittenes (phonetisch lang-gespanntes) [y] vor dem Labialnasal [m] gesprochen, wobei der sanfte Schnitt hier möglicherweise auf den in Fremdwörtern oft anzutreffenden Wortauslaut [y:m] zurückzuführen ist (vgl. *anonym*, *homonym*, *Synonym* usw.), jedoch ist die Wirkung des ursprünglichen sanft geschnittenen Nasalvokals auch nicht auszuschließen. Alles in allem ist zu sehen, dass bei dieser Art der Nasalvokalsubstitution kein systematischer ‚Silbenschnittumschlag‘ der Art (47) erfolgt.

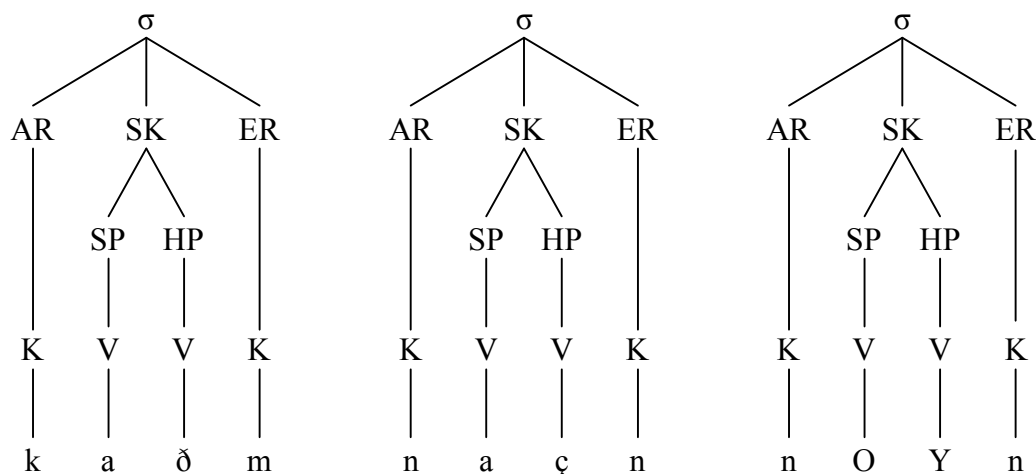
---

<sup>48</sup> Das zeigt auch die Unterrichtspraxis: Viele Deutschlerner ungarischer Muttersprache, in welcher Sprache keine Nasalvokale vorkommen und der Velarnasal einer noch strengeren distributionellen Einschränkung unterliegt, können den Unterschied zwischen einem Nasalvokal und der Sequenz Oralvokal + Velarnasal erst nach mehrmaligem Wiederholen wahrnehmen.

### 3.7 Silbenschnitt und Diphthonge

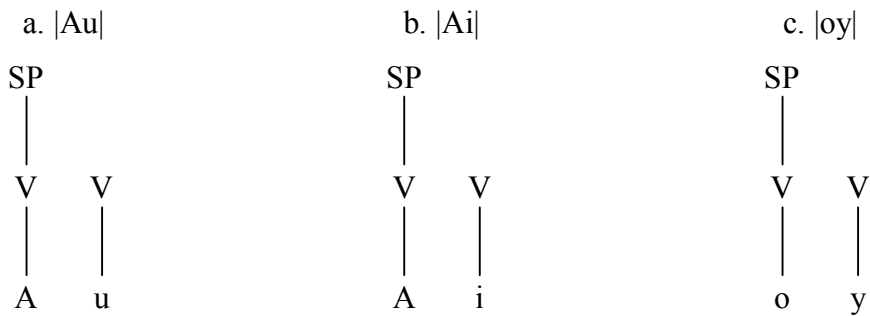
Von den in Kap. 2 in Anlehnung an die Klassifizierung von Becker (1998) eingeführten drei Diphthongtypen sind für den Silbenschnitt besonders die schließenden Diphthonge von Bedeutung, vgl. (5)a. Die von Moulton beobachtete und in Abschn. 3.1 besprochene phonotaktische Äquivalenz zwischen Langvokalen, Kurzvokal-Konsonant-Sequenzen und (schließenden) Diphthongen legt nahe, diese im Silbenkern zu lokalisieren, wobei der Kopfvokal die Sonoritätsposition, der Randvokal dagegen die Halteposition besetzt. Somit soll für die Wörter *kaum*, *nein* und *neun* die Silbenstruktur in (48) angenommen werden. Daraus geht hervor, dass nicht die schließenden Diphthonge selbst für den Silbenschnitt definiert sind, sondern nur ihr Kopfvokal: dieser kommt unter scharfem Schnitt vor, für den in diesem Fall nicht wie in den bisherigen Beispielen ein Konsonant, sondern vielmehr ein weiterer Vokal, der Randvokal der schließenden Diphthonge verantwortlich ist. Analog zur zugrunde liegenden Repräsentation unter (38)a lassen sich die ‚festen‘ Diphthonge des Gegenwartsdeutschen zugrunde liegend wie in (49) angeben. Diese Abbildungen zeigen, dass nach der hier vertretenen Silbenschnittauffassung für die subnukleare Struktur nur der Kopfvokal eine lexikalische Spezifizierung erhält. Der Randvokal ist zugrunde liegend nur auf der skeletalen Schicht, nicht aber auf der Ebene der subnuklearen Konstituenten spezifiziert – dafür, dass er in der Oberflächenstruktur mit der Halteposition im Silbenkern assoziiert ist, sorgt die universelle Silbenkernbedingung unter (39).

(48) Silbenstruktur der Wörter *kaum*, *nein*, *neun*



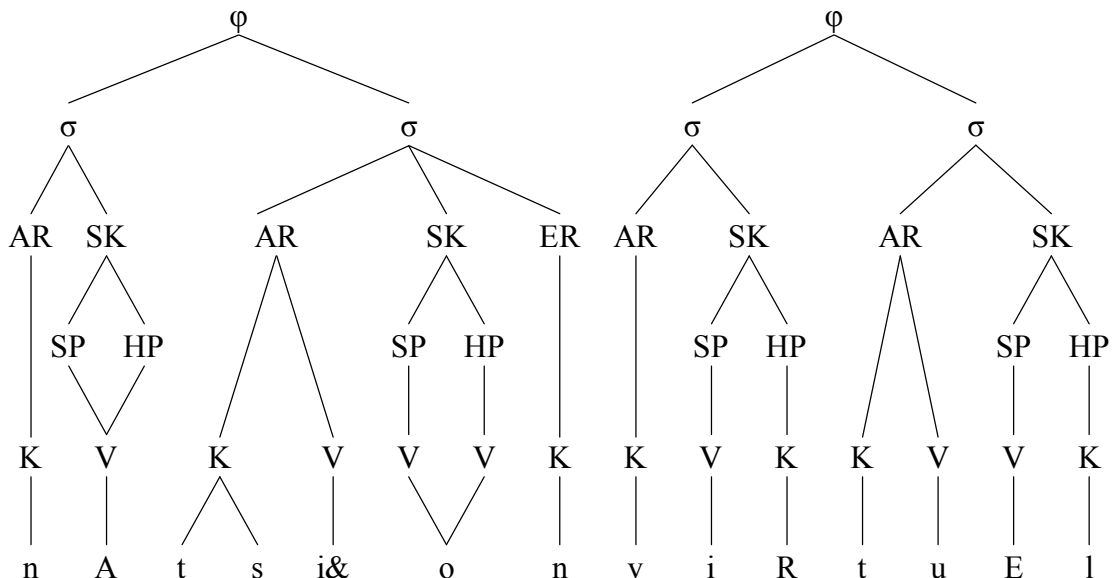


(49) Zugrunde liegende Repräsentation der der schließenden Diphthonge des Deutschen<sup>49</sup>



Während in den schließenden Diphthongen für den scharfen Schnitt des Kopfvokals der Randvokal verantwortlich ist, hat er in den steigenden Diphthongen (vgl. (5)b) keinerlei Einfluss auf die Schnitteigenschaften des Kopfvokals: in der zweiten Silbe von *Nation* steht der Kopfvokal unter sanftem, in der ebenfalls letzten Silbe von *virtuell* dagegen unter scharfem Schnitt. Das legt nahe, die steigenden Diphthonge außerhalb des Silbenkerns unterzubringen. Wir wollen nun den Randvokal dieser Diphthongen ähnlich wie Wiese (1996) und Becker (1998) als Teil des Anfangsrandes betrachten, und ihren Kopfvokal entweder nur mit der Sonoritätsposition oder mit beiden subnuklearen Positionen verbinden. Somit lassen sich zu den steigenden Diphthongen des Deutschen die zugrunde liegenden Repräsentationen in (51) angeben.

(50) Silbenstruktur der Wörter *Nation* und *virtuell*



(51) Zugrunde liegende Repräsentation der steigenden Diphthongen im Deutschen

<sup>49</sup> Zu den in der Abbildung vorkommenden IPA-Symbolen s. den Kommentar in Fn. 28.



die in den nicht-Silbenschnittsprachen mit demselben Vokal verbunden sind. In den Silbenschnittsprachen ist die unterschiedliche Besetzung dieser subnuklearen Konstituenten für die Unterscheidung zwischen den beiden Silbenschnitten verantwortlich: beim sanften Schnitt ist ein Vokal schon zugrunde liegend mit beiden Positionen, beim scharfen dagegen nur mit der Sonoritätsposition verbunden, in diesem Fall wird die Halteposition im Laufe der Silbifizierung mit dem postvokalischen Segment assoziiert. Nasalvokale können – wenn sie im Standarddeutschen überhaupt angenommen werden müssen – ähnlich wie sanft geschnittene Vokale analysiert werden, d.h. sie besetzen beide subnuklearen Positionen. Schließende Diphthonge sind ebenfalls im Silbenkern unterzubringen: der Kopfvokal ist dabei schon zugrunde liegend mit der Sonoritätsposition assoziiert, der Randvokal enthält dagegen keinerlei lexikalische Spezifizierung für die subnuklearen Konstituenten.

## **4. Die Reduktionssilbe im Deutschen**

Gegenstand des vorliegenden Kapitels ist der wohl eigentümlichste Vokal des Gegenwartsdeutschen, das Schwa. Die zentrale Idee der hier dargelegten Beschreibung ist die Annahme eines eigenen Silbentyps, der Reduktionssilbe, mit der viele der Eigentümlichkeiten dieses Vokals auf eine einfache Weise erfasst werden können. Das Kapitel ist wie folgt gegliedert. In einem ersten Schritt (vgl. Abschn. 4.1) sollen die Besonderheiten dieses Vokals ermittelt werden. Darauf folgt die Behandlung der internen Struktur des Schwa in Abschn. 4.2. Abschn. 4.3 ist einer kritischen Besprechung einiger früherer Schwa-Beschreibungen gewidmet, die als Grundlage zur Herleitung der hier vorgeschlagenen Analyse in den Abschnitten 4.4 bis 4.7

dienen wird. Das Kapitel endet mit einer kurzen Zusammenfassung der Ergebnisse (vgl. Abschn. 4.8).

### **4.1 Das Schwa-Phänomen im Deutschen**

Wie oben besprochen zeigt das Schwa eine Fülle von Besonderheiten, welche die phonetische Realisierung, die phonologische Wertung, das phonotaktische Verhalten dieses Vokals sowie seine Alternationsfähigkeiten betreffen. Im Folgenden sollen diese Besonderheiten im Einzelnen untersucht werden.

### 4.1.1 Phonetik des Schwa

Das Schwa ist phonetisch gesehen ein nichtrunder mittlerer Zentralvokal, es wird also in der absoluten Indifferenzlage des Artikulationsraumes gebildet, vgl. Abb.1. Bei diesem Segment handelt es sich demnach eigentlich nicht um einen vollwertigen, sondern vielmehr um einen *unbestimmten* (Trubetzkoy 1939), *minimalen Vokal* (Giegerich 1987), um *minimale Vokalität* also (Vennemann 1991). Eine adäquate Schwa-Beschreibung soll diesem Sachverhalt unbedingt entgegenkommen.

### 4.1.2 Schwa und Akzent

Ein anderes ebenfalls wichtiges Charakteristikum dieses Vokals ist, dass er nur in unbetonter Stellung vorkommt, wie dies die Beispiele unter (2)a auch zeigen. Eine Silbe mit Schwa ist jedoch nicht nur unbetont, sondern im Allgemeinen unbetonbar (vgl. Wiese 1988: 140). Daten wie die in (52), in denen ein Schwa doch unter Akzent vorkommt und in Abhängigkeit vom Silbenschnitt als langes gespanntes [e:] (sanfter Schnitt) oder als kurzes ungespanntes [E] (scharfer Schnitt) realisiert wird, stellen im Grunde genommen keine echten Gegenbeispiele dar und sprechen somit nicht für die Akzentuierbarkeit einer Schwa-Silbe. Dabei handelt es sich nämlich um den metalinguistischen Gebrauch des Akzents (vgl. Vennemann 1991a: 211), um Kontrastakzent also, bei dem auf die sprachliche Form selbst fokussiert wird, in welchem Fall die im Lexikon vergebenen Akzente vom Satzakzent überschrieben werden können (vgl. Wiese 1988: 166). In einer lexikalisch-phonologischen Beschreibung des Deutschen sollte man also die generelle Unbetonbarkeit einer Schwa-Silbe für das Lexikon festhalten, und ihre exzeptionelle Betonung in der postlexikalischen Phonologie zulassen (vgl. ebda).

(52) *einé, nicht einén, bé- und éntladen*

(53) a. *lébend – lebéndig*

b. *Itálien – Italiéner(in) – italiénisch*

c. *Charákteer – Charaktére*

Dass eine Silbe mit Schwa im Lexikon unbetonbar ist, bedeutet, dass beim Schwa keine systematische Akzentvariation der Art *Musik–Músiker–musikálisch* (vgl. (9)) zu beobachten ist. Betrachtet man die Daten unter (53), so findet man jedoch genau das Gegenteil.

Suffigierung mit *-ig*, *-er*, *-isch*, bzw. Hinzufügung der Pluralendung *-e* bewirkt in allen drei Fällen Verschiebung des Akzents auf die letzte Stammsilbe, die ein Schwa enthält, welches aufgrund der Kontrastakzentuierung unter (53)a erwartungsgemäß in Abhängigkeit von der Silbenstruktur als scharf geschnittener (phonetisch kurzer ungespannter, wie in (53)a) oder eben sanft geschnittener (phonetisch länger gespannter wie in (53)b–c) Vokal realisiert wird. Doch sprechen diese Daten m.E. nicht für die lexikalische Betonbarkeit von Schwa-Silben. Ein erstes Argument dagegen kommt von der Anzahl dieser Variation: möglicherweise finden sich außer den Beispielen unter (53) keine weiteren Fälle, sodass hier keineswegs vom systematischen Wechsel zu sprechen ist. Die fraglichen Daten sind jedoch nicht nur im Hinblick auf die Akzentverschiebung auf eine Schwa-Silbe, sondern auch aus anderen Gesichtspunkten phonologisch wie morphologisch exzeptionell. Zum Wortpaar *lebend* – *lebendig* ist anzumerken, dass nach Fleischer & Barz (1992: 258) *ig*-Suffigierung auf adjektivische Basis nur vereinzelt anzutreffen und äußerst unproduktiv ist, vgl. *voll* → *völlig*, *nieder* → *niedrig*. Außer *lebend* findet man auch kein weiteres Präsenspartizip, das das Suffix *-ig* erhalten würde. So sind im rückläufigen Wörterbuch von Mater (1970) neben *lebendig* nur zwei ähnliche Wörter zu finden: *elendig* und *behendig* (nach der neueren Rechtschreibung: *behändig*). Doch handelt es sich dabei bloß um eine formale Ähnlichkeit: der Stamm ist im ersten Fall ein Substantiv (*Elend*), im zweiten zwar ein Adjektiv (*behände*), aber kein Präsenspartizip. Beide Stämme haben darüber hinaus gemeinsam, dass die *ig*-Suffigierung an sie keine Akzentverschiebung auslöst. So bildet *lebendig* eine Ausnahme in der Wortbildung des Deutschen. Vielleicht sollte man dieser Gegebenheit dadurch Rechnung tragen, dass man *lebendig* nicht als Wortbildungsprodukt, sondern vielmehr als eigenständigen Lexikoneintrag betrachtet.

Was die Fälle unter (53)b und (53)c betrifft, so zeigen sie sicherlich eine systematische Akzentverschiebung und haben dadurch in der betonten Silbe ein als sanft geschnittenes (phonetisch langes gespanntes) [e:] realisiertes Schwa. Doch stellen sie Einzelfälle dar: bei der *er*-Suffigierung auf Ländernamen auf *-ien* (und sich daran anknüpfend bei der Suffigierung mit *-isch*) wird in der Basis die Endung *-en* (beim Herkunfsadjektiv auch *-i*) normalerweise getilgt: *Argentinien* – *Argentinier* – *argentinisch* (vgl. Fleischer & Barz 1992: 155). (53)b ist möglicherweise eine einzigartige Ausnahme. Und ähnlich verhält es sich mit (53)c: Substantive auf *-er* bilden den Plural mit dem Nullsuffix, *e*-Plural bei einem solchen Nomen ist eine Ausnahme. Trotz dieser bei der Wortbildung bzw. Nominalflexion beobachtbaren Exzeptionalität ist der auf die Schwa-Silbe fallende Akzent nicht ganz irregulär. So enden *er*- und *isch*-Bildungen dieser Art gewöhnlich auf einen trochäischen Fuß,

genauso wie die unregelmäßigen Pluralformen (also alle bis auf den *s*-Plural, vgl. Wiese 1996). Diese prosodisch-morphologische Beschränkung scheint im Deutschen so stark zu sein, dass sie die im übrigen ebenfalls sehr starke prosodische Beschränkung der lexikalischen Unbetonbarkeit der Silben mit Schwa ‚überschreibt‘. Es geht hier jedoch bloß um eine ausnahmsweise ‚Überschreibung‘, nicht jedoch um eine generelle ‚Außerkraftsetzung‘ der fraglichen Schwa-Beschränkung.

Wenn hier bisher über die exzeptionelle Schwasilben-Betonung die Rede war, müsste auch der umgekehrte Fall, d.h. die Reduktion eines Vollvokals zu Schwa kurz behandelt werden. Vollvokale können nämlich unter Akzentschwund in Abhängigkeit vom Sprechtempo und Sprechstil als Schwa realisiert werden. Dies gilt v.a. für die Vokale in Funktionswörtern, so kann /e:/ in *den* (über die Zwischenstufen [e] und [E]) bzw. /i:/ in *die* (über die Zwischenstufen [i] und [ç]) zu [W] reduziert werden, wobei es beim ersteren Wort sogar ausfallen kann, woraus sich der silbische Nasal [n1] resultiert (vgl. Meinhold 1973). Derselbe Prozess kann sich jedoch auch in Inhaltswörtern abspielen: *Metal* kann abhängig vom Sprechtempo und Sprechstil als [e], als [E] oder als [W] gesprochen werden. Bezüglich der Vollvokalreduktion in Inhaltswörtern ist zweierlei anzumerken. Erstens erweisen sich dabei die *e*-Vokale als gewissermaßen ‚reduktionsfähiger‘, d.h. sie werden häufiger zu Schwa reduziert als die anderen Vollvokale (vgl. Ramers 1988: 220, Becker 1998). Zweitens ist mit Vennemann (1991a: 212) ein spezifischer Fall der Schwa-Reduktion zu nennen, bei dem alle Vollvokale zwischen zwei rhythmisch prominenten Silben zu Schwa reduziert werden können (vgl. (54)). Eine adäquate Theorie des Schwa im Deutschen soll alle diese akzentmäßigen Besonderheiten erfassen können. Wir werden zeigen, dass der von uns entwickelte Ansatz dieser Anforderung durchaus entgegenkommen kann.

- (54) *Apparat* → *App[W]rat*  
*Molekül* → *Mol[W]kül*  
*Aspirin* → *Asp[W]rin*  
*Lokomotive* → *Lok[W]motive*  
*Känguruh* → *Käng[W]ruh*  
*Aphärese* → *Aph[W]rese*  
*Diozese* → *Di[W]zese*  
*amüsieren* → *am[W]sieren*

### **4.1.3 Schwa und Silbenschnitt**

Eine wiederum andere Schwa-Besonderheit betrifft die Diskrepanz zwischen Phonetik und Phonotaktik dieses Vokals im Hinblick auf die Silbenschnittopposition im Deutschen. Geht man nämlich im Einklang mit dem im vorausgehenden Kapitel Gesagten davon aus, dass auch in unbetonten Silben Silbenschnittunterschiede bestehen, so kann Folgendes festgehalten werden: Scharf geschnittene Vokale sind solche, die nur in geschlossenen Silben vorkommen und phonetisch gespannt und kurz realisiert werden, während sanft geschnittene auch in offenen Silben stehen können und phonetisch ungespannt und in Abhängigkeit vom Akzent lang (wenn betont) bzw. kurz (wenn unbetont) erscheinen. Die oben angedeutete Diskrepanz zwischen Phonetik und Phonotaktik des Schwa lässt sich nun wie folgt formulieren. Phonetisch gesehen steht das Schwa den ungespannten (zentralisierten) Kurzvokalen nahe: Es ist ein Kurzvokal, der im Zentrum des Artikulationsraums gebildet wird (s. oben), sodass es durchaus als zentralisierter (ungespannter) Vokal aufgefasst werden kann; es wird außerdem kurz realisiert. Dies spricht für eine Einordnung des Schwa in die scharf geschnittene Vokalreihe. Phonotaktisch gesehen stellt das Schwa jedoch einen Vokal dar, der auch in offenen Silben vorkommt, was eine Unterbringung unter den sanft geschnittenen Vokalen sinnvoll macht. Dieser phonetisch-phonotaktische Widerspruch lässt sich nur auflösen, wenn man annimmt, dass das Schwa ein Vokal des Gegenwartsdeutschen ist, der überhaupt keine Schnitteigenschaften aufweist, also sozusagen außerhalb der Silbenschnittopposition steht.<sup>50</sup> In einer adäquaten Schwa-Beschreibung soll diese Gegebenheit ebenfalls ausgedrückt werden können.

### **4.1.4 Vorkommen des Schwa im Deutschen**

Auch das Vorkommen des Schwa zeigt Besonderheiten auf. Bei der Auflistung der Schwa-Vorkommen im Deutschen erscheint es zweckmäßig, von der Grobeinteilung Issatschenkos (1974) auszugehen. Der Autor unterscheidet nach der Alternationsfähigkeit dieses Vokals

---

<sup>50</sup> Wie das in Abschn. 3.6 dargelegt liegt auch bei den Nasalvokalen ein auf den ersten Blick ähnlicher Fall vor: auch diese Vokale zeigen nämlich ein widersprüchliches Verhältnis zum Silbenschnitt auf. Dieser Widerspruch lässt sich bei ihnen jedoch noch auflösen: nicht nur rein ‚zahlenmäßig‘ sprechen mehr Argumente für ihre Auffassung als sanft geschnittene Vokale, ihre Ungespanntheit kann außerdem auch als Folge der nasalen Artikulation betrachtet werden.

grundsätzlich zwischen zwei Typen: Schwa mobile und Schwa constans. Beim Ersteren handelt es sich um ein Schwa, das in Abhängigkeit von morphologischen und/oder prosodischen Faktoren eine Schwa-Zero-Alternation zeigt, vgl. (55). Schwa constans ist dagegen nicht tilgbar, ist immer präsent, vgl. (56).

- (55) *Atem* vs. *AtØmung*, *atØmen*, (*kurz*)*atØmig* usw.  
*Segen* vs. *SegØnung*, *segØnen*, *SegØner* usw.  
*Segel*, *segeln*, *Segelung* vs. *SegØler* usw.  
*Feuer*, *feuern*, *Feuerung* vs. *feuØre*, *feuØrig* usw.

- (56) *atmet*, *trocknen*, *schnelles*, *Hauses*, *Wiese*, *Stelle*, (*ich*) *gebe* usw.

**Schwa mobilia** erscheinen im Deutschen stets in Verbindung mit einem nachfolgenden Sonoranten, also einem Nasal [m, n] oder einer Liquide [l, R]. Das Vorkommen und Verhalten der Schwa mobilia wird in Tab. 13 zusammengefasst.

	Nasal		/l/		/R/	
	[W]	Ø	[W]	Ø	[W]	Ø
a. Stamm	<i>A<u>t</u>em</i> <i>Se<u>g</u>en</i>	–	<i>Se<u>g</u>el</i>	–	<i>Fe<u>u</u>er</i>	–
b. Verb						
i. Infinitiv <sup>51</sup>	–	<i>at<u>m</u>en</i> <i>se<u>g</u>nen</i>	<i>se<u>g</u>eln</i>	–	<i>fe<u>u</u>ern</i>	–
ii. 1.S.I.Präs.A	–	<i>at<u>m</u>e</i>	<i>se<u>g</u>ele</i>	<i>se<u>g</u>le</i>	<i>fe<u>u</u>ere</i>	<i>fe<u>u</u>re</i>
iii. 2./3.S.I.Präs.A	–	<i>at<u>m</u>e(s)t</i> <i>se<u>g</u>ne(s)t</i>	<i>se<u>g</u>el(s)t</i>	–	<i>fe<u>u</u>er(s)t</i>	–
c. Adjektiv	<i>eig<u>e</u>ne(n)</i>	<i>(eig<u>n</u>en)</i>	–	<i>dunk<u>l</u>e(n)</i>	<i>mager<u>e</u>(n)</i>	<i>(mag<u>r</u>en)</i> <i>(teu<u>r</u>en)</i>

Tab. 13 Schwa mobilia im Deutschen

Aus Tab. 13 geht folgendes hervor:

- a. In **monomorphemischen Wörtern**, d.h. in unsuffigierten Stämmen erscheint immer ein Schwa, u.z. unabhängig vom Typ des nachfolgenden Sonoranten. Auf der Oberfläche wird die Sequenz [W]+Sonorant im Falle der Nasale und des Laterals in Abhängigkeit vom

<sup>51</sup> Das hier Gesagte bezieht sich auch auf das Perfektpartizip.



Sprechtempo und -stil entweder als [Wn], [Wm], [Wl] (Explizitform) oder als silbischer Nasal bzw. Lateral [n1, m1, l1] (Reduktionsform) realisiert. [WR] erscheint auf der Oberfläche dagegen obligatorisch als vokalisches *r* [Ç].

- b. Die **Verbformen** zeigen schon ein etwas bunteres Bild. In Verbstämmen erscheint vor einem stammfinalen Nasal kein Schwa, in solchen mit stammauslautender Liquida ist das Schwa dagegen im Allgemeinen vorhanden. Eine Ausnahme bildet hierunter die 1. Person Singularform (vgl. bii): hier bestehen parallele Formen mit und ohne [W].<sup>52</sup>
- c. In den **Adjektivstämmen** mit auslautendem Nasal und *r* bleibt das Schwa immer erhalten, in denen mit stammfinalen Lateral dagegen nur in der endungslosen Form. Doch ist auch hier ein gewisses Ausmaß an Variation zu beobachten: die schwalose Form ist auch in mit Endungen versehenen Stämmen mit finalem Nasal und *r* nicht agrammatisch.<sup>53</sup>

**Schwa constans** erscheint im Deutschen entweder vor einem Konsonanten, oder ohne Folgekonsonanz. **Schwa constantia vor Konsonanz** treten in der Substantivdeklinaton, in der Adjektivdeklinaton, in der Konjugation, in der Wortbildung sowie in einigen wenigen Stämmen auf. Konstantes Schwa erscheint innerhalb der **Substantivdeklinaton** zwischen einem Stamm auf [s] oder [ts] und der Endung des Singular Genitiv der Maskulina und Neutra (vgl. (57)a). Fakultativ ist das Schwa dagegen in dieser Position nach betonter Silbe auf Konsonant (vgl. (57)b). Die Sequenz |WR| kommt außerdem als Pluralendung bei Maskulina und Neutra vor (vgl. (57)c). Schwa constans kommt in allen Endungen der **Adjektivdeklinaton** vor (vgl. (57)d). In der **Konjugation** tritt konstantes Schwa vor den beiden Endungen *-t* und *-st* nach Stämmen mit finalem Nasal und daher ohne stamminternes Schwa (vgl. (57)e), sowie in der Infinitiv- und 1./3. Person Pluralform auf, wenn der Stamm nicht auf die Sequenz [Wl]/[WR] endet (vgl. (57)f). In der **Wortbildung** erscheint ein Schwa

---

<sup>52</sup> Während Wiese (1988) in dieser freien Variation die Formen mit Schwa als primär betrachtet, nimmt Giegerich (1987) diesbezüglich nicht Stellung. Nach Helbig & Buscha (1989: 27) ist bei Stämmen mit finalem Lateral nur die schwalose Form möglich, bei denen mit finalem *r* dagegen beide, wobei die schwalose als umgangssprachlich gilt.

<sup>53</sup> Anzumerken ist dabei, dass nach Helbig & Buscha (1989: 302) bei auf *r* auslautenden Stämmen in fremden Adjektiven und in Adjektiven mit Diphthong nur die schwalose Form möglich ist. Allerdings waren unter den insgesamt 21 Belegen nach dem rückläufigen Wörterbuch von Mater (1970) lediglich 3 solche Fälle (*integer, teuer, sauer*) zu finden.

constans in bestimmten Ableitungssuffixen, so in *-chen* oder *-er* (vgl. (57)g). Schließlich sind einige Beispiele für Schwa constantia in **Stämmen** unter (57)h aufgelistet.

(57) Schwa constantia vor Konsonanz

- a. *des Hauses, des Kusses, des Schmerzes*
- b. *des Schiff(e)s, des Wind(e)s, des Teich(e)s*
- c. *Kinder, Würmer, Ränder, Irrtümer*
- d. *schönes, schönen, schönem, schöner*
- e. *atme(s)t, segne(s)t, rette(s)t, rede(s)t*
- f. *atmen, segnen, retten, reden, bauen, sehen*
- g. *Mädchen, Märchen, Seher, Verbraucher*
- h. *Abend, Tugend, Abenteuer, Dromedar*

**Schwa constantia ohne Folgekonsonanz** erscheinen als selbstständige Morpheme, als Teil selbstständiger Morpheme und in vielen Stämmen. Schwa als Morphem kommt in der **Substantivdeklinaton** als Pluralendung (vgl. (58)a) oder als Endung des Dativ Singular der Maskulina und Neutra vor, wenn deren Stamm auf eine betonte konsonantisch auslautende Silbe endet (vgl. (58)b). In der **Adjektivdeklinaton** erscheinen sie in allen drei Deklinatonstypen (vgl. (58)c), in der **Konjugation** z.B. als die Endung der 1. Person Singularform (vgl. (58)d), in der **Wortbildung** ferner als selbstständiges Suffix (vgl. (58)e). Sie kommen aber auch in den beiden Präfixen *be-* und *ge-*, bzw. in einer großen Fülle von Wörtern vor (vgl. (58)f, g).

(58) Schwa constantia ohne Folgekonsonanz

- a. *Tische, Ärzte, Stühle, Schränke, Boote, Hunde, Tage*
- b. *(dem) Volke, (dem) Balle, (dem) Buche, (dem) Kinde*
- c. *schöne Frau(en), das schöne Kind, eine schöne Frau*
- d. *(ich) leb(t)e, (ich)/(er) lebe, (ich)/(er) könnte, lebe!*
- e. *Stelle, Ausgabe, Aufnahme, Miete, Presse*
- f. *beenden, beweisen, gekommen, gefallen*
- g. *Wiese, Tante, Katze, Kante, Bande, bange*

### **4.1.5 Das Schwa-Phänomen: eine Zusammenfassung**

Wie aus den obigen Erörterungen hervorgeht, wirft der Vokal [ʷ] im Deutschen eine Fülle von Fragen auf, die wie folgt formuliert werden können:

(59) Interne Struktur des Schwa:

Wie lässt sich der neutrale phonetische Charakter des Schwa phonologisch erfassen?

(60) Schwa und Akzent

- a. Wie kann die Unakzentuierbarkeit des Schwa phonologisch erreicht werden?
- b. Wie sollen die Fälle der exzeptionellen Schwa-Betonung beschrieben werden?
- c. Wie lässt sich die Reduktion der Vollvokale zu Schwa ausdrücken?

(61) Schwa und Silbenschnitt

Wie kann man zum Ausdruck bringen, dass das Schwa außerhalb der Silbenschnittopposition des Deutschen steht?

(62) Schwa-Vorkommen

Wie sollen die Schwa-Vorkommen erfasst werden?

(63) Schwa-Alternationen

- a. Wie kann das Verhältnis des Schwa zum vokalischen *r* ausgedrückt werden?
- b. Wie kann das Verhältnis des Schwa zu den silbischen Sonanten erfasst werden?

Eine adäquate Schwa-Theorie soll zu all diesen Problemen eine Lösung bieten. Wie das weiter unten gezeigt wird, erweist sich die Theorie der Reduktionssilbe in dieser Hinsicht als ein angemessener Beschreibungsansatz.

## **4.2 Die interne Struktur des Schwa**

Der vorliegende Abschnitt konzentriert sich auf die minimale phonetische Ausprägung des Schwa und will somit auf die unter (59) formulierte Frage eine Antwort geben.

In den mit binären unstrukturierten phonologischen Merkmalen operierenden traditionellen phonologischen Beschreibungen kann dem minimalen phonetischen Charakter des Schwa nur indirekt Rechnung getragen werden. Das bedeutet, dass das Schwa wie alle anderen Vokale des Deutschen mit einem vollständigen Merkmalbündel repräsentiert wird, das Schwa-Bündel

enthält jedoch zumindest für die Merkmale der Ansatzrohraktivität ausschließlich negative Werte, die infolge der Tatsache, dass eine negative Merkmalspezifizierung immer das Fehlen einer artikulatorischen Eigenschaft (z.B. Bewegung) anzeigt, insgesamt auf die fehlende Ansatzrohraktivität, d.h. die phonetische Minimalität, hinweisen. So sind beispielsweise in Meinhold & Stock (1982, 82) oder Wiese (1996, 152) die Merkmale [±hoch], [±niedrig], [±vorn], [±hinten] und [±rund] im Schwa-Merkmalbündel jeweils mit ihrem negativen Wert vertreten, vgl. (64).

(64) Merkmalspezifizierung des Schwa bei Meinhold & Stock (1982) und Wiese (1996)

a. Meinhold & Stock (1982, 82)

+ vokalisch
– hoch
– niedrig
– vorn
– hinten
– rund
0 gespannt
0 lang

c. Wiese (1996, 152)

– konsonantisch
– hoch
– niedrig
– vorn
– hinten
– rund
– ATR
– lang

An der indirekten Markierung der phonetischen Minimalität des Schwa wird auch durch die Anwendung der Theorie der Unterspezifikation nichts geändert, nach der ein Segment für die vorhersagbaren Merkmale zugrunde liegend nicht spezifiziert werden und die betreffenden Merkmalswerte erst im Laufe der Derivation durch *default*-Regeln erhalten soll. Zwar kommt Wiese (1996) bei der Durchführung der Eliminierung der Redundanzen aus der Schwa-Repräsentation ausgehend von der Vollspezifizierung mit nur negativen Merkmalswerten unter (64)b zu dem Schluss, dass das Schwa phonologisch lediglich als „eine bloße prosodische Position“ (154) zu repräsentieren sei, diese prosodische Position erhält letzten Endes doch die fraglichen negativen Merkmalswerte. Somit kann die Unterspezifikationstheorie zu einer einfachen (ökonomischen) Beschreibung der deutschen zugrunde liegenden Vokale durchaus beitragen, doch bleibt in ihr die nicht unmittelbare Repräsentation der Neutralität des deutschen Schwa erhalten.

Eine noch einfachere Darstellung der internen Struktur der Vokalsegmente, die zugleich auch die phonetische Minimalität des Schwa und überhaupt die phonetische Komplexität der einzelnen Vokalsegmente im Deutschen ausdrücken kann, geht von der Annahme der Uniarität (Einwertigkeit) der phonologischen Merkmale aus. Die Idee der Einwertigkeit (vgl. etwa Lombardi 1991 oder Clements & Hume 1995) besteht darin, dass Merkmale nur auf das Vorhandensein, nicht aber auf das Fehlen der in ihnen ausgedrückten artikulatorischen Eigenschaften hinweisen sollen, was technisch durch die fehlenden Vorzeichen + oder – erfasst wird, sodass das Vorhandensein der fraglichen Eigenschaft allein durch die Anwesenheit des Merkmals in der internen Struktur des Segments angezeigt wird. Fasst man nun die oben genannten vokalischen Ansatzrohraktivitätsmerkmale uniär auf, so bedeutet [hoch] eine Hebung, [niedrig] eine Senkung, [vorn] eine Vorverlagerung, [hinten] eine

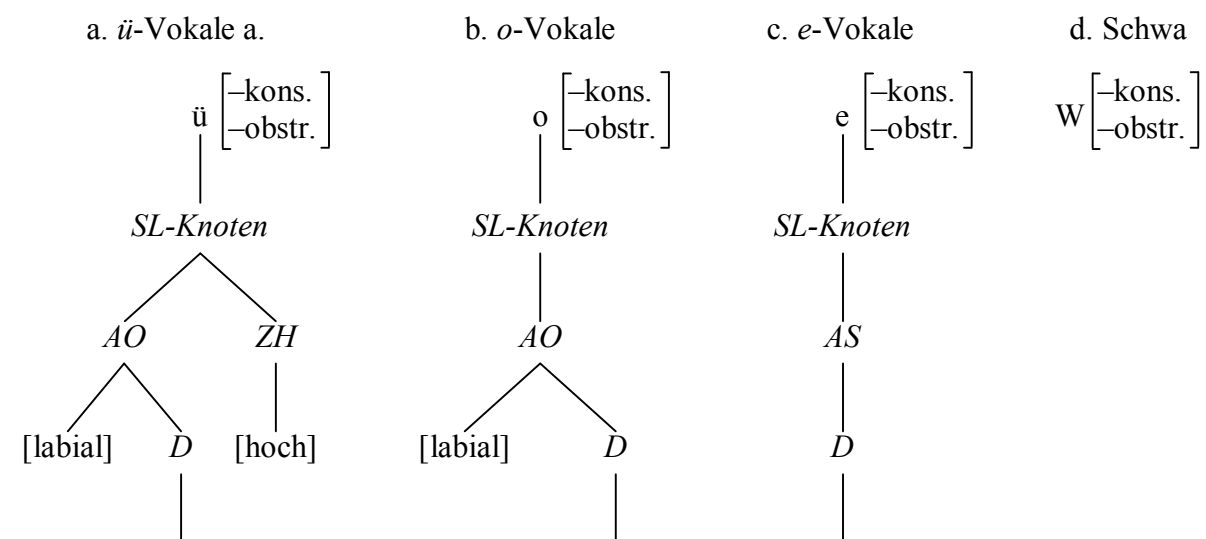
Rückverlagerung der Zungemasse aus seiner neutralen Position hinaus, und [rund] schließlich Lippenrundung, während eine fehlende Hebung, Senkung, Vorverlagerung, Rückverlagerung der Zungenmasse, bzw. eine fehlende Lippenrundung in dieser Theorie explizit nicht ausgedrückt werden können. Eine solche Darstellungsweise ist im Vergleich zu anderen Repräsentationsverfahren in zweierlei Hinsicht einfacher. Einerseits wird hier die Anzahl der Merkmale zwar nicht verringert, trotzdem können die einzelnen Vokalsegmente infolge der Beschaffenheit der Merkmale, dass sie nur auf eine einzige Kategorie referieren können, insgesamt einfacher (sprich: mit weniger Merkmalen) eindeutig bestimmt werden. So lässt sich ein *ü*-Vokal mit drei ([hoch vorn rund]), ein *o*-Laut dagegen nur noch mit zwei ([hinten rund]), ein *e*-Vokal lediglich mit einem Merkmal ([vorn]) spezifizieren, während zur eindeutigen Bestimmung des Schwa kein Merkmal angegeben werden braucht. Andererseits sind hier keine *default*-Regeln nötig, die den unspezifizierten Merkmalen die unmarkierten Werte zuweisen, wodurch die phonologische Beschreibung insgesamt vereinfacht wird, und die Repräsentationsmethode mühelos in die beschränkungs-basierten Theorien (etwa in die Optimalitätstheorie) übertragen werden kann. Die unmittelbare Erfassbarkeit der phonetischen Komplexität/Minimalität eines Vokalsegments wird schließlich dadurch erreicht, dass für die vertikale und horizontale Zungenbewegung sowie die Lippenrundung als zentrale Beschreibungs- und Klassifizierungskriterien genau diejenigen Vokale keine Merkmale erhalten, die angesichts des fraglichen Kriteriums als neutral betrachtet werden können, bei deren Bildung sich also der betreffende Artikulator (Zungenmasse vertikal, Zungenmasse horizontal bzw. Lippen) in seiner Ruhelage befindet. Somit weist die Anwesenheit eines vokalischen Merkmals immer auf eine Abweichung im Vergleich zur Ruhelage hin. Demnach ist ein *ü*-Vokal mit seiner dreifachen Abweichung von der Ruhelage (Senkung und Vorverlagerung der Zungenmasse, Rundung der Lippen) ziemlich komplex, ein *o*-Laut mit nur noch zwei Abweichungen von der Indifferenzlage (Rückverlagerung der Zunge und Rundung der Lippen) dagegen schon weniger komplex, ein *e*-Vokal mit nur noch einer einzigen solchen Abweichung (Vorverlagerung des Zungenkörpers) noch weniger komplex/mehr minimal, während ein Schwa als minimal zu betrachten ist: für seine Bildung ist die absolute Ruhelage der fraglichen Artikulatoren charakteristisch.

Ein Nachteil der Annahme einwertiger Merkmale besteht darin, dass sich mit ihr die ansonsten nur negativ bestimmbareren Vokalklassen (etwa: die nicht-hohen, die mittleren, die zentralen usw. Vokale) nicht explizit erfassen lassen, was bedeutet, dass in dieser Theorie keine phonologischen Prozesse beschrieben werden können, die solche Vokale betreffen. Dies bereitet für das Deutsche – soweit ich sehe – jedoch keine Probleme. Höchstens könnten in

dieser Hinsicht nur die *a*-Vokale als problematisch betrachtet werden, die phonetisch den zentralen Vokalen näher stehen, und aus diesem Grund im Allgemeinen als [–vorn –hinten] charakterisiert werden (vgl. etwa Meinhold & Stock 1982, Wiese 1996). Doch verhalten sich die deutschen *a*-Vokale – wie das bereits im vorausgehenden Kapitel ausführlich beschrieben wurde – phonologisch analog zu den hinteren Vokalen, was bei der Umlautbildung sowie der Ach-Ichlaut-Alternation sichtbar wird. Ein solcher ‚Zwiespalt‘ zwischen Phonetik und Phonologie eines Segments kann in der hier skizzierten Repräsentationsweise nicht zum Ausdruck gebracht werden. Da jedoch eine phonologische Repräsentation in erster Linie dazu dient, dass man mit ihrer Hilfe phonologische Prozesse ausdrückt, steht m.E. nichts im Wege, die *a*-Laute ähnlich wie in Becker (1998, 15) trotz ihrer phonetisch mehr zentralen Bildungsweise phonologisch als hintere Vokale anzusehen und mit dem Merkmalswert [hinten] zu versehen. Dies zeigt, dass die hier angewendete Merkmaltheorie nicht primär phonetisch, sondern vielmehr phonetisch-phonologisch motiviert ist.

Die hier vorgeschlagenen einwertigen Merkmale lassen sich natürlich auch merkmalsgeometrisch darstellen, für welche Repräsentation – wie darauf im einführenden Kapitel hingewiesen wurde – in der Fachliteratur zahlreiche Argumente genannt werden. Nimmt man als Grundlage die von Wiese (1996, 29) für das Deutsche vorgeschlagene Merkmalstruktur und ersetzt man die da vorkommenden binären Merkmale durch die oben genannten uniären, so lässt sich die interne Struktur der *ü*-, *o*- und *e*-Laute sowie des Schwa wie in (65) angeben.

(65) Interne Struktur der deutschen *ü*-, *o*- und *e*-Vokale sowie des Schwa<sup>54</sup>



<sup>54</sup> *SL-Knoten* steht für *Supralaryngal-Knoten*, *AO* für *Artikulationsort* und *D* für *Dorsal*. Anstelle von [rund] kommt hier als uniäres Merkmal [labial] vor.

[vorn]

[hinten]

[vorn]

Zusammenfassend kann festgehalten werden, dass mit der Annahme uniärer Merkmale im deutschen Vokalsystem die phonetische Minimalität des Schwa unmittelbar zum Ausdruck gebracht werden kann: Schwa wird demnach nur mit den beiden Oberklassenmerkmalen repräsentiert, und das zusammen mit den fehlenden Qualitätsmerkmalen, die folgerichtig auf die Zero-Ansatzrohraktivität hinweisen, ergibt in der Tat das, was Vennemann (1991) als *minimale Vokalität* bezeichnet.

### 4.3 Frühere Beschreibungen zum Schwa im Deutschen

Nach der Klärung der internen Struktur des Schwa im Deutschen wollen wir unsere Aufmerksamkeit auf die weiteren oben dargelegten Probleme dieses Vokals richten. Dazu sollen von der enormen Fülle der phonologischen Literatur zum Schwa, die Arbeiten in der strukturalistischen Phonologie (vgl. etwa Issatschenko 1974), solche in der klassischen generativen Phonologie (vgl. etwa Wurzel 1970, 1980, 1981, Kloeke 1982 oder Strauss 1982) bzw. in der nichtlinearen generativen Phonologie (vgl. Wiese 1986, 1988, 1996, Giegerich 1985, 1987, Féry 1991 oder Noske 1993) umfasst, in einem ersten Schritt drei nichtlineare generative Ansätze kurz zusammenfassen, die eine mehr oder weniger exhaustive Schwa-Beschreibung erstreben: Wieses (1988) und (1996), Giegerichs (1987) und Noskes (1993) Ansatz.

#### 4.3.1 Wiese (1988) und (1996)

Im Mittelpunkt des lexikalisch-phonologischen Ansatzes von Wiese (1988) und (1996) stehen drei Behauptungen:

- (66) a. Schwa ist „in allen seinen Kontexten und Erscheinungsformen ein im Lexikon zugewiesener Epenthesevokal“. (Wiese 1988: 140)
- b. Das Auftreten des Schwa hängt von zwei Faktoren ab (vgl. ebda):
- i. von silbischen Wohlgeformtheitsbedingungen und
  - ii. von der Interaktion zwischen Phonologie und Morphologie.



- c. Nicht alle deutschen Morpheme sind phonologisch „durch eine Kette konkreter Segmente spezifiziert“ (ebda).

(66)a wird durch die Formulierung der zweiteiligen Epentheseregeln in (67) Rechnung getragen (vgl. 144).<sup>55</sup> Der erste Teil der Regel (vgl. (67)a) fügt eine leere skeletale V-Position vor einem Segment ein, das die Silbifizierungsregeln nicht in die vorausgehende Silbe integrieren können. Der zweite Regelteil (vgl. (67)b) interpretiert dann dieses V als Schwa.

(67) a.  $\emptyset \rightarrow V / \text{___ } X ]_{\text{Wort}}$

- b. Ein leeres V wird mit Schwa assoziiert.

Die in (66)b formulierten Schwa-Epenthese bedingenden Faktoren bringt der Autor derart zum Ausdruck, dass er unterschiedliche Ausprägungen der Regel in (67) annimmt, die auf unterschiedlichen Ebenen des Lexikons angewendet werden. Die einzelnen Regel-Versionen nehmen dabei auf morphologische und phonologische Informationen Bezug, konkret auf die Wortklassenzugehörigkeit des Wortes sowie auf den stamm- bzw. wortfinalen Konsonanten. Der Autor nimmt für das Lexikon des Deutschen die Struktur in Tab. 14 an.

	<i>Morphologie</i>	<i>Phonologie</i>
<i>Ebene 1</i>	Derivation 1 → irreguläre Flexion ←	Wortakzent <b>W-Epenthese 1:</b> $\text{verb}[\dots[\text{Liquid}]]$
<i>Ebene 2</i>	Derivation 2 → Komposition ←	Kompositionsakzent <b>W-Epenthese 2–4:</b> $\text{Nom}[\ ]$ , $\text{Adj}[\dots[\text{Nasal}]]$ , $[\text{R}]$
<i>Ebene 3</i>	reguläre Flexion → ←	<b>W-Epenthese 5–7:</b> $\text{verb}[\dots[\text{Nasal}]]$ , $\text{Adj}[\ ]$ , $[\text{R}]$

Tab. 14 Struktur des deutschen Lexikons nach Wiese (1988: 152 und 165)

Die erste Fassung der generellen Schwa-Epenthese in (67) fügt in Verbstämmen auf Liquida auf Ebene 1 vor dieser ein Schwa ein, und ist dadurch, dass sie vor der Verbflexion angewendet wird, für die schwahaltigen Formen dieser Stämme während der Konjugation verantwortlich (vgl. Tab. 13 b.i–iii).<sup>56</sup> Die zweite Schwa-Epenthese greift auf Nomen, und führt die Schwa-Einfügung in diesen auf Ebene 2 (also ebenfalls vor der regulären

<sup>55</sup> Vgl. auch Wiese (1996: 243).

<sup>56</sup> Daraus folgt auch, dass bei den Doppelformen *segele/segle* bzw. *feuere/feure* erstere Form primär ist. Der Autor erklärt das Vorhandensein der letzteren Formen mit lexikalischer Variation.

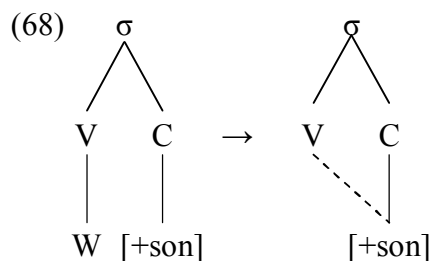
Flexion) unabhängig von der Qualität des nachfolgenden Sonoranten durch, woraus die un suffigierten schwahaltigen Nominalstämme (vgl. Tab. 13a), aber eigentlich auch das Diminutivsuffix *-chen* (vgl. (57)g) resultieren. Die dritte Ausprägung konzentriert sich auf Adjektivstämme auf Nasale, die also laut Regel auf Ebene 2, d.h. vor der Adjektivdeklinationsstufe ein Schwa erhalten. Das bringt den Unterschied zwischen *dunkle* und *trockene* zum Ausdruck: nur in Letzterem erscheint ein epenthetisches Schwa. Regel 4 behandelt alle Stämme auf /R/, also Nominalstämme wie *Vater*, *Feuer* (vgl. Tab. 13a), Adjektivstämme wie *locker*, *mager*, die in allen Erscheinungsformen (d.h. ob suffigiert oder nicht) mit stamminternem Schwa erscheinen wie *lockere*, *magere* (vgl. Tab. 13c), sowie Ableitungen mit dem Suffix /R/ (orthographisch: *-er*) wie *Leser*, *Seher* (vgl. (57)g). Die fünfte Regelfassung sorgt in Verben mit stammfinalelem Nasal dafür, dass ihre suffigierten Formen, wo nötig, ein Schwa erhalten. Dies ist der Fall bei Infinitiven (vgl. (57)f) und der 2./3. Person Singularformen (vgl. (57)e). Diese Regel ist insofern eine spezifische, als sie nicht nur auf den phonologischen Kontext ‚unsilbifizierbares finales X‘, sondern auch auf die prosodische Form der resultierenden Wortform Bezug nimmt, die bei den Infinitiven ein trochäischer Fuß ist (vgl. Wiese 1988). Version 6 ist dafür verantwortlich, dass alle un suffigierten Adjektive auf stammfinalelem Lateral sowie alle Adjektivendungen ein Schwa erhalten (vgl. (57)d). Schließlich wird die siebte Regel auf sämtliche Wörter mit finalelem *r* angewendet, vor das vorher im Lexikon kein Schwa eingefügt wurde, d.h. auf die Adjektive mit der Komparativendung /R/ (vgl. (57)d).

Vergleicht man diese kurze Zusammenfassung mit der Liste in (57), so sieht man, dass nach der Lexikonstruktur in Tab. 14 einige Fälle der Schwa-Epenthese nicht hergeleitet werden können. Schwa-Epenthese in *rette(s)t* und *rede(s)t* (vgl. (57)e) ist deswegen interessant, weil sie vom Autor thematisiert wird (vgl. 157): sie wird durch die Antigenimationsbeschränkung ausgelöst, nach der im Deutschen Geminaten grundsätzlich nicht erlaubt sind. Dass die zusammenfassende Darstellung in Tab. 14 keinerlei Hinweis auf diesen Fall hat, mag mit einem ‚Druckfehler‘ zusammenhängen: Sollte die fünfte Schwa-Regel auf keine phonologische Information (‚der Verbstamm soll auf einen Nasal enden‘) Bezug nehmen müssen, so würde sie auch diesen Fall erfassen. Anzumerken ist dabei, dass diese phonologische Bedingung sowieso überflüssig ist, da in liquidfinalen Verbstämmen Schwa-Epenthese schon auf Ebene 1 stattfindet. Schwa-Epenthese vor der Genitivendung *-s* in den Wörtern unter (57)a erklärt sich analog zu *rette(s)t* und *rede(s)t*: auch hier wird die Einfügung des Schwa durch die Adjazenz identischer Segmente ausgelöst. Da aber Wieses

Modell keine Schwa-Epenthese für Nomen auf Ebene 3 vorsieht, bleibt dieser Fall unerklärt.<sup>57</sup> Aus demselben Grund kann in diesem Modell die fakultative Schwa-Einfügung in den Genitivformen unter (57)b nicht abgeleitet werden, auch wenn der Grund gegeben ist (vgl. Wiese 1996: 108f): die bei der Pluralbildung aktive prosodische Bedingung, nach der irreguläre Pluralformen auf einen trochäischen Fuß enden müssen, wird auch auf diese Formen – wenn auch nur optional – angewendet. Dass in den Wörtern von (57)h Schwa-Epenthese stattfindet, kann der Autor nur zum Teil erklären, u.z. in den ersten beiden Wörtern, in denen der finale Konsonant als extrasilbisch erklärt wird. Die Schwa-Insertion in *Abenteuer* und *Dromedar* bleibt der Autor schuldig.

Dem epenthetischen Charakter des Schwa in den Wörtern von (58) wird in Wieses Ansatz dadurch Rechnung getragen, dass in diesen Stämmen bzw. Affixen anstelle des Schwa zugrunde liegend eine für die Silbenstruktur unspezifizierte leere skeletale Position X angenommen wird, die der Silbifizierungsalgorithmus unsilbifiziert lässt, sodass sie (auf der entsprechenden lexikalischen Ebene) die Schwa-Epenthese auslöst. Damit weist das Deutsche, das ansonsten eine verkettende (konkatenative) Morphologie hat, nach dem Autor eine Eigenschaft auf, die für die Sprachen mit nicht verkettender (non-konkatenativer) Morphologie (etwa Arabisch) charakteristisch ist, und stellt einen Mischtyp dar. Ein (relativ schwaches) Argument für diese Annahme betrifft die Flexionsendungen, die dadurch einheitlich(er) erfasst werden können (vgl. 160), ein anderes dagegen die exzeptionelle Betonung einer Schwa-Silbe: die Quantität des resultierenden Vokals hängt mit der leeren X-Position zusammen (vgl. 161).

Schließlich schlägt der Autor für die Herleitung der fakultativen silbischen Sonoranten eine postlexikalische Regel vor, die das Schwa tilgt und die leere V-Position mit dem Sonoranten im Endrand assoziiert.



<sup>57</sup> Vgl. auch die diesbezügliche Kritik Noskes (1993: 154, Fn.11).

Wieses Ansatz hat zahlreiche Vorteile. So kann er z.B. das Verhältnis zwischen Schwa und Akzent auf eine elegante Weise erfassen: die lexikalische Unbetonbarkeit des Schwa (vgl. (60)a) wird dadurch erreicht, dass die Wortakzentregeln zu einem Zeitpunkt Anwendung finden, wo Schwa als Epenthesevokal noch gar nicht vorhanden ist. Die exzeptionelle (postlexikalische) Betonung der Schwa-Silben (vgl. (60)b) lässt sich in Wieses Ansatz ebenfalls zum Ausdruck bringen: infolge der Kernbedingung von Wiese (1988) (s. oben) ist das Schwa in *eine* mit zwei, in *einen* dagegen mit einem X assoziiert, was erklärt, warum infolge der Zuweisung eines postlexikalischen Kontrastakzents an Stelle des Schwa in *eine* ein langes, in *einen* ein kurzes Vokalsegment entsteht. Dass dieser Vokal genau das mittlere vordere ungerundete /e/ ist, erklärt Wiese (1996) mit dem minimalen phonetischen Unterschied zwischen diesen Segmenten: /e/ ist das einzige Vokalsegment, das sich vom Schwa nur durch ein Merkmal (nämlich [+vorn]) unterscheidet. Die postlexikalische Reduktion von Vollvokalen zu Schwa (vgl. (60)c) wird in den genannten Werken von Wiese nicht thematisiert und könnte angesichts der in Wiese (1996) angenommenen Unterspezifikationstheorie nur schwer erfasst werden: die betreffende Regel sollte nämlich alle inhärenten Merkmalswerte von [e] in *Molekül* bzw. von [o] *Diozese* tilgen. Das Verhältnis des Schwa zur Vokalopposition in (6) wird in Wieses Beschreibung ebenfalls nicht besprochen. Die Annahme Wieses, dass das Schwa in der zugrunde liegenden Repräsentation nicht vorhanden ist, könnte nahelegen, dass dieser Vokal an dem bei Wiese als Quantitäts-, hier als Silbenschnittopposition beschriebenen Vokalgegensatz nicht teilnimmt. Doch erhält ein Schwa infolge Wieses Kernbedingung (s. oben) im Laufe der Derivation – wie das auch die oben bereits zitierten Beispiele *eine* und *einen* zeigen – mal eine für die Langvokale (den sanften Schnitt) mal eine für die Kurzvokale (den scharfen Schnitt) charakteristische prosodische Struktur, sodass die in (61) formulierte Anforderung gegenüber einer adäquaten Schwa-Theorie im Deutschen in diesem Ansatz keineswegs erfüllt ist. Wieses Modell ist grundsätzlich für die phonologische Erfassung der Schwa-Vorkommen (vgl. (62)) konzipiert. Jedenfalls als positiv zu bewerten ist daran die Bestrebung des Autors nach einer einheitlichen Beschreibung aller Schwa-Vorkommen. Doch ist Wieses Beschreibung – wie oben schon angemerkt – nicht ohne Probleme. Einen ersten Kritikpunkt hat Giegerich (1987: 450f) unter Berufung auf das von Kiparsky (1982: 63) formulierte Prinzip genannt, nach dem Lexikoneinträge exhaustiv silbifizierbar sein müssen, was erklären kann, warum zugrunde liegend keine Stämme etwa der Struktur |ktb| im Deutschen vorhanden sind. Wenn man dieses Prinzip ernst nimmt, so sollten es im Deutschen gar keine Stämme |A:tm|, |ze:gn|, |ze:g|, |vöndR| geben, es sei denn, der stammfinale Sonorant kann im Lexikon auf irgendeine Weise

silbifiziert werden (Giegerichs Lösungsvorschlag s. unten). Féry (1991: 66) und Noske (1993: 156) kritisieren an Wieses Modell die Annahme zugrunde liegend leerer X-Positionen. Féry ist der Ansicht, dass mit dieser Annahme nur schwer vom Eindruck wegzukommen ist, dass ein solcher extra Marker einfach ein nicht-epenthetisches Schwa vertritt. Schließlich wirft Noske (1993: 154) dem Modell Wieses u.a. vor, dass es infolge der extrem spezifischen Bedingungen zu viele idiosynkratische Eigenschaften enthalte, daher wäre eine Beschreibung vorzuziehen, die mit weniger Idiosynkrasien auskommt. Alles in Allem scheint es so, dass die Schwa-Vorkommen im Deutschen sich nicht einheitlich verhalten, was schwerwiegende Konsequenzen nach sich zieht, auf die wir im nächsten Abschnitt eingehen werden.

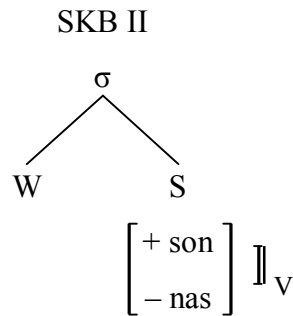
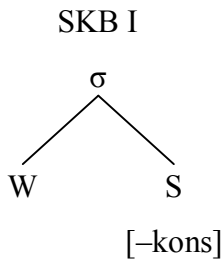
Schließlich kann Wieses Modell die beiden Schwa-Alternationen (vgl. (63)) adäquat erfassen: sie werden als postlexikalische Prozesse dargestellt, die silbischen Sonoranten werden aus der Sequenz Schwa+unsilbischer Sonorant mit einer fakultativen postlexikalischen Schwa-Tilgungsregel, das vokalische *r* dagegen mit einer obligatorischen postlexikalischen Schwa-Tilgungsregel und einer sich daran anschließenden *r*-Vokalisierungsregel hergeleitet.

### **4.3.2 Giegerich (1987)**

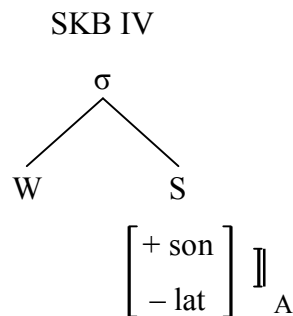
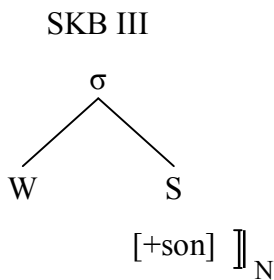
Giegerichs (1987) Beschreibung zeigt viele Ähnlichkeiten mit Wieses Ansatz: ihr liegt die Theorie der lexikalischen Phonologie zugrunde, abgesehen von der Handhabung der Schwa-Epenthese nimmt er eine ähnliche Struktur des deutschen Lexikons an wie Wiese, der Autor betrachtet Schwa durchgängig als Epenthesevokal, Schwa-Vorkommen werden in Interaktion mit der Morphologie und Phonologie hergeleitet. Anders als Wiese leitet er aber Schwa mobilia und Schwa constantia (bei ihm: prosodisches und Flexions-Schwa) auf unterschiedliche Art her. Zur Ableitung des prosodischen Schwa geht er von dem oben zitierten Prinzip Kiparskys aus, und lässt im Lexikon silbische Sonoranten zu, da nur unter dieser Annahme zugrunde liegendes |A:tm|, |ze:gl|, |ze:gl|, |fA:tR| voll silbifiziert werden können. Der unter den Schwa mobilia beobachtbaren Schwa-Zero-Alternation wird er mit der Annahme unterschiedlicher Silbikernbedingungen gerecht, die zum Ausdruck der Interaktion der Schwa-Epenthese mit der Morphologie und Phonologie dienen, und auf den unterschiedlichen Ebenen des Lexikons in Stämmen unterschiedlicher Wortklassenzugehörigkeit unterschiedlichen Sonoranten die Silbischkeit erlauben, vgl. (69) (nächste Seite).

(69) Silbenkernbedingungen (SKB) nach Giegerich (1987: 464)

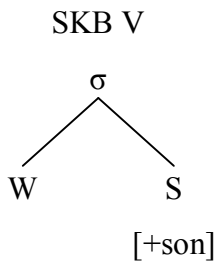
a. Ebene 1



b. Ebene 2



c. Ebene 3



SKB I gibt den Normalfall an. Mit der Lockerung dieser starken Beschränkung durch SKB II können die Formen *segeln*, *segelt*, *feuern*, *feuert* usw. im Gegensatz zu *atmen*, *atmet*, *segnen*, *segnet* hergeleitet werden, diese SKB leistet somit genau das, wozu Wieses erste Schwa-Epenthese-Regel dient. In der lexikalischen Variation *segele/segle* bzw. *feure/feure* betrachtet Giegerich im Gegensatz zu Wiese die schwalose Form als primär, was mit der von ihm angewendeten metrischen Silbentheorie zusammenhängt: Silblichkeit eines Sonoranten ist eigentlich nicht absolut zu verstehen, sondern mit der Besetzung einer metrisch starken Position gleichzusetzen. Wird einem solchen metrisch starken Sonoranten ein weiteres metrisch starkes Segment, etwa das Suffix *-e* hinzugefügt, so erfolgt automatisch eine Uminterpretation der metrischen Verhältnisse und der Sonorant wird zum Anfangsrand gezählt. Die schwahaltigen Formen werden bei Giegerich durch eine frühere Anwendung der Schwa-Epenthese-Regel (s. unten) erzeugt.

SKB III findet ebenfalls seine Entsprechung im Wiese'schen Modell: sie ist genauso wie die zweite Epenthese-Regel Wieses für die korrekten Formen *Atem*, *Segen*, *Segel*, *Feuer* usw. verantwortlich. SKB IV leitet die schwahaltigen Adjektivformen (besser: die mit silbischen Sonoranten) *eigen(e)* und *mager(e)* im Gegensatz zu *dunkel*, aber *dunkle* ab. In dieser Hinsicht ist eine Differenz zwischen den beiden Beschreibungen festzustellen: in Wieses Ansatz erfolgt dieselbe Ableitung durch zwei Regeln (Regel 3 und 4). Dadurch, dass bei Giegerich Wieses 4. Regel, die Schwa-Epenthese vor *r* in allen Wortarten durchführt, keinerlei Entsprechung findet, bleibt der Autor die Silbischkeit des *r* in Wörtern mit dem Ableitungssuffix *-er* schuldig. Die liberalste SKB V schließlich sorgt für die Silbifizierung aller bisher nicht silbifizierten Sonoranten und sichert damit, dass alle Wortformen das Lexikon mit einer vollständigen Silbentruktur verlassen. Dies ist – soweit ich sehe – nur bei den un suffigierten Adjektiven mit stammfinalen Lateral (so z.B. *dunkel*) der Fall, trotzdem stört die sehr allgemeine Formulierung nicht, da sie gerade deswegen nicht zu Übergeneralisierungen führt, da alle anderen Sonoranten bereits silbifiziert worden sind. Somit deckt SKB V teilweise Wieses 6. Schwa-Regel ab. Nun ist zu sehen, dass nicht allen Wiese'schen Regeln der dritten Lexikonebene Silbenkernbedingungen entsprechen. Dies ist jedoch damit zu erklären, dass Giegerich das Schwa in *atmen*, *atme(s)t*, *trockenen* und *dunkler* nicht als prosodisch betrachtet und mit je einer Regel beim Flexions-Schwa behandelt. Vor der Besprechung dieses Themas soll noch auf eine weitere Eigenschaft des prosodischen Schwa eingegangen werden. Da ein lexikalisch-silbischer Sonorant auf der Oberfläche nicht unbedingt silbisch erscheint, muss Giegerich im Lexikon eine Epenthese-Regel annehmen, die vor einem metrisch starken Sonoranten eine leere V-Position einfügt, was zur Folge hat, dass dieser in den Endrand verdrängt wird. Die leere V-Position wird bei Giegerich erst in der postlexikalischen Komponente mit Schwa assoziiert.

Die anderen Schwa-Vorkommen (Flexions-Schwa, Schwa mobile) leitet der Autor durch einzelne Regeln ab, die ein unspezifiziertes skeletales X in bestimmte morphologische bzw. phonologische Kontexte einfügen. So wird das Schwa vor der Genitivendung *-s* (vgl. (57)a) durch Anwendung der Regel in (70) eingefügt. Die fakultative Schwa-Insertion in der Genitivform von *des Schiff(e)s* (vgl. (57)b) erklärt der Autor durch die Ausdehnung der Schwa-Epenthese-Regel in *(dem) Manne* (vgl. (58)b) auch auf die Singular Genitivform, vgl. (71). Dass alle Adjektivendungen obligatorisch auf Schwa enden (vgl. (57)d und (58)c), wird mit der Annahme der Regel in (72) Rechnung getragen, die somit Wieses 6. und 7. Regel leistet. Für das epenthetische Schwa in den Verbformen *atme(s)t*, *rette(s)t*, *segnen*, *reden* (vgl. (57)e-f) ist die Regel in (73) verantwortlich. Die Endung der 1. Person Singular (vgl. (58)d)

wird mit (74) eingefügt. Plural-, sowie Wortbildungsschwa (vgl. (58)a, e) werden von Giegerich zwar nicht thematisiert, sind aber durch eine Regel analog zu (74) durchaus handhabbar. Von den restlichen Schwa-Vorkommen unter (57) und (58) werden Wörter wie *Abend* und *Tugend* (vgl. (57)h) ähnlich wie bei Wiese mit der Extrametrikalität des finalen *d* und dann mit prosodischem Schwa erklärt. Das Diminutivsuffix *-chen* (vgl. (57)g) enthält möglicherweise ebenfalls prosodisches Schwa. Wie oben bemerkt, kann das Modell Wörter mit dem Derivationsuffix *-er* (vgl. (57)g) nicht erfassen. Die übrigen Fälle (vgl. (57)g und (58)f–g) stellen nur dann keine Probleme für Giegerichs Modell dar, wenn der allgemein epenthetische Charakter des Schwa aufgegeben wird.

(70) Schwa-Regel zur Genitivendung *-s* (vgl. Giegerich 1987: 462)

$$\emptyset \rightarrow X / \left[ \begin{array}{l} [-\text{son}] \\ +\text{kor} \\ +\text{kont} \end{array} \right] \text{---} \left[ \begin{array}{l} X \\ | \\ [s] \end{array} \right] \text{]}_N$$

(71) Schwa-Regel zur Dativendung *-e* (vgl. Giegerich 1987: 462)

$$\begin{array}{c} \swarrow \quad \searrow \\ S \quad \quad W \\ \emptyset \rightarrow X / \quad \sigma C \text{]}_N \text{---} \text{]}_N \end{array}$$

(72) Schwa-Regel zu den Adjektivendungen (vgl. Giegerich 1987: 463)

$$\emptyset \rightarrow X / \text{]}_A \text{---} (X) \text{]}_A \\ | \\ [+k\text{ons}]$$

(73) Schwa-Regel zu den Verbalendungen *-t*, *-st*, *-en* (vgl. Giegerich 1987: 460)

$$\emptyset \rightarrow X / \left\{ \begin{array}{l} [-\text{son}] \quad [+nas] \\ \left[ \begin{array}{l} -\text{son} \\ +\text{kor} \\ -\text{kont} \end{array} \right] \end{array} \right\} \text{]}_V \text{---} \left[ \begin{array}{l} X \\ \swarrow \quad \searrow \\ (s) \quad t \end{array} \right] \text{]}_V$$

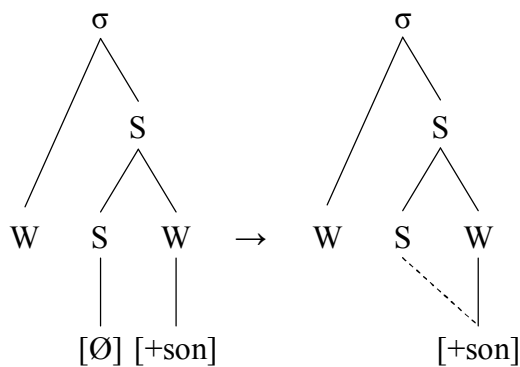
(74) Schwa-Regel zur Verbalendung *-e* (vgl. Giegerich 1987: 460)

$$\emptyset \rightarrow X / \text{]}_V \text{---} \text{]}_V$$

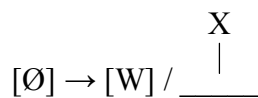


Schließlich sei noch auf die postlexikalischen Schwa-Prozesse in Giegerichs (1987) Beschreibung eingegangen. Wie die Regeln (70) bis (74) zeigen, ist das Flexions-Schwa im Lexikon in Form einer leeren skeletalen Position präsent. Nicht anders ist es mit dem prosodischen Schwa – wie oben besprochen. Daher bedarf es in der postlexikalischen Komponente zweier Prozesse, die diese leere skeletale Position interpretieren. Dazu schlägt der Autor die Regeln (75) und (76) vor, von denen Erstere fakultativ in Abhängigkeit vom Sprechtempo und Sprechstil, Letztere dagegen obligatorisch angewendet wird.

(75) Postlexikalische silbische Sonoranten (vgl. Giegerich 1987: 467)



(76) Schwa-Assoziation (vgl. Giegerich 1987: 467)



Wie eingangs gesagt, zeigt Giegerichs Ansatz viele Ähnlichkeiten mit Wieses Vorschlag zur Beschreibung des Schwa im Deutschen. Somit kann man in Bezug auf die unter (60) und (61) formulierten Anforderungen genau dasselbe festhalten, was bereits bei der Besprechung des Wiese'schen Modells gesagt wurde. Auch der Ansatz von Giegerich wurde primär für die Erfassung der unterschiedlichen Schwa-Vorkommen und der Schwa-Alternationen konzipiert und unterscheidet sich gewissermaßen von der Wiese'schen Beschreibung, wobei angemerkt werden soll, dass diese Unterschiede nicht gerade vom Primat der Theorie Giegerichs zeugen. Eine der größten Schwachstellen des Giegerich'schen Ansatzes betrifft die Annahme silbischer Sonoranten im Lexikon. Auch wenn dadurch dem von Kiparsky (1982) formulierten Grundprinzip der zugrunde liegenden Repräsentation genüge getan wird, führt sie zu einer erheblichen Komplikation und m.E. unplausiblen Derivation: die silbischen Sonoranten müssen in einem ersten Schritt noch im Lexikon durch ein epenthetisches Schwa, oder besser: durch eine epenthetische X-Position ‚verdünnt‘ werden, um dann – auch wenn fakultativ – in der postlexikalischen Komponente eine Reassoziaton zu durchlaufen, als deren

Ergebnis wiederum silbische Sonoranten entstehen. Das Modell scheint somit von einer Reduktionsform ausgehend zuerst die Explizitform und daraus wieder die Reduktionsform herzuleiten. Wünschenswert wäre m.E. nur die eine Richtung, u.z. möglichst die: Explizitform → Reduktionsform. Ebenfalls als etwas problematisch empfinden wir an manchen Stellen die Trennung zwischen prosodischem und Flexions-Schwa, die sich auch in der Derivation widerspiegelt. Manche Flexions-Schwas sind einfach nicht idiosynkratisch, sondern gehorchen phonologischen Bedingungen. So scheinen uns die Regeln (70) und (73) teilweise dieselbe Gegebenheit zum Ausdruck zu bringen: Schwa wird aus phonologischen Gründen, nämlich um Geminat zu vermeiden, vor einem Segment eingeschoben. Mit der (auch technisch) einheitlichen Behandlung der Schwa-Epenthese kann dieser Zusammenhang adäquat zum Ausdruck gebracht werden.

### **4.3.3 Noske (1993)**

Schließlich sei noch kurz Noskes (1993) Ansatz zum Schwa im Deutschen skizziert, um zu zeigen, wie eine alternative Beschreibung des Schwa im Deutschen, die keine morphologischen Informationen berücksichtigt, auszusehen, bzw. mit welchen schwerwiegenden Konsequenzen sie zu rechnen hat.

Noske (1993), der – wie oben gezeigt – an Wieses Modell dessen Komplexität kritisiert, macht den Versuch, ein Konzept zu entwickeln, das auch ohne morphologische Informationen auskommt, und Schwa als epenthetisches Segment allein unter Berufung auf die Silbenstrukturbedingungen sowie auf einige weitere Annahmen herleiten kann. Der Grundgedanke der Silbifizierung bei Noske beruht auf der Direktionalität, nach der dieser Prozess von rechts nach links durch eins-zu-eins-Assoziation zwischen skeletalen Einheiten und Silbenpositionen erfolgt, wobei die Silbe im Deutschen die drei subsilbischen Konstituenten Onset, Nukleus und Koda enthält. Seine Analyse geht ferner von der Annahme aus, dass Silbifizierung an zwei Stellen in der Phonologie stattfindet: im Lexikon, wo sie im Gegensatz zu Wiese (1988) und Giegerich (1987) postzyklisch erfolgt, sowie in der postlexikalischen Phonologie. Außerdem nimmt er an, dass nicht alle Schwas im Deutschen als Ergebnis von Epentheseeregeln anzusehen sind, ein Teil der Schwa-Vorkommen ist schon in Form von leeren skeletalen Positionen vorhanden. Schließlich können nach dem Autor Liquide im Deutschen – nicht aber Nasale – schon im Lexikon silbisch sein.

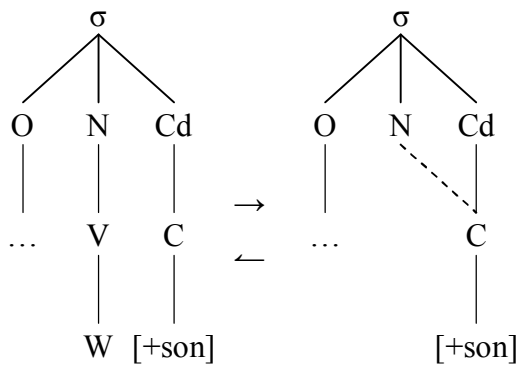
Bei der Silbifizierung der Stämme *Atem*, *Segen*, *Segel* und *Vater* (vgl. Tab. 13a), die zugrunde liegend kein Schwa enthalten, wird der stammfinale Sonorant zur Koda gerechnet, das vorletzte Segment dagegen zum Onset, da sie nicht silbisch sein können. Der leere Nukleus wird mit einem leeren V auf der skeletalen Ebene assoziiert, das auf der Oberfläche als Schwa realisiert wird. Im Falle von *atmen*, *atme(s)t*, *segeln*, *segel(s)t*, *feuern*, *feuer(s)t* (vgl. Tab. 13bi, biii) bildet das finale Segment die Koda, der Unterschied zwischen stammfinalen Nasalen und Liquiden ergibt sich aus der oben genannten Annahme, dass stammauslautende Liquide im Gegensatz zu Nasalen silbisch sein können. Somit werden Liquide mit dem Nukleus, Nasale dagegen mit dem Onset verbunden, was bei letzteren Sonoranten einen leeren Nukleus ergibt, der V-Epenthese und später Schwa-Assoziation auslöst. Die lexikalische Variation in der 1. Person Singular (vgl. *segele/segle* bzw. *feuere/feure* in Tab. 13bii) kann Noskes Modell nicht ausdrücken: die postzyklische Silbifizierung erzeugt in beiden Fällen die schwalose Form. Die unsuffigierten Adjektivstämme *eigen*, *dunkel* und *mager* (vgl. Tab. 13c) könnten analog zu den phonologisch ähnlichen Nominalstämmen (vgl. Tab. 13a) behandelt werden, was sich im Falle des stammfinalen Laterals als eine adäquate Beschreibung erweist, da dadurch die Schwa-Zero-Alternation zwischen *dunkel* vs. *dunkle(n)* auf eine prägnante Weise zum Ausdruck gebracht wird. Doch erscheinen die Adjektivstämme auf einen Nasal bzw. *r* schon problematisch, da in diesen das Schwa nicht nur in der endungslosen, sondern auch in der mit einer Endung versehenen Form vorhanden ist. Als einziger Ausweg bietet sich hier, soweit ich sehe, in diesen Adjektivstämmen schon zugrunde liegend ein Schwa, besser eine leere skeletale V-Position anzunehmen, was für die nasalfinalen Adjektivstämme vom Autor auch vorgeschlagen wird, wenn ich das richtig verstehe (vgl. 166).

Ein Teil der Schwa constantia erklärt sich mit der oben erwähnten Annahme Noskes über die zugrunde liegenden leeren V-Positionen, die auf der Oberfläche als Schwa realisiert werden. Dies betrifft alle Schwas unter (58), aber nicht nur das, wie das oben schon gezeigt wurde und gleich noch gezeigt wird. Die Fälle in (57) können in Noskes Ansatz wie folgt beschrieben werden. Schwa-Epenthese vor der Genitivendung *-s* (vgl. (57)a–b) ist – wie das aus den bisherigen Ausführungen deutlich hervorgeht – prosodisch bedingt: im Falle von *des Hauses* wird sie von der Antigeminationsbeschränkung ausgelöst, bei *des Schiff(e)s* dagegen von der fakultativen Bedingung des trochäischen Fußes flektierter Substantive. Ich denke, Noskes Konzept ist durchaus in der Lage, auch diese Fälle zu erfassen: das Genitiv-*s* wird in *des Hauses* mit der Koda assoziiert, das stammfinale *s* dagegen mit dem Onset, da Geminaten im Deutschen zu vermeiden sind, entsteht hier ein leerer Nukleus, der V-Epenthese auslöst

und schließlich als Oberflächen-Schwa erscheint. Andererseits werden bestimmte prosodisch-morphologische Bedingungen auch in Noskes Ansatz berücksichtigt, und genau das erklärt die fakultative Schwa-Epenthese in *des Schiff(e)s*. In den Adjektivendungen (vgl. (57)d) nimmt der Autor schon zugrunde liegend ein Schwa an, da es auch nach Adjektivstämmen erscheint, in die es durchaus integriert werden könnte (vgl. *schnelles, genauen, hohen*). Die Fälle in (57)e (*atme(s)t, segne(s)t* usw.) wurden bereits oben bei der Behandlung der Schwa mobilia besprochen. Die Infinitivformen (vgl. (57)f) wurden teilweise schon ebenfalls thematisiert, doch muss Noske – ähnlich Wiese und Giegerich – zur Ableitung der korrekten Form mit Schwa *bauen* und *sehen* die oben bereits besprochene Bedingung einführen, nach der Infinitivformen auf einen trochäischen Fuß enden müssen (vgl. 163). Das Schwa im Diminutivsuffix *-chen* (vgl. (57)g) kann aus prosodischen Gründen erscheinen, d.h. infolge eines leeren Nukleus zwischen dem finalen mit der Koda assoziierten *n* und dem mit dem Onset verbundenen Ichlaut, *-er* als Derivationsuffix (vgl. (57)g), das von Noske nicht thematisiert wird, könnte man analog zu der von ihm diskutierten Komparativendung *-er* analysieren, die nach dem Autor schon zugrunde liegend ein Schwa enthält. Die Motivation für diese Analyse ergibt sich aus Wörtern mit stammfinalen Vokal, bei denen das Suffix *-r* durchaus in die letzte Stammsilbe eingebaut werden könnte (vgl. *Seher* vs. *sehr*). Dem stammfinalen Schwa in den Wörtern *Abend* und *Tugend* (vgl. (57)h) könnte man ähnlich wie in den anderen beiden besprochenen Ansätzen durch Extrasilbizität Rechnung tragen, in *Abenteuer* lassen sich beide Schwach einfach herleiten. In *Dromedar* muss dagegen schon zugrunde liegend eine leere V-Position angenommen werden.

Schließlich sei noch auf die von Noske (1993) angenommenen postlexikalischen Prozesse einzugehen. Da die Stämme mit finalelem Liquid und finalelem Nasal in unterschiedlicher Form das Lexikon verlassen, Erstere nämlich mit silbischen (mit dem Nukleusknoten assoziierten) Sonoranten, Ldagegen mit der Sequenz Schwa+Sonorant, beschreibt der Autor die postlexikalische Variation zwischen silbischen Sonoranten und Schwa+Sonorant-Sequenzen mit einer bidirektionalen Regel, die einerseits einen silbischen Sonoranten in die Sequenz Schwa+Sonorant, andererseits die Sequenz Schwa+Sonorant in einen silbischen Sonoranten verwandelt, vgl. (77).

(77) Postlexikalische Variationsregel nach Noske (1993: 170)



Der große Vorteil des Ansatzes von Noske (1993), der sich noch stärker als die anderen beiden Ansätze auf die unterschiedlichen Schwa-Vorkommen konzentriert, besteht in seiner konzeptuellen Einfachheit: er ist im Vergleich zu Wieses (1988) oder Giegerichs (1987) Analyse in der Tat weniger komplex, mit Hilfe lediglich einiger Grundannahmen gelingt nämlich dem Autor eine relativ ausführliche Beschreibung durchzuführen. Doch ist diese Beschreibung nicht unproblematisch. Die wichtigsten Probleme lassen sich m.E. auf dieselbe Quelle, nämlich auf die Grundannahme Noskes zurückführen, dass im deutschen Lexikon auch Liquide silbisch sein können.<sup>58</sup> Mit dieser Annahme lassen sich die im verbalen Bereich beobachtbaren Variationen – wie oben gezeigt – auf eine prägnante und elegante Weise beschreiben. Doch erweist sich diese Annahme im Bereich der Adjektive eher als Nachteil, denn dort verhalten sich nicht die liquidfinalen Stämme exzeptionell, sondern die nichtlateralen. Der einzige Ausweg aus diesem Problem, der sich in dieser Rahmentheorie anbietet, und nach dem Adjektivstämmen mit finalelem nichtlateralem Sonoranten zugrunde liegend ein Schwa haben, ist jedoch sehr problematisch. Mit ihm geht nicht nur die wichtige Generalisierung verloren, dass sich *Segen* und *trocken*, *Wunder* und *mager* phonologisch ähnlich verhalten, er kann zwar *segnen* und *wundern*, nicht aber *trocknen* und *(ab)magern* herleiten. Das andere große mit der lexikalischen Silbischkeit der Liquide zusammenhängende Problem betrifft die zuletzt beschriebene bidirektionale postlexikalische Regel. Diese kann m.E. die wichtigste Eigenschaft des fraglichen postlexikalischen Prozesses nicht erfassen, dass es sich dabei um einen vom Sprechtempo und Sprechstil abhängigen

<sup>58</sup> Die Motivation für diese Annahme, Liquide stehen höher auf der Sonoritätsskala, sie neigen folglich besser zur Silbischkeit (vgl. 157), ist m.E. nicht unplausibel. Doch sollte man dabei auch nicht außer Acht lassen, dass der sonorste Konsonant [R], der in den untersuchten Stämmen fast ausnahmslos silbisch erscheint, im Silbenkern obligatorisch vokalisiert wird, als könnte er trotz seines sehr hohen Sonoritätswerts als Konsonant nicht den Silbenkern bilden.

Reduktionsvorgang handelt, bei dem die schwahaltige Explizitform reduziert wird. Reduktion und Verstärkung erfolgen in Noskes Regel in beiden Richtungen.

Zusammenfassend lässt sich zu den oben skizzierten Ansätzen folgendes festhalten. Noskes Modell hat gezeigt, dass den Schwa-Vorkommen am ehesten beizukommen ist, wenn man die Interaktion zwischen Phonologie und Morphologie in die Beschreibung einbezieht. Dies spricht jedenfalls für eine lexikalisch-phonologische Behandlung des Problems. Um komplizierte Derivationen zu vermeiden, bei denen ein silbischer Sonorant zuerst in die Sequenz Schwa+unsilbischer Sonorant verwandelt wird, die dann in einem nächsten Schritt wieder zum silbischen Sonoranten geschwächt wird, scheint der einzig mögliche Weg das Verbot silbischer Sonoranten und die Annahme der Schwa-Epenthese im Lexikon zu sein, auch wenn dadurch zugrunde liegende Repräsentationen angenommen werden müssen, die nicht exhaustiv silbifizierbar sind. All das spricht für Wieses (1988) Beschreibung. Doch ist sie – wie gezeigt – auch nicht ohne Probleme. Ernst zu nehmen ist m.E. Férys (1991) und Noskes (1993) Kritik bezüglich der abstrakten zugrunde liegenden X-Positionen, für deren Annahme abgesehen davon, dass sie die Länge des unter exzeptioneller Betonung anstelle des Schwa erscheinenden *e*-Vokals gewährleisten, in der Mehrheit der Wörter, in denen sie angenommen werden müssen, nichts spricht. Somit kann die einzige Funktion dieser X-Position darin bestehen, die Allgemeingültigkeit der von Wiese (1988) postulierten Schwa-Epenthese im Deutschen zu sichern. Die Annahme zugrunde liegender Schwas im Deutschen hat jedoch zahlreiche schwerwiegende Konsequenzen, die v.a. die Schwa-Akzent-Beziehung betreffen. Zu untersuchen ist ferner besonders im Hinblick auf Noskes Analyse, ob bei allen von Wiese angenommenen im Lexikon zugewiesenen Schwas wirklich guter Grund für den epenthetischen Status besteht. Außerdem muss eine adäquate Schwa-Beschreibung im Deutschen selbstredend auch die Aspekte erfassen können, die oben bei der Behandlung des Wiese'schen Ansatzes bemängelt wurden, so v.a. das Verhältnis zwischen Schwa und der hier als Silbenschnittgegensatz analysierten Vokalopposition in (6). Der Rest dieses Kapitels ist der Erläuterung eines auf diesen Kritikpunkten basierenden Schwa-Ansatzes gewidmet.

#### **4.4 Zugrunde liegendes Schwa und Reduktionssilbe**

Wie im vorausgehenden Abschnitt besprochen erscheint Wieses Annahme von einer leeren X-Position auf der skeletalen Ebene dort, wo in der Oberflächenrepräsentation ein Schwa ohne Folgekonsonanz steht, im Lichte der Kritik Férys (1991) und Noskes (1993) als

unplausibel. Statt dessen sollte in den fraglichen Wörtern oder Affixen in irgendeiner Form ein zugrunde liegendes Schwa angenommen werden. Dies betrifft fast alle Schwa-Vorkommen unter (58)<sup>59</sup>: das finale Schwa in den Wörtern in (58)g, aber auch in gewissen Affixen, wie in der Pluralendung (vgl. (58)a), in der Adjektivendung (vgl. (58)c), in der Konjugationsendung in (vgl. (58)d), im Ableitungssuffix (vgl. (58)e), in den beiden Präfixen (vgl. (58)f), sowie das stamminterne Schwa in einigen wenigen Wörtern wie *Abenteuer* (vgl. (57)h).<sup>60</sup>

Wir haben oben auch angemerkt, dass eine solche Modifizierung des Wiese'schen Konzepts schwerwiegende Konsequenzen nach sich zieht. Das zweifelsohne wichtigste Problem betrifft die Beziehung zwischen Schwa und Akzent. Wie in Abschn. 4.1 ausgeführt, sind Silben mit Schwa im Lexikon unbetonbar. Dies lässt sich in einem Modell, in dem Schwa ausschließlich als Ergebnis von Epentheseeregeln entsteht, durch Regelordnung zum Ausdruck bringen: die Schwa-Epentheseeregeln kommen erst nach den Wortakzentregeln zur Anwendung, sodass Silben mit Schwa für die Akzentregeln ‚unsichtbar‘ sind. Wenn dagegen auch schon in der zugrunde liegenden Repräsentation Schwas in irgendeiner Form vorhanden sind, muss diese ‚Unsichtbarkeit‘ der schwahaltigen Silben für die Akzentregeln auf eine andere Weise erreicht werden. Ein möglicher Weg, den wir hier auch einschlagen wollen, beruht auf der von Vennemann (1991a) vorgenommenen Unterscheidung zwischen zwei Silbentypen im Deutschen: den Voll- und den Reduktionssilben. Während Vollsilben betont werden und nur Vollvokale enthalten können, sind Reduktionsvokale unbetonbar und können nur Schwa als Vokal haben. Der vorliegende Abschnitt ist einer näheren Beschreibung der Reduktionssilbe im Deutschen gewidmet.

Zunächst einmal ist zu fragen, ob sich die Reduktionssilbe im Deutschen nur durch die oben genannten beiden Eigenschaften auszeichnet, oder ob sie vielmehr über weitere Charakteristika verfügt, aufgrund derer sie den Vollsilben weiter gegenübergestellt werden kann. Ganz konkret interessiert uns dabei die Frage, ob man für die Reduktionssilbe dieselbe Struktur annehmen soll wie für die Vollsilbe (vgl. (35)). Da diese Struktur – wie in Abschn. 3.4 ausführlich besprochen – stark vom Silbenschnittkontrast geprägt ist, kommt bei der Beantwortung dieser Frage der oben in Abschn. 4.1 gemachten Beobachtung eine große Rolle

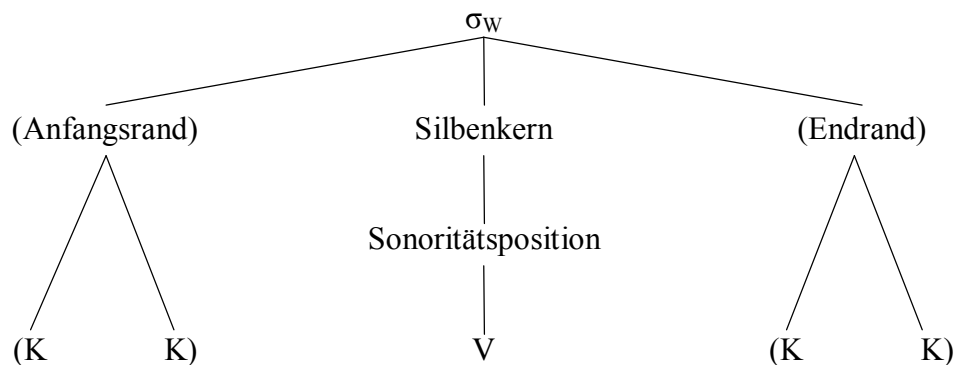
---

<sup>59</sup> Wie oben gezeigt, lässt sich das Schwa-Vorkommen in *(dem) Volke*, *(dem) Balle* usw. (vgl. (55)b) systematisch herleiten, s. auch Abschn. 4.5.

<sup>60</sup> Auf weitere zugrunde liegende Schwa-Vorkommen kommen wir noch in 4.5 zu sprechen.

zu, nach der das Schwa als silbenschnittneutraler Vokal aufzufassen ist. Daraus, dass das silbenschnittneutrale Schwa den einzig möglichen Vokal in Reduktionssilben darstellt, folgt, dass in der Reduktionssilbe kein Silbenschnittkontrast besteht. Da für den Silbenschnittkontrast die Halteposition innerhalb des Silbenkerns verantwortlich ist, kann man diesem Phänomen am besten gerecht werden, indem man in Reduktionssilben keine Halteposition annimmt. Somit hat dieser Silbentyp die Struktur in (78).<sup>61</sup> Diese Repräsentation der Reduktionssilbe entspricht m.E. durchaus der phonetischen Intuition. Mit Becker (1998) kann man davon ausgehen, dass scharf geschnittene Vokale infolge der Besetzung der – in seiner Terminologie – Implosionsposition durch die Folgekonsonanz als reduzierte (sprich: ungespannt realisierte) Vokale anzusehen sind. Nun unterscheidet sich das Schwa von diesen Vokalen dadurch, dass es keine Folgekonsonanz braucht, um reduziert (ungespannt) zu sein: es ist sozusagen ‚vom Hause aus‘ ein Reduktionsvokal. Diesem Sachverhalt wird durch die fehlende Halteposition in der Reduktionssilbenstruktur auf eine einfache und plausible Weise Rechnung getragen.

(78) Struktur der Reduktionssilbe im Deutschen<sup>62</sup>



Ein einfacher Vergleich der Struktur der Vollsilben in (35) mit der der Reduktionssilben in (78) zeigt, dass in ersterer im Endrand maximal ein Segment erscheinen kann, während die Anzahl der Endrandsegmente in den Reduktionssilben zwei ist. Dies ist unbedingt nötig, um Fälle wie *segeln* oder *feuern* silbifizieren zu können, in denen nach dem als Silbenkern dienenden Schwa zwei Konsonanten vorkommen, von denen der letzte nicht als extrasilbisch angesehen werden kann. Diese ‚Erweiterung‘ des Endrandes, die als eine weitere Eigenschaft

<sup>61</sup> Angemerkt werden soll, dass für Reduktionssilben dadurch die Struktur vorgeschlagen wird, die Becker (1998) für unbetonte Silben vorsieht, vgl. (29).

<sup>62</sup>  $\sigma_w$  steht für eine Reduktionssilbe, während für Vollsilben weiterhin das Symbol  $\sigma$  verwendet wird.

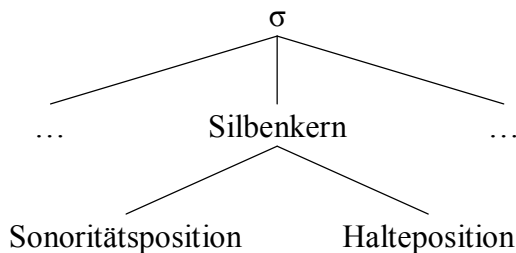


der Reduktionssilben im Deutschen betrachtet werden könnte, ist m.E. jedoch nicht störend, da sie sich – wie das unten in Abschn. 5.2 noch gezeigt wird – durchaus mit den Silbifizierungsregeln für den Endrand verträgt.

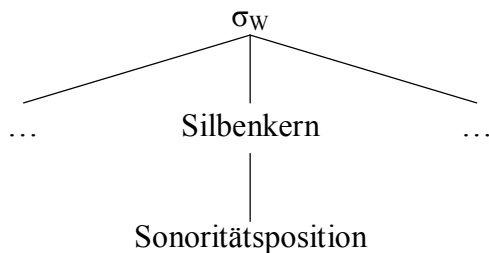
Die oben vorgeschlagene Struktur der Reduktionssilbe macht eine Modifizierung der in Abschn. 3.4 eingeführten universalen Silbenkernbedingung notwendig, die in der Formulierung unter (39) vorschreibt, dass der Silbenkern universell eine Sonoritäts- und eine Halteposition enthalten soll. Aus der hier in Anlehnung an Vennemann (1991a) vertretenen Silbentypologie folgt, dass eine solche Beschränkung nur für die Voll-, nicht aber für die Reduktionssilben gilt, was folgende Änderung der universalen Silbenkernbedingung erzwingt:

(79) **Silbenkernbedingung**

- a. Der Kern einer Vollsilbe enthält universell eine Sonoritäts- und eine Halteposition:



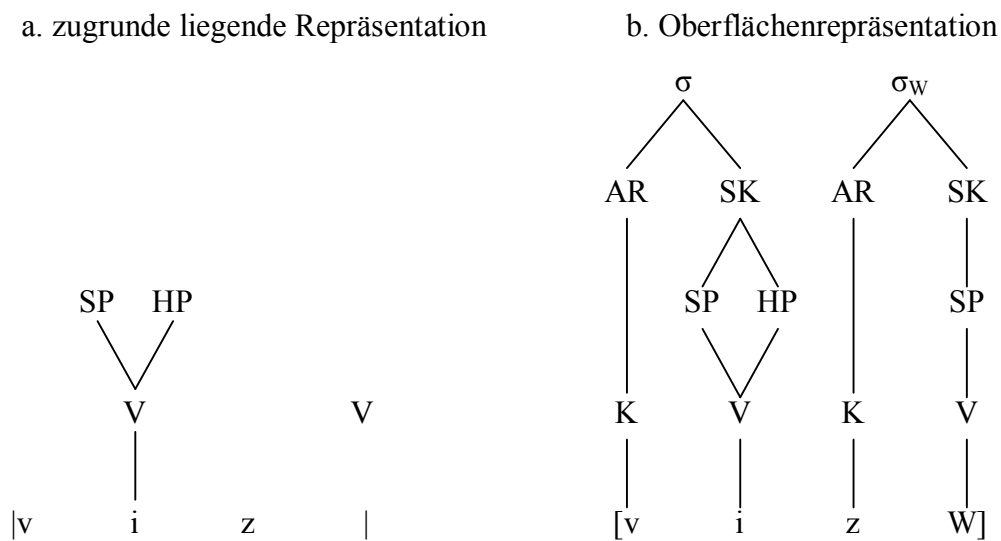
- b. Der Kern einer Reduktionssilbe enthält universell nur eine Sonoritätsposition:



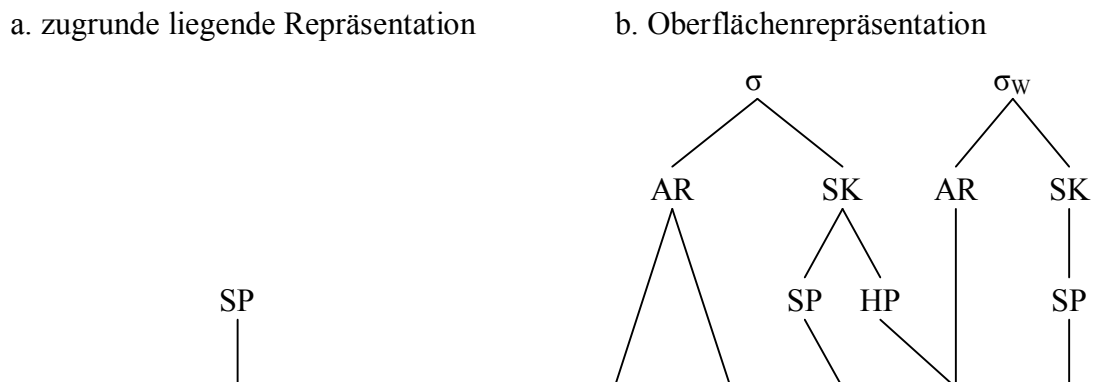
Schließlich soll noch die Frage nach der zugrunde liegenden Repräsentation des Schwa geklärt werden. Wir wollen uns Noske (1993) – und in gewisser Hinsicht auch Wiese (1988) – anschließen, und zugrunde liegendes wie lexikalisches bzw. postlexikalisches Schwa als eine leere skeletale Position repräsentieren, die dann später im Laufe der Derivation mit Schwa assoziiert wird. Die Wahl einer V-Position anstatt eines Segments [W] lässt sich wie folgt begründen. Schwa ist im Deutschen nicht bloß einer der Vokale, es ist der *default*-Vokal. Wenn im Lexikon ein Schwa eingefügt wird, wird in Wirklichkeit nicht willkürlich irgendein Vokal eingefügt, sondern vielmehr eine Stelle, die für die Silbifizierung unmittelbar zugänglich ist, mit der eine nicht silbifizierbare Segmentkette silbifiziert werden bzw. eine

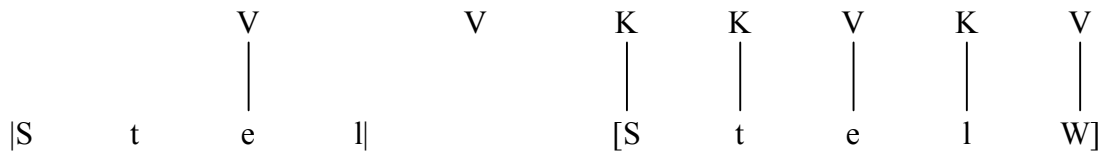
prosodisch-morphologische Struktur (etwa trochäischer Fuß) errichtet werden kann, also eine V-Position auf der skeletalen Ebene. Dass diese Position genau mit [W] assoziiert wird, hängt mit dem *default*-Charakter dieses Vokals im Deutschen zusammen. Ähnlich wird in der postlexikalischen Phonologie eine V-Position immer als [W] reinterpretiert (vgl. Abschn. 4.6). Analog dazu scheint der Annahme einer zugrunde liegenden leeren V-Position für das Schwa nichts im Wege zu stehen. Ein mögliches Argument gegen diese Repräsentation könnte sein, dass sie genauso abstrakt ist wie die in Wiese (1988) vorgeschlagene Darstellung, die sie ablösen sollte. Mit Noske (1993: 156) können wir dagegen einwenden, dass es einen großen Unterschied zwischen den beiden Repräsentationsvorschlägen gibt: während Wieses X-Positionen nie auf der Oberfläche erscheinen und dadurch wirklich abstrakt sind, werden unsere wie Noskes leere skeletale Positionen in der Derivation mit dem Neutralvokal [W] assoziiert. Somit erhalten die Wörter *Wiese* und *Stelle* folgende zugrunde liegende und Oberflächenrepräsentation:

(80) Zugrunde liegende (a) und Oberflächenrepräsentation (b) des Wortes *Wiese*



(81) Zugrunde liegende (a) und Oberflächenrepräsentation (b) des Wortes *Stelle*





## 4.5 Das lexikalische Schwa und die V-Epenthese

Im vorliegenden Abschnitt wollen wir die Schwa-Vorkommen beschreiben, die vorher nicht als zugrunde liegend erklärt wurden. Einen Teil dieser Schwa-Vorkommen werden wir ähnlich wie Wiese (1988) oder Giegerich (1987) als Ergebnis einer prosodischen Epenthese betrachten, bei der analog zur zugrunde liegenden Repräsentation des Schwa eine leere skeletale Position in eine prosodisch bestimmbare Umgebung eingefügt wird. Den Rest der Schwa-Vorkommen wollen wir dagegen als zugrunde liegender Teil der betreffenden Morpheme analysieren. Diese Abweichung von dem Wiese'schen bzw. Giegerich'schen Modell beruht auf der Grundidee, die Einfügung einer skeletalen Position nur da zuzulassen, wo sie zu prosodischen Funktionen dient, auch wenn dabei die einzelnen Prozesse morphologisch konditioniert sind. Dies ist der Fall z.B. in den unsuffigierten Stämmen *Atem*, *Segen*, sowie in den Infinitivformen *atmen*, *segnen*, nicht aber in der Adjektivendung in *trockenen*, *dunklen*. Damit zeigt unser Beschreibungsvorschlag eine gewisse Annäherung an Noskes (1993) Konzept: In der Tat wollen wir mit dem vorliegenden Modell Noskes oben zitierter Kritik an Wiese gerecht werden, und durch die Beseitigung eines Teils der „extrem spezifischen Bedingungen“ die Anzahl der „zu vielen idiosynkratischen Eigenschaften“ vermindern. Als Rahmentheorie nehmen wir die Lexikalische Phonologie, und schlagen mit Wiese (1988), Giegerich (1987) und Wiese (1996) als Grundstruktur des deutschen Lexikons die in Tab. 15 vor.

	<i>Morphologie</i>		<i>Phonologie</i>
<i>Ebene 1</i>	Derivation 1	→	Wortakzent
	irreguläre Flexion	←	
<i>Ebene 2</i>	Derivation 2	→	Kompositionsakzent
	Komposition	←	
<i>Ebene 3</i>	reguläre Flexion	→	
		←	

Tab. 15 Grundstruktur des deutschen Lexikons

Wie gesagt, gehen wir davon aus, dass die Epenthese einer leeren skeletalen Position nur dann stattfinden kann, wenn damit prosodische Konditionen erfüllt werden. Aufgrund des bisher Gesagten sehen wir zwei Grundtypen der V-Epenthese im Deutschen vor. Der erste ist identisch mit der von Wiese (1988) vorgeschlagenen Regel (vgl. (67)a). Diese Art der V-Epenthese werden wir als silbenstrukturelle V-Epenthese bezeichnen, da hier die leere skeletale V-Position zu dem Zweck eingefügt wird, dass ein nicht silbifizierbares Segment silbifiziert werden kann, vgl. (82). Diese Regel wird z.B. in den ansonsten nicht silbifizierbaren unsuffigierten Wortstämmen *Atem*, *Segen*, *Segel* und *Feuer* angewendet.

(82) Grundform der silbenstrukturellen V-Epenthese<sup>63</sup>

$$\emptyset \rightarrow V / \text{_____ } X ]_{\omega}$$

Neben (82) existiert im Deutschen jedoch auch ein anderer Typ der V-Epenthese, der in den zitierten Arbeiten zum Schwa im Deutschen fast ausnahmslos berücksichtigt, trotzdem nicht als eigene V-Epenthese anerkannt wird. Bei dieser Regel wird die leere skeletale V-Position nicht zum Zwecke der vollständigen Silbifizierung eines phonologischen Wortes eingefügt. Die Einfügung erfolgt hier vielmehr, um dadurch eine für eine bestimmte morphologische Form vorgeschriebene prosodische Struktur zu erstellen. Diese prosodische Struktur ist ausnahmslos der zweisilbige linksköpfige Fuß<sup>64</sup>, auf den bestimmte morphologische Formen, so der Infinitiv, enden müssen, vgl. (83). Diese Art der V-Epenthese wollen wir prosodisch-morphologische Epenthese nennen.

(83) Grundform der prosodisch-morphologischen V-Epenthese<sup>65</sup>

$$\begin{array}{c} \phi \\ \swarrow \quad \searrow \\ \sigma \quad \sigma_w \\ \quad \quad | \\ \emptyset \rightarrow V / \quad ] \text{_____ } (X)]_{\omega} \end{array}$$

Wichtig ist zu betonen, dass es sich bei (82) und (83) keineswegs um konkrete Regeln des Deutschen handelt, sondern lediglich um Regelmuster, die selber nicht unbedingt zur

<sup>63</sup> Das tiefgestellte Zeichen ‚ $\omega$ ‘ steht für das phonologische Wort.

<sup>64</sup> Der Trochäus scheint eine besondere Rolle in der deutschen Morphologie zu haben (vgl. Féry 1997).

<sup>65</sup> Das Zeichen ‚ $\phi$ ‘ steht für die Kategorie des Fußes.

Anwendung kommen, genauso wie in Wiesel (1988) Modell (67)a in dieser Form nie angewendet wird. Diese Muster sollen nur die elementaren Prozesse zusammenfassen, die im Deutschen bei der V-Epenthese wirksam sind. Die konkreten Ausprägungen dieser Grundformen, die sich voneinander durch verschiedene morphologische Charakteristika des phonologischen Wortes, die segmentale Assoziation der finalen X-Position oder eben im Falle von (83) durch das Fehlen dieser finalen X-Position unterscheiden, finden auf den einzelnen lexikalischen Ebenen Anwendung.

Zwischen den beiden elementaren Regeltypen sind zwei Unterschiede festzustellen. Einerseits betrifft die silbenstrukturelle V-Epenthese normalerweise Schwa mobilia, prosodisch-morphologische dagegen nur Schwa constatia. Andererseits finden die silbenstrukturellen Epentheseeregeln auf allen Ebenen des Lexikons Anwendung, prosodisch-morphologische jedoch nur auf Ebene 3. Im Folgenden sollen diese Regeln einzeln vorgestellt werden. Anschließend kommen wir auf die restlichen Schwa-Vorkommen zu sprechen.

Aufgrund des oben formulierten Prinzips der V-Epenthese sowie der oben in Abschn. 4.1 dargelegten Schwa-Vorkommen lassen sich im Deutschen folgende Fälle der silbenstrukturellen V-Epenthese bestimmen:

(84) Fälle der silbenstrukturellen V-Epenthese

- a. *segeln, segel(s)t, feuern, feuer(s)t* (vgl. Tab. 13b)
- b. *Atem, Segen, Segel, Feuer* (vgl. Tab. 13a)<sup>66</sup>
- c. *eigene, magere* (vgl. Tab. 13c)
- d. *dunkel, nobel, diskutabel, sensibel* (vgl. Tab. 13c)
- e. *des Hauses, des Kusses, des Schmerzes* (vgl. (57)a)
- f. *atme(s)t, segne(s)t, rette(s)t, rede(s)t* (vgl. (57)e)

Die ersten beiden Fälle (84)a und (84)b machen im Einklang mit Wiesel (1988) die Einführung zweier Regeln nötig: eine Ebene-1-Epenthese-Regel, die noch vor der Ebene-2-Derivation und der regulären Flexion auf Ebene 3 die V-Epenthese in Verbalstämmen auf Liquid durchführt; sowie eine Ebene-2-Epenthese-Regel, die nach der Ebene-2-Derivation aber vor der regelmäßigen Flexion auf Ebene 3 in allen Nominalstämmen eine leere V-Position

---

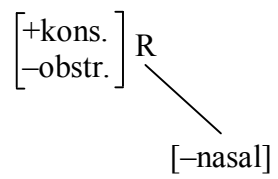
<sup>66</sup> Analog zu diesen Stämmen verhält sich das Diminutivsuffix *-chen* (vgl. (54)g): da *n* sonorer als der Ichlaut ist, können sie keinen komplexen Endrand bilden. Daher soll hier silbenstrukturelle Schwa-Epenthese erfolgen. *Abend* und *Tugend* (vgl. (58)h) lassen sich hier ebenfalls aufführen, wenn man mit Wiesel (1988) und Giegerich (1987) den finalen Obstruenten als extrasilbisch auffasst.

einfügt. Für erstere Regel schlagen wir die Form unter (85)a, für letztere die unter (85)b vor.<sup>67</sup> Anders als Wiese (1988) – und gewissermaßen im Einklang mit Giegerich (1987) – wollen wir die V-Epenthese in den Wörtern von (84)c nicht mit zwei, sondern mit einer einzigen Regel beschreiben (vgl. (85)c), durch deren Anwendung auf Ebene 2 vor der Adjektivdeklinaton auf Ebene 3 in den Adjektivstämmen auf einen nicht-lateralen Sonoranten eine leere V-Position eingefügt wird. Der Grund für nur eine Regel betrifft die unterschiedliche Behandlung des Schwa vor *r*, auf die weiter unten noch näher eingegangen wird. Die restlichen Fälle silbenstruktureller V-Epenthese in (84)d–f lassen sich mit einer einzigen Regel erfassen, die infolge der Heterogenität dieser Fälle keine besonderen Bedingungen benötigt und daher mit der Grundform in (82) identifiziert werden kann. Diese Regel fügt auf Ebene 3 vor jedem unsilbifizierten X eine leere V-Position ein. Sie zeigt wiederum einen Unterschied zu Wieses Ansatz und zugleich eine gewisse Annäherung an Giegerichs Konzept, ist jedoch m.E. beiden Vorschlägen überlegen, da ihre liberale Formulierung zum Ausdruck bringen kann, dass es sich bei allen Einzelfällen um das selbe Phänomen handelt: um die Ermöglichung der Silbifizierung. Zwar ist genau das auch Wieses Grundidee, er muss auf Ebene 3 trotzdem unterschiedliche Regeln formulieren, da er auch in den Adjektivendungen und in allen Wörtern vor *r* epenthetisches Schwa annimmt. Dazu kommt noch, dass die inadäquate Formulierung seiner fünften Regel (vgl. Tab. 14) – wie darauf oben hingewiesen wurde – die Fälle in (84)e und von (84)f Verbformen wie *rette(s)t* und *rede(s)t* mit ebenfalls silbenstrukturell bedingtem epenthetischem Schwa nicht erfassen kann. Auf der anderen Seite nimmt Giegerich (1987) auf Ebene 3 eine Silbenkernbedingung (vgl. SKB V (69)c) an, die in einem anderen theoretischen Rahmen dasselbe leistet wie unsere Ebene-3-Regel (82), doch infolge seiner Unterscheidung zwischen prosodischem und Flexions-Schwa muss er für die V-Epenthese in (84)e und (84)f extra Regeln einführen (vgl. (70) und (73)), sodass dadurch das Gemeinsame an all diesen Fällen (V-Epenthese zur Ermöglichung der Silbifizierung) verdeckt wird.

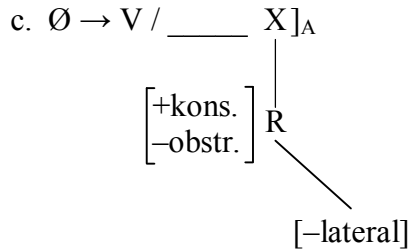
(85) Silbenstrukturelle V-Epentheseeregeln

$$\text{a. } \emptyset \rightarrow V / \text{ \_\_\_\_\_\_ X ]}_V$$

<sup>67</sup> Anzumerken ist bei der ersten Regel, dass durch die Übernahme des Wiese'schen Regel in der 1. Person Singular die schwahaltigen Formen *segele* und *feure* als primär betrachtet werden. Zur Herleitung der schwalosen Formen wollen wir mangels einer besseren Alternative ebenfalls Wieses Vorschlag von der lexikalischen Variation annehmen.



b.  $\emptyset \rightarrow \text{V} / \text{ \_\_\_\_\_\_ } \text{X}]_{\text{N}}$



Das oben formulierte Grundprinzip der V-Epenthese sowie die oben in Abschn. 4.1 besprochenen Daten legen folgende Fälle der prosodisch-morphologischen V-Epenthese nahe:

(86) Fälle der prosodisch-morphologischen V-Epenthese

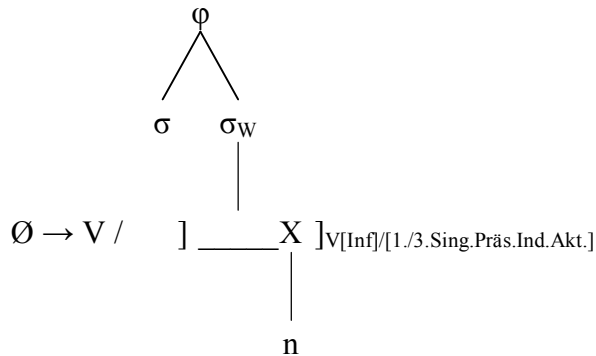
- a. *atmen, segnen, retten, reden, bauen, sehen* (vgl. (57)f)
- b. *(dem) Volke, (dem) Balle, (dem) Buche, (dem) Kinde* (vgl. (58)b)
- c. *des Schiff(e)s, des Wind(e)s, des Teich(e)s* (vgl. (57)b)

Zur Herleitung des Schwa in diesen Fällen erscheint die Annahme von drei prosodisch-morphologischen V-Epenthese-regeln als angebracht. Ort der Anwendung aller drei Regeln ist – wie eingangs besprochen – Ebene 3, da der in der Grundregel geforderte Trochäus ausschließlich in regulär flektierten Wortformen erscheint und reguläre Flexion auf Ebene 3 erfolgt. Für die V-Epenthese in den Wörtern unter (86)a schlagen wir die Regel in (87)a vor, die in der Infinitivform bzw. in der ersten und dritten Person Singular Indikativ Präsens Aktiv Form vor dem wortfinalen *n*, das daher als segmentaler (melodischer) Marker dieser Formen dient, eine leere V-Position einfügt, wenn die stamffinale Silbe metrisch stark ist, d.h. irgendeinen Grad an Betontheit hat. Dass V-Epenthese nur in den aufgelisteten Fällen, nicht aber in *segeln* oder *feuern* stattfindet, hängt genau mit der prosodischen Bedingung zusammen, dass die Suffixsilbe einer betonten Silbe folgen muss: da liquidfinale Verbstämme schon auf Ebene 1 via silbenstrukturelle V-Epenthese eine leere V-Position erhalten, kann Regel (87)a keinen Input mehr finden, die stamffinale Silbe ist nämlich eine Reduktionssilbe und daher lexikalisch unbetonbar. Das Schwa-Vorkommen in der Singular Dativ Form bestimmter Maskulina und Neutra (vgl. (86)b) legt die spezifische Version der prosodisch-morphologischen V-Epenthese-regel in (87)b nahe: hier wird allein zur Erfüllung der prosodischen Bedingung der finalen Trochäizität eine leere V-Position eingefügt. Ähnlich lässt sich die (fakultative) V-Epenthese in (86)c durch Anwendung der Regel (87)c herleiten, die in derselben prosodischen Umgebung vor finalelem *s* die leere V-Position einfügt.

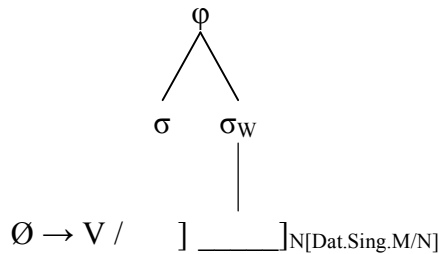


(87) Prosodisch-morphologische V-Epentheseeregeln<sup>68</sup>

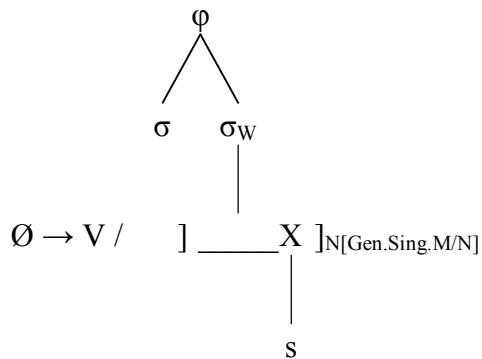
a. Infinitiv, 1./3. Person Singular Indikativ Präsens Aktiv (vgl. (57)f)



b. Dativ Singular bei Maskulina und Neutra (vgl. (58)b)



c. Genitiv Singular bei Maskulina und Neutra (vgl. (57)a)



Es sei angemerkt, dass auf den trochäischen Fuß am Ende der fraglichen Worformen auch von Wiese (1988), (1996) und von Giegerich (1987) in irgendeiner Form – als prosodische Bedingung (vgl. Wiese 1988: 157), als prosodisch-morphologische Schablone (*template*, vgl. Wiese 1996: 108ff) oder eben als Teil von Epentheseeregeln (vgl. Giegerich 1987: 459ff, bzw. die zitierten Regeln unter (71) und (73)) Bezug genommen wird. Neu an der hier vorgeschlagenen Beschreibung ist jedoch, dass von den drei genannten Einzelfällen ein

<sup>68</sup> Um komplizierte Formulierungen zu vermeiden, haben wir bei diesen Regeln auf eine merkmalsgeometrische Darstellung der mit dem finalem X assoziierten Segmenten verzichtet.

zweites Muster der V-Epenthese im Deutschen abstrahiert wird, was mit der Konzeptualisierung der Schwa-Epenthese zur Errichtung einer bestimmten prosodischen Struktur gleichgesetzt werden kann. Mit diesen konkreten Ausprägungen der oben formulierten beiden V-Epentheseeregeln in der Hand, ergibt sich folgende Struktur des deutschen Lexikons:

	<i>Morphologie</i>	<i>Phonologie</i>
<i>Ebene 1</i>	Derivation 1 →	Wortakzent
	irreguläre Flexion ←	<b>V-Epenthese 1: (85)a</b>
<i>Ebene 2</i>	Derivation 2 →	Kompositionsakzent
	Komposition ←	<b>V-Epenthese 2: (85)b, (85)c</b>
<i>Ebene 3</i>	reguläre Flexion →	<b>V-Epenthese 3: (82), (87)</b>
	←	

Tab. 16 Struktur des deutschen Lexikons und Ausprägungen der V-Epenthese

Kommen wir abschließend auf einige Fälle zu sprechen, die bisher nur erwähnt wurden, jedoch eine eigene Behandlung verdienen. Aus den obigen Ausführungen ergibt sich, dass Schwa in den Adjektivendungen hier Teil der zugrunde liegenden Repräsentation der fraglichen Endungen betrachtet wird. Dies stellt eine klare Abweichung von dem Wiese'schen und Giegerich'schen Modell dar, steht jedoch im Einklang mit Noskes Beschreibung (vgl. auch Féry 1991: 75f, die für die Repräsentation der Adjektivendungen ein prosodisches *Template* vorschlägt, in der diese mit einem eigenen Silbenknoten assoziiert sind). Eine eigene Behandlung verdienen auch die Schwa-Vorkommen vor *r*, die orthographisch mit der Graphemverbindung <er> wiedergegeben und auf der Oberfläche zumeist als vokalisches *r* [Ç] realisiert werden. Die Sequenz /WR/ kommt in Nominal- und Adjektivstämmen (vgl. (88)a und (88)b), als Pluralendung (vgl. (88)c), als Ableitungssuffix (vgl. (88)d), Komparativendung (vgl. (88)e) und als Deklinationsendung der Adjektive (vgl. (88)f) vor. Trotz der klaren orthographischen, phonetischen und teilweise lexikalisch-phonologischen (d.h. lexikalisch-repräsentationellen) Ähnlichkeiten ernährt sich dieses Schwa aus zwei Quellen. Während das Schwa in den Nominal- und Adjektivstämmen sowie in der Pluralendung als Ergebnis der V-Epenthese aufgefasst werden kann, muss es nach unserem Grundprinzip in dem Ableitungssuffix bzw. in der Komparativ- und Adjektivendung als zugrunde liegend betrachtet werden. In den ersteren Fällen kann das finale *r* nie mit dem vor dem Schwa stehenden Segment silbifiziert werden, sodass die Annahme einer silbenstrukturellen Schwa-Epenthese hier als durchaus gerechtfertigt erscheint. In den

letzteren Fällen ist das nicht mehr so: in *Seher*, *höher* und *hoher* könnte das finale *r* wohl in die erste Silbe integriert werden, wie dies die Beispiele *sehr*, *stör* und *Rohr* auch zeigen. Auch die prosodisch-morphologische V-Epenthese könnte hier nicht angewendet werden: in *Wanderer* und *trockener* haben wir vor der letzten eine weitere Reduktionssilbe. Alles spricht somit für die Annahme einer zugrunde liegenden V-Position in diesen Suffixen.

(88) [Ç]-Vorkommen im Deutschen

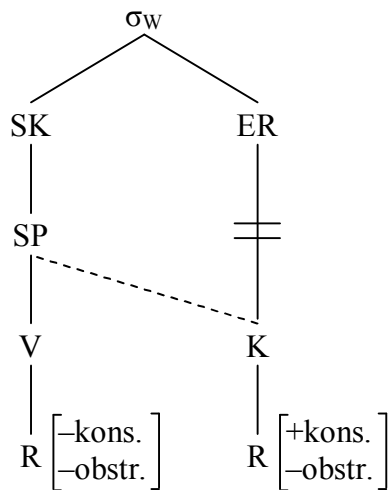
- a. *Vater*, *Bruder*, *Mutter*, *Schwester*, *Feuer* (vgl. Tab. 13a)
- b. *lecker*, *mager*, *sauber*, *sicher*, *teuer* (vgl. Tab. 13c)
- c. *Kinder*, *Würmer*, *Ränder*, *Irrtümer* (vgl. (57)c)
- d. *Verbraucher*, *Angeber*, *Käufer*, *Seher* (vgl. (57)g)
- e. *schöner*, *größer*, *trockener*, *höher* (vgl. (57)d)
- f. *schöner*, *großer*, *trockener*, *hoher* (vgl. (57)d)

## 4.6 Postlexikalische Schwa-Prozesse

In diesem letzten Abschnitt wollen wir uns mit phonologischen Prozessen beschäftigen, die sich in der postlexikalischen Phonologie abspielen und entweder ein zugrunde liegendes bzw. lexikalisch eingefügtes Schwa verändern oder ein Schwa einfügen. Ganz konkret interessieren uns dabei die Phänomene der Entstehung silbischer Sonoranten sowie der postlexikalischen Schwa-Reduktion und Verstärkung. Auf das vokalische *r*, das auch als Ergebnis postlexikalischer Prozesse entsteht, werden wir im nächsten Abschnitt ausführlicher eingehen.

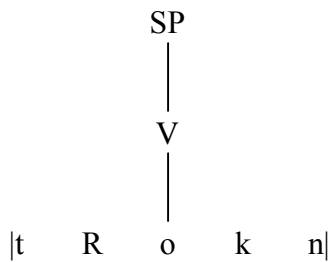
Wie oben besprochen, können Schwa+Sonorant-Sequenzen in Abhängigkeit vom Sprechtempo und -stil als silbischer Sonorant realisiert werden. Eine elementare Voraussetzung dieser Reduktion ist, dass das Schwa und der Sonorant zu derselben Silbe gehören. Somit kann in *trocken* ein silbischer Nasal erscheinen, in *trockene* muss jedoch die Sequenz Schwa+unsilbischer Nasal erhalten bleiben. Dieser fakultative Prozess lässt sich in unserem Modell durch die Verbindung des Endrandsonoranten mit dem Silbenkern und durch die sich parallel dazu ablaufende Tilgung der Assoziationslinie zwischen diesem Sonoranten und dem Endrand erfassen, vgl. (89).

(89) Fakultative Regel zur Herleitung silbischer Sonoranten

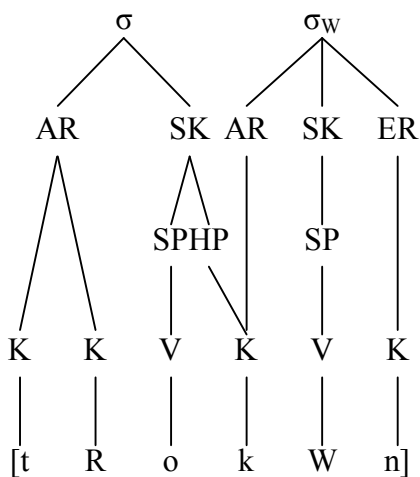


(90) Zugrunde liegende, explizite und reduzierte Oberflächenform von *trocken*

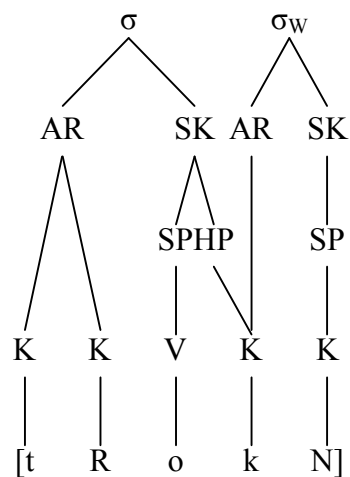
a. zugrunde liegende Form



b. explizite Oberflächenform



c. reduzierte Oberflächenform

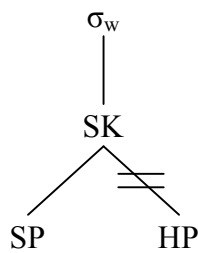


Bei der Ermittlung des Verhältnisses zwischen Schwa und Akzent in Abschn. 4.1 wurde darauf hingewiesen, dass dieser Vokal im Deutschen auch durch eine tempo- und stilabhängige Reduktion von Vollvokalen entstehen kann. Da in dem von uns vertretenen Ansatz Schwa ausnahmslos an einen bestimmten Silbentyp, nämlich die Reduktionssilbe

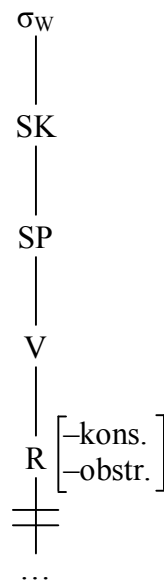
gebunden ist, können solche postlexikalischen Prozesse in zwei Phasen beschrieben werden. In einem ersten Schritt wird eine Vollsilbe in eine Reduktionssilbe verwandelt, anschließend wird der Kern der Reduktionssilbe mit dem Schwa assoziiert. Daraus, dass zwischen Voll- und Reduktionssilben der einzige Unterschied (von der unterschiedlichen Anzahl der Endrandsegmente einmal abgesehen) in der Anzahl der Silbenkernpositionen besteht (vgl. das in Abschn. 4.4 Gesagte), genauer: dass der Kern einer Vollsilbe eine Sonoritäts- und eine Halteposition besitzt, während der Kern einer Reduktionssilbe nur aus einer Sonoritätsposition besteht, folgt, dass die Reduzierung einer Vollsilbe durch Tilgung der Halteposition zu beschreiben ist (vgl. (91)a). Daran schließt sich die zweite Teilregel an, die die interne Struktur des Vokals in der postlexikalisch entstandenen Reduktionssilbe tilgt, wodurch nur die Oberklassenmerkmalspezifikation des Vokals übrigbleibt, die genau dem Schwa entspricht (vgl. (91)b).

(91) Regeln zur Schwa-Reduktion von Vollvokalen<sup>69</sup>

a. Postlexikalische Silbenreduktion



b. Schwa-Reduktion



Wir sind der Ansicht, dass es sich um die Silbenreduktionsregel in (91)a möglicherweise wiederum lediglich um ein Muster handelt, das zahlreiche Ausprägungen haben kann, die einer eigenen Untersuchung bedürfen. Wir wollen lediglich einen dieser Reduktionsprozesse kurz umreißen, nämlich die von Vennemann (1991a) beschriebene Reduktion, bei der jeder

<sup>69</sup> Tiefgestelltes ‚w‘ neben dem Silbenknoten soll hier ähnlich wie in Hall (1992) auf die metrische Schwäche der Silbe hinweisen (vgl. Fn.32).

beliebige Vollvokal zwischen zwei rhythmisch prominenten Silben zu Schwa reduziert wird, vgl. (92). Um Vennemanns Silbenreduktion zu erfassen, soll die Grundform wie in (93) modifiziert werden. (94) gibt zur Veranschaulichung die zugrunde liegende, explizite und reduzierte Oberflächenform des Wortes *Molekül* an.

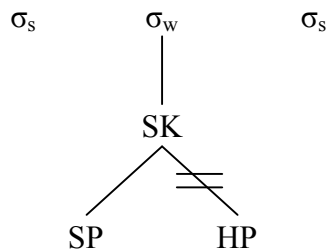
(92) Postlexikalische Schwa-Reduktion zwischen zwei prominenten Silben

*Molekül*: [moleky:l] → [molWky:l]

*Aspirin*: [aspiRi:n] → [aspWRi:n]

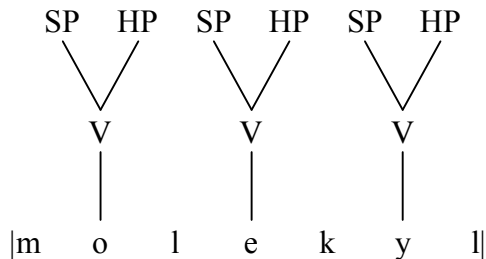
*Känguruh*: [kENguRu:] → [kENgWRu:]

(93) Rhythmisch bedingte postlexikalische Silbenreduktion

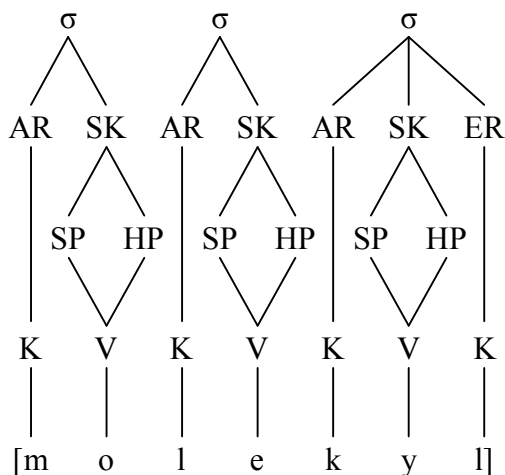


(94) Zugrunde liegende, explizite und reduzierte Oberflächenform von *Molekül*

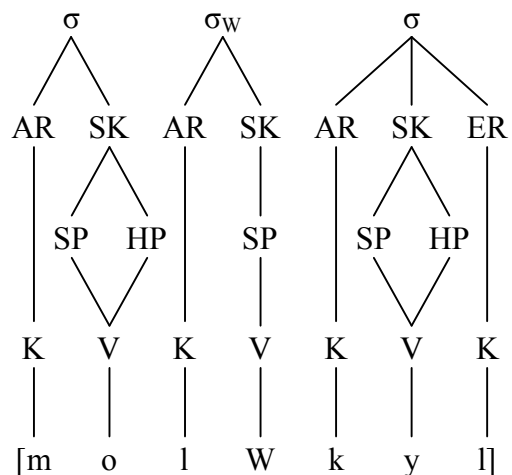
a. zugrunde liegende Form



b. explizite Oberflächenform



c. reduzierte Oberflächenform



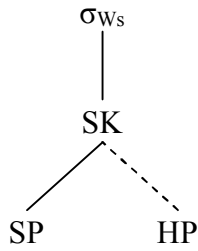
Wie ebenfalls in Absch. 4.1 besprochen, können Reduktionssilben postlexikalisch betont werden. Dabei handelt es sich um einen Kontrastakzent, mit dem auf die grammatische Form fokussiert wird, die dort gegebenen Beispiele seien hier zur Illustration wiederholt angeführt:

(95) *einé, nicht einén, bé- und éntladen*

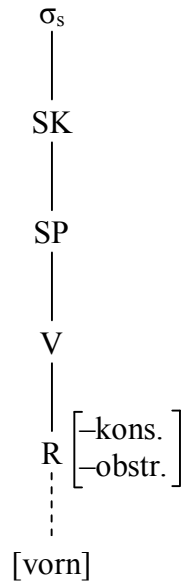
Diese exzeptionelle Akzentuierung einer Reduktionssilbe lässt sich in unserem Modell im Vergleich zur oben besprochenen Silbenreduktion genau umgekehrt beschreiben. Daraus, dass im Deutschen nur Vollsilben betont werden können, die obligatorisch einen verzweigenden Silbenaufbau haben, und Reduktionssilben, deren Kern lediglich aus einer einzigen Sonoritätsposition besteht, unbetonbar sind, muss man diesen Prozess als Verwandlung einer Reduktionssilbe in eine Vollsilbe, d.h. als Einfügung einer Halteposition in eine Reduktionssilbe auffassen. Die einschlägige HP-Einfügingsregel wird in (96)a formuliert. An der dadurch entstandenen Silbe müssen jedoch zwei Veränderungen durchgeführt werden, die die Anwendung von zwei weiteren Regeln nötig machen. Erstens muss an Stelle des Schwa ein *e*-Vokal erscheinen, da Schwa nur in Reduktionssilben vorkommen kann. Dies kann erreicht werden, indem dem Schwa als einem Vokal ohne interne Struktur die Merkmalspezifizierung der *e*-Vokale, d.h. das uniäre Merkmal [vorn] zugewiesen wird, vgl. (96)a. Dann muss man die eingefügte Intensitätsposition mit segmentalem Material füllen, und dadurch dem *e*-Vokal einen Silbenschnittwert zuweisen. Dabei stehen uns zwei Möglichkeiten zur Verfügung. Entweder wird die Halteposition mit dem Endrandkonsonanten oder mit dem Vokal selbst verbunden, im ersteren Fall resultiert daraus ein scharf geschnittenes *e*, das phonetisch als [E] realisiert wird, im letzteren dagegen ein sanft geschnittenes *e*, das in der phonetischen Realisierung als [e:] interpretiert wird. Die Regeln sind in (97) zu finden. Wichtig ist die Reihenfolge der beiden Regeln: wenn Regel (97)a keinen Input (d.h. keinen Konsonanten im Endrand) findet, wird (97)b angewendet. Zur Illustration steht unter (98) und (99) die zugrunde liegende, die unbetonte sowie die kontrastakzentuierte Oberflächenform der Wörter *eine* und *einen*.

(96) Regeln zur postlexikalischen Reduktionssilbenbetonung

a. HP-Einfügensregel

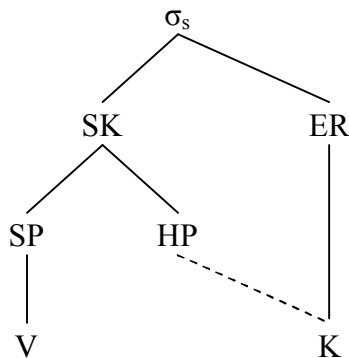


b. e-Assoziation

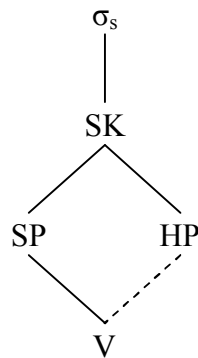


(97) Zuweisung des Silbenschnittwertes

a. Scharfer Schnitt

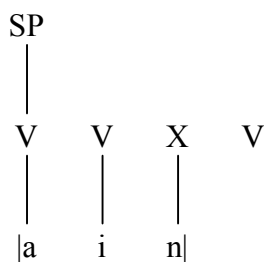


b. Sanfter Schnitt

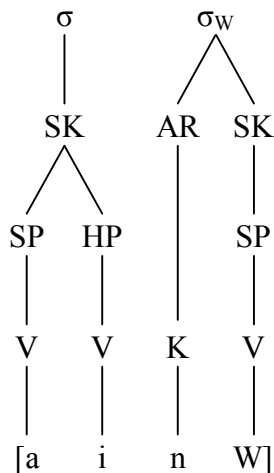


(98) Zugrundeliegende, unbetonte und kontrastakzentuierte OberflächenForm von *eine*

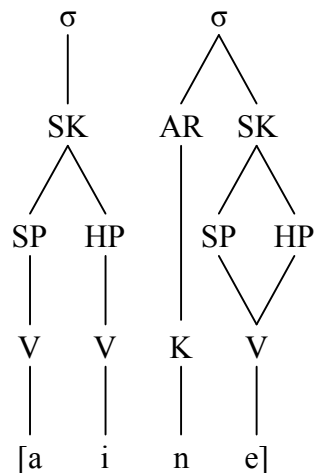
a. ZL-Form



b. unbetonte O-Form



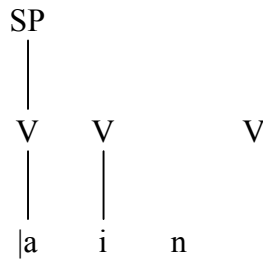
c. betonte O-Form



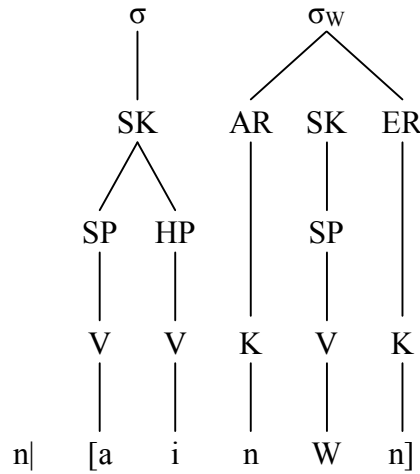


(99) Zugrundeliegende, unbetonte und kontrastakzentuierte OberflächenForm von *einen*

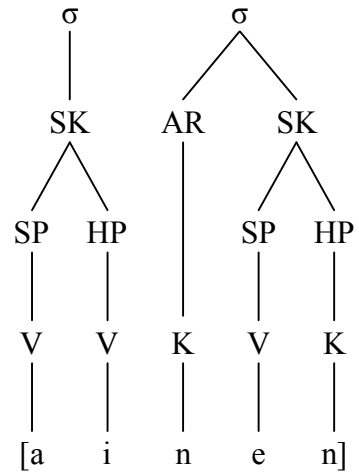
a. ZL-Form



b. unbetonte O-Form



c. betonte O-Form



## 4.7 Schwa und *r*-Vokal

Gegenstand des letzten Abschnitts des Kapitels zur Reduktionssilbe ist das Verhältnis des Schwa zum silbischen vokalischen *r* [Ç], das der Anforderung unter (63)a in Abschn. 4.1 entsprechend in einer adäquaten Schwa-Beschreibung des Gegenwartsdeutschen unbedingt behandelt werden muss, und das wie am Ende von Abschn. 4.5 besprochen einer eigenen Untersuchung bedarf. Bisher sind wir in Anlehnung an die gängigen generativ-phonologischen Analysen – so u.a. Wiese (1988) und (1996) – von der stillschweigenden Annahme ausgegangen, dass alle [Ç]-Vorkommen des Deutschen, das in (88) und nochmals unter (100) aufgelistet sind, das Lexikon als die Sequenz /WR/ verlassen.

(100) [Ç]-Vorkommen im Deutschen

- a. Nominalstämme: *Vater, Bruder, Mutter, Schwester, Feuer*
- b. Adjektivstämme: *lecker, mager, sauber, sicher, teuer*
- c. Pluralendung: *Kinder, Würmer, Ränder, Irrtümer*
- d. Ableitungssuffix: *Verbraucher, Angeber, Käufer, Seher*
- e. Komparativendung: *schöner, größer, trockener, höher*
- f. Adjektivendung: *schöner, größer, trockener, hoher*

Die Motivation für diese Annahme ergibt sich daraus, dass bei vielen dieser Schwa-Vorkommen ein systematischer Wechsel zwischen dem silbischen vokalischen *r* und der Sequenz Schwa + konsonantisches *r* zu beobachten ist – einerseits bei der Adjektivdeklinaton

(vgl. (101)b und e) und andererseits infolge einer Konversion bei der Konjugation und deverbale Wortbildung (vgl. (101)a–c und e). Wo eine solche Alternation nicht besteht (vgl. (101)d und f), beruht die Annahme der Sequenz /WR/ als lexikalischer Output auf der Analogie: Da für die alternierenden Fälle die phonologische Komponente sowieso einen Regelapparat zur Herleitung des silbischen vokalischen *r* aus der Sequenz Schwa + konsonantisches *r* hat, kann durchaus zugelassen werden, dass auch diese Wortformen das Lexikon mit der Sequenz /WR/ verlassen.

(101) Systematischer Wechsel zwischen [Ç] und [WR]

- a. Nominalstämme: *Feuer* [Ç] – *Feuerung* [WR], *Wunder* [Ç] – *wundere* [WR]
- b. Adjektivstämme: *lecker* [Ç] – *leckere* [WR], *sicher* [Ç] – *Sicherung* [WR]
- c. Pluralendung: (*Ränder* [Ç] – *rändere* [WR])
- d. Ableitungssuffix: –
- e. Komparativendung: *schöner* [Ç] – *Verschönerung* [WR], *größer* [Ç] – *vergrößere* [WR]
- f. Adjektivendung: –

Wie oben besprochen, wollen wir in unserem – im Vergleich zu Wiese (1988) oder Giegerich (1987) jedenfalls – ‚gemäßigt‘ abstrakten Beschreibungsansatz V-Epenthese nur dort zur Anwendung kommen lassen, wo sie zu prosodischen (Silbifizierbarkeit) oder prosodisch-morphologischen Zwecken (Erfüllung von prosodischen Mustern bestimmter morphologischer Formen) dient. Daraus haben wir den Schluss gezogen, dass das lexikalische Schwa in den Fällen (100)a–c als Ergebnis der Einfügung einer leeren V-Position zu betrachten ist und in diesen Wörtern zugrunde liegend folgerichtig nur ein konsonantisches *r* angenommen werden soll, während das Schwa, oder besser die leere V-Position in allen anderen Formen (vgl. (100)d–f) schon zugrunde liegend vorhanden ist. In *Seher*, *höher* und *hoher* wäre nämlich V-Epenthese unmotiviert, da ein bloßes konsonantisches *r* hier durchaus in die vorausgehende Silbe integriert werden könnte, in den Fällen (100)a–c bleibt dagegen das konsonantische *r* ausnahmslos unsilbifiziert, was V-Epenthese erzwingt<sup>70</sup>, wobei die

---

<sup>70</sup> Als einzige Problemfälle könnten die Nominal- bzw. Adjektivstämme mit einem Diphthong vor dem silbischen vokalischen *r*, so z.B. *Feuer*, *Steuer*, *teuer*, *sauer* usw. betrachtet werden. Hier könnte im Prinzip das konsonantische *r* in die vorausgehende Silbe integriert werden, da es ein weniger sonores Segment darstellt als der vorausgehende Randvokal des Diphthongs. Dass auch hier V-Epenthese erfolgt bzw. *r* als Silbenträger funktioniert, erklären Wiese (1988), Giegerich (1987) und Féry (1993) mit dem minimalen

Einfügung der leeren V-Position – wie in Abschn. 4.5 beschrieben – in Abhängigkeit von der Wortklassenzugehörigkeit des fraglichen Stammes auf der ersten (Verben), auf der zweiten (Nomen und ein Teil der Adjektive) bzw. auf der dritten Ebene (die restlichen Wörter) erfolgt, vgl. Tab. 16 bzw. die Regeln (85), (82).

Wie gesagt, werden alle leeren V-Positionen – unabhängig davon, ob sie schon zugrunde liegend vorhanden sind oder erst durch V-Epenthese eingefügt werden – am Ende der lexikalischen Derivation mit der für das Schwa charakteristische Merkmalstruktur unter (65)d assoziiert, sodass alle Wortformen unter (100) mit der Sequenz /WR/ in die postlexikalische phonologische Komponente gelangen. Dort wird die fragliche Sequenz vokalisiert, wenn die beiden Segmente – wie dies das jeweils erste Glied in den Worpaaren unter (101) zeigt – zur gleichen Silbe gehören. Die Veränderung /WR/ → [Ç] lässt sich in zwei Schritten herleiten: Einerseits soll das Schwa aus dem Silbenkern getilgt bzw. *r* mit dem Silbenkern verbunden werden, was naturgemäß mit der ‚Losbindung‘ dieses Segments vom Endrand einhergehen muss. Andererseits soll das konsonantische *r* vokalisiert werden. Vor der Formulierung dieser beiden Regeln scheint es jedoch angebracht, einen Exkurs auf die in Kap. 2 eingeführte andere Ausprägung des vokalischen *r*, auf das unsilbische vokalische *r* zu machen. Die Daten unter (102) zeugen nämlich von einem systematischen Wechsel zwischen diesem Vokal und dem konsonantischen *r*, welcher Wechsel sich ebenfalls als *r*-Vokalisierung beschreiben lässt. Somit stellt sich die Frage, ob man die beiden phonologischen Prozesse mit einer einheitlichen Regel erfassen kann, m.a.W. ob sich nachweisen lässt, dass sie in demselben Kontext stattfinden.

(102) Systematischer Wechsel zwischen konsonantischem und unsilbischem vokalischem *r*

*Tier* [Ç&] – *Tiere* [R], *Tür* [Ç&] – *Türen* [R], *stur* [Ç&] – *sture* [R]...

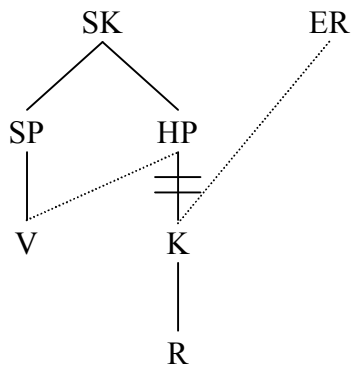
Gewöhnlich geht man davon aus, dass das unsilbische vokalische *r* nur nach tautosilbischem Langvokal, nicht aber nach tautosilbischem Kurzvokal vorkommt, so wird ein zugrunde liegendes konsonantisches *r* in *wir* vokalisiert, in *wirr* bleibt es dagegen auch in der Oberflächenrepräsentation als Konsonant erhalten. Dieser Sachverhalt legt für die hier vertretene Silbenschnittauffassung den Schluss nahe, die Veränderung konsonantisches *r* →

---

Sonoritätsunterschied zwischen dem Gleitlaut (diphthongischen Randvokal) und *r*. Nicht auszuschließen ist jedoch, dass das Deutsche keine Triphthonge, d.h. tautosilbische Vokal-Vokal-Vokal-Verbindungen zulässt, was durch die Formulierung eines Anti-Triphthongs-Filters zum Ausdruck gebracht werden könnte, der V-Epenthese auf den entsprechenden lexikalischen Ebenen auslöst.

unsilbisches vokalisches *r* – ähnlich wie Becker (1998) – nur im Endrand, nicht aber in der Halteposition (bei Becker (1998) in der Implosion) zuzulassen. Damit ist nun der gemeinsame Kontext für die beiden *r*-Vokalisierungen gegeben: ein konsonantisches *r* wird im Silbenendrand vokalisiert, das dadurch entstandene vokalische *r* bleibt in einer Vollsilbe auf der Oberfläche im Endrand erhalten, in einer Reduktionssilbe löst es dagegen die Tilgung des nuklearen Schwa aus und wird zum Silbenkern. Dieses auf den ersten Blick klare Bild wird jedoch gewissermaßen verzerrt, wenn man berücksichtigt, dass das oben genannte einfache orthoepische Prinzip (*r*-Vokal nach Langvokal – *r*-Konsonant nach Kurzvokal) von vielen Sprechern nicht eingehalten wird: [Ç&] in *wirr*, *dürr*, *Herr* usw. ist durchaus als normal zu bewerten, was nahelegen würde, dass *r*-Vokalisierung nicht nur im Endrand, sondern auch in der Halteposition erfolgt. Ob jedoch vor dem vokalischen *r* in diesen Wörtern in der Tat ein kurzer ungespannter Vokal vorkommt, ist nicht ganz klar. Wir wollen uns an dieser Stelle an Vater (1992) und Becker (1998) anschließen, und davon ausgehen, dass *r*-Vokalisierung in diesem Fall mit der Dehnung des vorausgehenden Vokals einhergeht, in *wirr*, *dürr*, *Herr* usw. also ein durch das nachfolgende *r* bewirkter Silbenschnittumschlag stattfindet, bei dem ein konsonantisches *r* aus der Halteposition in den Endrand verdrängt wird:

(103) Postlexikalischer Silbenschnittumschlag



Mit der Annahme einer solchen Regel steht nun nichts im Wege, den Kontext der einheitlichen *r*-Vokalisierungsregel als den Endrand der Silbe zu bestimmen. Allerdings muss vor der Formulierung der einschlägigen Regel die Merkmalstruktur der beiden betroffenen Segmente bestimmt werden. Ein erstes Problem bereitet dabei die enorme Vielfalt der konsonantischen *r*-Variationen, die von dem apikalen Vibranten über den uvular/velar-postdorsalen Approximanten bis hin zum uvularen Vibranten reichen (vgl. Schiller 1995). Will man die apikale *r*-Variante als dialektales Phänomen betrachten, wofür guter Grund besteht, so lässt sich für das Standarddeutsche ein konsonantisches *r* annehmen, das einen

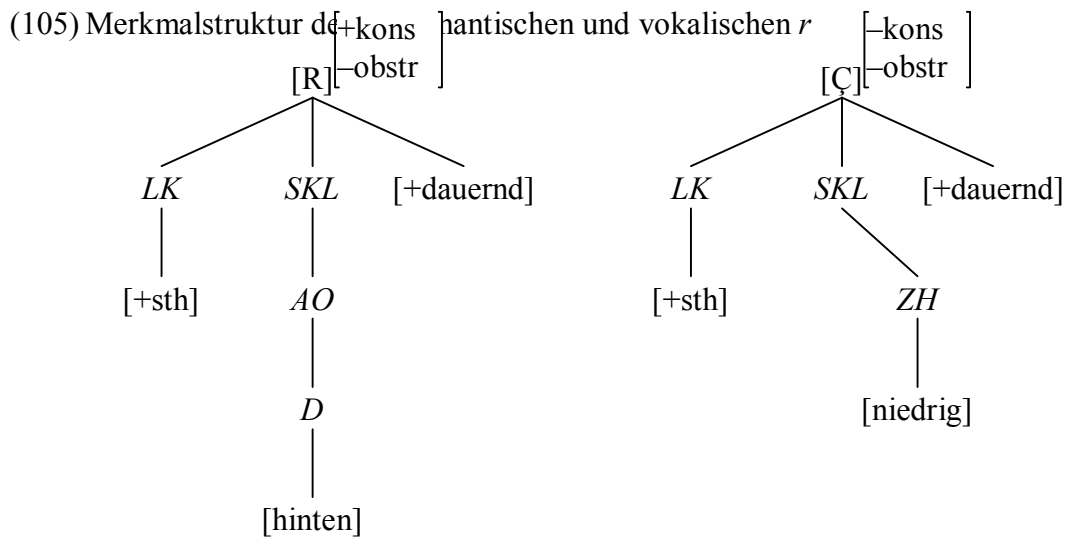
Sonoranten (darauf weisen seine phonotaktischen Eigenschaften hin) darstellt, der mit dem hinteren Teil des Zungenrückens im velar-uvularen Bereich gebildet wird. Vergleicht man dieses konsonantische sowie das vokalische *r*-Allophon miteinander, so findet man oft die Auffassung, dass sie voneinander nur minimal unterscheiden. So äußert sich Becker (1998) wie folgt:

Wenn man für Konsonanten ein artikulatorisches Hemmnis als charakteristisch ansieht und für Vokale die Einstellung eines Resonanzraumes [...], so kann man sagen, das vokalische /r/ ist das konsonantische, und zwar als Vokal gesehen. Das konsonantische /r/ ist das vokalische mit einem zusätzlichen Geräusch (Vibration oder Friktion), das in Schwächungspositionen, z. B. im Nukleus oder Endrand der Silbe, wegfallen kann. (Becker 1998: 154)

So wird z.B. von Hall (1992) oder Wiese (1996) zwischen dem konsonantischen und vokalischen *r* nur ein einziger Unterschied angenommen: Ersteres hat für das Merkmal den positiven, Letzteres den negativen Wert. Doch unterscheiden sich die beiden Ansätze in einem nicht unwesentlichen Punkt. Hall (1992) fasst das vokalische *r* als hinteren nichtrunden mittleren Vokal auf und verwendet dafür das IPA-Symbol [U], während es in Wiese (1996) als niedrigen nichtrunden Zentralvokal angesehen wird, wodurch das vokalische *r* mit den *a*-Vokalen gleichgesetzt wird. Zwar steckt in beiden Auffassungen eine gewisse Wahrheit, wir sind der Ansicht, dass sie aus einem phonetischen Standpunkt aus gesehen nicht ganz korrekt sind, genauso sind wir mit der oben zitierten Behauptung Beckers über die Verharmlosung der phonetischen Unterschiede zwischen den beiden *r*-Allophonen nicht ganz einverstanden. Die EMA-Untersuchung von Schiller & Mooshammer (1995) hat nämlich ergeben, dass im Allgemeinen ein silbeninitiales *r* mit einem höheren, ein silbenfinales dagegen mit einem niedrigeren Zungenkörper gebildet wird. Somit hat Hall (1992) gewissermaßen recht, wenn er davon ausgeht, dass das konsonantische *r* ein hinteres mittleres Segment darstellt, doch erscheint seine Einstufung des vokalischen *r*-Allophons als mittlerer hinterer Vokal als phonetisch unmotiviert. Von der Arbitrarität einer solchen Analyse zeugt auch Halls Argumentation. Er entscheidet sich nämlich deswegen für [U] im Gegensatz zu [Ç], da sich dieses Segment von dem uvularen [R] allein im Merkmal [konsonantisch] unterscheidet, was die Formulierung der *r*-Vokalisierungregel erleichtert, die dadurch nur eine Veränderung dieses Merkmals enthalten soll, während im Falle der Annahme eines niedrigen Vokals [Ç] bei der Herleitung des vokalischen *r* vom konsonantischen auch der Merkmalswert [niedrig] verändert werden sollte (vgl. Hall 1992: 20). Dieses Argument entbehrt naturgemäß jeder phonetischen Grundlage. Wieses (1996) Analyse hat dagegen den Vorteil, dass der Autor das



Ansonsten sind beide Segmente stimmhafte Sonoranten (vgl. den Merkmalswert [+stimmhaft] unterhalb des laryngalen Knotens *LK* bzw. die Oberklassenspezifizierung [–obstruent]).

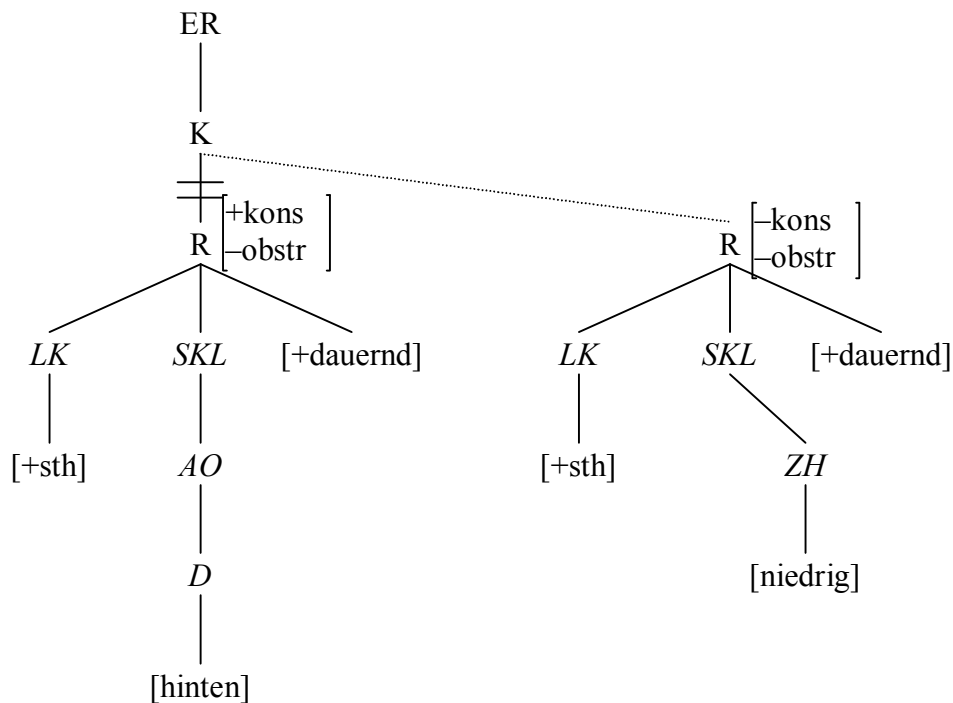


Der experimentalphonetischen Feststellung von der Senkung des hinteren Teils der Zungenmasse vom konsonantischen zum vokalischen *r* hin will dieser Repräsentationsvorschlag somit dadurch gerecht werden, dass für das konsonantische *r*-Allophon eine mittlere, für das vokalische eine niedrigere Zungenhöhe angenommen wird. Im mittleren Bereich kann jedoch der Zungenkörper eine beliebige vertikale Position einnehmen, wodurch die Frage nach der genauen Ausprägung der konsonantischen *r*-Variante offen gelassen wird: sie kann genauso ein Vibrant wie ein frikativähnlicher Sonorant oder Approximant sein. Auch das gewählte IPA-Symbol für das konsonantische *r* [R] soll hier – übrigens ähnlich wie bei Wiese (1996) – nicht phonetisch als Vibrant interpretiert, sondern eher als eine abstrakte Transkriptionsmöglichkeit für den (velar-uvularen) *r*-Konsonanten verstanden werden. Dass hier [Ç] als niedriger Zentralvokal aufgefasst wird, entspricht genau der phonetischen Beschaffenheit dieses Lautsegments, ermöglicht jedoch einigermaßen eine repräsentationelle Unterscheidung zwischen diesem Vokalsegment und den *a*-Vokalen, da Letztere infolge ihres phonologischen Verhaltens in unserer phonetisch-phonologisch motivierten Repräsentation als hintere Vokale aufgefasst werden, während für die phonologische Einstufung von [Ç] als Hinterzungenvokal nichts spricht.

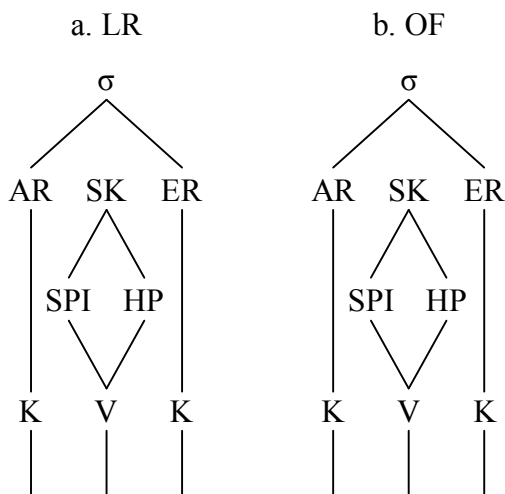
Mit dieser Merkmalsstruktur des konsonantischen und vokalischen *r*-Allophons sowie des oben ermittelten Kontextes der *r*-Vokalisierung lässt sich die fragliche phonologische Regel wie in (106) formulieren. Diese Regel kann die betreffende Veränderung (ein hinterer mittlerer Sonorant wird zum zentralen niedrigen Vokal) infolge der Konventionen der

autosegmentalen Regelformulierung dadurch zum Ausdruck bringen, dass unterhalb des Endrandes der gesamte Merkmalbaum des [R] durch die Merkmalstruktur des [Ç] ersetzt wird, da hier gleichzeitig zur Veränderung der Zungenstellungen auch ein Oberklassenwechsel erfolgt. (107) enthält als Illustration der Arbeitsweise der *r*-Vokalisierungsregel die lexikalische und Oberflächenrepräsentation des Wortes *Bier*. Aus diesen Abbildungen geht auch hervor, dass sich die öffnenden Diphthonge des Deutschen (vgl. (5)c) im Vergleich zu den steigenden (vgl. (5)b sowie (50) und (51)) spiegelbildlich darstellen lassen: Der Kopfvokal besetzt dabei immer den Silbenkern, der Randvokal (das vokalische *r*) bildet dagegen den Endrand.

(106) *r*-Vokalisierungsregel



(107) Lexikalische Repräsentation (LR) und Oberflächenform (OF) von *Bier*

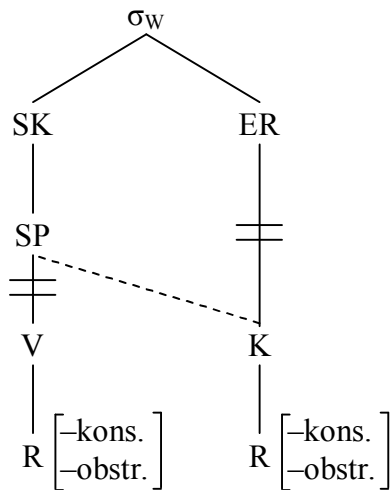




b i R            b i Ç

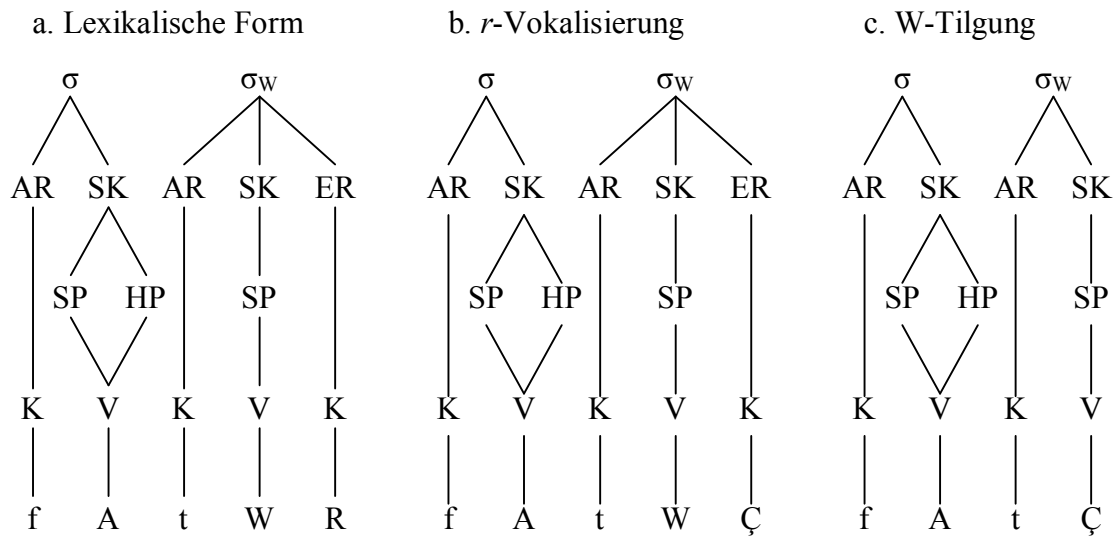
Zur Herleitung eines silbischen vokalischen *r* ist auch eine weitere Regel nötig, die das Schwa, das die einzige subnukleare Position einer Reduktionssilbe besetzt, durch den im Endrand stehenden *r*-Vokal ersetzt. Dies lässt sich nach den Konventionen der autosegmentalen Regelformulierung dadurch erfassen, dass der mit dem Endrand assoziierte Vokal mit dem Silbenkern verbunden wird, was mit der Tilgung der Assoziationslinie zwischen Endrand und Vokal bzw. zwischen Silbenkern und Schwa einhergeht:

(108) Schwa-Tilgung vor tautosilbischem *r*



Aus der Abbildung ist zu sehen, dass in der Regelformulierung der mit dem Endrand verbundene Vokal für seine segmentale Beschaffenheit (Merkmalstruktur) nicht spezifiziert ist. Dies ist deswegen so, da in dieser Position nur das durch die Anwendung von (106) entstehende vokalische *r* vorkommen kann. Anzumerken ist, dass diese Regel vor der fakultativen Regel (89) angewendet werden muss, der für die Herleitung der silbischen Sonoranten verantwortlich ist. Eine solche Regelordnung ermöglicht nämlich eine einfache Formulierung der fraglichen fakultativen Regel. Da das konsonantische *r* schon vorher vokalisiert wird, kann man als Input für die Regel unter (89) lediglich die im Endrand stehenden Sonoranten angeben. In (109) wird die Interaktion der *r*-Vokalisierungs- und der Schwa-Tilgungsregel am Beispiel *Vater* veranschaulicht.

(109) Lexikalische Repräsentation (a), *r*-Vokalisierung (b) und Schwa-Tilgung (c) in *Vater*



## 4.8 Zwischenbilanz

Der Gegenstand des vorliegenden Kapitels waren die verschiedenen Besonderheiten des Schwa des Gegenwartsdeutschen, die wir zusammen als das Schwa-Phänomen bezeichnet haben. Unser Ausgangspunkt war dabei, dass das Konzept der Reduktionssilbe zusammen mit einigen weiteren Annahmen in der Lage ist, für all diese Eigentümlichkeiten einen adäquaten Beschreibungsrahmen zu bieten.

Die erste Besonderheit dieses Vokals, sein neutraler phonetischer Charakter (vgl. die Frage in (59)) haben wir mit der Annahme uniärer phonologischer Merkmale unterhalb des supralaryngalen Knotens im Strukturbaum erfasst. Die Einwertigkeit dieser Merkmale erlaubt, dass die neutrale Stellung der Lippen, der vertikalen und horizontalen Zungenbewegung mit fehlenden Merkmalswerten ausgedrückt wird. Dadurch erhält das Schwa nicht nur zugrunde liegend, sondern während der gesamten Derivation außer den Oberklassenmerkmalen für Vokale überhaupt keine Merkmalspezifizierung, was insgesamt seine minimale Vokalität widerspiegelt.

Die Unakzentuierbarkeit der schwahaltigen Silben (vgl. die Frage unter (60)a) wurde hier mit der Annahme eines eigenen Silbentyps, der Reduktionssilbe erreicht, die sich gegenüber den Vollsilben infolge der fehlenden Halteposition im Silbenkern unversell durch generelle Unbetonbarkeit auszeichnet. Der exceptionellen Betonung der Schwa-Silben zu metalinguistischen Zwecken (vgl. die Frage unter (60)b) wurde unsere Analyse durch die Annahme einer postlexikalischen Regel gerecht, die eine Reduktionssilbe unter

(postlexikalischem) Akzent in eine Vollsilbe verwandelt, indem sie im Silbenkern eine Halteposition einfügt. Diese Position wird dann mit dem ersten Segment im Endrand oder mangels eines solchen Segmentes mit dem Schwa selbst verbunden, woraus im ersteren Fall ein scharf geschnittener, im letzteren ein sanft geschnittener Nukleusvokal resultiert. Da jedoch Vollsilben universell nur Vollvokale (und folgerichtig kein Schwa) enthalten können, folgt auf diesen Prozess die Verstärkung des Schwa zu dem ihm artikulatorisch am nächsten stehenden Vollvokal, also zu einem *e*-Vokal. Die vom stil- und tempoabhängige Reduktion von Vollvokalen (vgl. die Frage unter (60)c) wurde hier mit einem der exzeptionellen Schwa-Silben-Betonung entgegengesetzten postlexikalischen Prozess beschrieben, bei dem eine Vollsilbe durch Tilgung der Halteposition in eine Reduktionssilbe verwandelt wird, auf die die Tilgung der Merkmalstruktur des Nukleusvokals folgt, da Reduktionssilben nur Schwa als Vokal enthalten können.

Dass das Schwa außerhalb der Silbenschnittopposition steht (vgl. die Frage unter (61)), wurde hier dadurch erfasst, dass das Schwa zugrunde liegend nur mit einer skelettalen V-Position repräsentiert ist, während die an dieser Opposition teilnehmenden Vokale des Deutschen für die erste oder für beide subnuklearen Silbenpositionen spezifiziert sind. Damit das Schwa auch im Laufe der Derivation keine Schnitteigenschaften erhält, wird davon ausgegangen, dass der Kern einer Reduktionssilbe nur den Tochterknoten Sonoritätsposition dominiert.

Die einzelnen Schwa-Vorkommen (vgl. die Frage unter (62)) wurden hier im Rahmen der Lexikalischen Phonologie beschrieben. Dabei sind wir davon ausgegangen, dass die meisten Schwa constantia in Form einer leeren skelettalen V-Position schon zugrunde liegend vorhanden sind, während die restlichen Schwas auf den einzelnen lexikalischen Ebenen durch V-Epenthese erzeugt werden, wobei die Einfügung einer V-Position entweder zu prosodischen oder zu prosodisch-morphologischen Zwecken dient. Auf die zugrunde liegenden wie durch die V-Epenthese eingefügten leeren skelettalen Positionen werden vom Silbifizierungsalgorithmus Reduktionssilben aufgebaut, diese Positionen werden am Ende der lexikalischen Derivation mit den beiden Oberklassenmerkmalen assoziiert.

Die Alternation zwischen der Sequenz Schwa + konsonantisches *r* und vokalischem *r* (vgl. die Frage unter (63)a) – wie überhaupt die Vorkommen des silbischen vokalisches *r* – wurde hier mit zwei sukzessiven postlexikalischen Regeln beschrieben, von denen die erste ein im Endrand stehendes konsonantisches *r* vokalisiert und die zweite dieses vokalisches *r* mit dem Silbenkern verbindet, was selbstredend mit der Tilgung der Assoziation zwischen Schwa und Sonoritätsposition bzw. zwischen *r*-Vokal und Endrand einhergeht. Exkursmäßig sind wir

auch auf die Herleitung des unsilbischen vokalischen *r* eingegangen und gezeigt, dass die beiden *r*-Vokalisierungsprozesse mit einer einheitlichen Regel erfasst werden können.

Schließlich haben wir die freie Variation zwischen der Sequenz Schwa + unsilbischer Sonorant und silbischem Sonoranten (vgl. die Frage unter (63)b) durch die Annahme einer fakultativen postlexikalischen Regel erfasst, die den Sonoranten im Endrand unter gleichzeitiger Tilgung der Assoziationslinie zwischen ihm und dem Endrand sowie dem Schwa und dem Silbenkern mit der Sonoritätsposition verbindet.

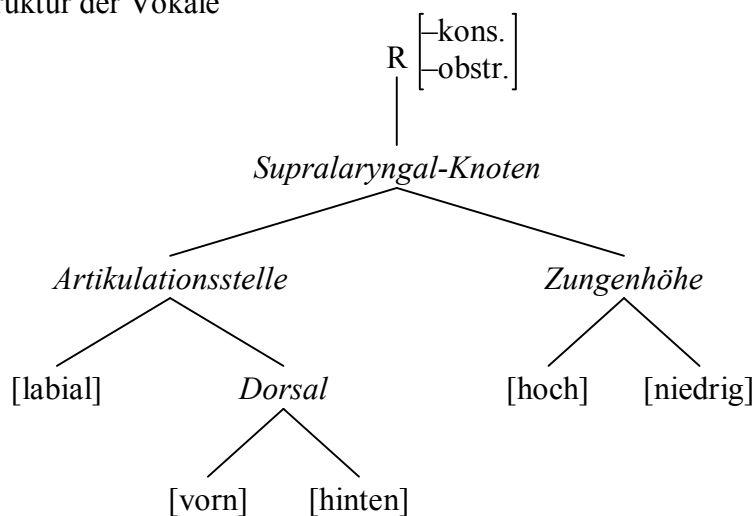
## 5. Zusammenfassung und Ausblick

### 5.1 Das Vokalsystem des Gegenwartsdeutschen

Aufgrund des in den beiden thematischen Kapiteln der vorliegenden phonologischen Untersuchung zum deutschen Vokalismus Gesagten lässt sich das Vokalsystem des Gegenwartsdeutsche wie folgt zusammenfassen.

Im heutigen Deutsch kann man nach den drei qualitativen Kriterien Lippenrundung, vertikale und horizontale Zungenbewegung 8 Vokalklassen unterscheiden. Diese Vokalklassen lassen sich mit den 5 einwertigen Merkmalen [labial], [hoch], [niedrig], [vorn] und [hinten] beschreiben, die sich im Merkmalbaum unter dem supralaryngalen Knoten wie in (110) unterbringen lassen. Eine tabellarische Übersicht über diese Klassifizierung ist in Tab. 17 gegeben. (111) enthält die Merkmalbäume für die einzelnen Vokalqualitäten des Deutschen.

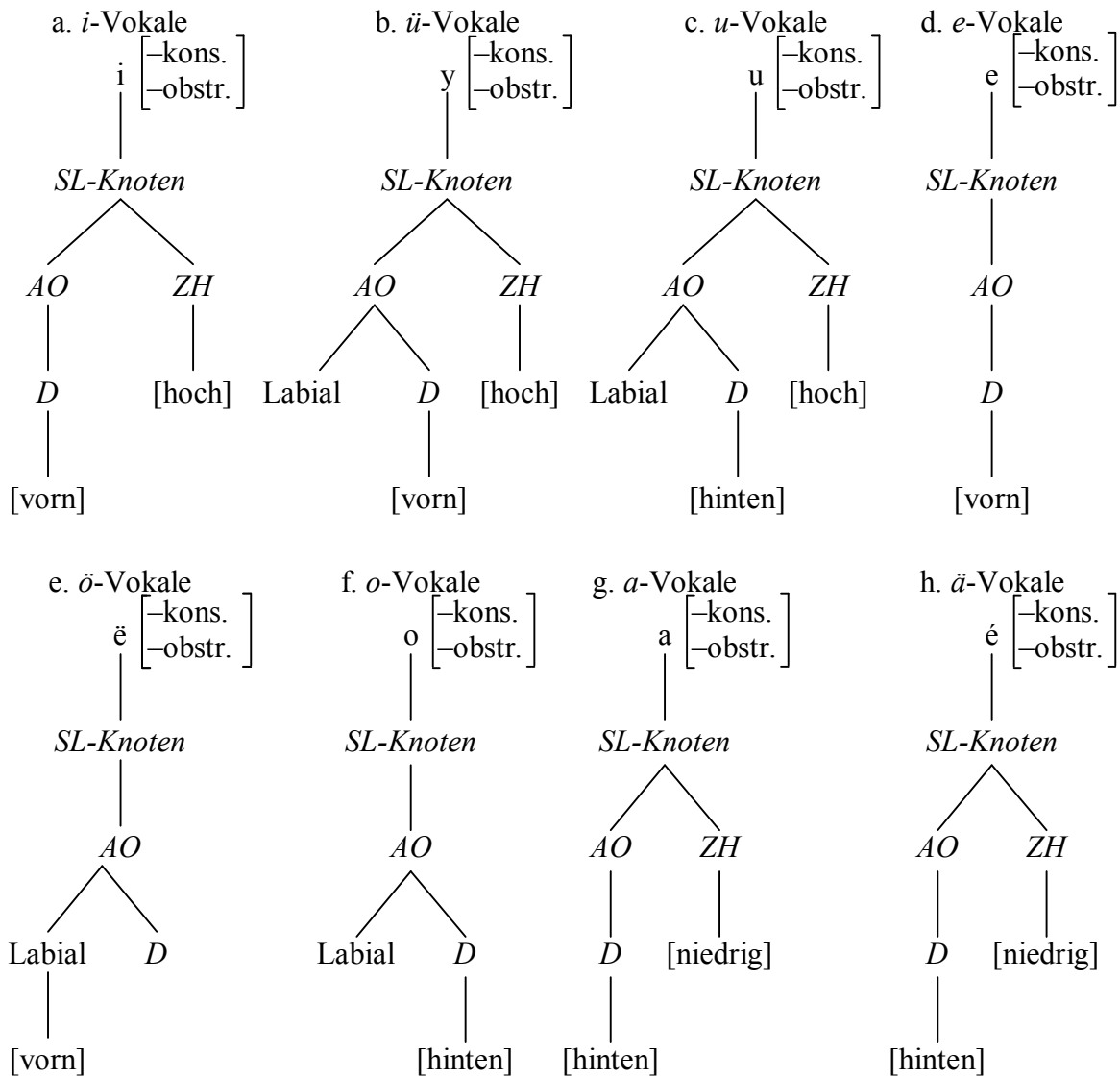
(110) Interne Struktur der Vokale



<i>Knoten</i>		<b>i</b>	<b>ü</b>	<b>e</b>	<b>ö</b>	<b>ä</b>	<b>u</b>	<b>o</b>	<b>a</b>
<i>Artikulationsort</i>	<i>Dorsal</i>	vorn	vorn	vorn	vorn	vorn	hinten	hinten	hinten
			labial		labial		labial	labial	
<i>Zungenhöhe</i>		hoch	hoch			niedrig	hoch		niedrig

Tab. 17 Qualitative Klassifizierung der deutschen Vokale mit uniären Merkmalen

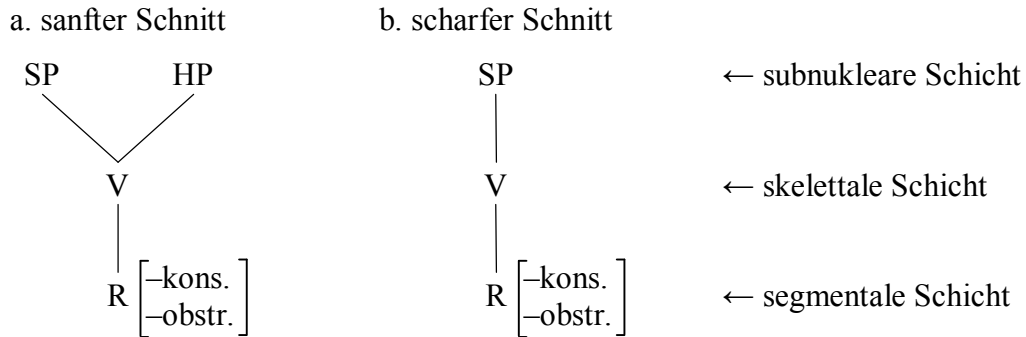
(111) Interne Struktur der Vollvokale des Gegenwartsdeutschen



Neben dieser qualitativen Unterscheidung weist das Deutsche auch eine prosodische Opposition auf, die als Silbenschnittgegensatz aufgefasst werden kann. Dabei wird ein Teil der Silbenstruktur schon in der zugrunde liegenden Repräsentation angenommen, dieser Teil entspricht den subnuklearen Konstituenten Sonoritätsposition und Halteposition. Ein sanft geschnittener Vokal ist dabei zugrunde liegend mit beiden universellen Bestandteilen des Kerns einer Vollsilbe assoziiert, ein scharf geschnittener jedoch nur mit der Sonoritätsposition, vgl. (112). Ein Vokal unter sanftem Schnitt wird phonetisch durch gespannte (dezentralisierte) Artikulation und (zumindest unter Akzent) länger, einer unter scharfem Schnitt durch ungespannte (zentralisierte) Artikulation und (unabhängig vom Akzent) kürzer realisiert. Gespanntheit und Dauer sind daher aus einem strikt phonologischen Gesichtspunkt keine phonologischen Größen und müssen folgerichtig als Teil der

aus keine phonologischen Größen und müssen folgerichtig als Teil der phonetischen Realisierung gesehen werden.

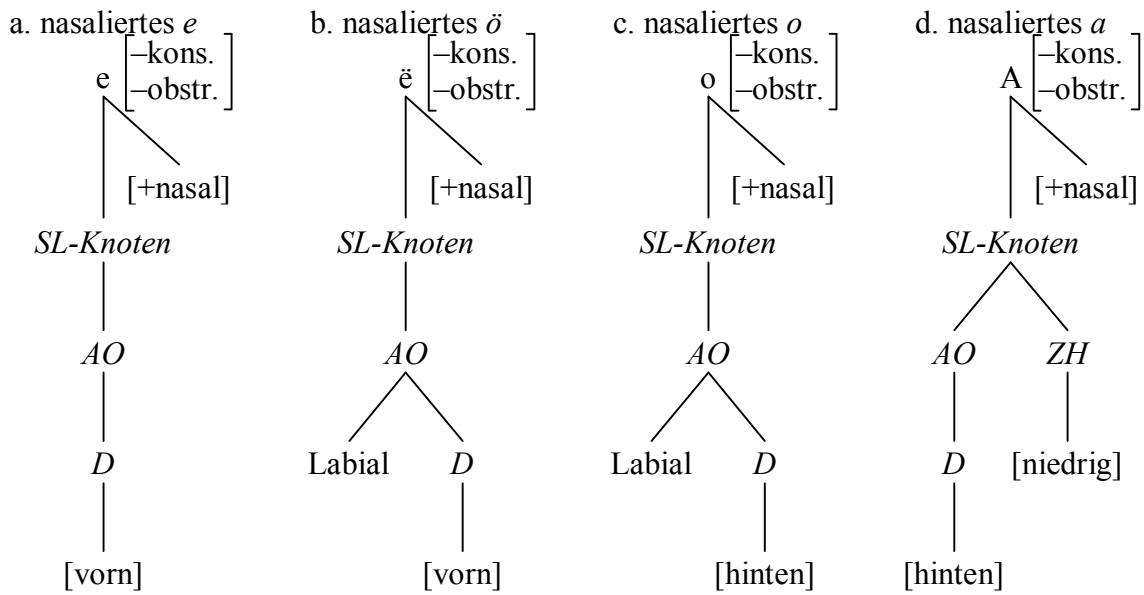
(112) Repräsentation der Silbenschnitte



Für alle diese 8 Vollvokale mit einem prosodischen Kontrast ist eine orale Bildungsweise charakteristisch. Im Deutschen können jedoch auch vier Nasalvokale vorkommen, die alle sanften Schnitt zeigen: Sie besetzen zugrunde liegend sowohl die Sonoritäts- als auch die Halteposition und werden unter Akzent länger realisiert. Ihre einzige aus der phonetischen Realisierung gesehene atypische artikulatorische Eigenschaft, ihre Ungespanntheit kann als Folge der auf die Vokalartikulation reduzierend auswirkenden Nasalität (Senkung des Velums) betrachtet werden. Zur Integration der vier Nasalvokale ins deutsche Vokalsystem bietet sich das zur Klassifizierung der Sonoranten verwendete Merkmal  $[\pm\text{nasal}]$  an. Dabei wären Segmente  $[\text{+nasal}]$ , bei deren Bildung infolge des gesenkten Velums der Zugang zum Nasenraum frei wird, sodass bei oraler Öffnung ein Teil, bei oralem Verschluss der gesamte Luftstrom durch den Nasenraum entweichen kann; alle anderen Segmente wären als  $[\text{-nasal}]$  spezifiziert.<sup>72</sup> Somit würden die Nasalvokale des Deutschen die Struktur in (113) erhalten. Jedoch ist das Vorkommen der nasalierten Vokale im Gegenwartsdeutschen äußerst eingeschränkt. Von dem Großteil der Mitglieder der deutschen Sprachgemeinschaft werden sie durch die Folge eines entsprechenden Oralvokals und eines Nasalkonsonanten ersetzt, sodass sie nur noch von einer kleinen v.a. sozial und vom Bildungsgrad her bestimmbar Minderheit der deutschen Sprecher und von ihnen auch nur in einer nicht gerade bedeutenden Anzahl von Bildungswörtern verwendet werden.

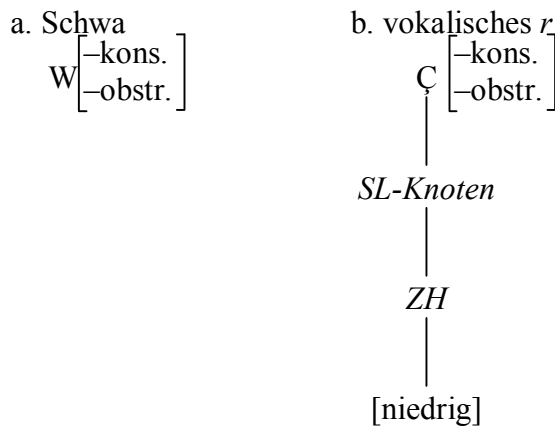
<sup>72</sup> Die Zweiwertigkeit dieses Merkmals etwa im Gegensatz zur Einwertigkeit von  $[\text{vorn}]$ ,  $[\text{hinten}]$  usw. lässt sich damit begründen, dass bei diesem auch der negative Wert eine phonologisch relevante natürliche Klasse bezeichnet.

(113) Interne Struktur der Nasalvokale



Von den neben den Vollvokalen existierenden beiden Reduktionsvokalen des Gegenwartsdeutschen weist das Schwa eine minimale interne Struktur auf: es ist nur für die beiden Oberklassenmerkmale spezifiziert. Das vokalische *r* weicht von ihm nur unwesentlich ab, es hat nämlich außerdem noch das Merkmal [niedrig], das für seinen *a*-Klang verantwortlich ist:

(114) Interne Struktur der Reduktionsvokale des Gegenwartsdeutschen

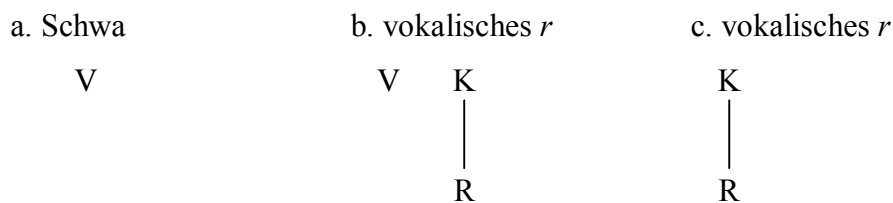


Das Schwa kommt in gewissen Wörtern zugrunde liegend in Form einer leeren skelettalen V-Position vor (vgl. (115)a), in anderen wird diese V-Position im Laufe der lexikalischen Derivation zu prosodischen oder prosodisch-morphologischen Zwecken eingefügt. Die leere V-Position erhält am Ende der lexikalischen Derivation die obige minimale interne Struktur. Das vokalische *r* kommt in gewissen Wörtern zugrunde liegend als Verbindung einer leeren skelettalen V-Position mit einem nachfolgenden konsonantischen *r* vor (vgl. (115)b), in den



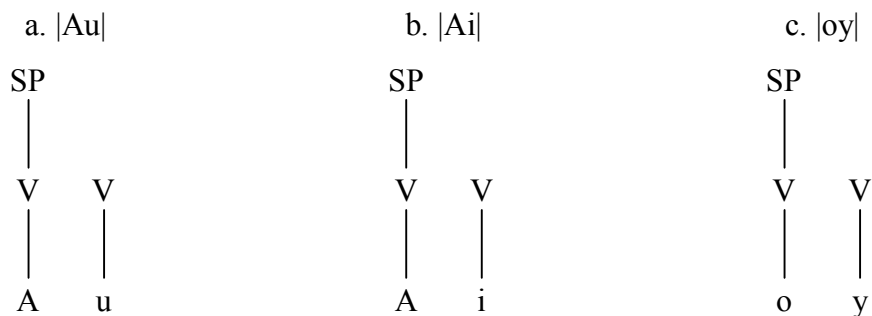
anderen erscheint in der zugrunde liegenden Repräsentation lediglich der Konsonant |R| (vgl. (115)c). Im ersteren Fall wird die leere skeletale V-Position am Ende der lexikalischen Derivation als Schwa interpretiert, die dadurch entstandene tautosilbische Sequenz wird dann in der postlexikalischen Komponente vokalisiert. Im letzteren Fall wird im Lexikon vor dem *r*-Konsonanten eine leere V-Position eingefügt, wenn |R| in die vorausgehende Silbe nicht integriert werden kann, die V-Position wird mit dem Schwa assoziiert, und die zur gleichen Silbe gehörende Sequenz /WR/ wird zum vokalischen *r*. Ansonsten gelangt der *r*-Konsonant in die postlexikalische Komponente, wo er im Endrand vokalisiert wird.

(115) Zugrunde liegende Repräsentation der Reduktionsvokale des Gegenwartsdeutschen



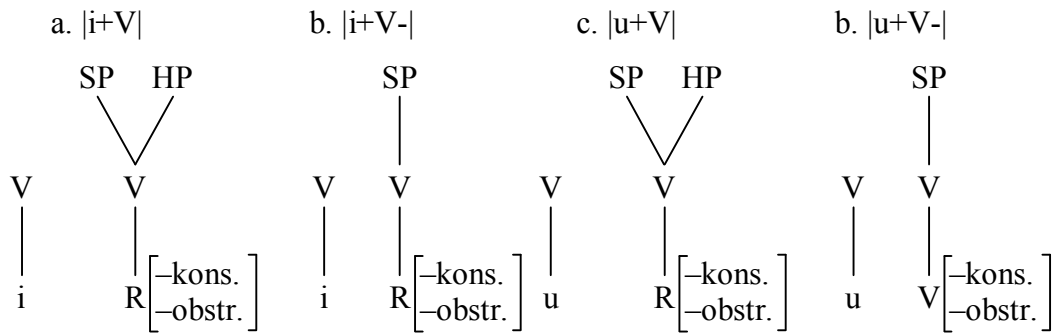
In allen drei Diphthongtypen des Deutschen ist der Kopfvokal schon zugrunde liegend für den Silbenschnitt im Sinne von (112) spezifiziert, die Randvokale erhalten dagegen keine zugrunde liegende Silbenschnittmarkierung. Im Falle der schließenden Diphthonge besetzt der Kopfvokal stets die Sonoritätsposition, während der Randvokal nur auf der skeletalen Ebene repräsentiert ist (vgl. (116)) und im Laufe der lexikalischen Derivation mit der Halteposition des fraglichen Silbenkerns assoziiert wird.

(116) Zugrunde liegende Repräsentation der der schließenden Diphthonge des Deutschen



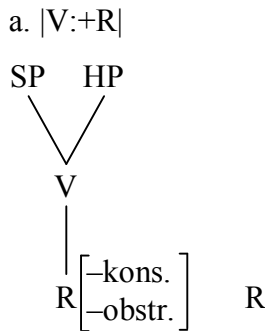
Bei den steigenden Diphthongen ist der Kopfvokal nur mit der Sonoritätsposition oder sowohl mit der Sonoritäts- als auch mit der Halteposition verbunden (vgl. (117)). Der Randvokal wird bei der Silbifizierung mit dem Silbenanfangsrand assoziiert.

(117) Zugrunde liegende Repräsentation der steigenden Diphthongen im Deutschen



Die öffnenden Diphthonge des Deutschen werden zugrunde liegend als die Verbindung eines sanft geschnittenen Vokals und eines nachfolgenden *r*-Konsonanten (vgl. (118)) dargestellt. Durch die Vokalisierung dieses konsonantischen *r* entsteht in der postlexikalischen Komponente der öffnende (phonetische) Diphthong.

(118) Zugrunde liegende Repräsentation der öffnenden Diphthonge des Deutschen



## 5.2 Exkurs: Überlegungen zur Silbenstruktur im Deutschen

In diesem kurzen exkursmäßigen Abschnitt wollen wir zwei Themen ansprechen, die zwar mit dem eigentlichen Anliegen der vorliegenden Untersuchung nicht unmittelbar zusammenhängen, sich jedoch für die Silbenschnittphonologie als relevant erweisen. Das erste davon wurde im Laufe der Arbeit schon mehrmals erwähnt und betrifft die Silbifizierung im Deutschen. Das andere, die Auslautverhärtung, wird als Paradebeispiel für einen silbenabhängigen phonologischen Prozess betrachtet, weswegen er in den einzelnen Arbeiten zum Silbenschnitt oft als eine Art Probe des Ansatzes verwendet wird.

## 5.2.1 Silbenstruktur und Silbifizierung im Deutschen

Im Folgenden soll die Frage geklärt werden, wie die Struktur der Voll- und Reduktionssilbe im Deutschen mit Regeln hergeleitet wird. Aufgrund der bisherigen Erörterungen bezüglich des Silbenschnitts und der Reduktionssilbe lässt sich für das Gegenwartsdeutsche folgender Silbifizierungsalgorithmus formulieren:

(119) Silbifizierungsalgorithmus für das Deutsche

### a. Silbenkernregeln (SKR)

SKR 1 Bilde einen Vollsilbenkern oberhalb eines für den Silbenschnitt spezifizierten Vokals. Verbinde dabei die Sonoritätsposition und die Halteposition mit dem Silbenkern.

SKR 2 Ist der Vollvokal zugrunde liegend nur mit der Sonoritätsposition verbunden, bilde eine Halteposition unterhalb des Silbenkerns und verbinde sie mit dem postvokalischen Segment.

SKR 3 Bilde einen Reduktionssilbenkern mit einer Sonoritätsposition oberhalb eines nicht prävokalischen für den Silbenschnitt nicht spezifizierten Vokals. Verbinde dabei die V-Position mit der Sonoritätsposition.

### b. Anfangsrandregeln (ARR)

ARR 1 Bilde einen Anfangsrand (AR) vor dem Silbenkern.

ARR 2 Verbinde unter Berücksichtigung der Wohlgeformtheitsbedingungen alle skeletalen Positionen mit dem Anfangsrand.

### c. Endrandregeln (ERR)

ERR 1 Bilde einen Endrand (ER) vor dem Silbenkern.

ERR 2 Verbinde unter Berücksichtigung der Wohlgeformtheitsbedingungen alle skeletalen Positionen unmittelbar mit dem Endrand.

### d. Silbenknotenprojektion (SKP)

Verbinde Anfangsrand (AR), Silbenkern (SK) und Endrand (ER) mit dem Silbenknoten  $\sigma$ .

Aus (119) geht hervor, dass in einem ersten Schritt der Silbenkern gebildet wird. Bei der Silbenkernbildung werden die Vollsilben vorgezogen: zuerst wird der sanfte und erst dann der scharfe Schnitt mit dem Silbenkern verbunden. Die dritte Silbenkernregel dient zur Bildung des Kerns einer Reduktionssilbe. Besonders wichtig ist dabei die Bedingung *nicht prävokalisch*, mit der vermieden werden soll, dass aus den Randvokalen der steigenden

Diphthonge eigene Reduktionssilben gebildet werden (die Randvokale der schließenden Diphthonge werden schon via SKR 2 mit der Halteposition eines Vollsilbenkerns assoziiert). Auf die Silbenkernbildung folgt zuerst die Maximierung des Silbenanfanrandes und schließlich wird der Rest der skeletalen Positionen mit dem Endrand verbunden.

## 5.2.2 Silbenstruktur und Auslautverhärtung

Die Auslautverhärtung als eines der charakteristischsten Silbenphänomene des Deutschen betrifft die Stimmhaftigkeit der Obstruenten. Wie die Beispiele unter (120) zeigen, werden die Obstruenten, die im Silbeninlaut stimmhaft (vgl. die Spalte b.) vorkommen, im Silbenauslaut (vgl. die Spalten a. und c.) stimmlos realisiert. Dass dabei die stimmhaften Segmente zugrunde liegend anzunehmen sind, zeigen Daten, in denen im Silbeninlaut stimmlose Obstruenten erscheinen (vgl. etwa *Typen*, *bunte*, *Werke*, *Schafe*, *heiße* usw.).

(120) a.	b.	c.
<i>Stab</i> [p]	<i>Stäbe</i> [b]	<i>buchstäb.lich</i> [p]
<i>Grund</i> [t]	<i>Gründe</i> [d]	<i>gründ.lich</i> [t]
<i>Tag</i> [k]	<i>Tage</i> [g]	<i>täg.lich</i> [k]
<i>brav</i> [f]	<i>brave</i> [v]	<i>Brav.heit</i> [f]
<i>Kreis</i> [s]	<i>Kreise</i> [z]	<i>Kreis.lauf</i> [s]

In dem hier vertretenen Silbenschnittansatz lässt sich eine Regel formulieren, in der die Entstimmlichung eines zugrunde liegend stimmhaften Obstruenten nicht von einer Silbenkonstituente (Auslaut, d.h. Silbenendrand) gesteuert wird, sondern vielmehr vom rechten Silbenrand. Das wird deutlich, wenn man auch die wenigen Wörter betrachtet, in denen die fragliche Veränderung auch in der Halteposition stattfindet (vgl. etwa *weg*, *Job* usw.). Die einschlägige Regel findet sich unter (121).

(121) Auslautverhärtungsregel (AVR)

[+obstruent] → [-stimmhaft] / \_\_\_<sub>σ</sub>

### 5.3 Ausblick: weiterführende Fragen

In der vorliegenden Untersuchung haben wir uns zum Ziel gesetzt, durch die Einbettung des Konzepts des Silbenschnitts und der Reduktionssilbe einen theoretischen Rahmen zu schaffen, die auch eine Beschreibung der bisher eher vernachlässigten Aspekte der Vokale des Gegenwartsdeutschen ermöglicht. Trotz der Absicht der Vollständigkeit konnten wir im Rahmen dieser Untersuchung rein aus Platzgründen jedoch nicht auf alle Probleme dieses Gegenstandsbereichs detailliert eingehen. Die hier nicht behandelten Themen sind phonetischer, kontrastiv-typologischer und phonologischer Natur. Bei den phonetischen Problemen sind an erster Stelle die Mangelhaftigkeiten der als experimentalphonetischer Grundlage dieser phonologischen Analyse dienenden Untersuchung Spiekermanns (2000) zu nennen. Da der Autor nur betonte Vokale in trochäischen Formen untersuchte, wäre unbedingt notwendig, den Silbenschnitt auch auf die restlichen Vorkommen der deutschen Vokale auszudehnen. Was den kontrastiv-typologischen Aspekt betrifft, so sollte man u.a. untersuchen, inwiefern sich die nicht-Silbenschnittsprachen nach der hier vertretenen Auffassung beschreiben lassen. Besonders interessant und wichtig erscheint dabei die Repräsentation der Diphthonge in diesen Sprachen. Und schließlich konnten hier naturgemäß nicht alle silbenabhängigen phonologischen Prozesse thematisiert werden, von denen das Wichtigste aller Wahrscheinlichkeit nach die Akzentzuweisung betrifft. Ferner hat sich bei der Konzipierung dieser Arbeit an zahlreichen Stellen gezeigt, dass die gewählte regelbasierte Rahmentheorie unerwünschte Nebeneffekte mit sich bringt. Sowohl bei den Voll- als auch bei den Reduktionsvokalen kamen viele Phänomene vor, für die ein *constraint*-basierter Ansatz, etwa die Optimalitätstheorie eine angemessenere Beschreibung bieten könnte. Alle diese Probleme bedürfen natürlich weiterer Untersuchungen.

## Literatur

- Alfonso, P. J. & Baer, T. (1982): Dynamics of vowel articulation, *Language and Speech* 25. 151-173.
- Becker, T. (1996a): Die Aufhebung des Vokallängengegensatzes in unbetonten Silben der deutschen Standardsprache. *Deutsche Sprache* 24,3. 268-281.
- Becker, T. (1996b): Zur Repräsentation der Vokallänge in der deutschen Standardsprache. *Zeitschrift für Sprachwissenschaft* 15,1. 3-21.
- Becker, T. (1998): *Das Vokalsystem der deutschen Standardsprache*. Frankfurt/Main: Lang.
- Becker, T. (2002): Silbenschnitt und Silbenstruktur in der deutschen Standardsprache der Gegenwart. In: Auer, P., Gilles, P. & Spiekermann, H. (Hrsg.): *Silbenschnitt und Tonakzente*. Tübingen: Niemeyer, 87–102.
- Bohn, O.-Sch., Flege, J. E., Dagenais, P. A. & Fletcher, S. G. (1992): Differenzierung und Variabilität der Zungenpositionen bei der Artikulation deutscher Vokale. In: Hess, W. & Sendlmeier, W. F. (Hrsg.): *Beiträge zur angewandten und experimentellen Phonetik* (= *Zeitschrift für Dialektologie und Linguistik*. Beiheft 72). 1-26.
- Bolla, K. & Valaczkai, L. (1986): *Német beszédhangok atlasza*. [Atlas deutscher Sprachlaute] Budapest: MTA Nyelvtudományi Intézet.
- Chomsky, N. A. & Halle, M. (1968). *The Sound Pattern of English*. New York & Evanston & London: Harper & Row.
- Clements, G. N. (1985): The Geometry of Phonological Features. *Phonology Yearbook* 2, 225–252.
- Clements, G. N. (1987): Phonological feature representation and the description of intrusive stops. *Papers from the Twenty-third Meeting, Chicago Linguistic Society, vol. 2*, 29–50. Chicago: Chicago Linguistic Society, University of Chicago.
- Clements, G. N. & Hume, E. V. (1995): The Internal Organisation of Speech Sounds. In: Goldsmith, J. (ed.): *The Handbook of Phonological Theory*. Cambridge: Blackwell. 245–306.
- Clements, G. N. & Keyser, S. Jay (1983): *CV Phonology. A Generative Theory of the*

- Syllable*. Cambridge/London: The MIT Press.
- Delattre, P. & Hohenberg, M. (1968): Duration as a cue to the tense/lax distinction in German unstressed vowels. *International Review of Applied Linguistics in Language Teaching* 6. 367-390. Zitiert nach dem Wiederabdruck in: Delattre, P. (ed.): *Studies in comparative phonetics. English, German, Spanish and French*. Heidelberg: Groos. 39–62.
- Essen, O. von (1962): Trubetzkoy's „fester“ und „loser Anschluß“ in experimentalphonetischer Sicht. *Proceedings of the 4th International Congress of Phonetic Sciences*. The Hague: Mouton. 590-595.
- Féry, C. (1991): German schwa in prosodic morphology. *Zeitschrift für Sprachwissenschaft* 10,1. 65-85.
- Féry, C. (1997): The Trochaic Ideal. *Sprachtheorie und germanistische Linguistik* 5, 7–16.
- Fischer-Jørgensen, E. (1941): Los og fast tilslutning. *Nordisk Tidsskrift for Tale og Stemme* 5. Jahrgang 3. 41-69.
- Fischer-Jørgensen, E. (1969): Untersuchungen zum festen und losen Anschluß. In: Hyldgaard-Jensen, K. & Steffensen, S. (Hrsg.): *Kopenhager Germanistische Studien* 1. Kopenhagen: Akademisk Forlag. 138-164.
- Flege, J. E. (1988): Using visual information to train foreign-language vowel production. *Language Learning* 38, 365–407.
- Fleischer, W. & Barz, I. (1992 ): *Wortbildung in der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Fliflet, A. L. (1962a): Gespannte und ungespannte Vokale. *Studia Linguistica* 16, 24–28.
- Fliflet, A. L. (1962b): Einige Bemerkungen über Anschluß und Silbe. *Proceedings of the 4th International Congress of Phonetic Sciences* 1961. The Hague: Mouton. 610-615.
- Giegerich, H. J. (1985): *Metrical phonology and phonological structure. German English*. Cambridge: Cambridge University Press.
- Giegerich, H. J. (1987): Zur Schwa-Epenthese im Standarddeutschen. *Linguistische Berichte* 112. 449-469.
- Hall, T. A. (1992): *Syllable Structure and Syllable-Related Processes in German*. Tübingen: Niemeyer.
- Heike, G. (1972): Quantitative und qualitative Differenzen von /a(:)/-Realisationen im Deutschen. *Proceedings of the 7th International Congress of Phonetic Sciences* 1971, The

- schen. *Proceedings of the 7th International Congress of Phonetic Sciences 1971*, The Hague/Paris: Mouton, 725–729.
- Helbig, G. & Buscha, J. (1989): *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Leipzig: Enzyklopädie.
- Hinderling, R. (1982): Das nhd. Phonem /ä:/ in synchroner Sicht und die Problematik der Aussprachenormierung. *German Life and Letters* 35, 287–295.
- Ickelsamer, Valentin (1534<sup>3</sup>): *Teutsche Grammatica: daraus einer von im selbs mag lesen lernen, mit allem dem, so zum Teutschen lesen u. desselben Orthographiam mangel u. uberflus, auch anderem vil mehr zu wyssen gehört[...]*. Nürnberg.
- IPA (1949): *The Principles of the International Phonetic Association*. London.
- Issatschenko, A. (1974): Das 'Schwa mobile' und 'Schwa constans' im Deutschen. In: Engel, U. & Grebe, P. (Hrsg.): *Sprachsystem und Sprachgebrauch. Festschrift für Hugo Moser zum 65. Geburtstag*. Düsseldorf: Schwann, 142–171.
- Jakobson, R. & Halle, M. (1956): *Fundamentals of language*. The Hague: Mouton.
- Jakobson, R., Fant, G. & Halle, M. (1951; 1963<sup>3</sup>): *Preliminaries to Speech Analysis*. Cambridge, Mass: MIT Press.
- Jespersen, O. (1932): *Lehrbuch der Phonetik*. Leipzig & Berlin: Teubner.
- Jessen, M.; Marasek, K.; Schneider, K. & Clahßen, K. (1995): Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. *Proceedings of the 13th International Congress of Phonetic Sciences*. 428–431 (auch in: *Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (AIMS) der Universität Stuttgart 2*, Heft 2, 141–146).
- Jessen, M. (1998): *Phonetics and Phonology of Tense and Lax Obstruents in German*. Amsterdam: Benjamins.
- Jones, D. (1917): *An English Pronouncing Dictionary*. London.
- Jørgensen, H. P. (1969a): Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur ihrer Formantfrequenzen. *Phonetica* 19, 217–245.
- Jørgensen, H. P. (1969b): Über den Intensitätsverlauf beim sogenannten losen und festen Anschluß im Deutschen. In: Hyldgaard-Jensen, K. & Steffensen, S. (Hrsg.): *Kopenhager*



- Germanistische Studien* 1. Kopenhagen: Akademisk Forlag. 165-186.
- Kiparsky, P. (1982): Lexical Morphology and Phonology. In: Yang, I.-S. (ed): *Linguistics in the Morning Calm*. Seoul: Hanshin, 3–91.
- Kiparsky, P. (1985): Some Consequences of Lexical Phonology. *Phonology Yearbook* 2, 83–138.
- Kloeke, WUS van Lessen (1982): *Deutsche Phonologie und Morphologie. Merkmale und Markiertheit*. Tübingen: Niemeyer.
- Kohler, K (1995): *Einführung in die Phonetik*. Berlin: Schmidt.
- König, W. (1989): *Atlas zur Aussprache des Schriftdeutschen in der Bundesrepublik Deutschland*. München: Huber.
- Krech, E.-M. (1982): *Großes Wörterbuch der deutschen Aussprache*. Leipzig: Bibliographisches Institut.
- Lehiste, I. (1970): *Suprasegmentals*. Cambridge, Mass. & London: MIT Press.
- Lenerz, J. (2000): Zur sogenannten Vokalopposition im Deutschen. *Zeitschrift für Sprachwissenschaft* 19.2, 167-209.
- Lenerz, J. (2002): Silbenstruktur und Silbenschnitt. In: Auer, P., Gilles, P. & Spiekermann, H. (Hrsg.): *Silbenschnitt und Tonakzente*. Tübingen: Niemeyer, 67–86.
- Lindau, M. (1978): Vowel features. *Language* 54. 541-563.
- Maack, A. (1949): Der Einfluß der Betonung auf die Lautdauer deutscher Sonanten. *Zeitschrift für Phonetik* 3, 341–356.
- Maack, A. (1951): Die Variation deutscher Sonanten. *Zeitschrift für Phonetik* 5, 287–340.
- Maas, U. (1999): *Phonologie. Einführung in die funktionale Phonetik des Deutschen*. Opladen/Wiesbaden: Westdeutscher Verlag.
- Maas, U. & Tophinke, D. (1993): *Looser und fester Anschluß*. Versuch der Neubewertung einer abgelegten Kategorie. In: Schmidt-Radefeldt, J. & Harder, A. (Hrsg.): *Sprachwandel und Sprachgeschichte. Festschrift für Helmut Lüdtke zum 65. Geburtstag*. Tübingen: Narr. 133-151.
- MacKay, I. R.A. (1977): Tenseless in Vowels: an Ultrasonic Study. *Phonetica* 34. 325-351.
- MacNeilage, P. F & Sholes, G. N. (1964): An Electromyographic Study of the Tongue during

- Vowel Production. *Journal of Speech and Hearing Research* 7. 209-232.
- Mangold, M. (1990): *DUDEN 6. Das Aussprachewörterbuch*. Mannheim etc.: Dudenverlag.
- Mater, E. (1970): *Rückläufiges Wörterbuch der deutschen Gegenwartssprache*. Leipzig: Enzyklopädie Verlag.
- Mayerthaler, W. (1974): *Einführung in die generative Phonologie*. Tübingen: Niemeyer.
- McCarthy, J. C. (1988): Feature Geometry and Dependency: A Review. *Phonetica* 43, 84–108.
- Meinhold, G. (1973): *Deutsche Standardsprache. Lautschwächungen und Formstufen*. Jena: Friedrich-Schiller-Universität Jena.
- Meinhold, G. & Stock, E. (1982): *Phonologie der deutschen Gegenwartssprache*. Leipzig: VEB Bibliographisches Institut.
- Meyer, E. A. (1913): Das Problem der Vokalsprannung. *Die neueren Sprachen* 21. 65-86 und 145-171.
- Mohanan, K. P. (1982): *Lexical Phonology*. Ph.D. Dissertation, MIT.
- Moulton, W. G. (1956): Syllabic nuclei and final consonant clusters in German. In: Halle, M., Lunt, H.G. & McLean, H. (eds.): *For Roman Jakobson*. The Hague: Mouton, 372–381.
- Moulton, W. G. (1961): Zur Geschichte des deutschen Vokalsystems. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 83, 1–35.
- Moulton, W. G. (1962): *The Sounds of English and German*. Chicago: The University of Chicago Press.
- Narahara, Y. & Shimoda, H. (1977): Akustisch-phonetische Studie über die deutschen Vokale, in: Okamoto, T. (Hrsg.): *Beiträge zur Phonetik des Deutschen*. Hamburg: Buske, 1-37.
- Neppert, J. & Pétursson, M. (1996): *Elementarbuch der Phonetik*. Hamburg: Buske.
- Nespor, M. & Vogel, I. (1986): *Prosodic Phonology*. Dordrecht: Foris.
- Noske, R. (1993): *A Theory of Syllabification and Segmental Alternation. With Studies on the Phonology of French, German, Tonkawa and Yawelmani*. Tübingen: Niemeyer.
- Pilch, H. (1966): Das Lautsystem der hochdeutschen Umgangssprache. *Zeitschrift für Mundartforschung* 33, 247–266.
- Pompino-Marschall, B. (1995): *Einführung in die Phonetik*. Berlin: de Gruyter.

- Ramers, K. H. (1988): *Vokalquantität und -qualität im Deutschen*. Tübingen: Niemeyer.
- Ramers, K.-H. & Vater, H. (1992): *Einführung in die Phonologie*. Köln: Gabel.
- Raphael, L. J. & Bell-Berti, F. (1975): Tongue Musculatur and the Feature of Tension in English Vowels. *Phonetica* 32. 61-73.
- Rausch, A. (1972): Untersuchungen zur Vokalartikulation im Deutschen, in: Kelz, H. & Rausch, A. (Hrsg.): *Beiträge zur Phonetik des Deutschen*. Hamburg: Buske, 35–82.
- Reis, M. (1974): *Lauttheorie und Lautgeschichte*. München: Finke.
- Restle, D. (1998): *Silbenschnitt - Quantität - Kopplung. Zur Geschichte, Charakterisierung und Repräsentation der Anschlußprosodie unter dem Blickwinkel einer Oszillationssibentheorie*. Diss. Universität München.
- Sagey, E. (1986): *The Representation of Features and Relations in Nonlinear Phonology*. Ph.D. Dissertation, MIT.
- Sanders, W. (1972): Hochdeutsch /ä/ – „Ghostphonem“ oder Sprachphänomen? *Zeitschrift für Dialektologie und Linguistik* 39, 37–58.
- Schuhmacher, W.W. (1972): Beitrag zur Bestimmung des physiologischen Korrelates des deutschen Vokalgegensatzes. *Linguistics* 90. 35-78.
- Sendlmeier, W. F. 1981: Der Einfluß von Qualität und Quantität auf die Perzeption betonter Vokale des Deutschen, *Phonetica* 38, 291-308.
- Sievers, E. (1901): *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre indogermanischer Sprachen*. Leipzig: Breitkopf & Härtel.
- Smith, T. & Hirano, M. (1968): Experimental investigations of the muscular control of the tongue in speech. *UCLA Working Papers in Phonetics* 10. 145-156.
- Spiekermann, H. (2000): *Silbenschnitt in deutschen Dialekten*. Tübingen: Niemeyer.
- Strauss, S. L. (1982): *Lexical phonology of English and German*. Dordrecht: Foris.
- Szulc, A. (1966): The phonemic status of NHG [E:]. *Kwartalnik Neofilologiczny* 13, 425–429.
- Techmer, F. (1884): Naturwissenschaftliche Analyse und Synthese der hörbaren Sprache. *Internationale Zeitschrift für allgemeine Sprachwissenschaft* 1. 69-170.
- Törkenczy, M. (1994): A szótag. [Die Silbe] In: Kiefer, F (Hrsg.): *Strukturális magyar nyelvtan 2. Fonológia. [Strukturelle Grammatik des Ungarischen 2. Phonologie]*. Budapest: Akadémiai Kiadó, 279–392.

Akadémiai Kiadó, 279–392.

Trubetzkoy, N. S. (1939): *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.

Valaczkai, L. (1981): *Lauttafeln zur kontrastiven Phonetik Ungarisch-Deutsch*. Szeged: JATE.

Vennemann, T. (1991a): Syllable structure and syllable cut prosodies in modern Standard German. In: Bertinetto, P. M.; Kenstowicz, M. & Loporcaro, M. (eds.): *Certamen Phonologicum II: Papers from the 1990 Cortona Phonology Meeting*. 211-243.

Vennemann, T. (1991b): Skizze der deutschen Wortprosodie. *Zeitschrift für Sprachwissenschaft* 10,1. 86-111.

Vennemann, T. (1992): *From quantity to syllable cut: on so-called lengthening in the Germanic languages*. Ms. Universität München.

Vennemann, T. (1994): Universelle Nuklearphonologie mit epiphänomenaler Silbenstruktur. In: Ramers, K. H., Vater, H. & Woder, H. (Hrsg.): *Universale phonologische Strukturen und Prozesse*, Tübingen: Niemeyer. 7-54.

Wängler, H.-H. (1960): *Grundriß einer Phonetik des Deutschen mit einer allgemeinen Einführung in die Phonetik*. Marburg: Elwert.

Wängler, H.-H. (1964): *Atlas deutscher Sprachlaute*. Berlin: Akademie Verlag.

Wiese, R. (1986): Schwa and the Structure of Words in German. *Linguistics* 24, 695–724.

Wiese, R. (1988): *Silbische und lexikalische Phonologie. Studien zum Chinesischen und Deutschen*. Tübingen: Niemeyer.

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

Wodarz, H.-W. & Wodarz-Magdics, K. (1971): Beiträge zu einer kontrastiven Phonetik des Deutschen und Ungarischen. *Phonetica* 24, 116–124.

Wurzel, W. U. (1970): *Studien zur deutschen Lautstruktur*. Berlin: Akademie Verlag.

Wurzel, W. U. (1980): Der deutsche Wortakzent: Fakten – Regeln – Prinzipien. *Zeitschrift für Germanistik* 1. 299-318.

Wurzel, W. U. (1981): Phonologie: Segmentale Struktur. In: Heidolph, K. E.; Flämig, W. & Motsch, W. (Hrsg.): *Grundzüge einer deutschen Grammatik*. Berlin: Akademie Verlag. 898-993.