

Silbe und Sonorität
in Sprache und Gehirn

Inaugural-Dissertation
zur
Erlangung des Grades eines Doktors der Philosophie
in der
Fakultät für Philologie
der
Ruhr-Universität Bochum

vorgelegt von
Judith Meinschaefer
aus Bochum

Gedruckt mit Genehmigung der Fakultät für Philologie der Ruhr-Universität Bochum.

Referent: Prof. Dr. Helmut Schnelle
Korreferent: Prof. Dr. Onur Güntürkün
Tat der mündlichen Prüfung: 19.11.1998

Danksagung

Diese Dissertation wurde durch ein Promotionsstipendium der Deutschen Forschungsgemeinschaft sowie durch ein Promotionsstipendium der Wilhelm-und-Günter-Esser-Stiftung gefördert. Sie hätte in dieser Form nicht entstehen können ohne das Graduiertenkolleg "Kognition, Gehirn und Neuronale Netze" an der Ruhr-Universität Bochum, in dem ich mir vorher unbekannte wissenschaftliche Gebiete und Arbeitsweisen entdeckt habe, und das mir eine herausfordernde soziale und wissenschaftliche Umgebung geboten hat.

Zu tiefem Dank verpflichtet bin ich meinen beiden Doktorvätern Helmut Schnelle und Onur Güntürkün, für die stetige Ermutigung, für eine richtungsweisende Betreuung, dafür, dass sie so viele ihrer Ideen mit mir geteilt haben, und dafür, dass sie Wissenschaft mit Begeisterung verbinden.

Technische und statistische Unterstützung habe ich von Susanne Sterbing und von Markus Hausmann bekommen. Für geduldiges Zuhören, für kritische Fragen, für stetige Diskussionsbereitschaft und für umfangreiches und kurzfristiges Korrekturlesen danke ich Ido Iurgel, Carmen Kelling, Bruce Mayo, Rainer Osswald, Christiane Stange und Susanne Stricker. Mein besonderer Dank gilt meinen Eltern, die die Entstehung dieser Arbeit in jeder Phase und in jeder Hinsicht unterstützt haben.

Inhaltsverzeichnis

0	Einleitung	1
0.1	Einführung	1
0.2	Gliederung dieser Arbeit	2
1	Fragen und Methoden zur Erforschung der neuronalen Grundlagen der menschlichen Sprachfähigkeit	5
1.1	Einleitung	5
1.2	Funktionale Lateralisierung der Sprachfähigkeit	6
1.2.1	Lateralisierte Reizdarbietung	6
1.2.2	Der Natrium-Amytal-Test	7
1.3	Funktionale Lokalisierung der Sprachfähigkeit	8
1.3.1	Elektrophysiologische und bildgebende Verfahren	8
1.3.2	Läsionsstudien	13
1.3.3	Elektrophysiologische Stimulation	15
1.4	Untersuchung der funktionalen Dynamik von neuronalen Prozessen	17
1.5	Zusammenfassung	22
2	Silbe und Sonorität in der phonologischen Theorie	25
2.1	Einleitung	25
2.2	Die Silbe als Struktureinheit der Sprache	26
2.2.1	Einleitung	26
2.2.2	Die Silbe als Konstituente in phonotaktischen Beschränkungen	27
2.2.3	Die Konstituenten der Silbe	29
2.2.4	Zusammenfassung	31
2.3	Sonoritätshierarchie und Sonoritäts-Folge-Prinzip	32
2.4	Sonorität, Silbizität und Oberklassenmerkmale	36
2.4.1	Einleitung	36
2.4.2	CHOMSKY & HALLE (1968)	38
2.4.3	SELKIRK (1984)	41
2.4.4	CLEMENTS (1990)	44
2.4.5	Zusammenfassung	49
2.5	Zusammenfassung	51
3	Phonetische Korrelate der phonologischen Sonorität	53
3.1	Einleitung	53
3.2	Akustische Korrelate von Silbe und Sonorität	54
3.2.1	Schallfülle und Sonorität bei SIEVERS und JESPERSEN	54
3.2.2	Akustische Korrelate wahrgenommener Sonoritätsmaxima	58
3.3	Artikulatorische Korrelate von Silbe und Sonorität	61
3.3.1	Sonorität bei SAUSSURE	61
3.3.2	Sonorität in einem Modell der artikulatorischen Dynamik	64
3.3.2.1	Prosodie und Artikulationsdynamik	64

3.3.2.2	Sonorität und Akzent	65
3.3.2.3	Sonorität und Silbizität	68
3.4	Sonorität, Silbenstruktur und prosodische Struktur	69
3.5	Zusammenfassung	75
4	Silbe und Sonorität in der Sprachwahrnehmung	77
4.1	Einleitung	77
4.2	Überlegungen zu einem psycholinguistischen Modell der Sprachwahrnehmung	78
4.2.1	Probleme der Sprachwahrnehmung	78
4.2.2	Lexikalische und prälexikalische Repräsentation	80
4.2.3	Verarbeitungsstrategien	82
4.2.4	Zwei Annahmen zur Sprachverarbeitung	85
4.3	Rhythmische Einheiten der Sprache als Segmentationseinheiten	86
4.3.1	Einleitung	86
4.3.2	Silbenzählende Sprachen	88
4.3.3	Akzentzählende Sprachen	93
4.3.4	Universale und sprachspezifische Segmentationsstrategien	95
4.3.5	Zusammenfassung	98
4.4	Sonorität in der Sprachwahrnehmung	99
4.4.1	Einleitung	99
4.4.2	Warum sollte die Sonoritätskontur verarbeitet werden?	100
4.4.3	Wie könnte die Sonoritätskontur verarbeitet werden?	103
4.5	Zusammenfassung	106
5	Funktionale Lateralisierung der Sprachwahrnehmung	107
5.1	Einleitung	107
5.2	Hemisphärenspezialisierung für sprachliche und nichtsprachliche kognitive Funktionen	107
5.2.1	Funktionen der linken Hemisphäre	107
5.2.2	Funktionen der rechten Hemisphäre	109
5.2.3	Sprachfunktionen der rechten Hemisphäre	110
5.3	Theorien zur Hemisphärenspezialisierung	114
5.3.1	Einleitung	114
5.3.2	Hypothese einer 'funktionalen' Lateralisierung	115
5.3.3	Die linke Hemisphäre ist spezialisiert für die Verarbeitung von akustischer Information in schneller Folge	116
5.3.4	Globale und lokale Informationsverarbeitung	117
5.3.5	Die Ortsfrequenz-Hypothese	118
5.3.6	Zusammenfassung	119
5.4	Lateralitätseffekte in der Sprachwahrnehmung	120
5.4.1	Einleitung	120
5.4.2	Lateralitätseffekte in der Verarbeitung von segmentalen Merkmalen	121
5.4.3	Lateralitätseffekte in der Verarbeitung von suprasegmentalen Merkmalen	124
5.4.3.1	Satzprosodie	125
5.4.3.2	Wortakzent	130
5.4.3.3	Lexikalischer Ton	131
5.4.3.4	Emotionale Prosodie	131
5.4.4	Zusammenfassung	132
5.5	Lateralitätseffekte in der Sprachwahrnehmung und Theorien zur Hemisphärenspezialisierung	132
5.6	Ist die Verarbeitung der Sonoritätskontur einer Äußerung lateralisiert?	135
5.7	Zusammenfassung	140

6	Lateralitätseffekte in der Verarbeitung der Sonoritätskontur einer Äußerung	141
6.1	Einleitung	141
6.2	Allgemeines zur Methode	142
6.2.1	Probanden	142
6.2.2	Verfahren	143
6.2.3	Dichotisches Hören	144
6.3	Experiment 1	145
6.3.1	Methode	145
6.3.1.1	Stimuluswörter	145
6.3.1.2	Herstellung des Stimulusmaterials	147
6.3.1.3	Versuchsaufbau	149
6.3.2	Ergebnisse	152
6.3.3	Diskussion	156
6.4	Experiment 2	163
6.4.1	Methode	163
6.4.1.1	Stimuli	163
6.4.1.2	Versuchsaufbau	166
6.4.2	Ergebnisse	168
6.4.3	Diskussion	171
6.5	Zusammenfassung	173
7	Schluss	175
7.1	Zusammenfassung	175
7.2	Ausblick	177
8	Literaturverzeichnis	181

Abbildungsverzeichnis

Abbildung 6-1. Dichotisches Hören.	145
Abbildung 6-2. Silbenstruktur der Stimuli, dargestellt am Wortpaar 5: greift – gereift.	146
Abbildung 6-3. Sonoritätskontur der Stimuluswortpaare, dargestellt am Wortpaar 5: greift – gereift.	147
Abbildung 6-4. Breitbandspektogramm, akustisches Signal und Einhüllende des Wortes ‘greift’.	148
Abbildung 6-5. Breitbandspektogramm, akustisches Signal und Einhüllende des Wortes ‘gereift’.	148
Abbildung 6-6. Versuchsaufbau in Experiment 1.	150
Abbildung 6-7. Ereignisse und Zeitverlauf in einem Durchgang in Experiment 1.	150
Abbildung 6-8. Mittelwerte der Variablen ‘Treffer’ je Ohr nach Geschlechtergruppen.	156
Abbildung 6-9. Versuchsaufbau Experiment 2.	167
Abbildung 6-10. Ereignisse und Zeitverlauf in einem Durchgang in Experiment 2	168

Abkürzungsverzeichnis

C	Konsonant
V	Vokal
σ	Silbe
O	Silbenanlaut (onset)
N	Silbennukleus (nucleus)
C	Silbenauslaut (coda)
.	Silbengrenze (.Ra.be.)
+	Morphemgrenze (Tisch+e)
'	Hauptakzent
,	Nebenakzent
*	indiziert einen ungrammatischen Ausdruck
[]	bezeichnet einen Ausdruck in phonetischer Notierung
//	bezeichnet ein Segment
[_N]	indiziert, dass ein Segment mit der Silbennukleusposition assoziiert ist
REA	Rechts-Ohr-Vorteil (<i>right-ear-advantage</i>)
LEA	Links-Ohr-Vorteil (<i>left-ear-advantage</i>)
RHD	rechtshemisphärisch geschädigte Probanden (<i>right-hemisphere damaged subjects</i>)
LHD	linkshemisphärisch geschädigte Probanden (<i>left-hemisphere damaged subjects</i>)

Dissertation Abstract

The human language faculty can be described from various perspectives: from a linguistic viewpoint, from a psychological viewpoint, and from a neurobiological viewpoint. Starting from a single aspect of the linguistic (phonological) description of language, namely the category syllable and the feature sonority, this study integrates linguistic, psychological and biological knowledge about speech perception to show that the syllable structure or the sonority contour of an utterance is processed by the right hemisphere of the brain. This hypothesis has been empirically tested using the method of dichotic listening.

It is well known that words and utterances are not just sequences of isolated segments. The segments are themselves integrated into higher phonological units, i.e. syllables, whose structure is uniform across all natural languages. Syllables are defined by a syllabic nucleus or peak, usually a vowel, and a syllabic margin on each side of the nucleus, usually one or more consonants. The regularities of the proper sequencing of segments within a syllable is informally stated in the Sonority Sequencing Principle. Phonologically speaking, the integration of segments into syllables corresponds to a cyclic modulation of sonority, or to a sonority profile.

In articulatory phonetics, the syllabic structuring of sound sequences results in a continuous modulation of the degree of vocal tract openness; in acoustic phonetics, it results in a continuous modulation of the parameters intensity and periodicity. Thus, syllabic nuclei (in the phonological sense) correspond both to intensity peaks in the signal and to the articulatory phase of maximal opening of the vocal tract.

The role of the syllable in speech perception has been addressed in various recent psycholinguistic studies. Could it be the case that a hypothetical syllable-based processing strategy might be steered by the phonological feature 'sonority' or by its acoustic correlates, and that it might exploit the speaker's knowledge about the crosslinguistically uniform structure of the syllable, i.e. about the Sonority Sequencing Principle? Psychoacoustic and psycholinguistic research has shown that the listener uses information about the number and location of intensity peaks and information about the degree

of periodicity and duration of a given segment of the signal to infer the number of syllabic peaks or syllables in a word or utterance. Thus, it may be assumed that human listeners do make use of the acoustic correlates of the phonological feature sonority to reconstruct a syllable representation based on an incoming speech signal.

Further weight could be given to this assumption if such a sonority-based perceptual strategy could be localized in the brain. It is well known that language is lateralized, with the perception of consonant phonemes lateralized to the left hemisphere, and the perception of certain aspects of suprasegmental or prosodic characteristics of an utterance lateralized to the right hemisphere. Given the intermediary nature of syllable structure, standing between phonemic structure on the one hand and prosodic structure on the other hand, it is an interesting question whether the perception of syllabic structure on the basis of the sonority contour is lateralized to the left or to the right hemisphere.

The hypothesis that the perception of the number of syllabic peaks in a word is lateralized to the right hemisphere of the brain was tested in a dichotic listening experiment, in which German word pairs differing in number of syllables, but otherwise similar in phonemic content, were dichotically presented. The experimental results for the male subjects support the hypothesis of a right-hemispheric processing of syllable structure. Furthermore, the results support the hypothesis of a greater hemispheric lateralization in male subjects as compared to female subjects.

The finding of a left-ear-advantage for the perception of words, instead of the often demonstrated right-ear-advantage, supports the hypothesis that besides the well documented left-hemispheric phonemic processing strategy, subjects also make use of a differentially lateralized syllabic processing strategy. Finally, this study shows that the phonological feature sonority is not only relevant for theories of linguistic competence, but may also play an important role in speech and language processing.

0 Einleitung

0.1 Einführung

Die Sprache ist eine Eigenschaft des Menschen als biologischem Organismus. Die Sprache ist aber auch ein komplexes System aus einer Menge von Ausdrücken und den diesen Ausdrücken zugrundeliegenden Strukturen. Schließlich hat die Sprache auch eine äußere Form, die unter akustischer, artikulatorischer oder graphischer Perspektive betrachtet werden kann. Daher bildet die menschliche Sprache den Untersuchungsgegenstand so verschiedener Disziplinen wie der theoretischen Sprachwissenschaft, der Phonetik, der Psychologie und der Neurowissenschaft. Die theoretische Sprachwissenschaft beschäftigt sich dabei eher mit den Strukturen, welche den sprachlichen Ausdrücken zugrundeliegen und die Phonetik mit der akustischen und artikulatorischen Form der Ausdrücke. Die Psycholinguistik untersucht das Sprachverhalten und die Neurowissenschaft schließlich die Vorgänge im Gehirn, die mit dem Sprachverhalten korreliert sind. Um eine umfassende Theorie der Sprache als biologischer Eigenschaft, als menschlichem Verhalten und als abstraktem Sprachsystem zu entwickeln, ist es nötig, dass die Beschreibungen der menschlichen Sprachfähigkeit, die von diesen drei Disziplinen geliefert werden, aufeinander bezogen werden und zu einem Gesamtbild der Sprache als biologisch realisierter Fähigkeit der menschlichen Spezies integriert werden. SCHNELLE (1991) beschreibt ein solches Forschungsprogramm mit den folgenden Worten:

[Es] geht [...] um die Entwicklung einer konkreten Sprachwissenschaft, die vor allem auch die Sprachakte des Menschen und die Prozesse ihrer Realisierung im Organismus systematisch mit den bisherigen Kenntnissen über die den Sprachakten zugrundeliegenden Strukturen und die in ihren Formen hervorgebrachten und vernommenen sprachlichen Äußerungen verbindet. (SCHNELLE 1991:119)

Eine solche Vorgehensweise bedeutet also, dass die psychologischen Verhaltensbeschreibungen und die neurowissenschaftlichen Beschreibungen der Prozesse im Gehirn auf die abstrakten Strukturbeschreibungen der theoretischen Sprachwissenschaft bezogen werden müssen. Im Rahmen eines solchen Forschungsprogrammes sieht sich die vorliegende Studie. In dieser Arbeit wird ein Aspekt der linguistischen Beschreibung lautsprachlicher Strukturen herausgegriffen, d.h. die phonologische Kategorie Silbe und das phonologische Merkmal Sonorität, und, nachdem der phonologische Status dieser Konzepte geklärt wurde, nacheinander auf eine phonetische Beschreibung akustischer

Signale und artikulatorischer Gesten, auf eine psycholinguistische Beschreibung des Sprachverhaltens und schließlich auf eine neuropsychologische Beschreibung der funktionalen Lateralisierung der Sprache im Gehirn bezogen. Hieraus wird die These abgeleitet, dass die Verarbeitung der phonologisch-phonetischen Sonoritätskontur einer Äußerung rechtshemisphärisch lateralisiert ist. Diese These wurde in zwei dichotischen Sprachwahrnehmungsexperimenten empirisch getestet.

0.2 Gliederung dieser Arbeit

Im ersten Kapitel dieser Arbeit wird ein Überblick über die Methoden gegeben, die zur Erforschung der Sprache im Gehirn zur Verfügung stehen. Diese Ausführungen beschreiben den größeren Zusammenhang, in den diese Arbeit einzuordnen ist. Im zweiten und dritten Kapitel wird geklärt, was unter dem phonologisch-phonetischen Merkmal Sonorität zu verstehen ist, und wie dieses Merkmal für die Zwecke dieser Untersuchung definiert werden soll. Im zweiten Kapitel wird gezeigt, dass Sonorität unter einer phonologischen Perspektive als Verbindungsglied zwischen segmentaler Struktur und Silbenstruktur konzipiert werden kann (CLEMENTS 1990). Im dritten Kapitel werden die akustischen und artikulatorischen Korrelate des phonologischen Merkmals Sonorität betrachtet. In diesem Kapitel wird insbesondere eine Konzeption des Merkmals Sonorität referiert (BECKMAN, EDWARDS & FLETCHER 1992), die Sonorität nicht nur auf die segmentale Struktur und die Silbenstruktur, sondern auch auf die Akzentstruktur bezieht.

Das vierte Kapitel dieser Arbeit beschäftigt sich mit der Kategorie Silbe und dem Merkmal Sonorität in einem Modell der Sprachwahrnehmung. In diesem Kapitel werden zunächst einige Studien zur Funktion der Silbe für die Sprachwahrnehmung zusammengefasst. Ausgehend davon wird die Annahme entwickelt, dass bei der Wahrnehmung einer Äußerung auch die Sonoritätskontur dieser Äußerung verarbeitet wird und dass diese Sonoritätsrepräsentation die Grundlage für die Rekonstruktion der Silbenstruktur bilden kann. Indem das Merkmal Sonorität in diesem Kapitel auf ein psycholinguistisches Modell des Sprachverhaltens bezogen wird, ist die Voraussetzung dafür geschaffen, die Sprachverarbeitungsfunktion 'Wahrnehmung der Sonoritätskontur einer Äußerung' mit Aktivierungsprozessen über funktionalen Einheiten im Gehirn zu korrelieren.

Eine Möglichkeit, Sprachverarbeitungsmodule mit funktionalen Einheiten des Gehirns zu korrelieren, bildet die Untersuchung der funktionalen Lateralisierung einer Verarbeitungsleistung. Im fünften und sechsten Kapitel wird daher die Frage erörtert, ob

die Verarbeitung der Sonoritätskontur einer Äußerung lateralisiert ist. Im fünften Kapitel werden Studien zur funktionalen Lateralisierung der Verarbeitung von segmentalen und suprasegmentalen Merkmalen und Kategorien referiert. Im Anschluss daran wird die Hauptthese dieser Arbeit entwickelt und begründet, dass die Verarbeitung des suprasegmentalen Merkmals Sonorität rechtshemisphärisch lateralisiert ist.

Im sechsten Kapitel werden zwei dichotische Sprachwahrnehmungsversuche beschrieben, in denen die These überprüft wurde, dass sich für die Verarbeitung der Sonoritätskontur eine rechtshemisphärische Spezialisierung nachgewiesen werden kann. Im ersten Experiment, in dem die Aufgabe des Probanden in der Identifikation deutscher Wörter bestand, konnte ein Geschlechtsunterschied in der Lateralität für die Verarbeitung der Sonoritätskontur demonstriert werden, mit einer stärkeren Lateralisierung für die männlichen Probanden und einer eher bilateralen Aktivierung für die weiblichen Probanden. Die Versuchsaufgabe des zweiten Experimentes bestand in der Detektion von dreisilbigen sinnlosen Silbenfolgen in einem dichotisch präsentierten Sprachstimulus. Hier konnte kein Ohr-Vorteil, d.h. keine funktionale Lateralisierung für die Testaufgabe, nachgewiesen werden. Diese Versuchsergebnisse werden vor dem Hintergrund der Annahme diskutiert, dass sprachliche Äußerungen auf unterschiedliche Weise und mittels unterschiedlicher Verarbeitungsstrategien wahrgenommen werden können.

Im letzten Abschnitt werden die Ergebnisse dieser Studie zusammengefasst und vor dem Hintergrund offener Fragen diskutiert.

1 Fragen und Methoden zur Erforschung der neuronalen Grundlagen der menschlichen Sprachfähigkeit

1.1 Einleitung

In diesem Kapitel werden verschiedene Methoden vorgestellt, mit denen die neuronalen Grundlagen der menschlichen Sprachfähigkeit untersucht werden können. Diese Methoden sind zur Untersuchung von jeweils verschiedenen Fragestellungen geeignet und haben jede für sich ihre Möglichkeiten und Grenzen. Im folgenden soll daher gezeigt werden, dass die Wissenschaft zu einem umfassenden Bild der ‘Sprache im Gehirn’¹ nur dann gelangen kann, wenn sie ausgehend von einem Teilaspekt des Untersuchungsgegenstandes, d.h. der menschlichen Sprachfähigkeit, diesen nacheinander unter verschiedenen Fragestellungen und mit verschiedenen Methoden betrachtet, um dann die Ergebnisse zu einem Gesamtbild zu integrieren. In der vorliegenden Studie wird eine dieser Fragestellungen herausgegriffen, d.h. die Frage nach der funktionalen Lateralisierung der Sprache und mit einer dieser Methoden, d.h. mit der Methode des dichotischen Hörens, untersucht. Der Sinn der Ausführungen in diesem Kapitel besteht darin, den allgemeinen wissenschaftlichen Zusammenhang aufzeigen, in den die vorliegende Studie einzuordnen ist. Im folgenden werden zunächst Methoden betrachtet, mit denen sich die funktionale Lateralisierung der Sprache im Gehirn erforschen lässt (1.2), sodann Methoden, die Aufschluss über die funktionale Lokalisierung der Sprache in einzelnen Arealen des Kortex geben können (1.3) und schließlich Methoden, die sich zur Untersu-

¹ Im folgenden ist wiederholt vom Untersuchungsgegenstand ‘Sprache im Gehirn’ die Rede. Diese Redeweise von der ‘Sprache im Gehirn’ möge als *Abkürzung* für die Frage danach verstanden werden, welche Prozesse im Gehirn ablaufen, wenn Menschen Sprache verstehen und hervorbringen und welche Eigenschaften des Gehirns dem Menschen das Sprechen erst ‘ermöglichen’ (vgl. LENNEBERG 1972:72).

chung der funktionalen Dynamik einzelner Neuronen oder Neuronengruppen eignen (1.4).

1.2 Funktionale Lateralisierung der Sprachfähigkeit

1.2.1 Lateralisierte Reizdarbietung

Wenn man Aspekte der menschlichen Sprachfähigkeit mit Aktivierungsprozessen über funktionalen Einheiten im Gehirn korrelieren will, so kann zunächst die Frage gestellt werden, welche Hemisphäre für eine Teilfähigkeit spezialisiert ist. In bezug auf die auditive Sprachwahrnehmung kann man diesen Aspekt mit der Methode des dichotischen Hörens untersuchen, in Bezug auf die visuelle Sprachwahrnehmung (d.h. in Bezug auf die Lesefähigkeit) mit der Methode der tachistoskopischen Präsentation von geschriebenen Wörtern.

Beim dichotischen Hören² werden dem Probanden zwei verschiedene Sprachsignale gleichzeitig präsentiert, je eines auf jedem Ohr (vgl. Abbildung 6-1). Da die Hörbahn gekreuzt ist und die Signale bei beidohriger Stimulation vom rechten Ohr über die kontralaterale Verbindung in die linke Hemisphäre und vom linken Ohr in die rechte Hemisphäre weitergeleitet werden (KIMURA 1961, MILNER, TAYLOR & SPERRY 1968), lässt sich aus einem Verarbeitungsvorteil für das eine oder das andere Ohr (d.h. die Stimuli, die dem einen Ohr präsentiert werden, werden schneller und mit weniger Fehlern verarbeitet als diejenigen, welche dem anderen Ohr präsentiert werden) auf die Spezialisierung einer Hemisphäre im Vergleich zur anderen schließen.

Bei der tachistoskopischen Darbietung³ wird der Stimulus nur in einer Gesichtsfeldhälfte, entweder in der rechten oder in der linken, gezeigt. Die Sehbahn verläuft so, dass die Information aus einer Gesichtsfeldhälfte (d.h. aus der einen Retinahälfte beider Augen) nur in die kontralaterale Hemisphäre projiziert wird, d.h. die Information aus der rechten Gesichtsfeldhälfte gelangt nur in die linke Hemisphäre, aus der linken Hälfte nur in die rechte Hemisphäre (KOLB & WHISHAW 1993:61). Werden die Reize, die in der einen Hälfte dargeboten wurden, nun schneller und mit weniger Fehlern verarbeitet als Reize, die in der anderen Hälfte dargeboten wurden, so lässt sich daraus eine Spezialisierung derjenigen Hemisphäre ableiten, die kontralateral zur bevorzugten Gesichtsfeldhälfte liegt. Wichtig sind für diese Methode jedoch sehr kurze Reizdarbietungszeiten um 100 ms, um zu verhindern, dass die Blickausrichtung geändert und der

2 Für einen Überblick vgl. HUGDAHL (Hrsg.) 1988, s.u. 5.4.2, 6.2.3.

3 Für eine Einführung vgl. SPRINGER & DEUTSCH 1995:81ff, 51ff, WALLECH, LINS & HERMANN 1997:277f.

Reiz fokussiert wird, d.h. ins Zentrum des Gesichtsfelds gelangt. Die Möglichkeit, visuelle (Sprach-)Reize lateralisiert darzubieten, ohne auf solche sehr kurzen Darbietungszeiten beschränkt zu sein, bietet die Verwendung einer 'Z-Linse' (ZAIDEL 1975),⁴ die so konstruiert ist, dass das Bild des Blickfeldes nur auf die eine Hälfte der Retina eines Auges fällt, so dass auf diese Weise die gesamte einfallende visuelle Information nur in eine Hemisphäre projiziert wird.

Sowohl mit dichotischer Präsentation von gesprochenen Wörtern als auch mit tachistoskopischer Präsentation von geschriebenen Wörtern lässt sich eine stabile Überlegenheit (d.h. schnellere Verarbeitung der Stimuli und eine geringeren Fehlerrate) der linken im Vergleich zur rechten Hemisphäre für die Verarbeitung von Wörtern oder Sätzen demonstrieren. Daneben wurden in einer Vielzahl von Experimenten Lateralitätseffekte für die Verarbeitung einzelner phonologischer und semantischer Merkmale untersucht.⁵ Zu bedenken ist, dass mit diese beiden Methoden der lateralisierten Reizdarbietung zerebrale Asymmetrien für die auditive und visuelle Sprachperzeption, aber nicht für die Sprachproduktion untersucht werden kann.

1.2.2 Der Natrium-Amytal-Test

Beim Natrium-Amytal-Test (für einen Überblick vgl. JONES-GOTMAN, ROULEAU & SNYDER (Hrsg.) 1997), der nach seinem Erfinder Juhn WADA auch Wada-Test genannt wird, wird das Barbiturat Natrium-Amytal in eine der beiden Karotisarterien injiziert, die das Gehirn mit Blut versorgen. Auf diese Weise wird eine der Hemisphären für einige Minuten anästhesiert und es kann während dieser Zeit getestet werden, zu welchen Ausfällen der Sprachfähigkeit (oder anderer Fähigkeiten) die Betäubung der rechten oder der linken Hemisphäre führt, d.h. es können Aussagen darüber gemacht werden, für welche Teilfähigkeiten der Sprachfähigkeit eine voll funktionsfähige rechte oder linke Hemisphäre notwendig ist. Der Natrium-Amytal-Test wird zur präoperativen Dia-

4 ZAIDEL, E.: „Auditory language comprehension in the right hemisphere following cerebral commissurotomy and hemispherectomy: a comparison with child language and aphasia“, in: Caramazza, A. & Zurif, E.B. (Hrsg.): *Language Acquisition and Language Breakdown*. Baltimore: John Hopkins University Press, zitiert nach Springer & Deutsch 1995:60.

5 Mit dichotischem Hören wurden insbesondere Lateralitätseffekte für die Verarbeitung segmentaler und suprasegmentaler phonologischer Merkmale untersucht (s.u. 5.4), vgl. auch TARTTER 1988. Was die Vergleichbarkeit der Ergebnisse von dichotischen und tachistoskopischen Studien angeht, so ist zu beachten, dass die Schriftsprache nur einen Bruchteil der Merkmale gesprochener Sprache notiert, nämlich nur die distinktiven oder bedeutungsunterscheidenden Eigenschaften der Sprache, aber nicht die suprasegmentalen phonologischen Merkmale wie Wortakzent und Satzintonation, deren Verarbeitung nur mit akustischen Reizen getestet werden kann. In visuellen Halbfeldstudien wurden vor allem Hemisphärenunterschiede für die semantische Verarbeitung von Wörtern untersucht (für einen Überblick vgl. CHIARELLO 1998).

agnostik verwendet, um die 'sprachdominante' Hemisphäre eindeutig zu identifizieren und mögliche Folgen der Hirnoperation für die kognitive Leistungsfähigkeit des Patienten zu kalkulieren. Die Anwendung dieser Methode ist, wie auch die Mehrzahl der im folgenden aufgeführten Methoden, auf Patienten beschränkt, die sich auf Grund einer schon bestehenden Hirnanomalie (Tumoren, Epilepsie) einer Hirnoperation unterziehen müssen.

Fähigkeiten, deren Lateralisierung mit dieser Methode im Rahmen präoperativer Diagnostik üblicherweise getestet werden, umfassen die Benennung von Gegenständen auf Bildern, Zählen, Aufzählen der Wochentage oder der Buchstaben des Alphabets usw., d.h. die Fähigkeit zur Sprachproduktion als Ganzes wird getestet. Grundsätzlich erlaubt diese Methode auch die Untersuchung der Lateralisierung von Teilkomponenten der Sprachfähigkeit, z.B. orthographische und phonologische Verarbeitungsstrategien beim Lesen. Beispielsweise konnte gezeigt werden, dass nach rechtsseitiger Injektion beim lauten Lesen von visuell präsentierten Wörtern eher orthographische Fehler (d.h. Verwechslung von Buchstaben), nach linksseitiger Injektion eher phonologische Fehler (d.h. Verwechslung von Lauten) gemacht wurden (FEDIO ET AL. 1997:107). Grenzen werden der Komplexität möglicher Testaufgaben aber durch die zur Verfügung stehende Zeit von nur wenigen Minuten gesetzt. Ferner orientiert sich die Anwendung der Methode an der klinischen Notwendigkeit, die sprachdominante Hemisphäre zu lokalisieren.

1.3 Funktionale Lokalisierung der Sprachfähigkeit

1.3.1 Elektrophysiologische und bildgebende Verfahren

Eine der wichtigsten Neuerungen der Methoden der Hirnforschung stellt die Entwicklung der elektrophysiologischen und der funktionellen bildgebenden Verfahren dar. Diese Verfahren messen elektrophysiologische (EEG und MEG) oder metabolische (PET und fMRI, SPECT) Vorgänge im Gehirn, die zeitgleich mit bestimmten Verhaltensweisen (z.B. Fixieren eines Punktes, Erkennen eines Wortes, Verstehen eines Satzes) auftreten (für eine Einführung vgl. KISCHKA, WALLESCH & WOLF 1997).

Bei einem **EEG** (Elektroenzephalogramm) wird kortikale elektrische Aktivität mithilfe von Elektroden an der Schädeloberfläche abgeleitet (vgl. WALLESCH & DEUSCHL 1997, COLES & RUGG 1995). Die durch einen Reiz (oder ein anderes Ereignis) ausgelösten Veränderungen im EEG nennt man ereigniskorrelierte Potentiale (ERPs, *event-related potentials*). Einzelne ereigniskorrelierte Potentiale sind wegen der starken Hintergrundaktivität im EEG nicht sichtbar, lassen sich aber durch die Mitte-

lung vieler Signale aus verschiedenen Testdurchgängen gegenüber der Hintergrundaktivität selektiv verstärken. In ein einzelnes von der Schädeloberfläche abgeleitetes ERP geht immer auch die Aktivität weiter entfernter Neuronengruppen ein. Daher lassen sich die Neuronengruppen, die ein ERP erzeugt haben, nur ungenau lokalisieren. Hinzu kommt das Problem, dass die Aktivität von Neuronengruppen nur dann mit Elektroden auf der Schädeloberfläche erfasst werden kann, wenn die generierenden Neuronen eine gewisse geometrische Konfiguration relativ zueinander und zur Schädeloberfläche aufweisen.

Während ein EEG die elektrischen Potentiale mißt, die von Neuronengruppen erzeugt werden, misst ein **MEG** (Magnetoenzephalogramm) die magnetischen Felder, die im Zusammenhang mit neuronaler Aktivität entstehen (vgl. WALLESCHE & DEUSCHL 1997, PAPANICOLAOU 1995). Der Vorteil von EMFs (*evoked magnetic fields*) im Vergleich zu ereigniskorrelierten Potentialen besteht darin, dass die signalerzeugenden Neuronengruppen für erstere mit größerer Exaktheit lokalisiert werden können als für letztere; auf der anderen Seite ist die Durchführung eines MEGs (bzw. die Einrichtung und Wartung der nötigen Technik) erheblich aufwendiger als die Durchführung eines EEGs.

Mit der Positronenemissionstomographie (**PET**) lassen sich regionale Veränderungen des Blutflusses im Gehirn (*regional cerebral blood flow*, rCBF) messen, die mit Veränderungen, d.h. mit der Verstärkung oder Inhibition der neuronalen Aktivität einhergehen (für eine detaillierte Darstellung vgl. POSNER & RAICHLE 1994):

Emission tomography produces an image of the distribution of a previously administered radionuclide in any desired section of the body. Positron emission tomography (PET) uses the unique properties of the annihilation radiation that is generated when positrons are absorbed in matter that is a highly faithful representation of the spatial distribution of the radionuclide at a selected plane through the tissue. [...] PET has been used in humans to measure brain blood flow, [...] metabolism of glucose and oxygen (RAICHLE 1994:335).

Auch mit **fMRI** (*functional magnetic resonance imaging*, Funktionelle Kernspintomographie) lassen sich regionale Veränderungen des Blutflusses im Gehirn messen:

By manipulating the atoms (usually of hydrogen) in a magnetic field, scientists can line them up just as needles of a compass line up in the earth's magnetic field. When radiowave pulses are applied to slices of the human brain whose atoms have been so aligned, the sample emits detectable radio signals that are characteristic of the number of atoms present. In addition, the signals are affected by the chemical environment of the atoms. (POSNER & RAICHLE 1997²:228)

Während mit PET der erhöhte regionale Blutfluss als erhöhte Zahl an vorher injizierten radioaktiven Teilchen gemessen wird, misst fMRI erhöhten regionalen Blutfluss als Veränderung der Resonanzeigenschaften körpereigener Substanzen.

Ähnlich wie PET funktioniert schließlich auch **SPECT** (*single photon emission computed tomography*), eine Methode, deren Aussagekraft in Experimenten zur Lokalisierung der neuronalen Grundlagen kognitiver Fähigkeiten derzeit noch untersucht wird:

Die SPECT verwendet radioaktiv markierte Substanzen (*Tracer*), die nach intravenöser Injektion die Blut-Hirn-Schranke leicht passieren und im Hirngewebe gespeichert werden. In Hirnarealen mit hoher neuronaler Aktivität ist der regionale Blutfluss [...] erhöht und dadurch auch die Ablagerung von Tracersubstanz. [...] [So] kann durch einmalige Gabe des Radiopharmakons ein [...] stabil bleibendes Abbild eines Durchblutungszustandes zum Injektionszeitpunkt gewonnen werden. (KISCHKA & RIHS 1997:295f)

PET, SPECT und fMRI werden ‘bildgebende’ Verfahren – *brain-imaging techniques* – genannt, weil die Ergebnisse der Messungen bildlich als Veränderungen in bestimmten Bereichen des Gehirns dargestellt werden können. Dabei muss bedacht werden, dass diese Verfahren nicht unmittelbar Bilder ‘erzeugen’, sondern dass die Ergebnisse der Messvorgänge erst nach gründlicher Aufbereitung der gemessenen Signale und ihrer anschließenden Kombination mit strukturellen anatomischen Daten als funktionale Veränderungen oder Ereignisse im Gehirn bildlich dargestellt werden können.

Ein erstes Problem dieser experimentellen Techniken besteht dabei darin, dass die Messdaten mehrerer Versuchspersonen⁶ gepoolt werden müssen, um Rauschen in den Daten zu minimieren, welches durch aufgabenunabhängige Vorgänge im Gehirn hervorgerufen wird und um die aufgabenrelevante Erhöhung des rCBFs zu verstärken.⁷ Da sich einzelne Gehirne erstens in der Anatomie unterscheiden und da zweitens anzunehmen ist, dass sich verschiedene Personen auch in der funktionalen Lokalisierung der neuronalen Grundlagen kognitiver Fähigkeiten unterscheiden,⁸ birgt dieses Verfahren der Mittelung der Daten für verschiedene Versuchspersonen stets das Risiko, dass die Areale, in denen sich aufgabenbezogene Aktivitätserhöhung findet, nicht exakt aufeinander bezogen werden,⁹ so dass die relevante Aktivitätserhöhung durch die Mittelung reduziert statt verstärkt werden kann.

Eine zweite Schwierigkeit ergibt sich daraus, dass sich zu jedem Zeitpunkt Aktivierung in allen Arealen des Gehirns zeigt, da wir immer etwas hören, sehen, denken, fühlen, so dass jede Untersuchung, die eine Erhöhung des rCBFs im Zusammenhang

6 Da erstens alle drei Verfahren im Vergleich z.B. zu Reaktionszeitexperimenten sehr zeitaufwendig sind und da zweitens im Falle von PET- und SPECT-Experimenten die Probanden einer gewissen radioaktiven Strahlenbelastung ausgesetzt sind, ist es nicht möglich, einen Probanden so oft zu testen, als dass allein die Daten einer einzigen Versuchsperson genügen würden.

7 Die Mittelung der Signale ist auch bei der Verwendung elektrophysiologischer Verfahren nötig, welche allerdings zusätzlich zur Mittelung über verschiedene Versuchspersonen eine Mittelung der Signale aus verschiedenen Testdurchgängen einer einzigen Versuchsperson nötig machen.

8 Daneben können sich Probanden auch in den Strategien unterscheiden, die sie bei der Durchführung komplexer Aufgaben anwenden. Dies ist ein Problem, das sich auch für die experimentelle Arbeit mit Reaktionszeit- oder anderen Paradigmen stellt, da Strategieunterschiede zwischen den Probanden zu erheblicher Inhomogenität der Resultate in einer Probandengruppe führen können. Dieser Aspekt der Variation der Verarbeitungsstrategien wird in den folgenden Kapitel wiederaufgegriffen und ist auch für die hier beschriebenen Experimente zur funktionalen Lateralisierung der Verarbeitung der Sonoritätskontur von großer Bedeutung, s.u. 4.3.4, 5.4.4, 6.3.3.

9 Die Messdaten für jede einzelne Versuchsperson werden auf ein Standardgehirn bezogen, danach werden die Messdaten für alle Probanden gepoolt (vgl. POSNER & RAICHLE 1994:70).

mit einer einzigen Komponente der Wahrnehmung oder Kognition erforschen will, einen definierten Referenzzustand des Gehirns braucht. Als Referenzzustand wird eine Aufgabe gewählt, die sich von der experimentellen Aufgabe möglichst nur in einer Komponente unterscheidet, nämlich in derjenigen Komponente, deren neuronale Grundlagen lokalisiert werden sollen. Das Messergebnis für die Referenzaufgabe wird dann von dem Messergebnis für die experimentelle Aufgabe subtrahiert und das Ergebnis wird als die Aktivierung interpretiert, welche durch den zusätzlichen kognitiven Aufwand hervorgerufen wird, den die experimentelle Aufgabe im Vergleich zur Referenzaufgabe erfordert (Subtraktionsmethode).

So können z.B. die Aktivierungsmuster, die für die Verarbeitung von (bedeutungslosen) Pseudo-Wörtern ermittelt wurden, subtrahiert werden von den Aktivierungsmustern, die für die Verarbeitung von (bedeutungsvollen) Wörtern gemessen wurden. Dabei wird angenommen, dass sich der zweite Prozess nur in einem Merkmal, nämlich in der zusätzlichen semantischen Verarbeitung, vom ersten Prozess unterscheidet. Das Messergebnis wird als Aktivierung der neuronalen Basis für die Komponente 'semantische Verarbeitung' interpretiert, wohingegen die Aktivierungsmuster, die sich für beide Aufgaben zeigen lassen (und die daher nach der Bildsubtraktion wegfallen), auf andere, beiden Aufgaben gemeinsame Komponenten, z.B. visuelle oder phonologische Verarbeitung, zurückgeführt werden (POSNER & RAICHLE 1994:84).

Die Schwierigkeit der Subtraktionsmethode liegt vor allem in der Auswahl von geeigneten experimentellen und Referenz-Aufgaben, denn die Ergebnisse einer Messung mit PET, SPECT oder fMRI sind stets nur so aussagekräftig wie das Modell der Sprachfähigkeit (oder eines anderen Moduls der Kognition) und ihrer Teilkomponenten, welches der Konzeption von Referenzaufgabe und experimenteller Aufgabe zugrundeliegt. So könnte man gegen die oben aufgeführte Schlussfolgerung, dass die Aktivität, welche nach Subtraktion der Aktivierungsmuster für die Verarbeitung von visuell präsentierten Pseudo-Wörtern von den Aktivierungsmustern für die Verarbeitung von Wörtern übrigbleibt, die zusätzliche semantische Verarbeitung oder den Zugriff aufs Lexikon widerspiegelt, z.B. einwenden, dass Pseudo-Wörter und Wörter auf ganz unterschiedliche Weise verarbeitet werden können – erstere erfordern nämlich möglicherweise (auch wenn sie visuell präsentiert werden) eine analytische phonologische Verarbeitungsstrategie, letztere können aber auch allein auf der Grundlage visueller Merkmale unmittelbar als Wörter identifiziert werden, werden also möglicherweise gar nicht phonologisch verarbeitet.

Das Problem, dass für die Durchführung von experimenteller und Referenz-Aufgabe möglicherweise ganz andere Strategien angewendet werden, auch wenn sich beide Aufgaben auf den ersten Blick allein in einer Komponente unterscheiden, ist lös-

bar, wenn der Konzeption des experimentellen Paradigmas eine ausreichend differenzierte linguistische und psycholinguistische Analyse vorangeht. Um zu vermeiden, dass verschiedene Probanden unterschiedliche Strategien anwenden, müssen die einzelnen Verarbeitungsschritte, die eine experimentelle Aufgabe erfordert, genau expliziert werden und die Versuchsaufgabe muss so konzipiert werden (z.B. durch die Manipulation der verwendeten Sprachreize, so dass sie sich von natürlichsprachigen Ausdrücken unterscheiden), dass die Probanden gezwungen sind, eine bestimmte Verarbeitungsstrategie anzuwenden. Andernfalls liefert ein Experiment mit einem bildgebenden Verfahren zwar Daten, aber die postulierte Beziehung dieser Messdaten zur Kognition bleibt anfechtbar. Diese Schwierigkeit der Auswahl des Paradigmas kann als eine der Ursachen dafür gelten, dass Experimente mit bildgebenden Verfahren, die dieselbe Komponente der Sprachfähigkeit lokalisieren wollen, häufig zu wenig übereinstimmenden und zum Teil inkompatiblen Ergebnissen kommen.¹⁰

Dennoch lässt sich sagen, dass die Untersuchungen mit bildgebenden Verfahren meist zu ganz ähnlichen Ergebnissen kommen wie klinische Studien:¹¹ der gesamte perisylvische Kortex der 'sprachdominanten' Hemisphäre ist an der Sprachverarbeitung beteiligt (BINDER ET AL. 1997), bei der Verarbeitung von Einzelwörtern werden posteriore Bereiche [Wernicke'sches Zentrum] mehr als anteriore Bereiche [Broca'sches Zentrum] aktiviert (ZATORRE ET AL. 1996), wobei sich aber auch Aktivität im frontalen Kortex finden lässt (BINDER ET AL. 1997, DÉMONET ET AL. 1992), die syntaktische Verarbeitung aktiviert vor allem frontale Bereiche (CAPLAN 1995) und Aufgaben, bei denen Einheiten phonologischer Repräsentationen extrahiert und manipuliert werden müssen,¹² aktivieren ebenfalls frontale Bereiche (ZATORRE ET AL. 1996).

Vergleicht man elektrophysiologische und bildgebende Verfahren, so lässt sich folgendes festhalten: Bildgebende Verfahren messen den regionalen Blutfluss im Gehirn und erlauben eine feine räumliche Auflösung (fMRI bis zu 1 mm², KISCHKA & RIHS 1997:311, SPECT um 1 cm², *ibid*:304), aber nur eine begrenzte zeitliche Auflösung im Bereich einiger Sekunden bis Minuten (*ibid*:304, 307). Elektrophysiologische Verfahren messen im Gegensatz dazu die elektrische Aktivität im Gehirn und haben zwar eine sehr feine zeitliche Auflösung im Bereich von einigen Millisekunden (KISCHKA & RIHS 1997), aber eine durch die Zahl der Messelektroden und andere Faktoren begrenzte räumliche Auflösung. EEG und MEG werden deshalb häufig ver-

10 Zu diesem Schluss kommt POEPEL (1996) in einem Vergleich einiger Studien, welche die 'phonologische Verarbeitung' mit PET lokalisiert haben.

11 Für einen vergleichenden Überblick vgl. POSNER & RAICHLE 1994:125ff.

12 Beispiel: der Proband muss entscheiden, ob sich zwei Sprachstimuli reimen.

wendet, um den Zeitverlauf der neuronalen Aktivität¹³ zu messen, die mit sehr komplexen Verhaltensleistungen korreliert ist, z.B. zur Untersuchung des Zeitverlaufs der Verarbeitung eines syntaktisch oder semantisch inkongruenten Satzes.¹⁴ PET und fMRI hingegen werden verwendet, um die neuronalen Grundlagen verschiedener Verhaltensleistungen mit hierarchischer Komplexität vergleichend im Gehirn zu lokalisieren, z.B. zur Lokalisierung von phonologischer Wortverarbeitung im Vergleich zu semantischer Wortverarbeitung (vgl. POSNER & RAICHLE 1994). Ein umfassendes Bild von den Vorgängen im Gehirn lässt sich letztlich am besten mit einer Kombination beider Arten von Verfahren gewinnen: bildgebende Verfahren können zur Lokalisierung der neuronalen Korrelate eines mentalen Vorganges bzw. zu seiner Unterteilung in verschiedene, unterschiedlich lokalisierte Teilvorgänge dienen, deren zeitliche Ordnung und deren zeitliche Struktur zugleich mit elektrophysiologischen Methoden genauer untersucht wird.

1.3.2 Läsionsstudien

Schädigungen des Gehirns durch Schlaganfälle oder Blutgerinnsel, durch mechanische Einwirkung oder durch Tumoren, oder schließlich durch degenerative Erkrankungen des Nervensystems (z.B. die Alzheimer-Krankheit) führen zu einer Veränderung des Verhaltens des Patienten, z.B. zu Gedächtnisstörungen (Amnesien), zu Bewegungsstörungen (Apraxien), zu mehr oder weniger ausgeprägten Sprachstörungen (z.B. zu Aphasien), oder zu Erkennungsstörungen (Agnosien), die so spezifisch sein können, dass etwa nur die Fähigkeit, Gesichter zu erkennen, betroffen sein kann (Prosopagnosie). Die Tatsache, dass Hirnschädigungen zu Sprachdefiziten führen, ist seit langem bekannt und wurde schon im neunzehnten Jahrhundert systematisch untersucht und beschrieben.¹⁵

Diese Sprachdefizite oder Aphasien können auf verschiedene Weise und sehr differenziert klassifiziert werden;¹⁶ hier sollen nur die Broca-Aphasie und die Wernicke-Aphasie kurz angesprochen werden. Die Broca-Aphasie tritt auf nach einer Zerstörung des Broca-Areals im *gyrus frontalis inferior* (Brodmann-Areale 44 und 45) der sprachdominanten Hemisphäre (ZILLES & REHKÄMPER 1994²:302), und ist nach LEUNINGER (1990:21) u.a. durch mühevoll und verlangsamte Sprachproduktion und durch Agrammatismus, d.h. durch Störungen der Produktion syntaktischer Strukturen, ge-

13 Daher könnten diese Verfahren mit einigem Recht auch dem Absatz *Funktionale Dynamik neuronaler Prozesse* in diesem Kapitel zugeordnet sein.

14 Zur Verwendung von ERPs in der Erforschung der Sprachverarbeitung vgl. OSTERHOUT & HOLCOMB 1995

15 Zur Geschichte der Aphasologie vgl. ROCH-LECOURS, LHERMITTE & BRYANS 1983.

16 Für einen Überblick vgl. ROCH-LECOURS, LHERMITTE & BRYANS 1983, für eine Einführung siehe LEUNINGER 1990, DAMASIO 1992.

kennzeichnet. Die Wernicke-Aphasie wird durch eine Schädigung des Wernicke-Areals im hinteren Bereich des *gyrus temporalis superior* (okzipital der Areae 41/42) der ‘sprachdominanten’ Hemisphäre hervorgerufen (ZILLES & REHKÄMPER 1994²:226f), und zeichnet sich u.a. durch semantische Paraphasien und Neologismen, d.h. durch Störungen des Lexikons, aus (LEUNINGER 1990:41). Diese beiden Aphasien werden häufig auch dichotomisch entgegengesetzt als anteriore vs. posteriore Aphasie und als expressive vs. rezeptive Aphasie bezeichnet. Aphasien treten im allgemeinen nur nach Läsionen der sprachdominanten (d.h. bei den meisten Menschen der linken) Hemisphäre auf, Läsionen der rechten Hemisphäre hingegen verursachen häufig u.a. Defizite in Bezug auf die prosodischen Merkmale der Sprache, wie z.B. die Melodie eines Satzes (ROSS 1981, siehe unten 5.4.3.4).

Die Erforschung der neuronalen Grundlagen kognitiver Fähigkeiten in klinischen Studien basiert auf dem Prinzip der Korrelierung von mehr oder weniger begrenzten Hirnschädigungen mit selektiven Defiziten einzelner Komponenten der Sprachfähigkeit.¹⁷ Aus dem gleichzeitigen Auftreten von Hirnläsionen und Sprachdefiziten bei einem Patienten, oder besser, bei einer Patientengruppe, kann geschlossen werden, dass der zerstörte Teil des Gehirns *notwendig* ist für die gestörten Sprachfunktion, es kann daraus aber nicht geschlossen werden, dass der zerstörte Teil *hinreichend* ist für die gestörte Funktion. Dies bedeutet, dass klinische Studien zu der Erkenntnis führen können, dass eine zerstörte Hirnregion einen Teil der kortikalen Repräsentation einer Sprachfunktion bildet, aber nicht zu der Erkenntnis, dass die zerstörte Hirnregion die einzige neuronale Basis der defizitären Sprachfunktion ist. Daneben ist auch an dieser Stelle zu bemerken, dass klinische Studien nur so aussagekräftig sind, wie die linguistische Analyse, die hinter ihnen steht – je differenzierter die identifizierten sprachlichen Defizite, die mit Läsionen korreliert werden, desto größer ist die Wahrscheinlichkeit, dass sich sehr stabile *Korrelationen* zwischen Läsionstypen und sprachlichen Defiziten finden lassen.

Eine differenzierte linguistische Analyse der sprachlichen Defizite von Patienten mit Hirnschädigungen bedeutet insbesondere auch, dass die Sprachdaten, die ausgewertet werden, um solche Defizite zu diagnostizieren, mit subtilen experimentellen Paradigmen erhoben werden müssen. So kann die Analyse von aphasischer Spontansprache eine ausreichende Datenmenge zur Verfügung stellen, um z.B. eine phonetische und eine phonologische Ebene im Sprachproduktionssystem zu dissoziieren (Blumstein

¹⁷ Klinische Studien sind hier unter der Fragestellung ‘Funktionale Lokalisierung’ aufgeführt. Natürlich kann man klinische Studien auch durchführen, um Erkenntnisse zur Funktionalen Lateralisierung einzelner Sprachverarbeitungs- und Sprachproduktionskomponenten zu gewinnen, wenn man Patienten mit geschädigter linker und geschädigter rechter Hemisphäre vergleicht, s.u. 5.3.4.

1995b). Will man jedoch einzelne Komponenten des phonologischen Systems untersuchen, z.B. die Fähigkeit, die Silbenstruktur eines Wortes durch verschiedene phonetische Parameter zu kodieren, so müssen zur Analyse geeignete Äußerungen gezielt hervorgerufen werden. Sehr subtile experimentelle Paradigmen erfordert insbesondere die Untersuchung der Sprachverarbeitungsfähigkeit von hirngeschädigten Patienten, denn es ist leichter, die Fähigkeit zur Sprachproduktion zu untersuchen, nämlich anhand von phonologischen, syntaktischen und lexikalischen Fehlern in aktuellen Äußerungen, als die Fähigkeit zur Verarbeitung, in Bezug auf die sich Defizite ja nur in offensichtlichem Unverständnis zeigen, was aber vom Patienten durch eine Vielzahl von Strategien vermieden werden kann.

Diese methodischen Schwierigkeiten mögen zu den Gründen dafür gehören, dass BLUMSTEIN (1995a) zu folgendem Schluss kommt:

All aphasic patients display phonological, lexical, and syntactic deficits regardless of aphasia type and lesion localization, suggesting that the neural instantiation of these linguistic components requires complex neural structures that are represented throughout the left hemisphere. [...] In particular, language impairments in aphasia reflect processing disorders that cut across the linguistic components of the grammar including phonology, the lexicon, and syntax. Language impairments in aphasia are not selective with respect to a particular component of the grammar nor do they reflect impairments to particular representations. (BLUMSTEIN 1995a:364f)

Die Lösung liegt auch für BLUMSTEIN (1995a) in einem interdisziplinären und integrativen Vorgehen:

Future work requires greater integration of clinical and experimental findings with results from electrophysiological, metabolic, and neuroimaging studies with both neurologically intact and brain-damaged subjects (BLUMSTEIN 1995a:365)

Um aus der Korrelierung von sprachlichen Defiziten mit lokalisierten Hirnläsionen aussagekräftige Schlussfolgerungen über die neuronale Repräsentation der menschlichen Sprachfähigkeit im Gehirn ziehen zu können, müssen die Ergebnisse dieser Studien daher mit Erkenntnissen verbunden werden, die mit anderen methodischen Vorgehensweisen gewonnen wurden.

1.3.3 Elektrophysiologische Stimulation

Eine sehr interessante Methode zur Erforschung der Repräsentation der Sprache im Gehirn ist die elektrophysiologische Stimulation (PENFIELD & ROBERTS 1959, OJEMANN 1988, OJEMANN 1989), die zur präoperativen Diagnostik bei pharmakologisch unbehandelbaren Epilepsiepatienten angewendet wird, um die Folgen der Resektion eines Teils des Kortex, von dem aus die Anfälle sich ausbreiten, absehen zu können. Bei der elektrophysiologischen Stimulation wird über eine kleine Elektrode an einer Stelle des freiliegenden Kortex ein elektrisches Potential appliziert, das einen exzitatorischen oder

inhibitorischen Effekt auf das funktionelle System haben kann, in welchem stimuliert wird. Ein exzitatorischer Effekt kann sich z.B. darin zeigen, dass der Patient, der nur lokal anästhesiert wird, unwillkürliche Bewegungen vollzieht, wenn im motorischen Kortex stimuliert wird, oder Klänge oder Stimmen (ohne sensorischen Input) hört, wenn im auditorischen Kortex stimuliert wird. Hat die Stimulation hingegen einen inhibitorischen Effekt, so wird der Patient unterbrochen in der Tätigkeit, die er gerade ausführt, z.B. Sprechen, oder er ist nicht in der Lage, eine bestimmte Aufgabe durchzuführen, z.B. das Objekt auf einem Bild benennen (OJEMANN 1988:101).

Ebenso wie aus klinischen Studien lassen sich aus solchen Untersuchungen Aussagen darüber ableiten, welche Bereiche des Kortex notwendig sind für bestimmte Komponenten der Sprachfähigkeit: wenn die Applikation des elektrischen Potentials bei einem Patienten einen inhibitorischen Effekt¹⁸ in Bezug auf eine experimentelle Aufgabe, z.B. Objektbenennung, hat, so ist der Bereich, in welchem stimuliert wurde, *notwendig*, aber nicht *hinreichend* für die Teilfähigkeit, um die es geht. Mit der Methode der elektrophysiologischen Stimulation können grundsätzlich beliebige Komponenten der Sprache lokalisiert werden, doch ist hier, wie auch beim Natrium-Amytal-Test, zu bedenken, dass es sich um einen Eingriff ins Gehirn handelt, der den Zweck hat, Risiken notwendiger Operationen möglichst gering zu halten – daher ist es auch sinnvoll, eine so komplexe Fähigkeit wie Objektbenennung oder Vorlesen zu testen, für die sich zahlreiche notwendige Areale im perisilvischen Kortex (und nicht nur dort, vgl. OJEMANN 1988:107) identifizieren lassen, um den Aufwand zur Identifikation von sprachrelevanten Kortexarealen möglichst gering zu halten und es ist nicht vertretbar, diese Methode zur Überprüfung sehr differenzierter Modelle zur Lokalisierung verschiedener Komponenten der Sprachfähigkeit zu verwenden. Vergessen werden darf aber auf der anderen Seite auch nicht, dass z.B. die Fähigkeit zur Objektbenennung vermutlich sehr viel mehr als nur ‘eigentliche Sprachareale’ erfordert, nämlich auch Areale, in denen das ‘Wissen’ über diese Objekte gespeichert ist, d.h. Objektbenennung erfordert nicht nur intakte Sprache, sondern auch eine (ausreichend) intakte (visuelle, auditorische, somatosensorische etc.) Repräsentation dieser Objekte. Aus diesem Grund können Ergebnisse zur Lokalisierung ‘der Sprache’ im Gehirn mit der Methode der elektrophysiologischen

¹⁸ Kortikale Stimulation in Spracharealen hat niemals einen exzitatorischen Effekt, d.h. sie führt nicht zu unwillkürlichem Sprechen (PENFIELD & ROBERTS 1959, OJEMANN, G.A. (1983): „Brain organization for language from the perspective of electrical stimulation mapping“, *Brain and Language* 2, 101-120, zitiert nach CROSSON 1992:86) – Elektrische Stimulation des Thalamus der sprachdominanten Hemisphäre hingegen kann unwillkürliche Sprachäußerungen (Wörter, Phrasen oder Sätze) hervorrufen (SCHALTENBRAND, G. (1975): „The effects on speech and language of stereotactic stimulation in the thalamus and corpus callosum“, *Brain and Language* 2, 70-77, zitiert nach CROSSON 1992:86). Wiederholte Stimulation kann dabei dieselbe oder eine andere Äußerungen veranlassen.

Stimulation, die sich auf einige wenige Verhaltensmaße stützen, unter linguistischem Aspekt nur als bedingt aussagekräftig gelten.

1.4 Untersuchung der funktionalen Dynamik von neuronalen Prozessen

Die bisher vorgestellten neurowissenschaftlichen Methoden eignen sich zur Identifikation von (mehr oder weniger ausgedehnten) Bereichen des Gehirns, die an der Verarbeitung eines Sprachreizes oder an der Durchführung einer sprachbezogenen Aufgabe beteiligt sind. Dieses Forschungsprogramm der ‘funktionalen Lokalisierung’, in Bezug auf das die Wissenschaft, zumindest was die menschliche Sprachfähigkeit betrifft, noch am Anfang steht, liefert eine Theorie der funktionalen Systeme im Gehirn und im Idealfall auch ihrer Interaktion. Eine solche Theorie, die Hirnareale mit funktionalen Systemen identifiziert, kann aber nichts aussagen darüber, was *innerhalb* eines derart lokalisierten funktionalen Systems passiert. Um die interne Dynamik eines funktionalen Systems/Hirnareals zu erforschen, müssen die Bestandteile dieses Systems/Areals selbst zum Gegenstand der Untersuchung gemacht werden, d.h. die dynamische Interaktion der Neuronengruppen und Einzelneuronen muss in ihrer Funktion für das betreffende System untersucht werden.

Eine Möglichkeit, die funktionale Dynamik neuronaler Aktivierungsprozesse zu erforschen, bieten die nicht-invasiven elektrophysiologischen Verfahren mit der Auswertung von ERPs und EMFs – hierbei stellt die gegenwärtig nur ungenaue Lokalisierbarkeit der aktivitätserzeugenden Neuronengruppen aber noch ein großes Problem dar. Die Dynamik einzelner Neuronengruppen und Neuronen kann deshalb beim jetzigen Stand der Forschung nur invasiv, d.h. durch elektrophysiologische Ableitung in freiliegenden Bereichen des Kortex, untersucht werden. In sehr beschränktem Ausmaße ist dies beim Menschen möglich, nämlich, ähnlich wie auch die elektrophysiologische Stimulation, vor Kortexresektionen im zu entfernenden Hirngewebe; weitergehende und nicht unbedingt notwendige Eingriffe ins menschliche Gehirn aber sind nicht vertretbar. Eine dritte Möglichkeit schließlich liegt in der Untersuchung der funktionalen Eigenschaften einzelner Neuronen im Tierexperiment, wobei sich stets die Frage stellt, inwieweit Erkenntnisse über die neuronale Dynamik in funktionalen Systemen anderer Spezies auf den Menschen übertragbar sind. Diese Frage stellt sich insbesondere, wenn man die menschliche Sprachfähigkeit untersuchen will. Zwar finden sich sehr interessante Beispiele für Kommunikation und ‘Kommunikationsfähigkeit’ in der Tierwelt, aber keine andere Spezies verfügt über ein so komplex strukturiertes Sprachsystem wie

der Mensch. Angesichts der Unmöglichkeit, in großem Rahmen funktionale neuronale Dynamik am Menschen mit invasiven Methoden zu untersuchen, muss die Wissenschaft auch diese Möglichkeit der Erkenntnisgewinnung nutzen, um Hypothesen zur Repräsentation der Sprache im Gehirn zu bilden.

CREUTZFELDT, OJEMANN & LETTICH (1989a, b) führten mit dieser Methode eine Serie von Experimenten durch, in denen im Temporallappen der sprachdominanten Hemisphäre *single-unit*- und *multi-unit*-Ableitungen durchgeführt wurden. CREUTZFELDT ET AL. konnten u.a. Einzelneuronen identifizieren, deren Aktivierungsverhalten sie interpretierten als selektive Reaktion auf (i) bestimmte Phonemfolgen wie /kr/, /str/, (ii) auf das zweite Lexem in Komposita wie *airplane* oder *rattlesnake*, wobei gleichlange einfache Lexeme diese Aktivierung nicht hervorriefen, (iii) und schließlich auf mehrsilbige (drei- oder viersilbige) Wörter im Vergleich zu kürzeren Wörtern. Unter sprachwissenschaftlichem Blickwinkel könnte man spekulieren, dass CREUTZFELDT ET AL. (1989a) Signale von Neuronen ableiteten, die (ia) Silbenstrukturpositionen wie Anlaut oder Auslaut (spezifisch für bestimmte Phonemkombinationen) detektieren, (iia) von Neuronen, welche den Beginn der Reanalyse markieren, die einsetzt, wenn ein Lexem, das zunächst als einfaches Lexem geparkt wurde,¹⁹ und nach der Wahrnehmung des zweiten Lexems als erster Bestandteil eines zusammengesetzten Wortes reanalysiert werden muss und (iiia) von Neuronen, die Akzentmuster dreisilbiger (*grasshopper*, *hamburger* – 'σ σ σ,²⁰ *crocodile*, *rattlesnake* – 'σ σ σ) oder viersilbiger (*automobile*, *alligator* – ,σ σ'σ σ) Wörter repräsentieren, aber nicht das Akzentmuster einsilbiger (*word* – 'σ) oder zweisilbiger (*letter* – 'σ σ) Wörter. Hierzu kann noch bemerkt werden, dass ein- und zweisilbige Wörter einen einzigen Versfuß oder Takt bilden, drei- und viersilbige Wörter aber fast immer zwei Versfüße – diese Neuronen können also auch so etwas wie 'Füßigkeit' detektieren, d.h. phonologische Eigenschaften auf noch höherer Ebene, auf der Akzentebene.²¹ Besonders interessant für den Gegenstand der vorliegenden Untersuchung ist auch die folgende Beobachtung:

some neurons or multitunit activities which responded indiscriminately to every word of our word lists, or to almost all of them. *They could thus beat in the rhythm of speech*, at least for a limited period. (CREUTZFELDT, OJEMANN & LETTICH 1989a:472; Hervorh. v. m.)

Wenn Neuronen 'im Rhythmus der Sprache' feuern, dann feuern sie also 'phasenkoppelt' mit den Wörtern oder mit fast allen Wörtern. Dies könnte bedeuten, dass diese

¹⁹ Nach dem *Minimal Attachment Principle*, das in einem ersten Verarbeitungsdurchgang nur so viele Knoten wie unbedingt nötig rekonstruiert, was auch bedeutet, dass initiale Lexeme von Komposita zunächst als einfache Wörter interpretiert werden, vgl. FRAZIER 1995.

²⁰ σ repräsentiert eine Silbe, ' den Hauptakzent und , den Nebenakzent.

²¹ Zur prosodischen Hierarchie vgl. z.B. BLEVINS 1995:210, zu einer Definition der Kategorie *Fuß* vgl. WIESE (1996:56f): „a foot is a sequence of one or more syllables. In the latter case, the first syllable must carry stronger stress than the subsequent syllable in that foot.“

Neuronen phasengekoppelt mit wortinitialen akzentuierten Silben feuern, d.h. mit der rhythmischen Einheit des Englischen (s.u. 4.3.3). Daneben könnte es auch Neuronen geben, die zu jeder Silbe, oder zu jedem Silbengipfel, oder noch anders gesagt, zu jedem Sonoritätmaximum, einen Aktivitätsburst zeigen, möglicherweise auch in Abhängigkeit von der Akzentuierung, so dass sich für betonte Silben oder Vokale mehr und für unbetonte weniger Aktivität findet. Diese Neuronen könnten dann den Eingang liefern für Neuronen der Art, wie sie oben postuliert wurden, nämlich für Neuronen, die ganze Wortbetonungsmuster detektieren. Sprachwissenschaftliche Interpretationen dieser Art sind beim gegenwärtigen Stand der Forschung reine Spekulation – trotzdem können solche Spekulationen Hypothesen liefern, die im Modell oder im Experiment zu überprüfen sind.

Da die Stimuli vor der Versuchsreihe konzipiert werden mussten und da während der Durchführung keine Möglichkeit bestand, die Ergebnisse zu betrachten, verwendeten CREUTZFELDT ET AL. Sprachreize, die viele verschiedene akustische, morphologische und semantische Eigenschaften aufwiesen (CREUTZFELDT ET AL. 1989a:454). Dennoch soll hier angemerkt werden, dass die Ergebnisse dieser Studie noch aussagekräftiger sein könnten, wenn die verwendeten Stimuli in systematischerer Weise auf hypothetische Verarbeitungsschritte zu beziehen wären – wenn z.B. zum Test der Hypothese, dass manche Zellen ganze Phonemkomplexe repräsentieren, auch Stimuli verwendet worden wären, die diese Annahme systematisch überprüfen.

Um die Frage zu beantworten, mit welcher Art von funktionalen Einheiten und mit welcher Art von Aktivierung über diesen Einheiten im Gehirn die einzelnen Aspekte des menschlichen Sprachverhaltens, insbesondere der Sprachwahrnehmung, korreliert sind, müssen auch Erkenntnisse zur Repräsentation nichtsprachlicher akustischer Reize im auditorischen Kortex²² (und in anderen Arealen) anderer Spezies hinzugezogen werden. Neuronen im auditorischen System zeigen nicht nur eine frequenzselektive Aktivierung, sondern feuern auch in Abhängigkeit von der Intensität des Stimulus, d.h. sie reagieren vornehmlich auf akustische Reize mit einer ganz bestimmten Frequenz und mit einer ganz bestimmten Intensität, wobei mit steigender Intensität die Frequenzabstimmung immer breiter wird.²³ Daneben finden sich im auditorischen Kortex aber auch Zellen mit komplexeren Antworteigenschaften, z.B. Zellen, die auf Reize mit

22 Bevor ein neuronales Signal, das durch einen akustischen Reiz hervorgerufen wurde, den auditorischen Kortex erreicht, hat es bereits verschiedene Schaltstellen im *nucleus cochlearis*, *colliculus inferior* und im *corpus geniculatum mediale* passiert; für einen Überblick vgl. ZENNER 1994:119ff, vgl. auch POPPER & FAY (Hrsg.) 1991.

23 Für eine Einführung vgl. ZENNER 1994, HANDEL 1989, für einen Überblick zu neuronalen Karten im auditorischen Kortex vgl. SCHREINER 1995.

zwei Frequenzanteilen, d.h. mit zwei Harmonischen oder Formanten,²⁴ reagieren (NELKEN, PRUT, VAADIA, ABELES 1994 an Katzen), wie sie für die Identifikation von Vokalen relevant sind und Zellen, die selektiv auf Formanttransitionen²⁵ reagieren (HEIL, RAJAN & IRVINE 1992 an Katzen), wie sie auch für die Wahrnehmung von Konsonantenphonemen wesentlich sind (s.u. 5.4.2). Ebenso finden sich Zellen, die auf Amplitudenmodulationen von einer bestimmten Modulationsfrequenz (Periodizität des Signals) reagieren (für einen Überblick vgl. LANGNER 1992), welche für die Wahrnehmung der Intonation eines Sprachsignals von Bedeutung sind. Die selektive Reaktion auf Amplitudenmodulationen (CLAREY, BARONE & IMIG 1991, CREUTZFELDT 1983:207) könnte auch diejenige Eigenschaft sein, die dazu führt, dass manche Zellen „im Rhythmus der Sprache“ (CREUTZFELDT, OJEMANN & LETTICH 1989a:472) und vielleicht phasengekoppelt mit dem Silbenrhythmus,²⁶ feuern. Derartige Erkenntnisse zur Repräsentation akustischer Signale im auditorischen System sind von großer Bedeutung für die Beantwortung der Frage nach der Sprachwahrnehmung, d.h. nach den Verarbeitungsschritten, die aus einem akustischen Signal eine phonologische Repräsentation rekonstruieren. Aussagen über die Repräsentation der syntaktischen oder morphologischen Verarbeitung sprachlicher Äußerungen lassen sich daraus jedoch nicht ableiten.

Wo im Kortex könnte aber die Fähigkeit zur syntaktischen Verarbeitung lokalisiert sein und welche funktionalen Eigenschaften von Neuronen würde man hier erwarten? Vielleicht ließen sich auch im Broca'schen Zentrum, das für die Verarbeitung der Syntax der Sprache spezialisiert zu sein scheint (CAPLAN 1995), Neuronen der Art finden, wie CREUTZFELDT ET AL. (1989a) sie für die morphologische Verarbeitung postuliert haben – 'Inkongruenz-Detektions-Neuronen' oder 'Reanalyse-Initiations-Neuronen',²⁷ die mit starker Aktivität auf die Inkongruenz reagieren, die bei der Verarbeitung eines Holzwegsatzes²⁸ wie in (1) auftritt.

- (1) The horse raced past the barn fell.

24 Formanten sind die Frequenzbereiche des Spektrums, in denen sich Signalanteile mit besonders hoher Energie konzentrieren.

25 Formanttransitionen sind die Frequenzänderungen, die bei der Produktion einer Konsonant-Vokal-Silbe entstehen, wenn der Vokaltrakt von der Konfiguration, mit welcher der konsonantische Verschluss gebildet wird, in die vokalische Konfiguration übergeht.

26 REMEZ & RUBIN (1990) zeigen, dass langsame Amplitudenmodulationen der Wahrnehmung der Silbenzahl eines Sprachreizes zugrundeliegen, s.u. 4.4.3.

27 Für ein Netzwerk-Modell syntaktischer Verarbeitung, in welchem die syntaktische Reanalyse eines Holzweg-Satzes durch eine *alarm*-Zelle ausgelöst wird vgl. WILKENS 1997.

28 Für einen Überblick zur Verarbeitung von Holzwegsätzen (*garden path sentences*) und anderen syntaktischen Ambiguitäten vgl. HARLEY 1995:297ff.

Doch wie spezifische Neuronen würde man überhaupt erwarten? Würde man Neuronen erwarten, die ganze prototypische Satzmuster, z.B. Agens-Thema-Strukturen wie (2a-b) repräsentieren?

- (2a) [Agens Anna] trägt [Thema Berta]
 (2b) [Agens Anna] schiebt [Thema Berta]

Solche Neuronen würden jedesmal aktiv, wenn ein solcher Satz verarbeitet (d.h. geäußert oder verstanden) würde, sie würden aber nicht aktiv, wenn ein Satz mit einer Thema-Ziel-Struktur wie in (3a-b) verarbeitet würde.

- (3a) [Thema Anna] erreicht [Ziel] das Haus]
 (3b) [Thema Der Ball] trifft [Ziel] die Scheibe]

Dass es solche Neuronen geben mag, könnte man z.B. angesichts der Beobachtungen von Rizzolatti und seinen Mitarbeitern (RIZZOLATTI ET AL. 1996)²⁹ für möglich halten, die im prämotorischen Kortex (Area F5) von Affen (*macaca nemestrina*) Neuronen fanden, die sowohl dann feuerten, wenn der Affe selbst eine Handbewegung ausführte, als auch dann, wenn er beobachtete, dass jemand anderes diese Handbewegung ausführte. Diese Neuronen reagieren darüberhinaus selektiv nur bei der Ausführung bestimmter Handbewegungen, z.B. *Greifen* im Gegensatz zu *Festhalten*, oder *Greifen mit zwei Fingern* im Gegensatz zu *Greifen mit der ganzen Hand*. Die Autoren interpretieren diese Ergebnisse so, dass diese Neuronen Handlungen unabhängig vom Handelnden 'repräsentieren'. Ausgehend davon könnte man spekulieren, dass es im frontalen Kortex des Menschen nicht nur solche *Greifneuronen* geben könnte, sondern auch Neuronen, die Handlungen auf noch abstrakterer Ebene repräsentieren – nicht nur als *Drehen* vs. *Reißen*, sondern auch als Agens-Thema-Handlung im Gegensatz zu Thema-Ziel-Ereignis.

Ob es wirklich einzelne Neuronen gibt, die so komplexe Konzepte repräsentieren, darüber kann zum jetzigen Zeitpunkt keine Aussage gemacht werden – wahrscheinlich ist jedoch, dass in Bezug auf die Komponenten der menschlichen Sprachfähigkeit, die in den Bereich 'computationelles System' (vgl. CHOMSKY & LASNIK 1995) fallen, d.h. vor allem die Syntax, die Teilverarbeitungsschritte nicht einzelnen Neuronen mit sehr komplexen funktionalen Eigenschaften entsprechen, sondern eher verschiedenen interagierenden funktionalen Systemen. Erkenntnisse aus invasiven elektrophysiologischen Untersuchungen können daher zwar schon jetzt auf Peripherie-Fähigkeiten wie die auditive Wahrnehmung und auf die Fähigkeit zur Rekonstruktion einer phonologischen

²⁹ Zu unter sprachwissenschaftlicher Perspektive interessanten Interpretationen dieser Ergebnisse vgl. auch RIZZOLATTI & ARBIB 1998, FADIGA & GALLESE 1997.

Repräsentation aus dem akustischen Signal bezogen werden, bevor aber z.B. die Verarbeitung von Wh-Sätzen³⁰ elektrophysiologisch untersucht werden kann, müssen sehr viel detailliertere Modelle der einzelnen Verarbeitungsschritte in Sprachwahrnehmung und Sprachproduktion und der funktionalen Eigenschaften von Neuronen und ihrer Interaktion in komplexen Systemen entwickelt werden.

1.5 Zusammenfassung

Wie in den vorangegangenen Ausführungen deutlich wurde, gibt es eine Vielzahl von Methoden zur Erforschung der Sprache im Gehirn, von denen einige seit Jahrzehnten oder länger verwendet werden und es existieren zahlreiche Studien und Ergebnisse zu diesem Gegenstand. Die Resultate einiger dieser Methoden lassen sich leichter aufeinander beziehen als die anderer. So können Ergebnisse aus Studien mit bildgebenden Verfahren ohne weiteres mit Ergebnissen aus klinischen Studien verbunden werden und auch Erkenntnisse zur funktionalen Lateralisierung lassen sich mit bildgebenden Verfahren bestätigen. Eher isoliert bleiben aber z.B. die Ergebnisse elektrophysiologischer Verfahren: Aussagen zum Zeitverlauf der Verarbeitung, welcher sich in der Form der verschiedenen Komponenten der ERPs zeigt, lassen sich nicht so einfach mit den Erkenntnissen anderer, eher lokalisierend arbeitender, methodischer Zugänge zusammenfügen. Auch die Erkenntnisse, die mit elektrophysiologischer Stimulation gewonnen wurden, sind nur bedingt kompatibel mit den Ergebnissen aus klinischen Studien oder aus solchen mit bildgebenden Verfahren.³¹ Für einen Teil der Forschung gilt zudem, dass sich die Vorgehensweise empirischer Untersuchungen zur Sprache im Gehirn eher an den Untersuchungsmethoden und weniger an den kognitiven Modellen orientiert (und wegen der Komplexität der Untersuchungsmethoden auch orientieren muss), so dass mit einer Methode nacheinander verschiedene Komponenten der Sprachfähigkeit und anderer kognitiver Fähigkeiten wie Aufmerksamkeit und Gedächtnis untersucht werden, selten aber dieselbe Komponente der Sprachfähigkeit nacheinander mit verschiedenen Methoden.

Man kann sich dem Gegenstand ‘Sprache im Gehirn’ auch von der anderen Seite aus nähern, von der Sprache her. Unter einer solchen Zugangsweise kann man einen Aspekt linguistischer Strukturbeschreibungen herausgreifen, z.B. den phonologischen Aspekt, noch genauer den Bereich der Kategorie Silbe und des Merkmals Sonorität und

³⁰ Beispiele für Wh-Sätze: *Wer glaubst Du wird kommen? Chi credi que venga? Which folder do you keep the letters in?*

³¹ Zur Opposition von ‘Wernicke-Geschwind-Modell’ und seinem eigenen ‘elektrophysiologischen Modell’ vgl. OJEMANN 1988.

diesen Gegenstand dann mit den verschiedenen Techniken untersuchen. Auf diese Weise lassen sich unterschiedliche Erkenntnisse einer Komponente der Sprachfähigkeit gewinnen, deren biologische Grundlagen von allen Seiten beleuchtet werden können: ausgehend von der obersten Ebene der funktionalen Lateralisierung unter schrittweiser Verkleinerung des Blickwinkels auf einzelne Hirnareale, bis zur Betrachtung der funktionalen Eigenschaften einzelner Neuronen. Auf diese Weise könnte es eher möglich sein, die mit den verschiedenen Methoden gewonnenen Erkenntnisse in eine zusammenhängende Theorie der Sprache im Gehirn zu integrieren.

Die vorliegende Studie beschränkt sich auf einen kleinen Bereich der Sprache, nämlich auf die Kategorie Silbe und das Merkmal Sonorität als Teile der linguistischen Strukturbeschreibungen auf der Ebene der Phonologie. Diese Studie beschränkt sich aber auch auf eine einzige Fragestellung, nämlich auf die Frage nach der funktionalen Lateralisierung des Merkmals Sonorität. In den weiteren Ausführungen wird zunächst geklärt, welchen Status das Merkmal Sonorität in der phonologischen Theorie besitzt, welches seine phonetischen Korrelate sind und welche Funktion dieses Merkmal für die Sprachwahrnehmung erfüllen könnte. Im Anschluss daran wird die These entwickelt, dass für die Verarbeitung der Sonorität eine rechtshemisphärische Spezialisierung zu erwarten ist und in zwei dichotischen Sprachwahrnehmungsexperimenten empirisch getestet. Die Diskussion in diesem Kapitel soll dabei als Hintergrund dienen, um die weiteren Ausführungen in eine umfassendere Theorie der Sprache im Gehirn einordnen zu können.

2 Silbe und Sonorität in der phonologischen Theorie

2.1 Einleitung

Will man phonologische Strukturbeschreibungen auf den Prozess der Sprachverarbeitung und auf die Repräsentation der Sprachfähigkeit im Gehirn beziehen, so muss zunächst geklärt werden, welchen Status die Kategorien und Merkmale, die in den phonologischen Strukturbeschreibungen vorkommen, in der sprachwissenschaftlichen Theorie haben. Insbesondere muss geklärt werden, ob Kategorien und Merkmale für die Beschreibung sprachlicher Strukturen *elementar* sind, oder ob sie auf andere, einfachere Kategorien und Merkmale zurückgeführt werden können. Außerdem ist die genaue Kenntnis der Funktion, die diese Merkmale und Kategorien im Sprachsystem erfüllen, eine Bedingung dafür, sinnvolle Hypothesen über die Funktion dieser Merkmale und Kategorien für die Sprachwahrnehmung und Sprachproduktion zu bilden und sie schließlich mit funktionalen Einheiten im Gehirn und mit Aktivierungsprozessen über diesen Einheiten zu korrelieren.

Im zweiten Teil dieses Kapitels (2.2) wird daher zunächst geklärt, welche Funktion die *Silbe als Konstituente* für die Beschreibung phonologischer Strukturen hat und aus welchen Konstituenten sie selbst zusammengesetzt ist. Im dritten Teil (2.3) werden die Regularitäten der Anordnung einzelner Segmente in der Silbe betrachtet. Aus diesen Regularitäten wird das Konzept der *Sonorität* abgeleitet, mit dem sich die *Form der Silbe* verallgemeinernd beschreiben lässt. Im vierten Teil (2.4) schließlich wird der theoretische Status des Merkmals *Sonorität* in verschiedenen phonologischen Konzeptionen diskutiert. In diesem Abschnitt wird gezeigt, dass das Merkmal *Sonorität* im Rahmen einer Theorie der menschlichen Sprachfähigkeit nicht allein auf *segmentale* Merkmale, d.h. auf Eigenschaften von Einzellauten zurückzuführen ist, sondern nur mit Bezug auf die Kategorie Silbe als *suprasegmentaler* Strukturebene sinnvoll gekennzeichnet werden kann. Insbesondere wird gezeigt, dass sich nicht nur die *segmentalen*

Oberklassen in ihrer Sonorität unterscheiden, sondern dass auch *silbischen* und *unsilbischen* Segmenten ein unterschiedlicher *Sonoritätswert* zukommt. Ausgehend von dieser Annahme, dass das Merkmal Sonorität das Bindeglied zwischen den Strukturen auf segmentaler Ebene und den Strukturen auf suprasegmentaler Ebene bildet, wird in den nachfolgenden Kapiteln die Funktion der Sonorität für die Sprachwahrnehmung und ihre funktionale Lateralisierung im Gehirn betrachtet.

2.2 Die Silbe als Struktureinheit der Sprache

2.2.1 Einleitung

Welche Gründe gibt es für die Annahme, dass die Silbe eine elementare Konstituente einer phonologischen Theorie der Sprache ist?³² Ein erster Grund hierfür ist die intuitive Fähigkeit der Sprecher einer Sprache, Wörter in Silben zu zerlegen und die Zahl der Silben in einem Wort anzugeben: Sprecher des Deutschen werden übereinstimmend behaupten, dass das Wort *Ball* aus einer Silbe besteht und das Wort *Balken* aus zwei Silben; darüberhinaus werden sie darin übereinstimmen, dass im Wort *Balken* die Silben *Bal-* und *-ken* zu trennen sind. Die übliche Weise, einen Liedtext mit einer Melodie zu verbinden, ist die Zuordnung jeweils einer Silbe zu einem Ton; auch hierzu sind die Sprecher einer Sprache ohne weitere Instruktion in der Lage. Folgen von betonten und unbetonten Silben können ohne weiteres mit Bewegungen synchronisiert werden, z.B. in einem Abzählreim. Schließlich lassen sich in gesprochener Sprache auch *Versprecher*³³ beobachten, in denen ganze Silben hinzugefügt werden (1a) oder wegfallen (1b).

- (1a) Auf der .Ba.lu.stra.de. .spa.lu.zie.ren.^{34.35}
 (1b) .U.ni.ver.tät. statt .U.ni.ver.si.tät.³⁶

32 Diese Annahme ist in der früheren generativen Phonologie in Frage gestellt worden, vgl. z.B. CHOMSKY & HALLE 1968, die für das Englische keine Konstituenten von Silbendimension annehmen. Ein Grund hierfür mag sein, dass sich eine einheitliche *phonetische* Definition der Silbe bislang nicht hat formulieren lassen, vgl. LADEFOGED 1993²:143.

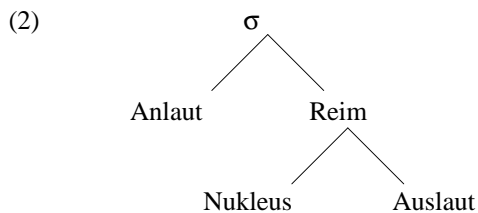
33 Zur Frequenz von silbenbezogenen Versprechern vgl. BERG 1992. BERG (1992:84) kommt in einem Überblick zur Rolle der Silbe in Versprechern, im Spracherwerb, in Sprachspielen und in aphasischer Sprache zu der Annahme, dass die Silbe im Vergleich zu Phonemen und Wörtern seltener von Fehlleistungen betroffen ist und folgert daraus, dass die Silbe (anders als das Wort oder Phonem) zwar eine „Struktureinheit“, aber keine „Inhaltseinheit“ in mentalen Repräsentationen von sprachlichen Ausdrücken sei. Für eine ähnliche Einschätzung vgl. auch FOWLER (1995:37). Zur Rolle der Silbe in der Sprachproduktion vgl. LEVELT 1989, LEVELT & WHEELDON 1994.

34 Zit. nach LEUNINGER 1996:43.

35 Punkte (.) bezeichnen Silbengrenzen.

36 Zit. nach BERG 1992:54.

Die Sprecher einer Sprache scheinen also eine intuitive Kenntnis davon zu haben, welche Einheiten der Sprache Silben sind. Im folgenden soll gezeigt werden, dass auch die linguistische Theorie die Kategorie Silbe benötigt, um die phonologischen Strukturen sprachlicher Äußerungen und Beschränkungen ihrer Form hinreichend zu beschreiben. Grundlage dieser Ausführungen ist das Modell der Silbe in (2), in welchem die Silbe in einem Anlaut und einem Reim zerfällt; letzterer setzt sich wiederum aus Nukleus und Auslaut zusammen. In den folgenden Abbildungen wird die Anlautkonstituente durch O (*onset*), der Nukleus durch N (*nucleus*) und die Auslautkonstituente durch C (*coda*) bezeichnet. Die intermediäre Konstituente Reim (*rime*) wird meist nicht repräsentiert.



2.2.2 Die Silbe als Konstituente in phonotaktischen Beschränkungen

Das Konzept *Silbe* wird in der phonologischen Theorie zur Formulierung einer ganz bestimmten Art von Beschränkungen der Form von phonologischen Strukturen benötigt: zur Beschreibung der *Phonotaktik* einer Sprache, d.h. zur Angabe der zulässigen und der unzulässigen Lautsequenzen. Zunächst lässt sich feststellen, dass in allen Sprachen Regularitäten in der Abfolge von Lauten existieren: es gibt in jeder Sprache mögliche Wörter und unmögliche Wörter. Mit dem Begriff *mögliches Wort* werden Lautfolgen bezeichnet, die einem Wort dieser Sprache entsprechen oder entsprechen könnten. So sind die Lautfolgen *trolpig* oder *friest* im Deutschen mögliche Wörter, aber nicht mit einer Bedeutung belegt. Der Begriff *unmögliches Wort* bezeichnet hingegen Lautfolgen, die in einer Sprache nicht mit einer Bedeutung belegt werden könnten. So sind die Lautfolgen *fpoklt*, *lkiefr* unmögliche Wörter des Deutschen. Wie aus (3a-c) zu ersehen ist, ist es nicht so, dass die Abfolge von Segmenten allein die Wohlgeformtheit eines Wortes bestimmt: ein Wort kann, obwohl es die Sequenzen /rk/ wie in **rkan*, /lp/ wie in **lpan* oder /nk/ wie in **nkie* enthält, wohlgeformt sein, wie *Werk*, *Tulpen*, *Tank*, je nachdem, ob diese Sequenzen zu einer einzigen Silbe gehören und je nachdem, an welcher Stelle in der Silbe diese Sequenzen stehen.

(3a) ³⁷	/. <u>k</u> ran./	* /. <u>r</u> kan./	/.we <u>r</u> .ken./	/.we <u>r</u> k./	* /. <u>w</u> er <u>k</u> ./
(3b)	/. <u>p</u> lan./	* /. <u>l</u> pan./	/.tu <u>l</u> .pen./	/.mi <u>l</u> p./	* /. <u>m</u> ipl./
(3c)	/. <u>k</u> nie./	* /. <u>n</u> kie./	/.pra <u>n</u> .ke./	/.ta <u>n</u> k./	* /. <u>t</u> ak <u>n</u> ./

(4) a.	σ	b.	* σ	c.	σ	σ	d.	σ
	O N C		O N C		O N C	O N C		O N C
	/k r a n/		/r k a n/		/w e r k e n/			/w e r k/

Diese Aussage lässt sich anhand der Daten in (4a-d) folgendermaßen präzisieren: Silben sind im Deutschen dann wohlgeformt, wenn sie aus zulässigen Silbenanlauten und zulässigen Silbenauslauten bestehen.³⁸ So bildet z.B. die Folge /rk/ einen wohlgeformten Auslaut (4d), aber keinen wohlgeformten Anlaut (4b); die umgekehrte Folge /kr/ hingegen ist ein wohlgeformter Anlaut (4a), aber kein wohlgeformter Auslaut. Werden die beiden Laute durch eine Silbengrenze voneinander getrennt, so sind beide Folgen möglich: /.wer.ken./ ebenso wie /.lock.re_r./ Um die phonotaktischen Regularitäten des Deutschen zu beschreiben, ist daher sowohl das Konzept der Silbe und der Silbengrenze als auch eine interne Strukturierung der Silbe in Anlaut, Nukleus und Auslaut nötig.

Wie aus (6a-b) und (7a-b) ersichtlich ist, bilden silbenstrukturbezogene phonotaktische Beschränkungen auch die Grundlage für die Alternationen in (5a-c).³⁹ Weil die Folge /tm/ im Deutschen keinen zulässigen Silbenauslaut bildet, muss die Lautfolge /a:tm/ als /.a:t[ə]m./ silbifiziert werden (5a, 6a-b). Die Lautfolge /atm+en/⁴⁰ hingegen kann als /.a:t.men./ silbifiziert werden, ohne phonotaktische Beschränkungen zu verletzen (5b, 7a-b), daher ist die Projektion eines zusätzlichen Nukleus wie in */.a:t[ə].men./ nicht nötig (und deshalb auch nicht zulässig⁴¹).

(5a)	Atem	* Atm
(5b)	atmen	* atemen
(5c)	Atmung	* Atemung

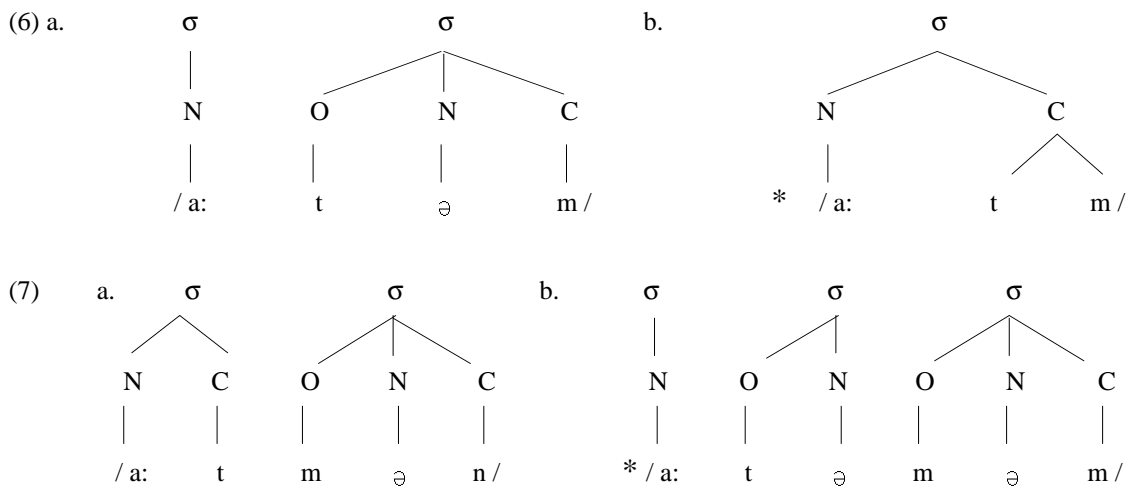
37 Punkte (.) bezeichnen Silbengrenzen, Sternchen (*) indizieren unzulässige Ausdrücke.

38 Darüberhinaus können im Deutschen nur Sonoranten (d.h. Vokale, Liquide wie /r/ und /l/ und Nasalkonsonanten wie /n/ und /m/) im Silbennukleus stehen.

39 Für eine Diskussion dieser Daten vgl. auch WIESE 1996:242ff, NOSKE 1993:141ff, HALL 1992:33ff, WIESE 1988:140ff.

40 + bezeichnet eine Morphemgrenze.

41 Leere Silbenstrukturpositionen wie der mit epenthetischem Schwa gefüllte Nukleus in *Atem* sind nur dann zulässig, wenn der zugrundeliegenden Struktur /atm/ im Rahmen einer Menge von Regeln oder Beschränkungen nicht auf eine andere Weise eine Oberflächenform zugeordnet werden kann. Die Unzulässigkeit unnötiger Epenthese wird z.B. in der optimalitätstheoretischen Beschreibung phonologischer Strukturen durch die Beschränkung FILL ausgedrückt (vgl. PRINCE & SMOLENSKY 1993:85ff).



Als letztes Beispiel für die Funktion der Silbe für die Formulierung von phonotaktischen Beschränkungen dienen die Daten in (8a-d) zur deutschen Auslautverhärtung (für eine Diskussion vgl. WIESE 1996:200ff).

(8a)	/lo[p]./	* /lo[b]./	/lo.[b]es./
(8b)	/win[t]. /	* /win[d]./	/win.[d]es./
(8c)	/krie[k]./	* /krie[g]./	/krie.[g]es./
(8d)	/ra[t]./	* /ra[d]./	/ra.[d]es./
(8e)	/ra[t]./		/ra.[t]es./

Wie aus diesen Daten ersichtlich wird, sind Obstruenten im Deutschen immer stimmlos, wenn sie am Silbenrand bzw. vor einer Silbengrenze stehen,⁴² im Silbenanlaut hingegen können sie stimmlos *oder* stimmhaft sein (8d-e). Eine phonologische Theorie, welche die Distribution von stimmhaften und stimmlosen Obstruenten beschreiben will, muss aus diesem Grunde die Kategorie *Silbe* miteinschließen.

2.2.3 Die Konstituenten der Silbe

Die vorangegangenen Ausführungen haben eine hierarchische interne Strukturierung der Silbe in Anlaut und Reim, Nukleus und Coda schon vorausgesetzt [s.o. (2)]. Dieses Modell der Silbe wurde eingeführt, weil die Differenzierung der drei Konstituenten Anlaut, Nukleus und Auslaut zur Darstellung phonotaktischer Regularitäten in der Silbe, um die es in den vorangegangenen Ausführungen ging, am besten geeignet sind.⁴³

⁴² Wie aus Formen wie /trü.[b]e./ – /trü[p].heit./, /le.[z]en./ – /le[s].bar./ hervorgeht, ist der kritische Kontext nicht das Wortende, wie aus den Daten in (8) geschlossen werden könnte, sondern tatsächlich die Silbengrenze.

⁴³ Die interne Struktur der Silbe kann auf verschiedene Weise konzipiert werden. LEVIN (1985) nimmt eine X-bar-ähnliche Struktur der Silbe als Projektion des Nukleus mit Coda-Komplement und An-

Externe, d.h. nicht linguistische, sondern psycholinguistische Evidenz für die Konstituente *Silbenanlaut* liefert z.B. die Versprecherforschung. So lassen sich häufig Versprecher beobachten, in denen Segmente zwischen den Anlautpositionen zweier Silben vertauscht (9a-b) oder im Anlaut einer vorhergehenden Silbe antizipiert (9c) werden.⁴⁴

- (9a) .Spau.̲bar.kas.se.⁴⁵
 (9b) .be.schlimmt. .stimm.⁴⁶
 (9c) .Ge.stenk.stät.te.⁴⁷

Warum die Existenz einer intermediären Konstituente *Reim*, welche Nukleus und Coda dominiert, angenommen werden muß,⁴⁸ soll im folgenden am Beispiel der Akzentuierung von englischen und lateinischen Wörtern erläutert werden.⁴⁹ Die Daten zum Wortakzent im Lateinischen⁵⁰ (10) und im Englischen⁵¹ (11) lassen sich folgendermaßen beschreiben: Die vorletzte Silbe trägt den Hauptakzent, wenn sie einen langen Vokal (10a), einen Diphthong (11a) oder einen Auslautkonsonanten [(10a) und (11a)] enthält; endet die vorletzte Silbe auf einen kurzen Vokal wie in (10b) und (11b), ist sie unbetont und die Antepenultima trägt den Hauptakzent.

- | | | |
|-------|----------------------------------|-----------------|
| (10a) | /re.'fe:.cit./ ^{52, 53} | /re.'fec.tus./ |
| (10b) | /ex.'is.ti.mo./ | /'vo.lu.cres./ |
| (11a) | /'a.ri.'z[ou].na./ | /'a.'gen.da./ |
| (11b) | /'Ca.na.da./ | /'A.'me.ri.ca./ |

Wie lassen sich die akzentbestimmenden Faktoren 'enthält einen langen Vokal', 'enthält einen Diphthong' und 'enthält einen Auslautkonsonanten' zusammenfassen? Dies ist

laut-Specifier an (N = Nukleus, N' = Reim und N'' = Silbe). KAHN (1980) setzt eine 'flache' Silbenstruktur ohne Reim-Konstituente an. Diese Konzeption bildet auch die Grundlage für das CV-Modell von CLEMENTS & KEYSER (1983), s.u. 2.4.4. Für einen Überblick zu verschiedenen Theorien der Silbenstruktur vgl. BLEVINS 1995:212ff, WIESE 1988:106ff.

44 Die experimentelle psycholinguistische Evidenz für die Strukturpositionen der Silbe ist weniger eindeutig. TREIMAN (1983) und PIERREHUMBERT ET AL. (1995) untersuchten die interne Struktur von Silben, indem sie Probanden „novel word games“ erlernen ließen, d.h. zu experimentellen Zwecken konstruierte sprachspielähnliche Prozeduren zur Ableitung neuer Wörter aus vorgegebenen Wörtern. Die Autoren kamen zu widersprüchlichen Schlussfolgerungen bezüglich der Angemessenheit des Anlaut-Reim-Modells als Erklärung für die Daten (vgl. auch TREIMAN ET AL. 1995).

45 Zit. nach LEUNINGER 1996:48.

46 Zit. nach LEUNINGER 1996:64.

47 Zit. nach LEUNINGER 1996:65.

48 SELKIRK (1982) begründet dieses Modell mit Vorkommensbeschränkungen für Segmente in der Reimkonstituente.

49 Diese kurze Diskussion von Silbengewicht und Akzent dient auch als Hintergrund für die Ausführungen in Kapitel 4, s.u. 4.3.3.

50 Daten nach KENSTOWICZ 1994:292.

51 Daten nach ROCA 1994:177.

52 Das Zeichen ' indiziert die betonte Silbe.

53 Der Doppelpunkt (:) indiziert einen langen Vokal, vgl. dt. /ro:te/ – /rotte/.

möglich, wenn man annimmt, dass Nukleus und Coda gemeinsam die Konstituente *Reim* bilden. Enthält der Reim *zwei* terminale Elemente, d.h. entweder einen langen Vokal oder Diphthong im Nukleus, oder einen kurzen Vokal im Nukleus und einen Konsonanten im Auslaut, dann kann diese Silbe den Hauptakzent tragen, andernfalls nicht. Diese strukturelle Komplexität des Reimes kann man als das *Gewicht* der Silbe bezeichnen: endet die Silbe auf einen kurzen Vokal, ist sie *leicht*, enthält sie einen langen Vokal oder Diphthong oder einen Auslautkonsonanten, ist sie *schwer*. Die Akzentregel im Englischen und Lateinischen lässt sich mit dem Konzept des Silbengewichts also kurz zusammenfassen: Die Pänultima (vorletzte Silbe) ist betont, wenn sie schwer ist, sonst ist die Antepänultima (vorvorletzte Silbe) betont. Es gibt demnach eine Konstituente in der Silbe, deren Gewicht entscheidend für das Gewicht der Silbe und damit für ihre Akzentuierung ist: der Reim. Die Schwesterkonstituente des Reimes, der Silbenanlaut, hat hingegen keine Funktion für das Gewicht der Silbe: ob eine Silbe mit einem, zwei oder drei Konsonanten beginnt, ist ohne Bedeutung für die Frage, ob sie schwer oder leicht ist und hat keinen Einfluss auf die Akzentuierung der Silben in einem Wort.⁵⁴

2.2.4 Zusammenfassung

Die vorangegangenen Ausführungen haben gezeigt, dass die Kategorie Silbe nötig ist, um die Sequenzierbarkeit von Segmenten, die Distribution von Segmenten wie stimmhaften und stimmlosen Obstruenten in einer Silbe und schließlich die Akzentuierung von Wörtern zu beschreiben. Ebenso wurde darauf hingewiesen, dass Sprecher einer Sprache ein intuitives Wissen über die Silbenstruktur von Äußerungen besitzen und dass sich in gesprochener Sprache *Versprecher* finden, die auf Fehlplanungen auf der Ebene der Silbenstruktur zurückzuführen sind. Hieraus muss geschlossen werden, dass die Silbe nicht nur eine elementare Konstituente einer phonologischen Theorie der Sprache ist, sondern auch eine noch genauer zu bestimmende Funktion für die Sprachproduktion und Sprachperzeption erfüllt. Der letzte Aspekt wird in den nachfolgenden Kapiteln wiederaufgegriffen. Hier geht es nun zunächst darum, wie sich die Form der Silbe beschreiben lässt.

⁵⁴ Das Konzept des Silbengewichts bildet den Ausgangspunkt für ein Modell der Silbe, in welchem die Silbe in Gewichtseinheiten oder *Moren* zerfällt (vgl. HYMAN 1985, HAYES 1989). Für eine Integration des Konzeptes *Sonorität* in eine moraische Silbenphonologie vgl. ZEC 1995.

2.3 Sonoritätshierarchie und Sonoritäts-Folge-Prinzip

Die Silbe als Konstituente phonologischer Strukturen steht zwischen der segmentalen Struktur und der metrischen Struktur. Silben sind diejenigen Strukturen, welche die Segmente zu größeren Einheiten zusammenzufassen, denen dann Akzente zugewiesen werden können. Wie oben gezeigt, zerfällt die Silbe selbst in die Konstituenten Anlaut und Reim, Nukleus und Coda. Zeichnet sich die Silbe darüberhinaus durch eine charakteristische Eigenschaft oder ein ihr eigenes tieferliegendes Formprinzip aus? Oder ist die Silbenstruktur nur eine Ebene der Strukturierung von Phonemfolgen, die aus den Spezifizierungen ihrer Konstituenten eindeutig abgeleitet werden kann? Um diese Frage zu beantworten, wird im folgenden aus den phonotaktischen Regularitäten des Deutschen das Konzept der *Sonorität* abgeleitet, mittels dessen das allgemeine Formprinzip der Silbe, das *Sonoritäts-Folge-Prinzip*, formuliert werden kann. Der Status des Merkmals *Sonorität* in verschiedenen phonologischen Theorien wird im nächsten Abschnitt diskutiert.

Zunächst lässt sich feststellen, dass die in (12a) aufgeführten Lautsequenzen im Silbenanlaut stehen können, während sie im Silbenauslaut (12b) nur in umgekehrter Reihenfolge vorkommen.

(12a)	t < r	Trog	(12b)	r < t	hart
	p < l	Plan		l < p	Alp
	k < n	Knief		n < k	Schrank
	f < l	Fliege		l < f	Schilf
	sch < r	Schreck		r < sch	harsch
(13a)	* p < t	* ptog	(13b)	* t < p	* tpog
	* f < ch	* fchilt		* ch < f	* chfilt
	* r < l	* rlack		* l < r	* lrack

Die Daten in (13a-b) weisen darauf hin, dass zwei Elemente *derselben* Oberklasse⁵⁵ [d.h. zwei Plosive, Frikative, Nasale, Liquide oder Vokale, s.u. (14)] in derselben Silbenstrukturposition nicht aufeinander folgen können, denn diese Lautsequenzen sind meist auch in umgekehrter Reihenfolge nicht zulässig. Diese Beobachtung kann vorläufig so interpretiert werden, dass die Reihenfolge der Segmente in einer zulässigen Silbe von ihrer Zugehörigkeit zu einer der fünf Oberklassen abhängt und zwar derart, dass im Silbenanlaut die Oberklassenelemente nur in einer bestimmten Reihenfolge und im Silbenauslaut nur in der umgekehrten Reihenfolge sequenziert werden können. Ausgehend

⁵⁵ Unter einer *Oberklasse* wird eine natürliche Klasse von Segmenten verstanden, welche dieselben Spezifizierungen für die Oberklassenmerkmale aufweisen (z.B. für die drei Oberklassenmerkmale [\pm konsonantisch], [\pm vokalisches], [\pm sonorant]).

von dieser Annahme lassen sich die Oberklassen und ihre Elemente in einer Hierarchie anordnen, so dass im Silbenanlaut Oberklassenelemente meist in aufsteigender Folge kombiniert werden können und im Silbenauslaut meist in absteigender Folge.⁵⁶

(14)	1.	Vokale	a, e, i, o, u
	2.	Liquidkonsonanten	l, r
	3.	Nasalkonsonanten	m, n, ŋ
	4.	Frikativkonsonanten	f, v, s, z, ʃ, ʒ, ç, x
	5.	Plosivkonsonanten	p, b, t, d, k, g

Diese Hierarchie kann man durch eine abstrakte⁵⁷ Eigenschaft erklären, indem man annimmt, dass die Elemente einer Oberklasse jeweils dieselbe Intensität dieser Eigenschaft aufweisen,⁵⁸ und dass sich die fünf Oberklassen hinsichtlich der Intensität dieser Eigenschaft paarweise unterscheiden, so dass die Oberklasse, die an unterster Stelle in der Hierarchie steht, am wenigsten⁵⁹ von dieser Eigenschaft hat, die zweitunterste Oberklasse in der Hierarchie am zweitwenigsten usw., bis zu den Vokalen, die in Relation zu den anderen Oberklassen am meisten von dieser Eigenschaft haben.

Die abstrakte Eigenschaft, mit welcher sich die Ordnung der Oberklassen in der Hierarchie erklären lässt, wird *Sonorität* genannt. Die Hierarchie, mit der sich die Reihenfolge der Segmente in der Silbe darstellen lässt, wird *Sonoritätshierarchie* genannt.⁶⁰ Das allgemeine, von einer Zuordnung zu den Oberklassen unabhängige Prinzip der Reihenfolge der Laute in der Silbe, das mit dem Begriff *Sonorität* formuliert werden kann, heißt *Sonoritäts-Folge-Prinzip* (im folgenden *SSP* für *Sonority Sequencing Principle*) und wird von CLEMENTS (1990) folgendermaßen formuliert:

56 Aber noch weitere Sequenzierungsbeschränkungen sind nötig: Folgen wie /,pfmla./ oder /,tsnla./ existieren (wenn überhaupt) nur in wenigen Sprachen. Solche Folgen könnten z.B. durch die Formulierung einer Bedingung ausgeschlossen werden, die festschreibt, dass die Sonorität im Anlaut von Laut zu Laut um einen bestimmten Grad ansteigen muss, bzw. dass aufeinanderfolgende Laute eine bestimmte Sonoritätsdifferenz aufweisen müssen.

57 *Abstrakt* soll hier im Sinne von ROCA (1994:134) verstanden werden: „phonological atoms are glued together by a prosodic fabric which is abstract to the extent that it cannot be heard directly“. Dass diese prosodische Organisation doch gehört werden kann, auch wenn sie nicht *bedeutungsunterscheidend* ist, wird in diesem und im nachfolgenden Kapitel gezeigt.

58 Oder eine identische Merkmalskonfiguration in einem zwei, drei oder vier Merkmale umfassenden Oberklassenmerkmal-System, s.u. 2.4.

59 Oder am meisten von dieser Eigenschaft, vgl. hierzu VENNEMANNNS Konzept der *konsonantischen Stärke* (VENNEMANN 1982, 1986, 1988), das in inverser Relation zur Sonorität steht.

60 Sonoritätshierarchie und Sonoritäts-Folge-Prinzip gehen auf SIEVERS (1901⁵) zurück und sind von JESPERSEN (1904), SAUSSURE (1972) und anderen aufgegriffen worden, vgl. hierzu auch die Ausführungen in Kapitel 3. Während dort die Frage nach den phonetischen Korrelaten der Sonorität diskutiert wird, geht es in diesem Kapitel um die Stellung des Merkmals *Sonorität* in der phonologischen Theorie.

Sonority Sequencing Principle: Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted. (CLEMENTS 1990:285)

Die Menge aller⁶¹ möglichen Silben kann auf diese Weise unter Rückgriff auf die Eigenschaft *Sonorität* und die durch diese konstituierte *Sonoritätshierarchie* der Laute charakterisiert werden.⁶² Gleichzeitig bietet der Begriff *Sonorität* die Möglichkeit, eine phonologische Definition der Silbe zu formulieren. Eine Silbe entspricht nach CLEMENTS' (1990) Auffassung einem *Sonoritätszyklus*:

syllables are normally characterized by a rise and fall in sonority which is reflected in the sonority scale values characterizing each of their segments. Sequences of syllables display a quasiperiodic rise and fall in sonority, each repeating portion of which may be termed a sonority cycle. (CLEMENTS 1990:299).

Nach Clements besteht eine Silbe also aus einem *Sonoritätsanstieg*, gefolgt von einem *Sonoritätsabstieg*. Ersterer entspricht dem *Silbenanlaut*, letzterer dem *Silbenauslaut*. Der *Silbennukleus* ist demnach der Punkt, an welchem der Anstieg in den Abstieg übergeht, d.h. das *Sonoritätsmaximum*. Besteht eine Silbe nur aus einem einzigen Laut, so bildet dieser ein lokales Sonoritätsmaximum und entspricht dem Silbennukleus.

Diese Kennzeichnung der Silbe als Sonoritätszyklus beschreibt auf der Basis des Merkmals Sonorität die Zuordnung von Segmenten zu Silbenstrukturpositionen. So sind lokale Sonoritätsmaxima, d.h. Segmente, die nur von weniger sonoranten Segmenten umgeben sind, immer dem Silbennukleus zugeordnet und Silbengrenzen fallen wegen des Prinzips des maximalen Anfangsrandes⁶³ in der Regel *vor* das Element mit der minimalen Sonorität.⁶⁴ Zweitens lässt sich diese Definition der Silbe als *Sonoritätszyklus* unter der Bedingung, dass ein eindeutiges (ein- oder mehrdimensionales) phonetisches

61 In einer Untersuchung zur Sequenzierbarkeit von Konsonanten in mehr als einhundert Sprachen kommt GREENBERG (1978) zu der Schlussfolgerung, dass „in relation to the peak of the syllable, combinations are favored in which sonants are closer to the peak than obstruents and in which voiced consonants are closer to the peak than unvoiced.“ (1978:270). Dies bedeutet, dass das Sonoritäts-Folge-Prinzip und die hier aufgeführte Sonoritätshierarchie zumindest als verallgemeinernde Beschreibung der Form möglicher Silben für sehr viele Sprachen der Welt Gültigkeit hat.

62 Neben dem SSP müssen weitere phonotaktische Restriktionen formuliert werden, z.B. solche, die Kombinationen von Segmenten mit identischer Artikulationsortspezifizierung wie in der Sequenz */tlu/ verbieten, d.h. Dissimilationsbeschränkungen, andererseits solche, die das SSP weiter eingrenzen, indem sie einen Sonoritäts-Anstieg von einem gewissen Grade fordern und Folgen wie */nru/ oder */mlu/ ausschließen, vgl. hierzu SELKIRK 1982, 1984, ebenso GIEGERICH 1992 für das Englische, WIESE 1996, HALL 1992 für das Deutsche.

63 Das *Prinzip des maximalen Anfangsrandes* besagt, dass ein intervokalischer Konsonant nicht dem Auslaut der ersten, sondern dem Anlaut der zweiten Silbe zugeordnet wird: VCV → V.CV (vgl. z.B. VENNEMANN 1986). Dieses Prinzip wird von CLEMENTS (1990:304) als *Sonority Dispersion Principle* formuliert, das besagt, dass in der ersten Halbsilbe, d.h. der Segmentfolge zwischen Silbengrenze und Nukleus, ein *maximaler* Sonoritätsanstieg bevorzugt wird, in der zweiten Halbsilbe, d.h. in der Segmentfolge zwischen Silbennukleus und Silbengrenze, ein *minimaler* Sonoritätsabfall.

64 Für Ansätze zur dynamischen computationellen Modellierung der Silbifizierung von Segmenten auf der Basis des Merkmals Sonorität vgl. GOLDSMITH 1993, LARSON 1990, GOLDSMITH & LARSON 1990.

Korrelat der Eigenschaft *Sonorität* identifiziert werden kann, auf den Prozess der Rekonstruktion von Silben aus einem akustischen Signal beziehen. Auf der Basis des phonetischen Korrelates könnten Sonoritätsminima erkannt und als Silbengrenzen repräsentiert werden, ebenso Sonoritätsmaxima als Silbennuklei.⁶⁵ Die Konzeption der Silbe als *Sonoritätszyklus* hat also den Vorteil, die Silbe mithilfe *einer einzigen Eigenschaft* von Lautfolgen zu definieren und so die Relation zwischen einerseits segmentalen Eigenschaften und andererseits Silbenstrukturpositionen auf der Basis dieser einen Eigenschaft erklärbar zu machen.

Was ist Sonorität jedoch genau? Weiter oben wurde Sonorität als eine Eigenschaft eingeführt, die sich scheinbar aus den phonotaktischen Regularitäten des Deutschen ableiten lässt. Dennoch finden sich im Deutschen Formen wie in (15a-b), auf deren Grundlage eine mit der vorher vorgestellten unvereinbare Hierarchie rekonstruiert werden müsste, nämlich eine, in welcher Frikative wie /s/ und /sch/ weniger sonorant sind als Plosive wie /p/ und /t/.

(15a)	.sprang.	.Strauch.
(15b)	.klopfst.	.rätst. ⁶⁶
(15c)	.Mast.	.Matz.

Das Verhalten von /s/ (und /sch/) im Deutschen (15c) widerspricht zudem der oben gemachten Annahme, dass die im Silbenanlaut stehenden Konsonantfolgen im Silbenauslaut nur in umgekehrter Reihenfolge vorkommen können, denn *Mast* ist ebenso wie *Matz* ein deutsches Wort. Diese Annahme bildete jedoch die Grundlage für die Ableitung einer *Hierarchie* von Oberklassen aus den phonotaktischen Regularitäten des Deutschen. Demnach lassen sich die phonotaktischen Beschränkungen des Deutschen mit der Annahme einer solchen Hierarchie nicht vollständig beschreiben.

Zudem sind für verschiedene Sprachen modifizierte Versionen der Sonoritätshierarchie abgeleitet worden (vgl. HALL 1992:64 und WIESE 1996:260 für das Deutsche), um sprachspezifischen Sequenzierungsbeschränkungen Rechnung zu tragen und es finden sich in vielen Sprachen der Welt Wörter, deren Form gegen das SSP verstößt.^{67,68}

⁶⁵ Welche Funktion die phonetisch-phonologische Eigenschaft Sonorität für die Rekonstruktion von Silben in der Sprachwahrnehmung erfüllen könnte, wird in Kapitel 4 diskutiert.

⁶⁶ Flektierte Formen wie *du rätst* bilden insofern einen Sonderfall in der Menge der möglichen Wörter des Deutschen, als in ihrem Falle immer ein Gleichgewicht geschaffen werden muss zwischen einer Tendenz zur Vermeidung von Allomorphie (d.h. der Tendenz, dass ein und dasselbe Morphem, z.B. das Morphem '2. Person Sg.' möglichst in allen Kontexten durch dieselbe Form /-st/ realisiert werden sollte) und einer Tendenz zur phonotaktischen Wohlgeformtheit der flektierten Formen (die z.B. dazu führen kann, dass ein Allomorph für vokalfinale Stämme und eines für konsonantfinale Stämme entsteht). Im Deutschen ist die Tendenz zur Vermeidung von Allomorphie zumindest im Verbalparadigma höherrangig als der Tendenz zur phonotaktischen Wohlgeformtheit.

⁶⁷ Für Beispiele aus verschiedenen Sprachen vgl. CLEMENTS 1990.

Die Tatsache, dass es nicht möglich ist, für nur eine einzige Sprache (z.B. für das Deutsche) aus den phonotaktischen Regularitäten eine eindeutige Sonoritätshierarchie des Lautinventars dieser Sprache abzuleiten, geschweige denn eine universale Hierarchie, welche die Phonotaktik aller Sprachen der Welt beschreibt, führt zu der Annahme, dass Sonorität noch etwas anderes ist als eine abstrakte Eigenschaft, die sich aus den inhärenten Eigenschaften von Segmenten (d.h. den Oberklassenmerkmalen) ableiten lässt. Will man die durch das SSP ausgedrückte Generalisierung aufrecht erhalten, so muss das Merkmal Sonorität noch auf eine andere Weise charakterisiert werden als durch seine Ableitung aus der Phonotaktik einer Sprache. Der Weg zu einer Lösung wurde bereits weiter oben impliziert, als die Stufen der Sonoritätshierarchie mit den Oberklassen identifiziert wurden. In einem ersten Schritt wurde angenommen, dass die Elemente einer Oberklasse einen identischen Sonoritätswert haben und dass verschiedene Oberklassen sich in ihrer Sonorität unterscheiden. Die Oberklassenmerkmale aber beschreiben Eigenschaften von Segmenten und werden mit Bezug auf phonetische Konzepte wie ‘Anwesenheit einer Formantenstruktur’, ‘Grad der Behinderung des Luftstromes’ usw. definiert (s.u. 2.4).

Lässt sich das Merkmal Sonorität also aus den Oberklassenmerkmalen ableiten? Um diese Frage zu beantworten, wird im nächsten Abschnitt die Beziehung zwischen Sonorität und anderen phonologischen Merkmalen genauer betrachtet.

2.4 Sonorität, Silbizität und Oberklassenmerkmale

2.4.1 Einleitung

Wie oben gezeigt wurde, erfüllt das Konzept der Silbe eine wichtige Funktion für die Beschreibung der Phonotaktik einer Sprache. Zudem lassen sich in fast allen Sprachen die Silben durch dasselbe Formprinzip, d.h. das Sonoritäts-Folge-Prinzip, kennzeichnen. Dennoch ist es so, dass die Silbenstruktur einer Äußerung vollständig vorhersagbar ist

⁶⁸ Diese eingeschränkte Gültigkeit der Sonoritätshierarchie und des SSP wirft auch die Frage nach dem theoretischen Status des SSP als phonologischem Prinzip auf. In der *generativen Phonologie* wird dieses Problem umgangen, indem das SSP nicht als eine Beschränkung über Oberflächen-Formen oder Phonetischen Formen konzipiert wird, sondern nur für die zugrundeliegende Ebene der *initialen Silbifizierung* gelten muss (vgl. CLEMENTS 1990). In eine beschränkungs-basierte phonologische Theorie wie die *Optimalitätstheorie* hingegen kann das SSP als universales Prinzip leicht integriert werden. Verstöße gegen dieses Prinzip, die sich in einzelnen Sprachen finden, machen nicht die durch das SSP (in optimalitätstheoretischer Terminologie *Nuclear Harmony Constraint*) ausgedrückte Verallgemeinerung ungültig, sondern ergeben sich durch die Interaktion von SSP und anderen, möglicherweise höherrangigen, Prinzipien, wie z.B. der Tendenz, dass Silben einen Anlaut haben müssen (*Onset Constraint*) (vgl. PRINCE & SMOLENSKY 1993).

aus den Segmenten, aus denen sie besteht, denn in den meisten Sprachen gibt es keine zwei Wörter, die sich nur in der Silbenstruktur, aber nicht in der segmentalen Struktur unterscheiden.⁶⁹ Um den Zusammenhang zwischen Silbe und Sonorität genauer zu klären, soll im folgenden noch einmal die Frage gestellt werden, ob die Silbe nur eine abgeleitete *Konstituente* ist, die eine bestimmte *Form* besitzt, jedoch keine andere *Funktion* erfüllt, als die Segmente nach phonotaktischen Prinzipien im Hinblick auf die phonologischen Strukturen auf anderen Ebenen zusammenzufassen, oder ob sich die Silbe durch eine eigenständige Struktureigenschaft auszeichnet, die sich nicht aus den inhärenten Eigenschaften der Segmente erklären lässt, d.h. durch die *Sonorität*.

Der Begriff *Sonorität* ist weiter oben mit dem der phonologischen Oberklasse in Zusammenhang gebracht worden: erstens wurde angenommen, dass sich die Elemente einer Oberklasse nicht in ihrer Sonorität unterscheiden und zweitens, dass sich die einzelnen Oberklassen im Grad ihrer Sonorität unterscheiden, so dass jede Stufe der Sonoritätshierarchie einer Oberklasse entspricht. Kann also das Merkmal Sonorität auf die segmentalen Oberklassenmerkmale zurückgeführt werden? Oder ist Sonorität ein Merkmal, das eine unmittelbare Eigenschaft der Silbenstruktur ist, so dass die Zuweisung von Sonoritätswerten an Segmente auch sekundär aus ihrer Distribution innerhalb der Silbe resultiert? Um die Frage zu klären, ob Sonorität auf inhärente segmentale Merkmale zurückgeführt werden kann, oder nicht ohne Bezug auf die Silbe zu definieren ist, werden im folgenden drei phonologische Konzeptionen der *Oberklassenmerkmale* und der Merkmale *Silbizität* und *Sonorität* betrachtet. In den nachfolgenden Ausführungen soll gezeigt werden, dass Sonorität nicht auf die segmentalen Oberklassenmerkmale zurückgeführt werden kann, wenn man die phonologische Sonorität mit einer phonetischen Eigenschaft korrelieren will. Vielmehr trägt auch das 'prosodische' Merkmal der Silbizität dazu bei, den Sonoritätswert eines Segmentes zu bestimmen. In den hierauf folgenden Ausführungen zu den Konzeptionen von CHOMSKY & HALLE (1968), SELKIRK (1984) und CLEMENTS (1990) liegt der Schwerpunkt dabei nicht darauf, die jeweiligen Modelle darzustellen, sondern es soll deutlich werden, dass in allen drei Konzeption ein Zusammenhang zwischen Sonorität und Silbizität einerseits und phonologischer und phonetischer Sonorität andererseits gegeben ist.

⁶⁹ Im Deutschen und im Englischen wird ein unbetontes Schwa häufig getilgt, z.B. in dt. *Laken* — /la.k[_{NN}]/ oder engl. *bottle* — /bot.t[_N]/ ([_N] indiziert, dass ein Segment im Nukleus steht). Auf diese Weise können quasi-homophone Wortpaare wie dt. *kann* und *Kann(e)n*, dt. *sein* und *sein(e)n* oder engl. *plight* und *p(o)lite* entstehen. Welche Silbenstruktur für eine Form wie *Kannen* anzusetzen ist, die Struktur /ka.[_{NN}]/, /kan.[_{NN}]/ oder /kan.[_{NN}n]/, wird hier nicht diskutiert. Die Beantwortung dieser Frage erfordert eine genauere artikulatorisch-phonetische und akustisch-phonetische Untersuchung dieser Formen und des Prozesses der Schwa-Elision im allgemeinen.

2.4.2 CHOMSKY & HALLE (1968)

CHOMSKY UND HALLE, die im *Sound Pattern of English* eine segmentbasierte Beschreibung des phonologischen Systems des Englischen durchführen, erwähnen das Merkmal *Sonorität* nur im Zusammenhang mit dem Wortakzent:

Suppose that we were to define a scale of 'sonority' in such a way that more heavily stressed vowels are greater in sonority than less heavily stressed vowels, and that all vowels are greater in sonority than consonants and boundaries (CHOMSKY & HALLE 1968:91)

Akzentuierte Vokale haben demnach mehr Sonorität als unakzentuierte Vokale, Konsonanten und Silbengrenzen haben minimale Sonorität. Mit *Sonorität* meinen CHOMSKY & HALLE hier also etwas anderes als der Sonoritätsbegriff, der aus der Sonoritätshierarchie abgeleitet wurde; ihr Sonoritätskonzept wird jedoch weiter unten (s.u. 3.2.2) wiederaufgegriffen. CHOMSKY & HALLE (1968:303) führen drei *phonetische*⁷⁰ Oberklassenmerkmale ein: die Merkmale [\pm sonorant]⁷¹, [\pm konsonantisch]⁷² und [\pm vokalisch]⁷³. Mit diesen drei Merkmalen werden fünf Oberklassen definiert: Vokale, Gleitlaute, Liquide, Nasalkonsonanten und Obstruenten. Die Funktion dieser Oberklassen beschreiben CHOMSKY & HALLE folgendermaßen:

Reduced to the most rudimentary terms, the behavior of the vocal tract in speech can be described as an alternation of closing and opening. During the closed phase the flow of air from the lungs is either impeded or stopped, and pressure is built up in the vocal tract; during the open phase the air flows out freely. This skeleton of speech production provides the basis for the major class features, that is, the features that subdivide speech sounds into vowels, consonants, obstruents, sonorants, glides, and liquids. Each of the three major class features – sonorant, vocalic, consonantal – focuses on a different aspect of the open-versus-closed phase. (CHOMSKY & HALLE 1968:302)⁷⁴

70 Im phonologischen Merkmalsystem definieren CHOMSKY & HALLE (1968:68) auf der Basis der zwei Oberklassenmerkmale [\pm vokalisch] und [\pm konsonantisch] vier Oberklassen: Vokale, Gleitlaute, Liquide und Konsonanten. Diese Einteilung in vier Lautklassen ist jedoch nicht geeignet, um auf die Sonoritätshierarchie bezogen zu werden, da Nasale mit den Plosiven und Frikativen eine Klasse bilden, obwohl sie sich phonotaktisch anders verhalten.

71 [\pm sonorant]: „Sonorants are sounds produced with a vocal tract cavity configuration in which spontaneous voicing is possible; obstruents are produced with a cavity configuration that makes spontaneous voicing impossible.“ (Chomsky & Halle 1968:302).

72 [\pm vocalic]: „Vocalic sounds are produced with an oral cavity in which the most radical constriction does not exceed that found in the high vowels [i] and [u] and with vocal cords that are positioned so as to allow spontaneous voicing; in producing nonvocalic sounds one or both of these conditions are not satisfied.“ (CHOMSKY & HALLE 1968:302).

73 [\pm consonantal]: „Consonantal sounds are produced with a radical obstruction in the midsagittal region of the vocal tract; nonconsonantal sounds are produced without such an obstruction.“ (CHOMSKY & HALLE 1968:302).

74 Die in diesem Zitat aufgezählten *sechs* Lautklassen sind nicht als Oberklassen, sondern als die durch die drei Oberklassenmerkmale definierten möglichen natürlichen Klassen zu verstehen. *Konsonanten* bilden in CHOMSKY & HALLES Oberklassensystem keine Oberklasse, sondern umfassen Liquide, Nasale und Oralkonsonanten. Sonoranten umfassen alle Oberklassen außer die Obstruenten bzw. Oralkonsonanten.

Die Oberklassen werden von CHOMSKY & HALLE also zunächst artikulatorisch gekennzeichnet und auf das grundlegende, sich immer wiederholende Muster der Artikulationsbewegungen bezogen, das den Silbenrhythmus konstituiert. Die Oberklassenmerkmale entsprechen nach dieser Auffassung den nicht-distinktiven, undifferenzierten Eigenschaften der Artikulationsbewegung. Gleichzeitig findet sich auch hier das Konzept des *Zyklus*: Während CLEMENTS (1990) die Silbe als *Sonoritätszyklus* beschreibt, beziehen CHOMSKY & HALLE die Oberklassenmerkmale auf den *Zyklus* des Öffnens und Schließens des Vokaltraktes. CHOMSKY & HALLE verwenden in diesem Kontext zwar an keiner Stelle den Begriff Sonorität, beziehen das Konzept der Oberklasse aber auf ähnliche Weise auf die Form der Silbe, wie das Konzept der Sonorität oben auf die Form der Silbe bezogen wurde.

Die Ausgangsfrage dieses Abschnittes ist die, ob das Merkmal Sonorität auf die Oberklassenmerkmale zurückzuführen, oder auch im Hinblick auf die Silbe zu definieren ist. Wie ist also die Beziehung zwischen Oberklassensystem und Silbe bei CHOMSKY & HALLE zu verstehen? Ist das Oberklassensystem so zu interpretieren, dass die Zugehörigkeit eines Segmentes zu einer Oberklasse nur von der Position in der Artikulationseinheit abhängt, an der es realisiert wird, d.h. im Silbennukleus oder am Silbenrand? Oder ist es eher so zu verstehen, dass die Elemente einer Oberklasse prototypischerweise, aber nicht notwendigerweise, in einer bestimmten Position in der Silbe realisiert werden? Sind die Oberklassen also an die Silbenstrukturpositionen gebunden, oder sind sie unabhängig von diesen definiert?

Um diese Frage zu beantworten, soll eine zweite Version des Oberklassenmerkmalsystems betrachtet werden, die CHOMSKY & HALLE (1968:354) später einführen. Dabei ersetzen sie das Merkmal [\pm vokalisch] durch das Merkmal [\pm silbisch]:

the feature 'vocalic' might be replaced by a feature 'syllabic' which would characterize all segments constituting a syllabic peak. Obstruents would by definition be excluded from forming syllabic peaks; vowels would normally be syllabic peaks, whereas the remaining sonorants – i.e., liquids, glides, nasal consonants – would normally be nonsyllabic, but could become syllabic under special circumstances (CHOMSKY & HALLE 1968:354)

Diese Revision hat zur Folge, dass das Oberklassensystem zwei weitere Elemente bekommt, nämlich die silbischen Liquide und die silbischen Nasale (CHOMSKY & HALLE 1968:354). Die Gleitlaute, die sich in der ersten Fassung des Oberklassensystems nur durch das Merkmal [– vokalisch] von den Vokalen unterschieden, erscheinen hier explizit als [– silbische] Entsprechungen der Vokale. Das Oberklassensystem von CHOMSKY & HALLE umfasst nun sieben Oberklassen: Vokale, silbische Liquide, silbische Nasale, Gleitlaute, nichtsilbische Liquide, nichtsilbische Nasale und Obstruenten (d.h. Frikative und Plosive). Dieses Oberklassensystem ist unmittelbar auf die Silbenstruktur bezogen, da alle Segmente, die in einer silbifizierten Lautkette im Silbennu-

kleus stehen, [+ silbisch] sind und alle Segmente, die am Silbenrand stehen, [– silbisch] sind. Dies bedeutet, dass die Zugehörigkeit der Vokale bzw. Gleitlaute, der Liquide und der Nasale zu der einen oder anderen Oberklasse von ihrer nuklearen oder nicht-nuklearen Position in einer Silbe abhängt.

Das obenaufgeführte Zitat „vowels would normally be syllabic peaks“ enthält aber noch einen anderen Gedanken: Vokale bilden normalerweise Silbennuklei, Liquide und Nasale stehen normalerweise am Silbenrand. Mit dieser Äußerung drücken CHOMSKY & HALLE gültige phonotaktische Verallgemeinerungen aus, doch wird darin das Merkmal [silbisch] in gewisser Weise von den beiden anderen Oberklassenmerkmalen [sonorant] und [konsonantisch] abgehoben, da dieser Aussage zufolge bestimmte Kombinationen aus diesen beiden Merkmalen, wie z.B. [+ sonorant, – konsonantisch] (d.h. Vokale) häufiger zusammen mit dem Wert [+ silbisch] vorkommen, andere wie [+ sonorant, + konsonantisch] (d.h. Nasale) häufiger mit dem Wert [– silbisch], wieder andere wie die Obstruenten zumindest im Englischen nie. Eine solche Konzeption von [silbisch] impliziert, dass den ‘normalerweise silbischen’ Lauten eine Eigenschaft zukommt, welche die ‘normalerweise nichtsilbischen’ Laute und (erst recht die niemals silbischen Obstruenten) nicht besitzen. Diese Eigenschaft scheint dieselbe oder eine ähnliche zu sein wie die, welche der Sonoritätshierarchie zugrundeliegt. Diese Frage wird in Abschnitt 2.4.4. wiederaufgegriffen.

Wie lassen sich diese Ausführungen im Hinblick auf die Ausgangsfrage zusammenfassen, ob das Merkmal Sonorität auf die Oberklassenmerkmale zurückzuführen oder nur unter Berücksichtigung der Silbenstruktur und der Silbenstrukturposition eines Segmentes zu definieren ist? Das phonetische Oberklassensystem von CHOMSKY & HALLE ist ausreichend differenziert, um die phonotaktischen Regularitäten des Englischen zu beschreiben, d.h. es kann auf die Stufen der Sonoritätshierarchie bezogen werden (abgesehen davon, dass Frikative und Plosive in eine einzige Oberklasse fallen). Das System der Oberklassen wird von CHOMSKY & HALLE auf die Form der Silbe als Zyklus von Öffnungs- und Verschlussbewegungen bezogen, was in Analogie zu CLEMENTS (1990) Konzeption dem Auf und Ab der Sonorität entspräche. Sonorität lässt sich in dieser Konzeption also prinzipiell auf Oberklassenmerkmale zurückführen, diese werden aber im Hinblick auf die Silbe als Einheit der Artikulation konzipiert und jedes der drei Oberklassenmerkmale [konsonantisch], [sonorant] und [vokalisch] wird auf eine Phase der Artikulation bezogen: den Verschlussmoment, die Öffnungs- bzw. Schließbewegung und den Moment der maximalen Offenheit.⁷⁵ Noch deutlicher wird die Korrelation von Oberklassensystem und Form der Silbe nach der Ersetzung des

⁷⁵ Vgl.: „Each of the three major class features [...] focuses on a different aspect of the open-versus-closed phase.“ (CHOMSKY & HALLE 1968:302)

Merkmals [vokalisch] durch [silbisch]. Bildet man aus den Oberklassen im System von CHOMSKY & HALLE also eine Sonoritätshierarchie (was auf der Basis von phonotaktischen Überlegungen möglich ist), so führt dies *nach* der Einführung des Merkmals [silbisch] dazu, dass man entsprechend der Verdoppelung der Klasse der Nasale und Liquide als silbische und nicht-silbische Nasale und Liquide, die sich auf diese Weise ergibt, auch annehmen muss, dass silbische Nasale eine höhere Stellung in der Hierarchie haben als nichtsilbische, silbische Liquide eine höhere als nichtsilbische usw., allgemein formuliert, dass silbische Segmente sonoranter sind als nichtsilbische. Dieser Aspekt erinnert an die oben gemachte Bemerkung, dass es ‘normalerweise silbische’ und ‘normalerweise nichtsilbische’ Segmente gibt, die sich in einer Eigenschaft unterscheiden, welche in Relation zur Silbenstruktur steht und möglicherweise mit dem Merkmal Sonorität identisch ist. Abschließend lässt sich also sagen, dass zumindest in einem Merkmalssystem, welches das Merkmal [silbisch] einschließt, Sonorität *nicht* unabhängig von der Stellung des Segmentes in der Silbe auf inhärente segmentale Merkmale zurückgeführt werden kann.

2.4.3 SELKIRK (1984)

Das Merkmal [\pm silbisch] bildet den Ausgangspunkt für SELKIRKS (1984) Kritik an dem Konzept der Oberklassenmerkmale, denn dieses Merkmal verdoppelt die Zahl der Oberklassen und ist doch gleichzeitig aus der Silbenstrukturposition eines Segmentes ableitbar:

the feature *syllabic* suggests itself as an especially obvious candidate for elimination from phonological theory. [...] whether a segment is ‘syllabic’ depends on its position in a syllable, not on any inherent phonological property of its own. (SELKIRK 1984:108)

Will man auf der Basis der Oberklassenmerkmale aber phonotaktische Regularitäten formulieren, so kann man auf das Merkmal [silbisch] nicht ohne weiteres verzichten. Zunächst muss darauf hingewiesen werden, dass sich Vokale wie /i/ und /u/ von Gleitlauten wie /w/ (wie engl. *we*) und /j/ (wie engl. *you*) nur darin unterscheiden, dass Vokale im Silbennukleus stehen, Gleitlaute aber am Silbenrand. Dies bedeutet, dass sich Vokale und Gleitlaute nur darin unterscheiden, dass erstere [+ silbisch] sind und letztere [– silbisch]. Die Folge /i/ + /a/ wird daher entweder als zwei Silben /i.a./ (wie in span. *día*) ausgesprochen, oder als Gleitlaut-Vokal-Folge /ja./ (wie in dt. *ja*). Andererseits ist es so, dass nur die [+ hohen] Vokale /i/ und /u/⁷⁶ als Gleitlaute /j/ und /w/ realisiert wer-

⁷⁶ Gleiches gilt für ihre gerundeten bzw. ungerundeten Entsprechungen, z.B. in frz. *huit* oder *cuisine* oder engl. *road*, *rye*.

den können, aber nicht die [– hohen] Vokale /a/, /e/ und /o/.⁷⁷ Daher müssen in einer Theorie der Phonotaktik, wenn man die Gleitlaute nicht explizit als [– silbische] Entsprechungen der [+ silbischen] Vokale aufführen will, die hohen Vokale als eigene Klasse aufgeführt werden, denn sie können, anders als die nicht-hohen Vokale, auch am Silbenrand stehen und [– silbisch] sein. SELKIRK schlägt also vor, die Theorie der phonotaktischen Beschränkungen auf der Basis einer Sonoritätshierarchie der Segmente zu konzipieren, anstatt sie auf der Grundlage der Zugehörigkeit von Segmenten zu Oberklassen zu formulieren:

My proposal is that the major class features be eliminated entirely from a theory of the phonotactics of the syllable [...] and that they be replaced in effect by the sonority hierarchy and the assignment of a sonority index to individual segments that reflects the niche they occupy in that hierarchy. [...] In other words, I propose that there is a single, n-ary feature, call it [n sonority], that is at play in language, where the feature specification n is the sonority index. (SELKIRK 1984:110)

Dies bedeutet also, dass jedem Segment ein Merkmalswert für das Merkmal *Sonorität* zugewiesen wird, der seine Position auf der Sonoritätsskala bzw. seine Stellung in der Sonoritätshierarchie bestimmt. Diese Hierarchie hat die Funktion, phonotaktische Beschränkungen zusammenzufassen. SELKIRK geht von der sehr differenzierten elfstufigen Sonoritätshierarchie in (16) aus, mit Werten von 0.5 bis 10 für [n Sonorität].

Die Frikative werden auf drei Stufen der Hierarchie verteilt, entsprechend der Klassifizierung als stimmlose Frikative, stimmhafte Frikative und Sibilanten. Auch die Vokalklassen erhalten entsprechend der Klassifizierung, die sich durch die Zungenhöhe ergibt, drei verschiedene Sonoritätswerte.⁷⁸

(16)	a	10
	e, o	9
	i, u	8
	r	7
	l	6
	m, n	5
	s	4
	v, z, ʃ	3
	f, v	2
	b, d, g	1
	p, t, k	0.5

Das Problem, dass nicht alle Vokale [– silbisch] sein können, wird in dieser Hierarchie gelöst, indem der natürlichen Klasse⁷⁹ der [+ hohen] Vokale ein niedrigerer Sonoritäts-

⁷⁷ Vgl. auch CHOMSKY & HALLE (1968:353ff) zur Motivation des Merkmals [silbisch], ebenso CLEMENTS & KEYSER 1983:98ff.

⁷⁸ Phonetisch lässt sich die Zungenhöhe, d.h. die Auslenkung der Zunge in Richtung Gaumen, als hoch, halbhoch und tief definieren, phonologisch durch die Merkmale [hoch] und [tief]: /a/ ist [– hoch, + tief], /e/ und /o/ sind [– hoch, – tief], /i/ und /u/ [+ hoch, – tief].

⁷⁹ *Natürliche Klassen* sind Klassen von Segmenten, die mindestens ein gemeinsames Merkmal haben, z.B. die Klasse der alveolaren Konsonanten ([+ alveolar]), die Klasse der Vokale ([+ vokalisches]), die

wert zugewiesen wird als der natürlichen Klasse der [– hohen] Vokale. Einigen anderen natürlichen Klassen von Segmenten, die auf der Basis distinktiver Merkmale definiert sind, z.B. der natürlichen Klasse der Segmente, die [+ nasal] sind, kommt gleichfalls derselbe Sonoritätswert zu:

it seems that members of certain natural classes of segments, defined in terms of nonmajor class features such as [continuant], [voice], [nasal], [high], etc. are so alike in sonority to make distinctions among them irrelevant for most descriptive purposes. (SELKIRK 1984:110)

Dennoch führt SELKIRK das Merkmal [n Sonorität] nicht auf die distinktiven Merkmale zurück, wie z.B. [+ hoch], [+ lateral], [– kontinuant] usw., obwohl dies für die von ihr angegebene Sonoritätshierarchie möglich wäre, da jede Stufe dieser Hierarchie einer natürlichen Klasse von Segmenten entspricht. Allgemein betrachtet haben Konzeptionen, die Sonorität auf *distinktive* Merkmale zurückführen,⁸⁰ gegenüber solchen, die mit *Oberklassenmerkmalen* operieren, den Vorteil, dass sie, ähnlich wie die Hierarchie von SELKIRK, mehr Sonoritätsstufen abbilden können und so die phonotaktischen Regularitäten einzelner Sprachen vollständiger beschreiben können.

SELKIRK ist jedoch der Auffassung, dass das Merkmal *Sonorität* weder auf Oberklassenmerkmale noch auf andere phonologische Merkmale zurückgeführt werden sollte. Der Grund hierfür ist, dass in ihrer Konzeption das Merkmale [n Sonorität] zwar ein phonologisches Merkmal ist und als solches in Implikationsbeziehungen zu anderen phonologischen Merkmalen steht,⁸¹ dass dieses phonologische Merkmal aber primär ein phonetisches Merkmal abbildet:

[n sonority] may be thought of as a feature representing the phonetic dimension of sonority [...] There is clearly a phonetic basis for it, probably corresponding in part to simple 'loudness'. (SELKIRK 1984:111)

SELKIRK vertritt demnach die Annahme, dass dem phonologischen Merkmal Sonorität ein phonetisches Korrelat entspricht und dass dieses phonetische Korrelat (akustischer und / oder artikulatorischer Natur) eine so wesentliche Funktion für die Sequenzierbarkeit von Segmenten in der Silbe erfüllt, dass das Merkmal Sonorität nicht auf natürliche Klassen zurückgeführt werden sollte, die durch unabhängig motivierte distinktive

Klasse der Frikativkonsonanten ([+ konsonantisch, – sonorant, + kontinuant]), usw. *Oberklassen* sind demnach natürliche Klassen von Segmenten, die dieselbe Merkmalswertkombination für die Oberklassenmerkmale haben.

⁸⁰ Vgl. GIEGERICH (1992:152) für das Englische und WIESE (1996:261) für das Deutsche, die eine sechsstufige Sonoritätshierarchie auf die Merkmale [obstruent], [nasal], [kontinuant], [konsonantisch] und [hoch] zurückführen, vgl. auch BLEVINS (1995:211), die eine neunstufige Hierarchie konzipiert. DOGIL & LUSCHÜTZKY (1989) führen das Merkmal Sonorität nicht auf Merkmalspezifikationen für phonologische Merkmale, sondern auf die geometrische Konfiguration des Merkmalsbaums für einzelne Lautklassen zurück.

⁸¹ Z.B. in Implikationsrelationen der Art [+ nasal] ⇒ [5 Sonorität], [– stimmhaft, – kontinuant] ⇒ [0.5 Sonorität]. Auf der Basis dieser Implikationsbeziehungen, die für alle n-Werte formuliert werden können, kann [n Sonorität] auf andere phonologische Merkmale zurückgeführt werden.

Merkmale konstituiert werden, sondern dass das Merkmal Sonorität seine eigene Berechtigung in einer Theorie der Phonotaktik der Silbe hat. SELKIRK liefert also zwei Argumentstränge: erstens argumentiert sie gegen eine Rückführung von Sonorität auf die phonologischen *Oberklassenmerkmale*, weil hierzu das Merkmal [silbisch] nötig ist, das aus der Silbenstruktur abgeleitet werden kann. Zweitens ist sie der Meinung, dass eine Erklärung von Sonorität nicht bei den phonologischen *distinktiven Merkmale* haltmachen sollte, weil die Annahme von Sonoritätsabstufungen nicht nur zur Formulierung von phonotaktischen Regularitäten geeignet ist, sondern darüber hinaus eine phonetische Basis hat.

SELKIRKS Ausführungen sind also so zu verstehen, dass sich Sonorität zwar durch eine Hierarchie der durch die distinktiven phonologischen Merkmale konstituierten natürlichen Klassen erklären ließe, dass aber zu bedenken ist, dass die phonologische Sonorität ein phonetisches Merkmal abbildet und daher nicht durch andere Merkmale erklärt werden sollte. Hierzu ist jedoch zu bemerken, dass sich auch Vokale und Gleitlaute in phonetischer Hinsicht unterscheiden und zwar sowohl in der Lautstärke als auch im Öffnungsgrad des Vokaltraktes. Nach FRY (1979:127) haben die Gleitlaute /j/ und /w/ im Englischen eine geringere Intensität als ihre vokalischen Entsprechungen /i/ und /u/. In artikulatorischer Hinsicht zeichnen sich Gleitlaute dadurch aus, dass sie im Vergleich zu Vokalen mit einer stärkeren Verengung im Vokaltrakt artikuliert werden (POMPINO-MARSCHALL 1995:193). Definiert man das phonologische Merkmal Sonorität also als Funktion eines phonetischen Merkmals, das der Lautstärke bzw. Intensität und dem Mundöffnungsgrad entspricht, so sollte der Tatsache Rechnung getragen werden, dass sich Gleitlaute und Vokale in phonetischer Hinsicht unterscheiden, indem ihnen ein unterschiedlicher phonologischer Sonoritätswert zugewiesen wird.

2.4.4 CLEMENTS (1990)

Im Gegensatz zu SELKIRK (1984), die die Auffassung vertritt, dass das Merkmal Sonorität nicht auf Oberklassenmerkmale zurückgeführt werden sollte, argumentiert CLEMENTS (1990) im *Sonority-Cycle*-Modell für eine Definition von Sonorität auf der Basis von unabhängig davon motivierten und definierten Oberklassenmerkmalen:

I will suggest that sonority is not a single, multivalued property of segments, but is derived from more basic binary categories, identical to the major class features of standard phonological theory (Chomsky & Halle 1968) supplemented with the feature 'approximant'. (CLEMENTS 1990:284)

Die vier Oberklassen Obstruent, Nasal, Liquid und Gleitlaut/Vokal definiert CLEMENTS (1990:292) mittels der drei Oberklassenmerkmale $[\pm \text{vokoid}]^{82}$, $[\pm \text{approximant}]^{83}$ und $[\pm \text{sonorant}]^{84}$. Diesen ‘echten’ Oberklassenmerkmalen („true major class features“, CLEMENTS 1990:295, s.u.) fügt er das Merkmal $[\pm \text{silbisch}]$ hinzu und rekonstruiert auf diese Weise die fünfstufige Sonoritätshierarchie in (17). Je mehr $[+]$ -Spezifikationen eine Oberklasse aufweist, desto höher ist ihr Rang in der Sonoritätshierarchie.

(17)⁸⁵

Obstruent	<	Nasal	<	Liquid	<	Gleitlaut	<	Vokal
–		–		–		–		+ silbisch
–		–		–		+		+ vokoid
–		–		+		+		+ approximant
–		+		+		+		+ sonorant

Gleitlaute und Vokale unterscheiden sich in diesem System wie bei CHOMSKY & HALLE (1968:354) nur im Merkmal $[\text{silbisch}]$. Das Merkmal $[\text{silbisch}]$ wird von Clements als silbenstrukturbezogenes Merkmal definiert, das einem Segment auf Grund seiner Stellung in der Silbe zukommt:

[syllabic] has a different status in feature representation than the true major class features, not functioning as a feature but as a prosodically defined position within the syllable (CLEMENTS 1990:295)

‘Syllabic’ can be interpreted as referring to the prosodic distinction between V and C elements of the timing tier (CLEMENTS 1990:292)

Diese Konzeption des Merkmals $[\pm \text{silbisch}]$ geht zurück auf CLEMENTS’ & KEYSERS *CV-Phonology* (1983). In dieser Theorie der Silbe liegt zwischen Silbenebene und segmentaler Ebene die CV-Ebene [vgl. (18)], die aus Folgen von Nuklei (V-Knoten) und Non-Nuklei (C-Knoten) besteht:⁸⁶

The elements of the CV-tier distinguish between syllable peaks and syllable non-peaks (or syllable margins). Specifically, any segment dominated by V is interpreted as a syllable peak, and any segment dominated only by C is interpreted as a non-peak. (CLEMENTS & KEYSER 1983:8)

In this respect the CV-tier can be seen as subsuming the function of the earlier feature category [syllabic]. (CLEMENTS & KEYSER 1983:10)

82 CLEMENTS (1990:293) konzipiert $[\pm \text{vokoid}]$ als Gegenteil von $[\pm \text{contoid}]$, d.h. $[+ \text{vokoid}]$ fasst „all central resonants“ zusammen (LAVER 1995:147). Das Merkmal $[\pm \text{contoid}]$ geht auf PIKE (1943) zurück.

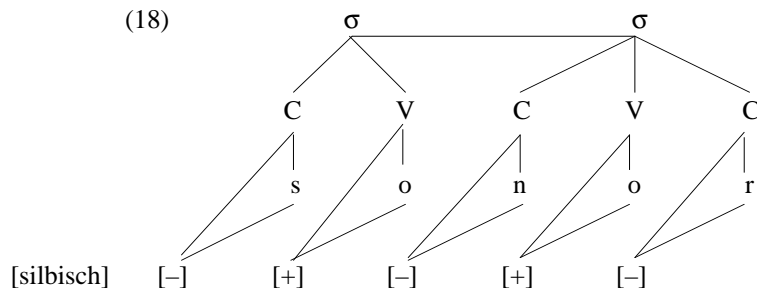
83 Ausgehend von LADEFOGED (1982, 1993²) wird $[+ \text{approximant}]$ definiert als „any sound produced with an oral tract stricture open enough so that airflow through it is turbulent only if it is voiceless“ (CLEMENTS 1990:293).

84 Zur Definition von $[\pm \text{sonorant}]$ vgl. CHOMSKY & HALLE (1968:302), s.o. 2.4.2, Anm.

85 Diese Darstellung ist eine Kombination aus zwei getrennten Darstellungen für silbische und nichtsilbische Segmente, die CLEMENTS aufführt (1990:292 und 294).

86 Diese Konzeption macht die Repräsentation der internen Struktur der Silbe als bestehend aus Anlaut, Reim, Nukleus und Coda überflüssig, weil diese Konstituenten aus der CV-Struktur abgeleitet werden können, vgl. CLEMENTS & KEYSER 1983:15ff.

Das Merkmal [silbisch] ist demnach ein Merkmal, das Segmenten auf Grund ihrer Assoziation mit einem C-Knoten oder V-Knoten auf der Ebene der Silbenstruktur zukommt. Je nachdem, wie eine Segmentkette einer CV-Struktur zugeordnet ist, wird ein Segment von einer V-Position dominiert und als Nukleus bzw. als [+ silbisch] interpretiert, oder von einer C-Position dominiert und als Silbenrand bzw. als [– silbisch] interpretiert [vgl. (18)]. Die Segmente, die von V-Knoten dominiert werden und [+ silbisch] sind, können demnach Vokale oder silbische Konsonanten sein, die Segmente, die von C-Knoten dominiert werden und [– silbisch] sind, sind Gleitlaute oder unsilbische Konsonanten.⁸⁷ Vokale und Gleitlaute bilden demnach *keine* verschiedenen Oberklassen, sondern unterscheiden sich auf dieselbe Weise, wie sich auch silbische und nichtsilbische Nasalkonsonanten unterscheiden.



Aus der Sonoritätshierarchie in (17) könnte man mit CLEMENTS also versuchen, das Merkmal [silbisch] herauszunehmen, da es in segmental-phonologischer Hinsicht überflüssig ist: ob ein Segment [+ silbisch] oder [– silbisch] ist, hängt allein davon ab, ob es mit einem C-Knoten oder einem V-Knoten assoziiert ist. Auf diese Weise gelangte man zu der vierstufigen Sonoritätshierarchie in (19), in der Vokale und Gleitlaute eine Klasse bilden und silbische und unsilbische Segmente nicht mehr unterschieden werden. Der Sonoritätswert eines Segmentes wäre in dieser Hierarchie unabhängig von seiner Silbizität bestimmbar, was deshalb plausibel wäre, weil erstens jedes Segment silbisch sein kann und weil die Silbizität eines Segmentes zweitens aus seiner Zuordnung zu einem C-Knoten oder V-Knoten ableitbar ist.

(19)

Obstruent	<	Nasal	<	Liquid	<	Gleitlaut/Vokal
–		–		–		+
–		–		+		+
–		+		+		+
						vokoid
						approximant
						sonorant

Die Elemente einiger Oberklassen belegen jedoch häufiger nukleare Positionen als die anderer Oberklassen. Vokale sind in allen Sprachen prototypische Silbennuklei. Silbi-

⁸⁷ In der in (17) aufgeführten vorläufigen Sonoritätshierarchie kommen zunächst keine silbischen Liquide oder Nasale vor, ob diese doch auch silbisch sein können.

sche Liquide und Nasale gibt es zwar im Deutschen und im Englischen, jedoch nur in Wörtern, die auch eine Variante mit unsilbischem Liquid bzw. Nasal und unbetontem Vokal aufweisen, wie z.B. *Teuf(e)l*, *Lapp(e)n*, *bottl(e)*. Silbische Obstruenten finden sich nur in sehr wenigen Sprachen. Daher fügt CLEMENTS hinzu, dass die 'Fähigkeit eines Segmentes, als Silbennukleus zu fungieren', d.h. [+ silbisch] zu sein, in Relation zu seinem Sonoritätswert bzw. seiner Oberklassenzugehörigkeit steht:

In principle, any segment can occupy the syllable peak, but the ability of a given segment to function as a syllable peak is related to its rank on the sonority scale. (CLEMENTS 1990:294)

Mit zunehmender Sonorität wächst demnach die Wahrscheinlichkeit, mit einer V-Position assoziiert und als [+ silbisch] spezifiziert zu sein. Außerdem ist die Klasse der [+ silbischen] Vokale nicht identisch mit der Klasse der [- silbischen] Gleitlaute,⁸⁸ zudem verhalten sich silbische Liquide und Nasale anders, als auf der Grundlage der Stellung der nichtsilbischen Liquide und Nasale in der Sonoritätshierarchie zu erwarten wäre (vgl. CLEMENTS 1990:294). Aus diesen Gründen muss nach CLEMENTS' Auffassung das Merkmal [silbisch] in die Sonoritätshierarchie integriert werden. CLEMENTS nimmt daher an, dass das Merkmal [silbisch] zwar zu den Oberklassenmerkmalen gehört, dass es aber dennoch einen anderen Status hat als [sonorant], [approximant] und [vokoid]:

the special status of the feature [syllabic] in the sonority scale: it is the only major class feature that cross-classifies all others. (CLEMENTS 1990:295)

Auf der Basis der Annahme, dass das Merkmal [silbisch] die anderen Oberklassenmerkmale kreuzklassifiziert, kommt CLEMENTS zu der zweidimensionalen Sonoritätshierarchie in (20): die erste Dimension wird von den Oberklassenmerkmalen, die zweite Dimension vom Merkmal [silbisch] konstituiert.

(20) ⁸⁹	[+ silbisch]	Obstruent	<	Nasal	<	Liquid	<	Vokal
		∨		∨		∨		∨
	[- silbisch]	Obstruent	<	Nasal	<	Liquid	<	Gleitlaut

Was bedeutet diese Zweidimensionalität im Hinblick auf das Merkmal Sonorität und die Stellung der Oberklassen in der Sonoritätshierarchie? CLEMENTS formuliert das so:

a single step to the right or up results in a one-degree increase in sonority rank. (CLEMENTS 1990:295)

Dies bedeutet, dass sich einerseits die 'echten' Oberklassen in ihrer Sonorität unterscheiden und dass andererseits silbische Segmente sonoranter sind als unsilbische Segmente. Zu bemerken ist, dass nicht ganz klar wird, ob eine totale Ordnung der Klassen

⁸⁸ Dieses Argument wird von CLEMENTS nicht explizit angeführt, vgl. aber SELKIRK (1984).

⁸⁹ Darstellung nach CLEMENTS 1990:295.

vorliegen soll, d.h. hier, ob eine Sonoritätsgleichheit mitzuverstehen ist, so dass z.B. ein silbischer Nasal den gleichen Sonoritätswert besitzt wie ein unsilbischer Liquid.

Die Ausführungen von CLEMENTS zum Konzept der Sonorität lassen sich folgendermaßen zusammenfassen:

- Sonorität ist auf andere Merkmale zurückzuführen und zwar einerseits auf die drei Oberklassenmerkmale, die vier Oberklassen konstituieren und andererseits auf das aus der CV-Struktur abgeleitete Merkmal [silbisch].
- Die Sonoritätsbeziehungen zwischen den Segmentklassen sind so konzipiert, dass erstens die Sonorität mit der Zahl der [+]Spezifizierungen für die drei Oberklassenmerkmale wächst und zweitens mit der Silbizität, insofern silbische Segmente sonoranter sind als unsilbische Segmente.

Daneben müssen auch in CLEMENTS' Modell zwei Bedeutungen des Konzeptes *Silbizität* hervorgehoben werden, die auf die wechselseitige Beziehung zwischen *Silbizität* und *Sonorität* hinweisen. Einerseits ist es so, dass die 'Fähigkeit eines Segmentes, als Nukleus zu fungieren', die man vielleicht 'potentielle Silbizität' nennen könnte, mit zunehmender Sonorität wächst, andererseits sind 'aktuell silbische' Segmente sonoranter als 'aktuell unsilbische' Segmente.⁹⁰

Die Ausgangsfrage dieses Abschnittes war, ob Sonorität sich auf segmentale Merkmale zurückführen lässt, oder ob Sonorität ein nichtsegmentales Merkmal ist, das nicht ohne Bezug auf die Silbenstruktur und die Silbenstrukturposition eines Segmentes definiert werden kann. In Clements Konzeption gilt beides: erstens ist Sonorität auf die Oberklassenmerkmale zurückzuführen, zweitens muss das aus der Silbe abgeleitete Merkmal [silbisch] hinzugenommen werden, wenn alle Sonoritätsbeziehungen zwischen den Segmenten in einer Silbe beschrieben werden sollen. Clements Ansatz also konzipiert das Merkmal Sonorität als Verbindungsglied zwischen *segmentalen* Oberklassenmerkmalen und *suprasegmentalen* Silbenstrukturposition.

⁹⁰ An dieser Stelle könnte man zu Recht fragen, ob auch unsilbifizierte Segmente existieren, bzw. ob es sinnvoll ist, über phonologische Repräsentationen zu reden und dabei von der Ebene der Silbenstruktur zu abstrahieren. Noch anders ausgedrückt: ist es überhaupt sinnvoll, anzunehmen, dass es nicht nur [+ silbische] und [- silbische] Nasale gibt, sondern auch Nasale ohne Spezifizierung für das Merkmal [silbisch]? Man könnte vielleicht annehmen, dass die Spezifizierung als [- silbisch] für alle Segmente die unmarkierte ist, oder dass die Spezifizierung [- silbisch] für alle [- vokoiden] Segmente die unmarkierte ist und die Spezifizierung [+ silbisch] für alle [+ vokoiden] Segmente: [- vokoid] ⇒ [- silbisch]. Die Beantwortung dieser Frage erfordert eine Klärung des Status des Merkmals [silbisch] in der phonologischen Theorie, die, wie aus CLEMENTS' Ausführungen, in denen [silbisch] mal als Oberklassenmerkmal, mal als 'prosodische Position' fungiert, ersichtlich ist, nicht leicht herbeizuführen sein wird.

2.4.5 Zusammenfassung

Wie die vorangegangenen Ausführungen gezeigt haben, können zwei Zugangsweisen zum Konzept der *Sonorität* unterschieden werden, die sich in den drei oben zusammengefassten Konzeptionen von Sonorität und Oberklassenmerkmalen jedoch vermischen. Die erste Zugangsweise gründet auf der Annahme, dass *Sonorität* ein primär *phonologisches Merkmal* ist, dessen Funktion darin besteht, *phonotaktische Regularitäten* zusammenzufassen. Der Sonoritätswert eines Segmentes wird in dieser Konzeption aus den Sequenzierungsmöglichkeiten der Einzellaute abgeleitet und ein eventuelles phonetisches Korrelat des phonologischen Merkmals Sonorität müßte gesondert ermittelt werden. In einer Variante versucht man, den Sonoritätsbegriff vollständig auf segmentale Merkmale zurückzuführen. Alternativ kann man, wie SELKIRK (1984) vorschlägt, ein skalares Merkmal [n Sonorität] verwenden, das als Abkürzung für segmentale Merkmalskonfigurationen dient, die natürliche Klassen von Segmenten mit demselben Sonoritätswert definieren, wie z.B. die Klasse der Nasalkonsonanten oder der hohen Vokale. Nutzt man dieses Konzept der Sonorität, um phonotaktische Regularitäten zu beschreiben, so benötigt man weder das Merkmal [silbisch], weil dieses aus der Silbenstruktur abgeleitet werden kann, noch die Annahme, dass silbische Segmente sonoranter sind als unsilbische.⁹¹ In dieser Konzeption ist in Folgen wie /ju./ oder /wi./, die aus zwei hohen Segmenten bestehen,⁹² der Silbennukleus nicht sonoranter als der Silbenanlaut, weil beide Segmente denselben inhärenten Sonoritätswert haben und sich nur in der Assoziation zu einer C- bzw. V-Position bzw. im Merkmal [silbisch] unterscheiden.⁹³ Die einzelsprachliche und sprachübergreifende Gültigkeit einer Sonoritätshierarchie, die auf diesem Sonoritätsbegriff basiert, hängt ab von der Regularität der Phonetik in einer Sprache und von der phonotaktischen Ähnlichkeit verschiedener Sprachen. Einerseits folgt aus dieser Konzeption, dass für verschiedene Sprachen mit unterschiedlichen phonotaktischen Beschränkungen verschiedene Sonoritätshierarchien re-

91 Diese letzte Annahme ist zur Beschreibung der phonotaktischen Regularitäten deshalb nicht nötig, weil ein Segment genau dann als Silbennukleus silbifiziert wird, wenn es ein lokales Sonoritätsmaximum bildet, d.h. von Segmenten mit geringerem Sonoritätswert umgeben ist und genau dann als Silbenrand, wenn es kein lokales Sonoritätsmaximum bildet, d.h. wenn ihm ein sonoranteres Segment vorangeht oder folgt, ohne dass eine Silbengrenze interveniert. Die Zuordnung zum Silbennukleus oder Silbenrand beeinflusst in dieser Konzeption daher nicht den Sonoritätswert eines Segmentes, sondern ist nur eine Konsequenz aus seinem Sonoritätswert relativ zu dem der umgebenden Segmente.

92 /i/ und /u/ sind [+ hoch]. Beide Segmente unterscheiden sich nur in der Lippenrundung und der Zungenlage: /i/ ist ungerundet und wird mit vorderer Zungenlage artikuliert, /u/ ist gerundet und wird mit hinterer Zungenlage artikuliert. /i/ und /j/ sowie /u/ und /w/ unterscheiden sich nur darin, dass /i/ und /u/ [+ silbisch] sind, d.h. Vokale und /j/ und /w/ [- silbisch], d.h. Gleitlaute.

93 Die Silbifizierung als /ju./ und /wi./ im Gegensatz zu /iw./ und /uj./ lässt sich nicht aus dem *Sonority Sequencing Principle*, wohl aber aus dem *Sonority Dispersion Principle* (s.o. 2.3, Anm.) ableiten.

konstruiert werden müssen, andererseits stellt z.B. das Verhalten der Sibilanten /s/ und /sch/ im Deutschen (s.o. 2.3) im Rahmen einer solchen Sonoritätskonzeption, die Sonorität nur taxonomisch aus der Phonotaktik ableitet, ein gravierendes Problem dar.

Die zweite Zugangsweise konzipiert *Sonorität* als primär *phonetisches Merkmal*, das auf die phonologische Struktur abgebildet werden muss. Sonorität wird entweder auf die Lautstärke bezogen, wie CLEMENTS (1990) und SELKIRK (1984) annehmen, oder es wird über die Ebene der Oberklassen auf die zyklische Artikulationsbewegung bezogen, wie CHOMSKY & HALLE (1968) vorschlagen.⁹⁴ Will man phonetische Sonorität auf ein phonologisches Merkmal abbilden, so müssen den Segmenten die phonologischen Sonoritätswerte so zugewiesen werden, dass das phonologische Merkmal Sonorität dieselben Sonoritätsklassen und dieselbe Sonoritätsrangordnung definiert wie das phonetische Merkmal Sonorität. Geht man davon aus, dass die phonetische Eigenschaft der Sonorität die Lautstärke oder Intensität eines Segmentes ist (SELKIRK 1984, CLEMENTS 1990, s.u. 3.2.1), so muss in dieser Konzeption daher die Annahme gemacht werden, dass silbische Segmente sonoranter sind als unsilbische, denn in phonetischer Hinsicht sind Vokale sonoranter (lauter) als Gleitlaute und silbische Liquide sonoranter (lauter) als unsilbische Liquide (vgl. PRICE 1980, s.u. 3.2.2). In dieser Konzeption bildet also in Folgen wie /ju./ oder /wi./ der Silbennukleus (d.h. der Vokal) relativ zum Silbenrand (d.h. zum Gleitlaut) ein Sonoritätsmaximum. Daher muss, wenn man diesen Begriff von Sonorität auf andere phonologische Merkmale zurückführen will, zusätzlich zu den Oberklassenmerkmalen das Merkmal [silbisch] miteinbezogen werden, weil auf andere Weise nicht ausgedrückt werden kann, dass silbische Segmente sonoranter sind als unsilbische Segmente. Wie in der ersten Konzeption besteht auch in dieser zweiten Sonoritätskonzeption die Funktion von Sonorität in der Erklärung phonotaktischer Regularitäten.

Die Konzeption von Sonorität, die in der vorliegenden Studie zugrundegelegt wird, ist die zweite, welche Sonorität als ein *primär phonetisches Merkmal* fasst, das unter Rückgriff auf die *segmentalen Oberklassenmerkmale* und auf das *suprasegmentale Merkmal [silbisch]* auf phonologische Merkmalsstrukturen *abgebildet* werden kann. Diese Konzeption wird deshalb übernommen, weil in dieser Studie der Schwerpunkt nicht auf der Beschreibung der Form sprachlicher Ausdrücke und ihrer Form liegt, sondern auf der Untersuchung der Zusammenhänge der menschlichen Sprachfähigkeit sowohl auf einer abstrakten linguistischen als auch auf einer konkreten phonetischen Ebene als auch auf einer neurowissenschaftlichen Ebene. In diesem Kontext ist daher ein Sonoritätskonzept, das eine Eigenschaft der akustisch-artikulatorischen

⁹⁴ Die Autoren reden jedoch nicht explizit von Sonorität in der hier intendierten Bedeutung.

Grundlage der Sprache auf phonologische Strukturen abbildet, geeigneter als eines, das allein phonotaktische Regularitäten erklären will.

Die Ausgangsfrage dieses Abschnittes war die Frage, ob Sonorität vollständig auf segmentale Merkmale zurückzuführen ist, oder ob das Merkmal Sonorität nur mit Bezug auf die Form der Silbe definiert werden kann. Im Rahmen des hier übernommenen Ansatzes muss angenommen werden, dass die Sonorität eines Segmentes in einem sprachlichen Ausdruck eine Funktion ist erstens aus der inhärenten Sonorität dieses Segmentes, die durch seine Oberklassenzugehörigkeit bestimmt wird und zweitens aus der Silbizität dieses Segmentes, die sich aus seiner Assoziation mit einer C-Position oder einer V-Position in der Silbenstruktur ergibt.⁹⁵ Auf der anderen Seite ist auch die Silbe mehr als eine Summe von Segmenten, denn der phonologische Sonoritätsverlauf einer Silbe ist in dieser Konzeption eine Funktion der inhärenten Sonoritätswerte der Segmente und ihrer Silbizitätspezifizierungen. Dieser phonologische Sonoritätsverlauf ist eine Abbildung des phonetischen Sonoritätsverlaufs einer Äußerung. Sonorität kann demnach als dasjenige phonologisch-phonetische Merkmal sprachlicher Ausdrücke gelten, das bei der Artikulation von Äußerungen die Integration einzelner Segmente zu ganzen Silben und bei der Wahrnehmung von gesprochenen Äußerungen die Rekonstruktion der phonologischen Silbenstruktur ermöglicht.

2.5 Zusammenfassung

In diesem Kapitel wurde gezeigt, dass es in der Sprache phonologische Strukturen und Beschränkungen der Form phonologischer Strukturen gibt, die sich ohne die Kategorie Silbe nicht beschreiben lassen. Hieraus wurde geschlossen, dass die Silbe eine elementare Konstituente in einer phonologischen Theorie sein muss. Im Anschluss daran wurde demonstriert, dass die Sequenzierung von Einzellauten in der Silbe einer Reihe von phonotaktischen Beschränkungen unterliegt, die sich mit dem Konzept *Sonorität* erklären lassen. Schließlich wurde gezeigt, dass dieses Merkmal Sonorität, mit dem die Form der Silbe in nahezu allen Sprachen beschrieben werden kann, nicht allein auf segmentale Merkmale zurückgeführt werden kann, wenn die Regularitäten der Silbe nicht nur in *phonologisch-phonotaktischer*, sondern auch in *phonetischer* Hinsicht erklärt werden soll. Sonorität ist vielmehr das Verbindungsglied zwischen *segmentaler* und *suprasegmentaler* Struktur der Sprache und zwischen *phonetischer Form* und *phonologischer*

⁹⁵ Während zu Beginn dieses Kapitels die Anlaut-Reim-Variante der Silbenstruktur eingeführt wurde, steht an dieser Stelle das CV-Skelett im Vordergrund. Beide Modelle sind miteinander kompatibel, obwohl die CV-Ebene die Anlaut-Reim-Struktur eigentlich überflüssig macht. Vgl. hierzu auch WIESE 1996:43ff.

Struktur auf der Segment- und Silbenebene. Die Frage nach den phonetischen Korrelaten des Merkmals Sonorität, deren Existenz in diesem Kapitel schon implizit vorausgesetzt wurde, wird im folgenden Kapitel erörtert.

3 Phonetische Korrelate der phonologischen Sonorität

3.1 Einleitung

Will man sprachwissenschaftliche Strukturbeschreibungen auf Vorgänge im Gehirn beziehen, so kann man die Frage untersuchen, welche neuronalen Prozesse mit der Rekonstruktion von linguistischen Strukturrepräsentationen aus dem Sprachsignal oder mit der Produktion von Sprachsignalen auf der Grundlage von linguistischen Strukturrepräsentationen korreliert sind. Hierzu muss geklärt werden, welche Beziehungen zwischen linguistischen Strukturrepräsentationen und Eigenschaften von akustischen Signalen bzw. Eigenschaften der sie hervorbringenden Artikulationsbewegungen bestehen. In diesem Kapitel wird daher die Frage nach den phonetischen Korrelaten des phonologischen Merkmals Sonorität diskutiert. Hierzu werden zunächst Konzeptionen besprochen, die sich auf die akustischen Korrelate der Sonorität beziehen (3.2) und zwar einerseits Ansätze, welche die akustischen Korrelate der Sonorität aus der Sonoritätshierarchie ableiten (3.2.1) und andererseits Ansätze, welche die akustischen Korrelate der Silbizität experimentell ermitteln (3.2.2). Im nächsten Abschnitt (3.3) werden die artikulatorischen Korrelate der Sonorität erörtert. Unter den Theorien, die sich auf die Artikulation der Sprache beziehen, wird insbesondere der Ansatz von BECKMAN, EDWARDS & FLETCHER (1992)⁹⁶ hervorgehoben (3.3.2), in welchem die Sonorität der Segmente auf den Akzent bezogen wird. Ausgehend von dieser Theorie wird schließlich gezeigt (3.4), dass das phonologisch-phonetische Merkmal Sonorität geeignet ist, sowohl auf den Prozess der Rekonstruktion von Silbenstrukturrepräsentationen als auch der Rekonstruktion von Akzentstrukturrepräsentationen aus dem akustischen Signal bezogen zu werden.

⁹⁶ Vgl. auch EDWARDS & BECKMAN 1988, BECKMAN 1996.

3.2 Akustische Korrelate von Silbe und Sonorität

3.2.1 Schallfülle und Sonorität bei SIEVERS und JESPERSEN

Die Beobachtung, dass Sprachlaute ihrer *Schallstärke* entsprechend geordnet werden können und dass eine so hergestellte Schallstärke-Ordnung in Bezug gesetzt werden kann zur Sequenzierung der Laute in der Silbe, geht auf SIEVERS (1901⁵:198ff) zurück.⁹⁷ Diese Schallstärke-Ordnung der Laute differenziert SIEVERS weiter in eine Ordnung der Laute nach der *Druckstärke*, mit der ein Laut bezogen auf die Expiration hervorgebracht wird und nach der *Schallfülle*, die ein Laut auf Grund seiner Art der Artikulation hat. Die *Schallstärke* eines Lautes ist demnach

das Mass der absoluten Lautheit jedes einzelnen Sprachschalles im einzelnen Falle und dies Mass ist [...] von zwei Factoren, der Druckstärke und der Schallfülle, abhängig. (SIEVERS 1901⁵:199f)

Eine Silbe ist nach SIEVERS gekennzeichnet durch die Kombination von Lauten mit unterschiedlicher Schallstärke:

In der [...] Silbe [...] muss nothwendig eine Abstufung der Schallstärke stattfinden, indem alle übrigen Laute der Silbe einem einzigen Laute untergeordnet werden. Dieser die Silbe beherrschende Laut heißt Sonant der Silbe (ist silbisch), die übrigen heißen die Consonanten der Silbe (sind unsilbisch). (SIEVERS 1901⁵:203f)

Silben lassen sich mit Sievers als „Schallmassen“ zwischen zwei „Minima der Schallstärke“ charakterisieren (SIEVERS 1901⁵:198):

(Es gilt) dass das Ohr des Hörers die zusammenhängende Rede subjectiv in gewisse Theilstücke zerlegt, [...] und dass diese Theilstücke das sind was man als Silben zu bezeichnen pflegt. (SIEVERS 1901⁵:198)

Die *Schallfülle* lässt sich wie folgt auf die Anordnung der Laute in der Silbe beziehen:

Beim Zusammentreffen mehrerer Laute muss [...] derjenige zum Sonanten werden, welcher [...] die grösste Schallfülle besitzt. (Es) gilt für die Consonanten untereinander: je näher dem Sonanten, umso größer muss die Schallfülle sein. Daher ist die Reihenfolge der Lautarten, welche einem Sonanten unsilbisch vorausgehen können, genau entgegengesetzt der Reihenfolge der Lautarten, welche dem Sonanten als Consonanten folgen können (SIEVERS 1901⁵:204)

In dieser Äußerung formuliert SIEVERS das *Sonoritäts-Folge-Prinzip*, welches im vorangehenden Kapitel diskutiert wurde: Silbennuklei sind Gipfel der phonetischen Schallfülle bzw. phonologischen Sonorität und zu den Rändern der Silbe fällt die Schallfülle bzw. Sonorität ab. Auf diese Weise gelangt SIEVERS zu der Schallfülle-Ordnung der Laute in (1).

- (1) offene Vokale > geschlossene Vokale
Liquide (r > l)

⁹⁷ CLEMENTS (1990:284) gibt als noch frühere Quelle an: WHITNEY, W.D. (1865): „The relation of vowel and consonant“, in: *Journal of the American Oriental Society*, vol. 8. Nachgedruckt in WHITNEY, W.D. (1874): *Oriental and Linguistic Studies*. Second Series. New York: Scribner.

Nasale
 Frikative (Spiranten)
 Explosivlaute (Plosive)

Diese Schallfülle-Ordnung leitet er aus Beobachtungen über die Kombinierbarkeit der Laute in der Silbe ab. An dieser Stelle verwendet SIEVERS nicht mehr wie oben den allgemeineren Begriff *Schallstärke*, sondern den spezielleren Begriff *Schallfülle*, welche einem Laut auf Grund seiner Art der Artikulation zukommt, d.h. es geht ihm in dieser *Hierarchie der Laute* um die *inhärenten* Eigenschaften der Sprachlaute, die ihre Kombinierbarkeit zu Silben bestimmen, es geht ihm aber nicht um die *aktuelle Druckstärke*, mit der sie in einer aktuellen Silbe hervorgebracht werden können.⁹⁸

JESPERSEN (1904) knüpft an SIEVERS' Schallfülle-Hierarchie der Laute an, die er wie SIEVERS als durch die Art der Artikulation der Laute gegeben konzipiert und auf ihre inhärente Lautstärke bezieht:

Die Schallfülle ist eine Resultante aus denselben Faktoren, die überhaupt sein ganzes Gepräge bedingen (JESPERSEN 1904:186)

Der sonorste Laut ist der, welcher – wenn alles andere gleich bleibt – im größten Abstand gehört werden kann. (JESPERSEN 1904:186)

Dies ist so zu verstehen, dass JESPERSEN die *Schallfülle* oder *Sonorität* eines Laut zunächst auf alle seine übrigen artikulatorischen oder akustischen Eigenschaften, phonologisch betrachtet also auf seine übrigen distinktiven und nicht-distinktiven Merkmale, vor allem aber die sich dadurch ergebene *Lautheit*, zurückführt. Insbesondere fügt JESPERSEN aber noch hinzu, dass „der Gebrauch oder Nichtgebrauch der Stimme die größte Rolle spielt“ (1904:186) und kommt so zu der differenzierteren Ordnung der Laute nach ihrer Schallfülle oder Sonorität in (2), die den Stimmhaftigkeitsparameter der Konsonanten miteinbezieht.

- (2)
- offene Vokale
 - halboffene Vokale
 - geschlossene Vokale
 - r-Laute
 - Laterale
 - Nasale
 - stimmhafte Frikative
 - stimmhafte Plosive
 - stimmlose Frikative

⁹⁸ In diesem Zusammenhang muss angemerkt werden, dass SIEVERS zwei Arten von Silben unterscheidet, nämlich *Schallsilben* als Folgen von Lauten, die sich um ein Schallfülle-Maximum gruppieren und *Drucksilben* als Folgen von Lauten, die sich um ein Druckmaximum gruppieren: „Schallsilben und Drucksilben können sich begrifflicher Weise decken, müssen es aber nicht und zwar können sowohl Lautfolgen, welche an sich eine Schallsilbe bilden können, expiratorisch in getrennte Silben zerlegt werden (vgl. z.B. zweisilbiges a-i mit dem Diphthongen ai) als umgekehrt Lautfolgen mit einheitlicher Expiration hervorgebracht werden, die nach der Abstufung der Schallfülle in mehrere Schallsilben zerfallen müssen.“ (SIEVERS 1871:203)

stimmlose Plosive

Die Silbe definiert JESPERSEN wie folgt:

In jeder Lautgruppe gibt es ebensoviele Silben als es deutliche relative Höhepunkte in der Schallfülle gibt. (JESPERSEN 1904:188)

und das Sonoritätsprinzip formuliert er so:

Man hat [...] den am meisten hervortretenden Laut der Silbe [...] den Sonanten genannt [...] Ich ziehe es vor, von Silbengipfel zu reden. [...] Zwischen einem gegebenen Laut und dem Silbengipfel werden nur Laute von derselben oder einer höheren Sonoritätsklasse geduldet. (JESPERSEN 1904:190f)

Was aber versteht JESPERSEN genau unter Schallfülle oder Sonorität und wie kann der Phonetiker diese messen? JESPERSEN vertritt die Auffassung, dass der Phonetiker die Korrelate wahrgenommener Silben messen muss:

Man wird sehen, dass in Wirklichkeit die Begriffe Silbe und Schallfülle korrelativ sind; ein objektives Maß für die Schallfülle haben wir jedenfalls noch nicht und wir müssen uns daher in gewissen Grenzfällen damit begnügen, einen Laut in verschiedenen Verbindungen mit anderen zu prüfen, um zu sehen, welche Silbenzahl für die unmittelbare Auffassung vorhanden ist. (JESPERSEN 1904:191)

Das Kriterium für die Existenz einer Silbe und damit auch eines Sonoritätsmaximums ist für JESPERSEN also die Wahrnehmung einer Silbe durch den Hörer. Die relative Schallfülle eines Lautes lässt sich nach dieser Auffassung ermitteln, indem dieser Laut in verschiedenen Verbindungen einem Hörer präsentiert wird, der ein Urteil darüber fällt, wieviele Silben wahrzunehmen sind. Hieraus kann die Zahl der Sonoritätsgipfel und die relative Sonorität der Laute zueinander abgeleitet werden. Dies bedeutet, dass JESPERSEN die Silbe und das Merkmal Sonorität nicht nur auf der Basis von phonetischen Eigenschaften definiert, sondern dass für ihn das entscheidende Kriterium die *Wahrnehmung einer Silbe* bzw. die *Wahrnehmung eines Sonoritätsmaximums* ist und dass ausgehend davon das objektive Maß der Sonorität und damit der Silbe erst noch ermittelt werden muss.

Wie lassen sich die Ausführungen von SIEVERS und JESPERSEN zusammenfassen? Zunächst muss festgehalten werden, dass in beiden Konzeptionen

- die Sequenzierung der Laute in der Silbe durch eine Eigenschaft der Laute in der Silbe *bedingt* ist, die man Schallfülle oder Sonorität nennen kann, und
- dass eine Ordnung der Laute entsprechend dieser Eigenschaft aus Beobachtungen über die Sequenzierbarkeit der Laute in der Silbe *abgeleitet* werden kann.⁹⁹

Weiter muss mit SIEVERS und JESPERSEN

⁹⁹ Hier findet sich die Art der Argumentation, die SAUSSURE (1972:89) mit den Worten „c'est là un cercle vicieux“ bedenkt.

- die inhärente Schallfülle oder Sonorität der Laute auf der einen Seite und wahrgenommene Silben (bei JESPERSEN) oder Schallstärkemaxima (bei SIEVERS) auf der anderen Seite unterschieden werden. Hier findet sich die Grundlage für eine Differenzierung zweier Arten von Sonorität, nämlich eine den Lauten inhärente Sonorität und eine Veränderung der Sonorität, die sich durch die Integration von Lautfolgen zu Silben ergibt. Dieser Gedanke liegt der Konzeption von Sonorität und Silbizität bei CLEMENTS (1990) zugrunde, die im vorangegangenen Kapitel diskutiert wurde (s.o. 2.4.4) und die in diesem Kapitel wiederaufgegriffen wird (s.u. 3.4).

Schließlich muss hervorgehoben werden, dass JESPERSEN ausgehend von

- SIEVERS' Konzeption der Schallstärke eines Lautes als „das Mass der absoluten Lautheit jedes einzelnen Sprachschalles im einzelnen Falle“ (SIEVERS 1901⁵:199f), also der aktuellen Lautheit eines Segmentes in einer Äußerung, die abhängig ist von seiner Schallfülle und von der Druckstärke, mit der es hervorgebracht wird,
- die Schallfülle eines Lautes auf seine inhärente oder relative Lautheit oder Intensität bezieht,¹⁰⁰ d.h. auf die Lautheit eines Lautes bei konstant gehaltenen Produktionsfaktoren, die sich nur aus den distinktiven Merkmalen seiner Artikulation ergibt.

Ordnet man jedoch z.B. die Laute des Englischen ihrer relativen Intensität¹⁰¹ nach (vgl. FRY 1979:127, vgl. auch LADEFOGED 1993²:246), so kommt man zu einer Hierarchie, die nur näherungsweise der Sonoritätshierarchie entspricht, die man für das Englische aufstellen muss, wenn man die Sonoritätshierarchie aus der Sequenzierung der Laute in der Silbe ableitet.¹⁰² Das Konzept der inhärenten oder intrinsischen Lautheit kann also nur als *ein* phonetisches Korrelat der Sonorität gelten und Sonorität ist nicht ausschließlich auf die inhärente Lautheit der Segmente zurückzuführen.

Die Konzeption von SIEVERS, die zwischen (aktueller) Schallstärke und (inhärenter) Schallfülle unterscheidet, ist im Zusammenhang mit JESPERSENS Konzept der *Silbe als Einheit der Wahrnehmung* von großem Interesse für die weitere Untersuchung der Sonorität, weil sich hieraus eine Möglichkeit ergibt, Schallstärke oder Sonorität zu messen: es muss nämlich untersucht werden, „welche Silbenzahl für die unmittelbare Auffassung vorhanden ist“ (JESPERSEN 1904:191) und zwar an Lautverbindungen, in denen ein und derselbe Laut einen *Silbengipfel* bilden oder nicht bilden kann. Wenn

¹⁰⁰ Diese Auffassung wird von Sprachwissenschaftlern häufig aufgegriffen, vgl. z.B. BLEVINS 1995, GIEGERICH 1992:132, ROCA 1994:154, WIESE 1996:258, für eine Kritik dieser Konzeption von phonetischer Seite vgl. HEIKE 1992.

¹⁰¹ FRY gibt die Intensität der Laute in Dezibel (dB) an, als Bezugswert dient ihm der Pegel des leisesten Phonems des Englischen, das stimmlose /θ/ wie in *thin*. Lautestes Phonem des Englischen ist das lange oder gespannte /u/ wie in *moon*, das er mit 29 dB (relativ zum leisesten Laut /θ/) verzeichnet.

¹⁰² Für eine Sonoritätshierarchie für das Englische vgl. GIEGERICH 1992:131.

- Die *Dauer* des Liquids ist der entscheidende Faktor für die Wahrnehmung des einen oder anderen Wortes: ein längerer Liquid führt zur Wahrnehmung des zweisilbigen, ein kürzerer Liquid zur Wahrnehmung des einsilbigen Wortes. Eine erhöhte Sprechrate (d.h. eine kürzere absolute Dauer des Signals) führt insgesamt zur häufigeren Wahrnehmung des einsilbigen Wortes. Die absolute Dauer des Liquids ist ein effizienterer Faktor als seine relative Dauer (im Vergleich zu anderen Signalabschnitten). Die große Bedeutung des Faktors Dauer für die Wahrnehmung der Silbenzahl wurde von PORTELE & HEUFT (1994) für deutsche quasihomophone Minimalpaare des Typs *sein* und *seinen* bestätigt.
- Bewegt sich die Dauer im Grenzbereich, so wird die *Amplitude* des Liquids genutzt: je lauter der Liquid, desto eher wird das zweisilbige Wort wahrgenommen. PRICE konnte darüberhinaus zeigen, dass ein wahrgenommener Silbengipfel nicht immer mit einem lokalen Schalldruckmaximum im Signal einhergeht.¹⁰⁵
- Die *Stimmeinsatzzeit* und die vorangehende Aspiration sind ebenfalls wichtige akustische Hinweise auf die Präsenz eines Sonoritätsmaximums: Je stimmhafter der Plosiv, desto eher wird das zweisilbige Wort, d.h. ein zusätzliches Sonoritätsmaximum, wahrgenommen. Wird die Aspiration durch Stille ersetzt, so muss die Stimmeinsatzzeit verkürzt werden (d.h. der Plosiv muss stimmhafter sein), um dasselbe zweisilbige Perzept hervorzurufen.

Hieraus muss geschlossen werden, dass einerseits die Lautheit oder Amplitude eines Segmentes keinen so großen Einfluss auf die Wahrnehmung eines Silbengipfels hat wie von SIEVERS und JESPERSEN angenommen wurde und dass andererseits alle drei untersuchten akustischen Dimensionen als phonetische Korrelate wahrgenommener Silbengipfel gelten können, wenn nicht sogar noch weitere Dimensionen, wie z.B. die spektrale Qualität eines Lautes, hinzugenommen werden müssen (was experimentell zu überprüfen ist). Dieser Ansatz von PRICE basiert also auf der Annahme, dass silbische Segmente sonoranter sind als nichtsilbische. Für das Mehr an Sonorität, das ein silbisches von einem nichtsilbischem /l/ unterscheidet, hat PRICE die phonetischen Parameter Dauer, Lautstärke und Periodizität ermittelt.

Stimm Lippen zeitgleich mit der Verschlusslösung ein, so hat man eine Stimmeinsatzzeit (engl. VOT für *voice onset time*) von 0 ms, setzt die Schwingung 50 ms nach der Verschlussöffnung ein, so hat man eine VOT von 50 ms. Setzt die Schwingung vor der Verschlusslösung ein, so ist die VOT negativ. PRICE (1980) verwendete Stimuli mit einer VOT zwischen 35 ms und 112 ms. Stimuli mit einer VOT von weniger als 35 ms werden nach PRICE als *ba* wahrgenommen. PRICE (1980) zeigte außerdem, dass sich die VOT des initialen /p/ in *plight* und *polite* (mit silbischem /l/ und elidiertem Schwa) nur um 10 ms für männliche und 20 ms für weibliche Sprecher unterscheidet.

¹⁰⁵ Zum selben Ergebnis kamen PORTELE & HEUFT (1994) für das Deutsche.

Eine ähnliche Vorgehensweise zur Messung der phonetischen Korrelate von phonologischer Sonorität schlägt auch CLEMENTS (1990) vor, der eine Analogie zwischen der phonologischen Kompositionalität des Merkmals Sonoritäts im Hinblick auf die Oberklassenmerkmale und einer Kompositionalität des phonetischen Korrelates der Sonorität postuliert: so wie phonologische Sonorität auf die Oberklassenmerkmale zurückgeführt werden kann, kann nach seiner Auffassung auch das phonetische Korrelat der Sonorität auf die phonetischen Korrelate der Oberklassenmerkmale zurückgeführt werden:

‘Sonority’ is a composite property of speech sounds which depends on the way they are specified for each of a certain set of features. Plus-specifications for any of these features have the effect of increasing the perceptibility or salience of a sound with respect to otherwise similar sounds having a minus-specification, for example by increasing its loudness [...], or making its formant structure more prominent. (CLEMENTS 1990:297)

Sonorität bedeutet für CLEMENTS also auf einer ersten Ebene soviel wie perzeptuelle Salienz. Akustische Eigenschaften, welche die Salienz eines Lautes vergrößern, sind zum Beispiel die Lautstärke oder die Periodizität bzw. die Anwesenheit einer Formantstruktur. Führt man diesen Ansatz von CLEMENTS weiter aus, so kann angenommen werden, dass auch die phonetischen Eigenschaften, die einen (unsilbischen) Gleitlaut vom entsprechenden (silbischen) Vokal unterscheiden, phonetische Korrelate des Merkmals Sonorität sind. Während sich Lautpaare finden, die außer im Merkmal [silbisch] in keinem anderen segmentalen Merkmal differieren, finden sich für die übrigen Oberklassenmerkmale wie [approximant] oder [vokoid], auf die CLEMENTS das phonologische Merkmal Sonorität zurückführt, jedoch keine Minimalpaare. Laute, die sich in einem dieser Oberklassenmerkmale unterscheiden, differieren stets auch in weiteren Merkmalen, so wie sich z.B. homorgane¹⁰⁶ [+sonorante] Nasale und [–sonorante] Obstruenten zusätzlich im Merkmal [nasal] unterscheiden.

Zusammenfassend lässt sich sagen, dass man, wenn man phonologische Sonorität auf phonologische Silbizität bezieht und die phonetischen Unterschiede zwischen silbischen und nichtsilbischen Segmenten als ein Mehr an Sonorität begreift, zu dem Ergebnis kommt, dass das phonologische Merkmal Sonorität zumindest in akustischer Hinsicht *mehrdimensionale* Korrelate hat. Ob sich diese mehrdimensionalen akustischen Korrelate auf eine allgemeine Eigenschaft der Artikulationsdynamik beziehen lassen, soll im nächsten Abschnitt erörtert werden.

¹⁰⁶ Konsonanten sind dann *homorgan*, wenn sie dieselbe Artikulationsortspezifizierung haben, z.B. bilabiales /b/ und bilabiales /m/, oder alveolares /d/ und alveolares /n/.

3.3 Artikulatorische Korrelate von Silbe und Sonorität

3.3.1 Sonorität bei SAUSSURE

Während SIEVERS und JESPERSEN die Sequenzierung der Laute in der Silbe auf ihre Schallfülle zurückführen, bezieht SAUSSURE (1972) die sich so ergebende Ordnung der Laute in der Silbe auf den Öffnungsgrad des Vokaltraktes, mit welchem sie artikuliert werden:

Quelle que soit la place d'articulation, elle présente toujours une certaine apertures, c'est-à-dire un certain degré d'ouverture entre deux limites qui sont: l'occlusion complète et l'ouverture maximale. Sur cette base, en allant de l'aperture minimale à l'aperture maximale, les sons seront classés en sept catégories (SAUSSURE 1972:70)

SAUSSURE differenziert ähnlich wie SIEVERS und JESPERSEN Okklusive, Frikative, Nasale, Liquide und geschlossene, halboffene und offene Vokale.¹⁰⁷ Sodann unterscheidet er öffnende und schließende Gesten:

quand on prononce un groupe *appa*, on perçoit une différence entre les deux *p*, dont l'un correspond à une fermeture, le second à une ouverture. (SAUSSURE 1972:79)

Ein intervokalisches /p/ zerfällt also nach seiner Auffassung in zwei Phasen der Artikulation: in eine Phase, in welcher der Mund nach der Artikulation des Vokals geschlossen wird (d.h. in welcher der bilabiale Verschluss gebildet wird) und in eine Phase, in der er in Vorbereitung auf den folgenden Vokal geöffnet wird (d.h. in welcher der Verschluss gelöst wird). Die erste Phase nennt SAUSSURE *Implosion*, die zweite *Explosion*:

On a appelé la fermeture *implosion* et l'ouverture *explosion*; [...] dans le même sens on peut parler de sons *fermants* et de sons *ouvrants*. (SAUSSURE 1972:80)

Schließlich definiert er die Konzepte Silbengrenze und Silbengipfel:

Si dans une chaîne de sons on passe d'une implosion à une explosion, on obtient un effet particulier qui est l'indice de la *frontière de syllabe* (SAUSSURE 1972:86)

Le son qui donne cette impression par son caractère de première implosive peut être appelé *point vocalique*. (SAUSSURE 1972:87)

Eine Silbengrenze fällt nach dieser Auffassung mit dem Moment des Übergangs vom Verschluss zur Öffnung zusammen, oder anders ausgedrückt, mit dem Moment einer Bewegungsgeschwindigkeit von Null. Dies bedeutet auch, dass SAUSSURE zufolge die Silbengrenze mitten in das /p/ fällt, welches dann entweder ambisilbisch ist,¹⁰⁸ oder als zwei Segmente aufgefasst werden kann.¹⁰⁹ Den Silbengipfel, den er *point vocalique*

¹⁰⁷ Die Konzeption von SAUSSURE wird von GRAMMONT (1933:99) aufgegriffen: „Une syllable est donc une suite d'ouvertures croissantes suivie d'une suite d'ouvertures décroissantes.“

¹⁰⁸ D.h. zu beiden Silben gehört, s.u. 4.3.2, Anm.

¹⁰⁹ SAUSSURE nimmt an, dass sich links und rechts einer Silbengrenze zwei /p/s finden, ein „son fermant“ und ein „son ouvrant“ (s.o.). In phonemischer Hinsicht handelt es sich aber um dieselben in-

nennt, identifiziert SAUSSURE mit dem ersten Segment nach dem Ende der Explosions-Phase, am Beginn der Implosions-Phase: mit dem Vokal, insofern nach dem Erreichen der vokalischen Artikulationskonfiguration der Vokaltrakt nicht weiter geöffnet wird. Analog zu der Annahme, dass die *Silbengrenze* an dem Punkt verläuft, an dem sich eine Bewegungsgeschwindigkeit von Null bei maximaler Geschlossenheit des Vokaltraktes findet, könnte man SAUSSURES Ausführungen so interpretieren, dass der *Silbengipfel* ebenfalls mit einem Minimum der artikulatorischen Dynamik zusammenfällt, jedoch bei maximaler Offenheit des Vokaltraktes. Eine Silbe schließlich entspricht in dieser Konzeption einem Zyklus aus einer Explosionsbewegung und einer Implosionsbewegung.

SAUSSURE versucht hier, die Silbe in einem Modell artikulatorischer Dynamik zu definieren. Der Vorteil dieser Konzeption scheint darin zu liegen, dass man es nur mit einem einzigen zu untersuchenden Parameter zu tun hat: mit dem Öffnungsgrad des Vokaltraktes. Ob diese Konzeption von Saussure letztlich geeignet ist, unter Rückgriff auf seine Silbendefinition *gemessene* Silben und *wahrgenommene* Silben zu identifizieren, d.h. ob *wahrgenommene* Silben immer mit *Explosions-Implosions-Zyklen* zusammenfallen und ob wahrgenommene Silbengipfel z.B. in den obengenannten Minimalpaaren wie *sein* und *seinen* stets dem Beginn einer Implosion-Phase entsprechen, bleibt phonetisch zu überprüfen.

An SAUSSURES Konzeption ist außerdem hervorzuheben, dass er der Dynamik der Artikulation gegenüber der Identität der Segmente eine gewisse Autonomie zugesteht, insofern er annimmt, dass das Phonem /p/, das Linguisten als invariante Einheit zu betrachten gewohnt sind, einerseits als eine Explosionsbewegung realisiert werden kann, nämlich am Silbenanfang und andererseits als Implosionsbewegung, nämlich am Silbenende. Eine mit dieser Auffassung verwandte Konzeption einer Autonomie der artikulatorisch definierten Sonorität gegenüber der phonemischen Identität der Segmente findet sich auch in einem neueren Ansatz zur artikulatorischen Phonetik, der weiter unten erörtert wird (s.u. 3.2.2).

Auch bei JESPERSEN findet sich ein ähnlicher Bezug auf die Dynamik der Artikulation und die Koordination von Öffnungs- und Verschlussgesten wie bei SAUSSURE. Jespersen macht hierzu die folgende Bemerkung:

Ebenso werden wir eine Vermehrung der Silbenzahl überall da erhalten, wo durch *ungenau* Artikulation ein wenn auch noch so kurzer Augenblick mit größerer Schallfülle in der Umgebung entsteht; das ist häufig der Fall in Verbindungen von [l] und [r] mit einem Konsonanten; wenn z.B. im englischen Henry *die Zunge* nach der Aussprache des [n] *eine Spur weiter als nötig* für das [r]

riante Einheit: „N’importe quel phonème peut être aussi bien implosif qu’explosif.“ (1972:81). Dennoch führt er wenig später aus: „Il faut donc dédoubler le tableau des phonèmes“ (1972:81). Wie diese Konzeption zu interpretieren ist, bleibt unklar.

herabgleitet und daher wieder zu diesem Laut hinaufgeht, so wird sogleich das Wort dreisilbig: [hen^əry] (JESPERSEN 1904:188, meine Herv.)¹¹⁰

Mit dieser Äußerung stellt JESPERSEN in seiner Konzeption den Bezug her zwischen den akustischen und artikulatorischen Korrelaten von Silbe und Sonorität: Momente der maximale Öffnung des Vokaltraktes entsprechen hiernach Abschnitten mit maximaler Schallfülle im akustischen Signal und werden als Silbengipfel wahrgenommen. Für JESPERSEN scheint in den Fällen, in denen durch eine ungenaue Artikulation eines Wortes ein lokales Öffnungsmaximum entsteht, ein eingeschobener Vokal, ein Schwa (/ə/ wie in *b[ə]reit*), artikuliert zu werden. Die Annahme, dass ein und dasselbe Segment verschiedenen Formen der artikulatorischen Dynamik entsprechen kann, d.h. ein Öffnungsmaximum bilden kann oder nicht bilden kann, je nachdem, ob es einen Silbengipfel konstituiert oder nicht, ist mit JESPERSENS Konzeption nicht kompatibel, denn er korreliert nur die *inhärente* Sonorität eines Segmentes mit dem Öffnungsgrad des Vokaltraktes, welche den Segmenten durch ihre distinktiven Merkmale zukommt (s.o. 3.2.1) und nicht die Sonorität eines Segmentes in einer silbifizierten Lautkette. Daher muss er annehmen, dass in diesen Fällen ein Vokal eingeschoben wird. Die Frage, ob in den obengenannten Wörtern, wenn sie auf die von JESPERSEN beschriebene Weise artikuliert werden, der Hörer tatsächlich eine zusätzliche Silbe wahrnimmt und die Frage, ob sich im hervorgebrachten akustischen Signal spektrale Hinweise auf einen eingeschobenen Vokal finden, kann experimentell überprüft werden. Für das Englische (PRICE 1980, s.o. 3.2.2) und für das Deutsche (PORTELE & HEUFT 1994, JANNEDY 1994)¹¹¹ hat sich jedoch gezeigt, dass die von JESPERSEN beschriebene Variation der Artikulationsdynamik, die sich in der Wahrnehmung des zweisilbigen Elementes eines Minimalpaares wie *plight* und *polite*, *sein* und *seinen* oder *breit* und *bereit* zeigt, nicht mit dem Einschub eines Vokals einhergehen muss, bzw. dass der Schwa-Laut in *polite*, *seinen* oder *bereit* ausfallen kann und trotzdem ein zweisilbiges Wort wahrgenommen wird.

Noch eine ganz andere Konzeption der Silbe aus artikulatorischer Perspektive muss kurz Erwähnung finden: die Silbentheorie von STETSON.¹¹² Nach STETSONS Auffassung entspricht eine Silbe einem Luftstoss (*chest pulse*), der durch die Aktivität der internen Zwischenrippenmuskeln hervorgebracht wird. Experimentelle Untersuchungen

¹¹⁰ JESPERSENS artikulatorisch-phonetische Terminologie bezieht sich nicht auf den Öffnungsgrad des Vokaltraktes, sondern auf die Zungenstellung. Eine gesenkte Zunge entspricht dabei einem geöffneten Vokaltrakt, eine hohe Zungenstellung (d.h. dicht am Gaumen) einem geschlossenen Vokaltrakt.

¹¹¹ JANNEDY, S. (1994): „Rate effects on German unstressed syllables“, *Working Papers in Linguistics, Ohio State University* 44:105-124, zitiert nach BECKMAN (1996).

¹¹² STETSON, R.H. (1951²): *Motor Phonetics*. Amsterdam: North Holland, zit. nach KOHLER (1976:81), vgl. auch ABERCROMBIE 1967:34ff und LAVER 1995:167f.

(LADEFOGED, DRAPER & WHITTERIDGE 1958,¹¹³ zit. nach KOHLER 1976:82) haben ergeben, dass für Wörter wie *pity* oder *around* nur ein Aktivitätsimpuls zu messen ist, wogegen für Wörter wie *sport* oder *stay*¹¹⁴ zwei Impulse erscheinen. In diesen Fällen ergibt sich also für STETSONS Konzeption der Silbe ein Unterschied zwischen *wahrgenommener* und *gemessener* Silbenzahl, ebenso wie sich in SIEVERS' Konzeption Unterschiede zwischen Schallsilben und Drucksilben finden (s.o. 3.2.1, Anm.). Daher ist STETSONS Konzeption nicht geeignet, wahrgenommene und gemessene Silben zu identifizieren.

3.3.2 Sonorität in einem Modell der artikulatorischen Dynamik

3.3.2.1 Prosodie und Artikulationsdynamik

Auch in verschiedenen neueren Ansätzen zur artikulatorischen Phonetik wird auf das Konzept der Sonorität Bezug genommen.¹¹⁵ Im folgenden wird eine dieser Theorien erörtert, in der sich eine ganz ähnliche Konzeption von Silbe und Sonorität wie bei SAUSSURE findet. Ausgehend von dieser Theorie soll dann ein umfassenderer Begriff der Sonorität entwickelt werden.

In einer Untersuchung¹¹⁶ zur Modifikation der artikulatorischen Dynamik¹¹⁷, die sich in einer Äußerung in Abhängigkeit von der Betonung¹¹⁸ einer Silbe, der finalen Längung¹¹⁹ einer Silbe und dem allgemeinen Sprechtempo vollzieht, machen

¹¹³ LADEFOGED, P, DRAPER, M.H. & WHITTERIDGE, D. (1958): „Syllables and stress“, *Miscellanea Phonetica* III. IPA London, 1-14.

¹¹⁴ Bei der Bewertung solcher Gegenbeispiele, deren Silbenzahl in STETSONS Silbentheorie nicht den Sprecherintuitionen entsprechend vorausgesagt wird, muss in Betracht gezogen werden, dass Wörter wie *stay* oder *sport* auch gegen das Sonoritäts-Folge-Prinzip verstoßen, weil /t/ bzw. /p/ nicht mehr Sonorität besitzt als das vorangehende /s/ (s.o. 2.3). Daher könnte man annehmen, dass es sich hier doch um zwei Silben handelt.

¹¹⁵ Vgl. LINDBLOM (1983), der die Sequenzierbarkeit der Laute in der Silbe, die durch die Sonoritäts-hierarchie bestimmt wird, auf ihre koartikulatorische Kompatibilität bezieht, für eine Kritik vgl. KEATING 1983. Vgl. auch SILVERMAN & PIERREHUMBERT (1990), die den Zeitverlauf des musikalischen Akzents (*pitch accent*) in verschiedenen phonologischen Kontexten zum Sonoritätsverlauf der Silbe in Bezug setzen, welchen sie ähnlich wie BECKMAN, EDWARDS & FLETCHER (1992) konzipieren.

¹¹⁶ Für eine Beschreibung des phonetischen Modells, in dessen Rahmen BECKMAN, EDWARDS & FLETCHER (1992) die Dynamik der Artikulation beschreiben, vgl. BROWMAN & GOLDSTEIN 1990.

¹¹⁷ Die Autoren zeichneten mit einem optoelektronischen Verfahren die Stellung des Kiefers auf und berechneten Punkte minimaler und maximaler Öffnung, die Bewegungsgeschwindigkeit sowie Punkte minimaler und maximaler Bewegungsgeschwindigkeit.

¹¹⁸ Die Begriffe *Betonung* und *Akzent* werden im folgenden in gleicher Bedeutung verwendet.

¹¹⁹ *Finale Längung* bezeichnet das Phänomen der Verlängerung der Dauer einer Silbe vor einer Sprechpause oder am Phrasenende.

BECKMAN, EDWARDS & FLETCHER (1992)¹²⁰ die Beobachtung, dass die Betonung einer Silbe die Artikulationsdynamik in umfassenderer Weise beeinflusst als das Phänomen der finalen Längung und als eine Verlangsamung des Sprechtempos.

Eine *Silbe* definieren BECKMAN, EDWARDS & FLETCHER (1992:72) zunächst als Sequenz aus einer Öffnungs- und einer Schließbewegung des Kiefers; diese Bewegungen oder *Gesten* wiederum entsprechen in ihrer Definition den Intervallen zwischen zwei Punkten mit einer Bewegungsgeschwindigkeit von Null, d.h. eine Silbe erstreckt sich von dem Moment, an dem eine Schließ- durch eine silbeninitiale Öffnungsbewegung abgelöst wird, bis zu dem Moment, wo eine erneute Öffnungs- die silbenfinale Schließbewegung ablöst. Der Silbengipfel entspricht dem Moment, an dem die silbeninitiale Öffnungs- in die silbenfinale Schließbewegung übergeht, d.h. dem Punkt mit maximaler Öffnung des Vokaltraktes.

Dieser Zyklus aus Öffnungs- und Schließbewegung, der einer Silbe entspricht, wird von den drei untersuchten Phänomenen folgendermaßen affiziert: *Finale Längung* und *langsames Sprechtempo* beeinflussen allein die Dauer der beiden Artikulationsgesten, indem Öffnungs- und Schließbewegung verlangsamt ausgeführt werden, mit dem Resultat, dass die Silbe insgesamt länger dauert. Die *Akzentuierung* einer Silbe hingegen beeinflusst sowohl die gesamte Dauer der Silbengeste als auch die Auslenkung des Kiefers und die Geschwindigkeit der Öffnungsbewegung, so dass die Öffnungsbewegung schneller ausgeführt wird,¹²¹ und relativ später in die Schließbewegung übergeht. Dies wiederum bedeutet, dass der Vokaltrakt weiter geöffnet wird und länger maximal geöffnet bleibt. Diese Beobachtung interpretieren BECKMAN, EDWARDS & FLETCHER (1992:84) so, dass die *Betonung* einer Silbe die *Sonorität* der Silbe *vergrößert*, während finale Längung und langsameres Sprechtempo nur die *Dauer* einer Silbe *verlängern*.

3.3.2.2 Sonorität und Akzent

Was verstehen BECKMAN, EDWARDS & FLETCHER (1992) in diesem Kontext unter dem Begriff Sonorität? Zunächst definieren sie Sonorität als phonologisches Merkmal, als inhärente Eigenschaft eines Segmentes:

We see the inherent phonological sonority of a segment as something analogous to the phonological specification of a tone, except that the intrinsic scale is derived from the manner features of a segment, as proposed by Clements (1990): stops have L (Low) sonority and open vowels have H (High) sonority. (BECKMAN, EDWARDS & FLETCHER 1992:84)

¹²⁰ Vgl. auch EDWARDS & BECKMAN 1988, BECKMAN 1996.

¹²¹ Die Schließbewegung hingegen wird nicht beschleunigt (BECKMAN, EDWARDS & FLETCHER 1992:75). Eine Beschleunigung der Schließbewegung würde den Vokaltrakt schneller schließen und so in der Terminologie der Autoren die Sonorität der Silbe verringern.

Diese Art von Sonorität nennen sie im folgenden *segmentale Sonorität*. BECKMAN, EDWARDS & FLETCHER beziehen sich in dieser Definition auf CLEMENTS' (1990) Sonoritätshierarchie, der Sonorität mit den Oberklassenmerkmalen korreliert und eine vierstufige Unterteilung in Obstruenten, Nasale, Liquide und Vokale annimmt (s.o. 2.4.4). In akustischer Hinsicht kann man CLEMENTS' Sonoritätshierarchie auf die Parameter Lautheit und Dauer beziehen, in artikulatorischer Hinsicht auf den Öffnungsgrad des Vokaltraktes.

Die Neuerung gegenüber CLEMENTS' segmentalem Sonoritätskonzept liegt darin, dass BECKMAN, EDWARDS & FLETCHER den Sonoritätswert eines Segmentes analog zum phonologischen Tonhöhenwert eines vokalischen Segmentes in gewisser Weise als Autosegment,¹²² d.h. als von den distinktiven Merkmalen unabhängige Merkmalspezifikation auf einer separaten Sonoritäts-Ebene¹²³ konzipieren. Diese neue Konzeption hat den Vorteil, dass die in der alten Konzeption dem Phonem inhärente Sonoritätsspezifikation (d.h. die segmentale Sonorität), die sich in der neuen Konzeption auf einer autosegmentalen Ebene befindet, von prosodischen Phänomenen¹²⁴ wie der Betonung affiziert werden kann:

just as prosodic effects on F_0 are specified by scaling H and L tones within an overall pitch range, prosodic effects on phonetic sonority are specified by scaling segmental sonority values within a sonority space. (BECKMAN, EDWARDS & FLETCHER 1992:84)

Hierzu ist zunächst zu erläutern, dass die beiden phonologischen Phänomene *Ton* und *Intonation* auf demselben akustischen Parameter beruhen, nämlich auf der Höhe der Grundfrequenz (F_0). *Ton* ist dabei eine Spezifizierung der Tonhöhe oder des Tonhöhenverlaufs, die auf einen Vokal, eine Silbe, ein Morphem oder ein Wort bezogen ist, *Intonation* ist eine Spezifizierung des Tonhöhenverlaufs, die auf eine ganze Äußerung bezogen ist (für einen Überblick vgl. LAVER 1995:450ff). Bei der Produktion einer Äußerung müssen sowohl die *Töne* wie auch die *Intonation* phonetisch, d.h. akustisch bzw. artikulatorisch realisiert werden. Diese Interaktion von Intonation und Tonhöhe kann

¹²² Eine Merkmalspezifikation kann dann als Autosegment gelten, wenn sie unabhängig von einem 'Phonem' (in traditioneller Bedeutung) auf einer eigenen *Ebene* (tier) repräsentiert wird (GOLDSMITH 1990:8): „In autosegmental representation [...] we posit two or more parallel tiers of phonological segments. Each tier itself consists of a string of segments, but the segments on each tier differ with regard to what features are specified in them“. Tonhöhenpezifizierungen gelten deshalb als Autosegmente, weil der Tonhöhenwert eines Vokals auch dann erhalten bleiben kann, wenn dieser Vokal getilgt wird. In diesem Fall wird der Tonhöhenwert mit einem anderen Vokal assoziiert, der dann u.U. zwei Töne, d.h. einen Konturton, trägt. Für eine Darstellung der *Autosegmentalen Phonologie* vgl. GOLDSMITH 1990.

¹²³ Unter einem *tier* (engl. Lage, Schicht) ist eine Repräsentations-Ebene für Autosegmente zu verstehen. Im folgenden wird der Begriff *Ebene* verwendet.

¹²⁴ Unter *Prosodie* werden hier allgemein Phänomene verstanden, die auf größere Einheiten als das Segment bezogen sind, z.B. Wortbetonung, Satzbetonung und Intonation.

als *Skalierung* der Tonspezifizierungen (H und L) durch die darübergelegte Tonhöhenbewegung konzipiert werden.¹²⁵

Auch die beiden phonologischen Phänomene Betonung¹²⁶ und Sonorität¹²⁷ beruhen tendenziell auf denselben akustischen Parametern, nämlich auf der Lautstärke und der Dauer eines Signalabschnittes von Segment- oder Silbendimension. BECKMAN, EDWARDS & FLETCHER (1992) schlagen daher vor, dass *segmentale Sonorität* und *Betonung* sich ähnlich verhalten wie *Ton* und *Intonation* und dass die Einflüsse der Betonung auf die segmentale Sonorität als *Skalierung* der segmentalen Sonorität durch die Betonung in einem Sonoritätsraum modelliert werden können. Anders gesagt, die Betonung vergrößert die segmentale Sonorität, insofern betonte Silben in artikulatorischer Hinsicht durch größere Auslenkung und länger andauernde Maximalöffnung gekennzeichnet sind und in akustischer Hinsicht länger und lauter sind. Ausgehend von diesen Überlegungen nehmen BECKMAN, EDWARDS & FLETCHER an, dass die segmentale Sonorität auf den zwei Parametern *Öffnungsgrad des Vokaltraktes* und *Dauer* basiert. Diese beiden Dimensionen spannen den zweidimensionalen¹²⁸ *Sonoritätsraum* auf:

The sonority space has two dimensions. One dimension is [...] overall vocal tract openness, which can be estimated from jaw height [...] The other dimension of the sonority space is time; a vertical increase in overall vocal-tract openness is necessarily coupled to a horizontal increase in the temporal extent of the vertical specification. (BECKMAN, EDWARDS & FLETCHER 1992:84)

Die beiden Dimensionen *Öffnungsgrad* und *Dauer* sind also miteinander *gekoppelt* und bilden gemeinsam das phonetische Korrelat der Sonorität. Wird die Merkmalspezifizierung eines Segmentes auf der Sonoritäts-Ebene von prosodischen Phänomenen wie Akzent affiziert, dann werden stets beide Parameter skaliert: in artikulatorischer Hinsicht die Geschwindigkeit der Öffnungsbewegung *und* ihre zeitliche Koordination mit der Schließbewegung und in akustischer Hinsicht die Dauer *und* die Lautstärke des einem Segment entsprechenden Signalabschnittes:

We understand lengthening associated with accent, then as part of an increase in overall sonority for the accented syllable's nucleus. The larger mean displacements of accented gestures reflect the vertical increase, and the later phasing represents the coupled horizontal increase. (BECKMAN, EDWARDS & FLETCHER 1992:84)

Zusammenfassend soll noch einmal herausgestellt werden, dass BECKMAN, EDWARDS & FLETCHER in ihrer Untersuchung zum Einfluss der prosodischen Struktur auf die Artikulationsdynamik das Konzept der *Sonorität* aufgreifen, um zu erklären, warum die

¹²⁵ Die Tonhöhe bzw. F_0 ist auch ein akustisches Korrelat des Akzentes, daher muss angenommen werden, dass nicht nur die Intonation, sondern auch der Akzent die Tonspezifizierungen skaliert.

¹²⁶ Als akustische Korrelate der Betonung können Tonhöhe, Lautstärke und Dauer gelten, vgl. LAVER 1995:512.

¹²⁷ Dies gilt nur, wenn man eine tentative phonetische Definition von *Sonorität* als Funktion der *Dauer* und der *Lautstärke* eines Signalsabschnittes schon voraussetzt.

¹²⁸ Im Gegensatz zum eindimensionalen *Tonraum*.

Betonung einer Silbe zwei artikulatorische Parameter gleichermaßen beeinflusst, während andere prosodische Effekte wie finale Längung oder Sprechtempo nur die Dauer beeinflussen. Diese Beobachtung erklären sie mit der Theorie einer Sonoritäts-Ebene, auf der die inhärente *Sonorität* jedes Segmentes spezifiziert ist. Diese segmentale Sonoritätsspezifikation kann in dieser Theorie von der Merkmalspezifikation auf einer höheren Ebene *skaliert* werden, nämlich vom prosodischen Merkmal *Betonung*, welches ganzen Silben zugeordnet wird. Die Skalierung der segmentalen Sonorität durch die Betonung zeigt sich in der phonetischen Interpretation der phonologischen Repräsentation als Skalierung der Parameter Offenheit des Vokaltraktes und Dauer der Silbengeste, in akustischer Hinsicht entspricht dieser Skalierung der segmentalen Sonorität eine längere Dauer und eine größere Lautstärke des Silbennukleus als Moment der maximalen Öffnung.

3.3.2.3 Sonorität und Silbizität

In der vorliegenden Studie geht es im Gegensatz zur Untersuchung von BECKMAN, EDWARDS & FLETCHER (1992) nicht um die Beziehung zwischen der *Akzentstruktur* und der segmentalen *Sonorität*, sondern um die Beziehung zwischen der *Silbenstruktur* und der segmentalen *Sonorität*. In den vorangegangenen Ausführungen wurde hervorgehoben, dass sich silbische Konsonanten von ihren unsilbischen Entsprechungen (bzw. silbische Vokale von unsilbischen Gleitlauten) in phonetischer Hinsicht durch ein Mehr an Lautstärke und Dauer unterscheiden (vgl. PRICE 1980, s.o. 3.2.2). Diese beiden akustischen Dimensionen Lautstärke und Dauer können wiederum als akustische Korrelate der aus der Sonoritätshierarchie abgeleiteten inhärenten Sonorität betrachtet werden (s.o. 3.2.1). Daher ist die Annahme gerechtfertigt, dass silbische Segmente sonoranter sind als unsilbische Segmente (vgl. CLEMENTS 1990, s.o. 2.4.4). Hieraus kann geschlossen werden, dass die Silbifizierung einer Segmentkette die Sonorität einzelner Segmente affiziert, denn nukleare Segmente (d.h. Silbennuklei) sind in phonologischer und phonetischer Hinsicht sonoranter als nicht-nukleare Segmente (d.h. Silbenränder). Im vorangegangenen Kapitel wurde diese Annahme so formuliert, dass die Sonorität eines Segmentes in einer silbifizierten Segmentkette eine Funktion seiner inhärenten oder segmentalen Sonorität und seiner Silbizitätsspezifizierung auf der CV-Ebene ist (s.o. 2.4.5).

Nimmt man mit BECKMAN, EDWARDS & FLETCHER an, dass der Sonoritätswert eines Segmentes auf einer Sonoritäts-Ebene spezifiziert ist und dass dieser Sonoritätswert von den Merkmalspezifikationen auf einer prosodischen Ebene skaliert werden kann, so könnte man weiter annehmen, dass auch die Merkmalspezifikation für ein

Segment auf dem CV-Skelett (s.o. 2.4.4), d.h. die Spezifizierung als [+ silbisch] oder als [– silbisch], die segmentale Sonorität skalieren kann. Diese Konzeption der Sonoritäts-Ebene von BECKMAN, EDWARDS & FLETCHER (1992) ist also geeignet, die Annahme, dass ein nukleares Segment sonoranter ist als ein nicht-nukleares Segment, in eine Theorie von Silbe und Sonorität zu integrieren, denn diese Konzeption ermöglicht eine Interaktion von segmentaler Sonorität und CV-Skelett oder Silbenstrukturrepräsentation. Die oben formulierte Annahme (s.o. 2.4.5), dass die phonologisch-phonetische Sonorität sowohl auf die Segment-Ebene als auch auf die Silben-Ebene bezogen ist und als Verbindungsglied zwischen den segmentalen Eigenschaften und den suprasegmentalen Eigenschaften eines sprachlichen Ausdrucks gefasst werden kann, ist demnach auch mit der Theorie von BECKMAN, EDWARDS & FLETCHER (1992) kompatibel. Welche Vorteile eine solche Theorie, welche die segmentale Sonorität, die Silbenstruktur und die Akzentstruktur auf diese Weise zueinander in Bezug setzt, darüberhinaus hat, soll im folgenden Abschnitt diskutiert werden.

3.4 Sonorität, Silbenstruktur und prosodische Struktur

BECKMAN, EDWARDS & FLETCHER (1992) konnten nachweisen, dass betonte Silben sich von unbetonten in artikulatorischer Hinsicht durch eine schnellere Öffnung und eine längere Dauer des Momentes mit maximaler Offenheit unterscheiden. Die artikulatorischen Korrelate der *Betonung* sind also dieser Untersuchung zufolge der *Öffnungsgrad* des Vokaltraktes und die *Dauer* der Öffnung. In akustischer Hinsicht unterscheiden sich betonte Silben von unbetonten u.a. dadurch, dass sie lauter und länger sind (vgl. LAVER 1995:512).¹²⁹

Über *Sonorität* hingegen wurde gesagt, dass sonorantere Segmente lauter sind als weniger sonorante (vgl. z.B. SIEVERS 1901, JESPERSEN 1904). Diese Aussage wurde damit begründet, dass die Sonoritätshierarchie ungefähr der Hierarchie der Segmente entspricht, die sich ergibt, wenn man sie nach ihrer intrinsischen *Lautheit* anordnet. Daneben gilt auch, dass sonorantere Segmente eine längere *Dauer* haben als weniger sonorante: Plosive sind kürzer als Frikative, Liquide sind kürzer als Vokale, usw.¹³⁰

¹²⁹ Betonte Silben haben auch eine höhere Tonhöhe als unbetonte. Im Englischen unterscheiden sich die Segmente in betonten Silben daneben in ihrer spektralen Qualität von Segmenten in unbetonten Silben, was damit zusammenhängt, dass, wie BECKMAN, EDWARDS & FLETCHER (1992) gezeigt haben, der Vokaltrakt bei der Artikulation von betonten Silben insgesamt weiter geöffnet wird, was seine Resonanzeigenschaften, und damit die spektrale Qualität der Segmente, verändert (vgl. LAVER 1995:512).

¹³⁰ Für eine Untersuchung zur durchschnittlichen Länge verschiedener Phoneme des Englischen vgl. CRYSTAL & HOUSE 1988.

SAUSSURE führte die Sonoritätshierarchie auf den *Öffnungsgrad* des Vokaltraktes zurück und nahm an, dass sonorantere Laute sich von weniger sonoranten durch einen höheren Grad der Öffnung des Vokaltraktes unterscheiden. Schließlich konnte PRICE (1980) zeigen, dass silbische Liquidkonsonanten, die sonoranter sind als nichtsilbische Liquide, von diesen auf der Basis ihrer Dauer und ihrer Lautstärke perzeptuell unterschieden werden.

Aus diesen Beobachtungen kann geschlossen werden, dass die akustischen und artikulatorischen Korrelate von segmentaler Sonorität, von Silbizität und schließlich von Betontheit mehr oder weniger identisch sind, nämlich Lautheit bzw. Öffnungsgrad des Vokaltraktes und Dauer eines Segmentes bzw. der Vokaltraktöffnung, die zu seiner Artikulation nötig ist. Diese Entsprechung der phonetischen Korrelate von Sonorität, Silbizität und Betontheit ist ein weiteres Argument dafür, dass es sinnvoll ist, eine segmentale Sonoritäts-Ebene zu postulieren, deren Merkmalspezifikationen von Silbenstruktur und Akzentstruktur skaliert werden können und die Annahme zu machen, dass sowohl die Zuordnung eines Segmentes zum Silbennukleus als auch die Betonung einer Silbe mit einer Vergrößerung der Sonorität des Segmentes bzw. der Silbe einhergeht. Diese Annahme ist auch mit einigen Beobachtungen von KOHLER (1986) zur Dauer von Silben und Segmenten im Deutschen vereinbar, der zeigte, dass die Dauer einzelner Segmente von ihrer Position in der Silbe ebenso beeinflusst wird wie von ihrer Position im Takt als einer Folge von betonten und unbetonten Silben. Auch diese Beobachtung lässt sich so interpretieren, dass die Silbenstruktur die Dauer der Segmente auf ähnliche Weise affiziert wie die prosodische Struktur.

Ein weiteres Argument für die Angemessenheit des Modells von BECKMAN, EDWARDS & FLETCHER (1992) und seiner Übertragung auf die Ebene der Silbenstruktur lässt sich aus den folgenden Überlegungen ableiten. Einerseits ist es so, dass die Betonung einer Silbe nicht nur die *vokalische Artikulationsgeste* beeinflusst, d.h. den Moment maximaler Offenheit des Vokaltraktes, sondern auch die *konsonantische Öffnungsgeste* – nicht nur dauert die maximale Öffnung länger an, sondern die Öffnungsbewegung wird auch schneller ausgeführt. Dies bedeutet, dass die Betonung nicht nur den Vokal als Nukleus der betonten Silbe affiziert, sondern auch den Anlaut der Silbe. Auf der anderen Seite wurde demonstriert (PRICE 1980), dass die Silbizität eines Segmentes nicht nur einer Änderung der Eigenschaften dieses Segmentes selbst entspricht, sondern darüberhinaus das vorangehende Segment in seiner Stimmeinsatzzeit affiziert, d.h. die zeitliche Koordination von Verschlussöffnung und Einsetzen der Stimmlippenschwingung beeinflusst. Hier findet sich also eine Parallelität zwischen Silbenstruktur und Akzentstruktur in Bezug auf die Domäne, auf welche die phonetischen Korrelate von Silbizität und Betontheit bezogen sind – nicht nur auf ein einziges

Segment bzw. auf den Silbennukleus, sondern auch auf vorangehende Segmente bzw. auf den Silbenanlaut.¹³¹ Diese Beobachtungen können nur durch ein phonologisches Modell erklärt werden, in dem einerseits die Spezifizierung als [+silbisch] auf dem CV-Skelett nicht nur das Segment beeinflusst, welches von diesem Merkmal unmittelbar dominiert wird (d.h. das nukleare Segment), sondern auch die vorangehenden (und eventuell die nachfolgenden) Segmente und andererseits die Akzentuierung einer Silbe nicht nur den Silbennukleus als Kopf der Silbe affiziert, sondern auch die vorangehenden Segmente.

Schließlich soll noch einmal hervorgehoben werden, dass die phonetischen Korrelate von Sonorität, Silbizität und Betontheit mehrdimensional sind: BECKMAN, EDWARDS & FLETCHER (1992) sprechen vom zweidimensionalen Sonoritätsraum, PRICE (1980) untersuchte zusätzlich zur Lautheit und Dauer der Segmente noch die Stimmhaftigkeit (als Stimmeinsatzzeit des vorangehenden Plosivkonsonanten) und kam zu dem Ergebnis, dass alle drei untersuchten Parameter die Wahrnehmung von Silbizität beeinflussen. Denkbar ist auch, dass die Stimmeinsatzzeit eines Konsonanten nicht nur mit der Silbizität bzw. Sonorität des nachfolgenden Segmentes, sondern auch mit der *Betontheit der Silbe* korreliert ist. Ebenso wie die Öffnungsbewegung in betonten Silben schneller ausgeführt wird als in unbetonten, könnte auch die Adduktionsbewegung¹³² der Stimmlippen in betonten Silben schneller als in unbetonten ausgeführt werden, was eine kürzere Stimmeinsatzzeit zu Folge hätte, da die Stimmlippen schneller zu schwingen begännen. Diese Frage erfordert weitere Untersuchungen der prosodischen Effekte auf die akustische und artikulatorische Realisierung von Segmenten. Auch könnte es so sein, dass in einer Untersuchung der phonetischen Korrelate der Silbizität noch andere akustische bzw. artikulatorische Parameter zu berücksichtigen sind, z.B. eine Veränderung der Tonhöhe in silbischen gegenüber unsilbischen Segmenten, oder eine Veränderung der spektralen Qualität. Diese beiden Parameter sind als akustische Korrelate der Betonung identifiziert worden; hieraus könnte man ableiten, dass es sich dabei auch um potentielle phonetische Korrelate der Silbizität handelt. Zuletzt ist noch zu bedenken, dass die phonetischen Korrelate der Betonung zu einem gewissen Grade sprachspezifisch sind: Lautstärke, Dauer und Tonhöhe werden von fast allen Sprachen genutzt, segmentale Qualität hingegen nur in einigen Sprachen, z.B. im Englischen

¹³¹ BECKMAN, EDWARDS & FLETCHER (1992) wiesen nach, dass die Betontheit einer Silbe die Verschlussgeste nicht beeinflusst. Ob die Silbizität eines Segmentes nicht nur vorangehende, sondern auch nachfolgende Segmente affiziert, muss noch untersucht werden.

¹³² Um durch die ausströmende Luft zum Schwingen gebracht zu werden, müssen die Stimmlippen in die geschlossene Position gebracht, d.h. *adduziert* werden (vgl. POMPINO-MARSCHALL 1995:31ff).

(LAVER 1995:512). Es ist daher möglich, dass auch die phonetischen Korrelate der Silbizität sprachspezifisch sind.

Auf der anderen Seite stellt sich die Frage, ob das *phonologische Modell* von BECKMAN, EDWARDS & FLETCHER (1992), in welchem die segmentale Sonorität durch die Silbenstruktur und die Akzentstruktur skaliert werden kann, überhaupt impliziert, dass diese Skalierungsfunktion auf der Ebene der *phonetischen Interpretation* auf dieselbe Weise realisiert werden muss, d.h. ob gleiche *phonologische Funktionen* auf verschiedenen Ebenen durch dieselben *phonetischen Parameter* ausgedrückt werden müssen. Möglich wäre es, dass stattdessen die Relation zwischen segmentaler Sonorität und Silbenstruktur einerseits und segmentaler Sonorität und Akzentstruktur andererseits auf der phonologischen Ebene zwar auf dieselbe Weise repräsentiert wird, dass aber für die beiden Ebenen Silbenstruktur und Akzentstruktur unterschiedliche Prinzipien der phonetischen Interpretation gelten. Die Beantwortung dieser Frage setzt eine explizitere Beschreibung des Prozesses der phonetischen Interpretation von (insbesondere prosodischen) phonologischen Strukturen voraus, als gegenwärtig zur Verfügung steht. Bis dahin soll als Arbeitshypothese gelten, dass sich dieselben oder möglichst ähnliche Prinzipien der phonologischen Strukturierung und der phonetischen Interpretation auf verschiedenen Beschreibungsebenen wie Segmentebene, Silbenstruktur, Akzentstruktur und Intonationskontur wiederfinden. Diese Arbeitshypothese soll auch deshalb aufrechterhalten werden, weil sich auf diese Weise ausgehend von bereits Bekanntem neue Hypothesen bilden lassen.

Schließlich muss noch darauf hingewiesen werden, dass in artikulatorisch-phonetischer Hinsicht das Konzept Sonorität, wie BECKMAN, EDWARDS & FLETCHER (1992) es fassen, auf die allgemeine Dynamik der Artikulation und auf die Art der Koordination der segmentalen Gesten bezogen ist, denn die Betonung einer Silbe verändert nicht die distinktiven Eigenschaften der Gesten, also nicht das, was ein /b/ von einem /g/ unterscheidet, sondern sie verändert die Art und Weise, wie diese Gesten ausgeführt werden: schneller, mit größerer Auslenkung und in größerem zeitlichen Abstand. Auf die gleiche Weise kann man sagen, dass Segmente, die sich in ihrer inhärenten Sonorität unterscheiden, sich auch in der groben Gestalt der Artikulationsbewegung unterscheiden, mit der sie hervorgebracht werden. Dies ist die Art der Kennzeichnung der Oberklassen, die CHOMSKY & HALLE (1968) anführen und die hier noch einmal wiederholt wird (s.o. 2.4.2):

Reduced to the most rudimentary terms, the behavior of the vocal tract in speech can be described as an alternation of closing and opening. [...] This skeleton of speech production provides the basis for the major class features, that is, the features that subdivide speech sounds into vowels, consonants, obstruents, sonorants, glides, and liquids. (CHOMSKY & HALLE 1968:302)

Identifiziert man die segmentale Sonorität mit den Oberklassen (s.o. 2.4), so kann man CHOMSKY & HALLES Begriff des „skeleton of speech production“ auf die Sonoritätskontur der Sprache beziehen. Die Sonoritäts-Ebene kann dann als die Ebene der Sprache gefasst werden, auf der die allgemeine Dynamik der Artikulation spezifiziert ist und die als das ‘Skelett der Sprache’, oder als die ‘Gestalt der Sprache’ bezeichnet werden kann. Die allgemeine Dynamik der Artikulation einer Sequenz von invarianten segmentalen Gesten würde in dieser Konzeption einerseits von Prozessen der Silbifizierung von Segmenten zu größeren Einheiten affiziert, die eine gegenseitige Anpassung der Artikulationsdynamik erfordern und andererseits von darüber ablaufenden Prozessen der Betonung einzelner Silben einer Äußerung gegenüber anderen, die einer Veränderung der Artikulationsdynamik einer betonten gegenüber einer unbetonten Silbe entsprechen würden.

Auch im Hinblick auf eine *Theorie der Sprachwahrnehmung* bietet die Annahme einer segmentalen Eigenschaft der Sonorität, welche die Basis der Silbenstruktur und die Basis der Akzentstruktur der Sprache bildet, einen großen Vorteil. Nimmt man an, dass der Hörer die zwei akustischen Parameter Lautheit und Dauer (und andere, wie z.B. Stimmhaftigkeit) verarbeitet, um eine Repräsentation der Sonoritätskontur eines Sprachsignals zu erzeugen, so kann diese Sonoritätskontur in einem ersten Schritt (und in einem kurzen Zeitbereich von Silbendauer) genutzt werden, um das Signal in Silben zu segmentieren, die Silbennuklei zu lokalisieren und schließlich eine Silbenstrukturrepräsentation zu erzeugen. In einem zweiten Schritt (und in einem längeren Zeitbereich über mehrere Silben hinweg) könnte der Hörer aus einer Repräsentation der relativen Sonorität der Silben eine Repräsentation der Akzentstruktur der Äußerung aufbauen. Dieser Gedanke wird im folgenden Kapitel wiederaufgegriffen (s.u. 4.4.2).

Die vorangegangenen Ausführungen haben erstens gezeigt, dass die Konzeption von BECKMAN, EDWARDS & FLETCHER (1992) zur Interaktion von Betonung und segmentaler Sonorität auch geeignet ist, die Interaktion von Silbenstruktur und segmentaler Sonorität in eine linguistische Theorie zu integrieren. Zweitens wurde hervorgehoben, dass segmentale Sonorität, Silbizität und Betonung auf denselben oder zumindest auf ähnlichen akustischen bzw. artikulatorischen Parametern beruhen. Was aber wird darüberhinaus gewonnen durch die Annahme einer Sonoritäts-Ebene, die von Silbenstruktur und Akzentstruktur gleichermaßen affiziert wird?

Unter einer solchen Sichtweise könnte man annehmen, dass die phonologisch-phonetische Form der Sprache aus zwei übereinandergelagerten Strukturierungsebenen besteht. Auf beiden Ebenen findet sich eine Variation von *Melodie* und *Sonorität* (oder

*Prominenz*¹³³) und beide Ebenen gliedern sich in Struktureinheiten. Die Struktureinheiten der oberen Ebene sind *Silben*, die Struktureinheiten der unteren Ebene sind *Segmente*. Auf der oberen Ebene findet sich eine Variation der Sprachmelodie, d.h. der Intonation und der Silbensonorität, d.h. ihrer Betontheit. Auf der unteren Ebene findet sich eine Variation der segmentalen Melodie,¹³⁴ d.h. der akustischen bzw. artikulatorischen Qualität der Segmente und der segmentalen Sonorität, d.h. der strukturellen Komplexität¹³⁵ und des Gewichts¹³⁶ der Silbe. Die Variation auf der unteren Ebene hat die Funktion, Wörter zu unterscheiden, die Variation auf der oberen Ebene hat die Funktion, morphologische, syntaktische oder pragmatische Strukturen zu repräsentieren, d.h. z.B. Morphem- und Phrasengrenzen zu signalisieren, Information hervorzuheben oder modale Eigenschaften der Äußerung zu kodieren.

Eine naheliegende Frage ist es, ob diese Art der Strukturierung von Lautäußerungen durch die Variation von Melodie und Sonorität zu den spezifischen Eigenschaften menschlicher Sprache gehört, oder ob sie sich auch in nicht-natürlichsprachlichen Kommunikationssystemen wiederfindet. Es ist denkbar, dass die lautlichen Kommunikationssysteme verschiedener Spezies sich darin unterscheiden, ob sie nur über *eine* Strukturebene verfügen, oder ob sie *zwei* Strukturebenen nutzen. Variationen von Tonhöhe und Lautstärke auf nur einer Ebene, d.h. der *Silbenebene*, lassen sich vermutlich in allen lautlichen Kommunikationssystemen finden, Variationen auf einer zweiten Strukturebene der *segmentalen* Melodie und Sonorität (d.h. Variationen in Melodie und Sonorität der Struktureinheiten, welche die erste Ebene konstituieren) vielleicht nur in den komplexeren Systemen wie der menschlichen Sprache, in Vogelgesängen¹³⁷ und bei einigen Primaten. Auch Neugeborene lernen zuerst, die Parameter Intonation und Lautstärkevariation (d.h. Melodie und Sonorität auf der oberen Ebene) zum Ausdruck

133 Zum Begriff der *Prominenz* vgl. LAYER (1995:450): „Syllables vary in their perceptual prominence. [...] Other things being equal, one syllable is more prominent than another to the extent that its constituent segments display higher pitch, greater loudness, longer duration or greater articulatory excursion from the neutral disposition of the vocal tract.“

134 Zum Begriff der segmentalen Melodie vgl. BLEVINS 1995.

135 Ein komplexer Anlaut, der aus drei Segmenten besteht, z.B. /pfl/, bildet einen langsameren Sonoritätsanstieg als ein einfacher Anlaut wie /b/, der aus nur einem Plosivkonsonanten besteht.

136 Eine geschlossene Silbe wie /bal/ in *Balken* ist schwere als eine offene Silbe wie /ba/ in *Backe*. Beide Silben unterscheiden sich in ihrem Sonoritätsverlauf: im ersten Fall findet sich ein langsamer Abfall, im zweiten Fall ein abrupter Abfall (vgl. hierzu auch das Silbenschnittkonzept, TRUBETZKOY 1977⁶:196f.). Das Gewicht einer Silbe, das hier in den Zusammenhang der Sonorität gestellt wird, üblicherweise aber auf die More als Gewichtseinheit bezogen wird (für einen Überblick vgl. ROCA 1994:177ff), interagiert in vielen Sprachen mit der Akzentebene, s.o. 2.3.2 und s.u. 4.3.3.

137 Kanarienvögel z.B. unterscheiden verschiedene Typen von Silben, die aus verschiedenen Mustern von Tonhöhen- und Lautstärkemodulation bestehen und die auf eine charakteristische Weise zu größeren Sequenzen angeordnet werden, welche so ihrerseits eine Modulation von Melodie und Sonorität erhalten (vgl. VALLET, KREUTZER & RICHARD 1992).

zu nutzen und sind erst viel später in der Lage, ihre Äußerungen auch auf der unteren Ebene von segmentaler Melodie und Sonorität zu strukturieren, d.h. Phoneme zu unterscheiden und die Silbenstruktur zu variieren (vgl. JUSZYCK 1995). Ob eine solche Konzeption von *Melodie* und *Sonorität* für die Untersuchung menschlicher und nicht-menschlicher lautlicher Kommunikation in einem allgemeineren Rahmen als der natürlichsprachlichen Phonologie geeignet sind, ist noch zu zeigen.

Diese letzten Ausführungen hatten vor allem den Zweck, das Konzept der Sonorität in einen allgemeineren Rahmen zu stellen und zu verdeutlichen, dass Sonorität nicht nur ein abstraktes phonologisches Merkmal ist, welches aus der Sonoritätshierarchie der Segmente abgeleitet werden kann, sondern zu den in jeder lautlichen Kommunikation variierten Eigenschaften gehört. Diese Auffassung der Sonoritätsvariation einer Äußerung als Konstante in lautsprachlicher Kommunikation kann als weiteres Argument für die Annahme dienen, dass die Sonoritätskontur einer Äußerung bei der Wahrnehmung von Sprache genutzt wird, um das Signal zu strukturieren und zu segmentieren (s.u. 4.4).

3.5 Zusammenfassung

Will man die phonetischen Korrelate der phonologischen Sonorität messen, so muss man sie zunächst definieren. In diesem Kapitel wurden drei verschiedene Möglichkeiten erörtert, Sonorität zu definieren, um ihre phonetischen Korrelate zu ermitteln. Erstens kann man Sonorität als dasjenige Merkmal fassen, welches die Ordnung der Laute in der *Sonoritätshierarchie* bestimmt, die aus Beobachtungen über die phonotaktischen Regularitäten in der Sprache abgeleitet wurde (SIEVERS 1901⁵, JESPERSEN 1904, SAUSSURE 1972). In diesem Zusammenhang wurde phonologische Sonorität auf die inhärente Lautheit der Segmente und den Öffnungsgrad des Vokaltraktes bei ihrer Artikulation bezogen. Die Sonoritätshierarchie kann nicht allein auf diese zwei Eigenschaften zurückgeführt werden, es ist jedoch prinzipiell möglich, eine Funktion zu definieren, die auf der Basis von gegebenen akustischen und / oder artikulatorischen Eigenschaften jedem Segment (oder jedem Signalabschnitt in einem akustischen Signal) den richtigen Sonoritätswert zuordnet. Auf diese Weise sollte es auch möglich sein, einen Algorithmus zu entwickeln, der einem gegebenen Sprachsignal Sonoritätswerte zuweist, aus denen dann eine Silbenstrukturrepräsentation erzeugt wird, in welche schließlich erkannte Phoneme eingefügt werden könnten. Dass auch der Hörer den Sonoritätsverlauf nutzen könnte, um das Signal in Silben zu segmentieren und um Silbenkerne zu lokalisieren, wird im folgenden Kapitel gezeigt.

Zweitens kann man *Sonorität* auf *Silbizität* beziehen und die akustischen Parameter ermitteln, die dazu führen, dass ein Segment, welches silbisch oder nicht-silbisch sein kann, als *silbisch* und damit als *Sonoritätsmaximum* wahrgenommen wird. Auf diese Weise kann ein Mehr an Sonorität bei gleichbleibenden segmentalen Eigenschaften gemessen werden. PRICE 1980 zeigte, dass es die akustischen Eigenschaften Dauer und Lautstärke des silbischen Segmentes und Stimmeinsatzzeit des vorangehenden Plosivkonsonanten sind, welche die Wahrnehmung als Sonoritätsmaximum beeinflussen.

Drittens kann man die *Sonorität* von Silben auf ihre *Betontheit* beziehen (BECKMAN, EDWARDS & FLETCHER 1992). Betonte Silben sind nach dieser Auffassung sonoranter als unbetonte. Als artikulatorische Parameter der so definierten Sonorität wurden Mundöffnungsgrad und Dauer der maximalen Öffnung ermittelt. Auf akustischer Seite entspricht der Betontheit oder dem Mehr an Sonorität eine größere Lautstärke und eine längere Dauer des Signalabschnittes mit maximaler Lautstärke. Auf der Basis dieser Beobachtung konzipieren BECKMAN, EDWARDS & FLETCHER ein phonologisches Modell, in welchem die segmentale Sonorität auf einer Sonoritäts-Ebene spezifiziert ist und von Merkmalspezifizierungen auf der prosodischen Ebene skaliert wird. Diese Analogie von Sonorität und Silbenstruktur auf der einen Seite und Sonorität und Akzentstruktur auf der anderen Seite wird im nächsten Kapitel wiederaufgegriffen (s.u. 4.4).

Auf der Grundlage dieser drei Definitionsweisen von Sonorität lassen sich demnach dieselben phonetischen Korrelate der phonologischen Sonorität ermitteln: Dauer und Lautstärke bzw. Mundöffnungsgrad. Aus diesem Grund, so wurde argumentiert, ist es sinnvoll, mit BECKMAN, EDWARDS & FLETCHER (1992) eine Sonoritäts-Ebene anzunehmen, auf der die segmentale Sonorität spezifiziert ist, die erstens von der Silbizifizierung der Segmente affiziert wird, insofern silbische Segmente sonoranter sind als unsilbische und zweitens von der Akzentstruktur einer Äußerung, insofern betonte Silben sonoranter sind als unbetonte. Schließlich wurde darauf hingewiesen, dass eine solche Hypothese auch im Rahmen einer Theorie der Sprachwahrnehmung sinnvoll sein kann, weil auf diese Weise die Wahrnehmung von Silbenstruktur und Betonung auf eine gemeinsame akustische und symbolische Basis zurückgeführt wird. Die möglichen Funktionen des phonologischen Merkmals Sonorität für die Sprachwahrnehmung werden im folgenden Kapitel erörtert.

4 Silbe und Sonorität in der Sprachwahrnehmung

4.1 Einleitung

Will man linguistische Repräsentationen, d.h. Beschreibungen der Struktur von Äußerungen, auf Beschreibungen von dynamischen Aktivierungsvorgängen im Gehirn beziehen, wie diese Studie es für Silbe und Sonorität anstrebt, so müssen die *statischen* linguistischen Repräsentationen zunächst auf *dynamische* Repräsentationen bezogen werden: auf Beschreibungen der Struktur von Äußerungen in einer zeitlichen Reihenfolge (SCHNELLE 1991).¹³⁸ Aussagen über die Dynamik linguistischer Repräsentationen werden in psycholinguistischen Modellen gemacht, die dynamische Verhaltensbeschreibungen umfassen, d.h. die Aussagen machen über die Verarbeitungsschritte, die beim Sprechen oder bei der Verarbeitung von Gesprochenem vollzogen werden. Dieses Kapitel beschäftigt sich aus diesem Grund mit der Rolle der Kategorie *Silbe* und des Merkmals *Sonorität* in psycholinguistischen Modellen der Sprachwahrnehmung. Wie die vorangegangenen Ausführungen gezeigt haben, ist das phonetisch-phonologische Merkmal Sonorität prinzipiell geeignet, als Grundlage für die Rekonstruktion einer Silbenstruktur aus dem akustischen Signal zu fungieren. Daher wird in den folgenden Ausführungen nur die Wahrnehmungsseite der Sprachfähigkeit betrachtet, nicht aber die Produktionsseite.

Zu Beginn dieses Kapitels (4.2) werden einige Überlegungen zu einem psycholinguistischen Modell der Sprachwahrnehmung angestellt. Zunächst (4.2.1) werden

¹³⁸ Die zeitliche Dimension, die hier eingeführt wird, ist nicht die zeitliche Dimension, die sich durch die lineare Abfolge der Segmente, Silben und Wörter in einer Äußerung ergibt. Hier ist vielmehr gemeint, dass auch die Verarbeitung einer einzigen Äußerungseinheit Zeit kostet, weil diese Äußerungseinheit eine Strukturierung auf verschiedenen Ebenen aufweist, die zwar gleichzeitig gegeben sein können, aber möglicherweise nacheinander verarbeitet werden. Die Dynamik der Repräsentationen, von der hier die Rede ist, ist also die zeitliche Ordnung der Rekonstruktion einer segmentalen Repräsentation, einer Akzentstruktur, einer lexikalischen oder morphologischen Repräsentation usw.

zwei grundlegende Probleme der Sprachwahrnehmung beschrieben. Diese Erläuterungen dienen als Hintergrund für die nachfolgenden Ausführungen zur Funktion der Silbe als Einheit der Segmentation (s.u. 4.3). Sodann (4.2.2) wird die Frage diskutiert, in welcher Ebene einer Theorie der Sprachwahrnehmung phonologisch-phonetische Repräsentationen sinnvoll integriert werden können. Schließlich (4.2.3) wird gezeigt, dass aus der Variabilität, die sich in den phonologischen Systemen verschiedener Sprachen findet, geschlossen werden muss, dass sprachliche Äußerungen auf sehr unterschiedliche Weise verarbeitet werden können. Diese Erläuterungen schaffen die Voraussetzungen für die Diskussion von universalen und sprachspezifischen Verarbeitungsstrategien in der Sprachverarbeitung (s.u. 4.3.4). Danach werden (4.3) einige Studien zur Funktion der Kategorie *Silbe* in der Sprachwahrnehmung referiert. Diese Ausführungen sollen zeigen, dass die Silbe nicht nur eine notwendige Konstituente einer *phonologischen* Theorie der Sprache ist, sondern auch in einer *psycholinguistischen* Theorie der menschlichen Sprachfähigkeit nicht fehlen darf. Ausgehend von der im letzten Kapitel gemachten Annahme, dass das Merkmal Sonorität die Grundlage für die Rekonstruktion einer Silbenstruktur- und einer Akzentstrukturepräsentation aus dem akustischen Signal bilden könnte, wird im Anschluss daran (4.4) die Annahme formuliert, dass bei der Verarbeitung von gesprochener Sprache auch die Sonoritätskontur einer Äußerung verarbeitet wird.

4.2 Überlegungen zu einem psycholinguistischen Modell der Sprachwahrnehmung

4.2.1 Probleme der Sprachwahrnehmung

Sprachwissenschaftliche Beschreibungen der Struktur von Äußerungen beziehen sich auf diskrete Einheiten: u.a. auf Sätze, Wörter und einzelne Laute. Auch wenn wir schreiben, repräsentieren wir einen Laut mit einem Buchstaben, Wörter trennen wir durch Leerstellen voneinander und Sätze beenden wir mit einem Punkt. Wenn wir aber sprechen, bewegen wir viele verschiedene Artikulatoren (Lippen, Zunge, Gaumensegel, Stimmlippen usw.) gleichzeitig und bilden auf diese Weise Folgen von Konstellationen der Artikulatoren im Artikulationstrakt, von denen eine jeweils für einen Laut charakteristisch ist. Natürlich realisieren wir diese Folgen von Konstellationen als fließende Bewegung: wir führen beim Sprechen kontinuierliche Artikulationsbewegungen durch, ohne innezuhalten oder in eine neutrale Position zurückzukehren und daher ist jede Konstellation der Artikulatoren zu einem bestimmten Zeitpunkt beeinflusst durch ihre

Konstellation zu vorhergehenden Zeitpunkten und zu nachfolgenden Zeitpunkten. Diese Tatsache wird *Koartikulation* genannt.¹³⁹

Hieraus ergeben sich bei der Sprachwahrnehmung als Rekonstruktion linguistischer Beschreibungen der diskreten Struktur der zu verarbeitenden Äußerung zwei Probleme: das *Invarianzproblem* und das *Segmentationsproblem* (vgl. MILLER & EIMAS 1995, NYGAARD & PISONI 1995). Auf der akustischen Seite zeigt sich die fließende Koartikulation der einzelnen Segmente in einem kontinuierlichen Sprachsignal ohne Pausen, in welchem die akustische Information, die genutzt wird, um einen Laut zu erkennen, über mehr als einen Signalabschnitt verteilt ist und darüberhinaus in Abhängigkeit von den umgebenden Lauten variiert. Das *Invarianzproblem* (genauer: Problem der *fehlenden* Invarianz) bezieht sich dabei auf die Tatsache, dass je nach Kontext ein und demselben Phonem verschiedene akustische Muster entsprechen können, da eben der lautliche Kontext die Konstellation des Vokaltraktes bei der Artikulation eines gegebenen Phonems beeinflusst, was auf akustischer Seite unterschiedlichen Signalen entspricht.¹⁴⁰ Das *Segmentationsproblem* (auch: Diskretheitsproblem oder Problem der fehlenden Diskretheit) bezieht sich hingegen auf die Tatsache, dass die Grenzen einzelner Laute, Silben oder Wörter aus dem akustischen Signal nicht zu rekonstruieren sind, da die Artikulationsbewegungen, welche diesen Einheiten entsprechen, fließend ineinander übergehen. Beide Probleme, Invarianzproblem und Segmentationsproblem, ergeben sich aus dem hohen Grad an motorischer Effizienz, mit dem sprachliche Artikulationsbewegungen ausgeführt werden.

Einerseits ist es so, dass die Koartikulation von Phonemen es uns ermöglicht, sehr schnell zu sprechen und die daraus resultierende redundante¹⁴¹ Verteilung der akustischen *cues* für die Lauterkennung im Signal ermöglicht es uns, Sprache sehr schnell und störungsresistent zu verstehen. Andererseits führt dies dazu, dass es keine eins-zu-eins-Entsprechung von artikulatorischen oder akustischen Einheiten und den Einheiten der linguistischen Repräsentation geben kann. Ein Modell der Sprachwahrnehmungsfähigkeit muss daher erstens Aussagen darüber machen, wie die Sprecher einer Sprache das kontinuierliche Sprachsignal für die weitere Verarbeitung überhaupt erst in Einheiten segmentieren und zweitens darüber, wie sie, nachdem sie ein längeres Signal an

¹³⁹ Für einen Überblick zu Koartikulationsphänomenen in gesprochener Sprache vgl. LAVER 1995:149ff, FARNETANI 1997, zum Deutschen vgl. KOHLER 1977:213ff.

¹⁴⁰ Nicht nur die Koartikulation der Segmente führt zu fehlender Invarianz im Signal; auch andere Faktoren wie Sprechtempo und Sprecheridentität spielen hier eine Rolle.

¹⁴¹ Die redundante Verteilung der akustischen Information, die zur Identifikation eines einzelnen Phonems verarbeitet werden kann, ergibt sich daraus, dass die Konstellation des Vokaltraktes zu einem gegebenen Zeitpunkt nicht nur Information zur Identität des gerade artikulierten Phonems, sondern auch der vorangehenden und nachfolgenden Phoneme enthält.

verschiedenen Stellen segmentiert haben, diese Segmente trotz der fehlenden Invarianz der entsprechenden akustischen Muster als Folgen von invarianten linguistischen Einheiten, d.h. Phonemen, Silben oder Wörtern, erkennen. In diesem Kapitel wird gezeigt, dass die Integration der Kategorie Silbe und des Merkmals Sonorität in ein Modell der Sprachwahrnehmung einen Beitrag zur Lösung des Segmentierungsproblems leistet.

4.2.2 Lexikalische und prälexikalische Repräsentation

In diesem Kapitel sollen die abstrakten linguistischen Konzepte Silbe und Sonorität auf ein dynamisches Modell der Sprachverarbeitung bezogen werden. Daher muss zunächst die Frage geklärt werden, auf welcher Ebene einer psycholinguistischen Theorie der Sprache die abstrakten phonologisch-phonetischen Repräsentationen integriert werden können. Zu diesem Zweck werden zunächst die beiden Prozesse Sprachwahrnehmung und Worterkennung differenziert, welche zwei Komponenten der Spracherkennungsfähigkeit konstituieren. Mit *Spracherkennungsfähigkeit* wird die Komponente der *Sprachfähigkeit* bezeichnet, die aus einem akustischen Signal eine symbolische Repräsentation rekonstruiert. Der Prozess der Rekonstruktion einer symbolischen Repräsentation aus dem akustischen Signal kann unter zwei Blickwinkeln betrachtet werden: Erstens als Rekonstruktion einer *prälexikalischen symbolischen*¹⁴² *Repräsentation* aus dem akustischen Signal, d.h. einer phonologischen Repräsentation, die Phoneme, Silben oder segmentale Merkmal umfasst,¹⁴³ oder zweitens als Rekonstruktion einer *lexikalischen Repräsentation*, d.h. als Repräsentation der im Signal erkannten Wörter (MCQUEEN & CUTLER 1997, vgl. auch HARLEY 1995:41, 45f). Theoretische Ansätze, welche die Frage nach der Form der prälexikalischen Repräsentationen in den Mittelpunkt stellen, werden häufig mit dem Begriff *Sprachwahrnehmungsforschung* (*speech perception*, vgl. z.B. NYGAARD & PISONI 1995) bezeichnet; Ansätze, die sich mit dem Prozess des Zugriffs aufs Lexikon beschäftigen, fallen unter dem Begriff *Worterkennung* (*spoken word recognition*, vgl. z.B. CUTLER 1995).¹⁴⁴ Die vorliegende Studie fällt

¹⁴² Prälexikalische Repräsentation können auch als nichtsymbolische Repräsentationen konzipiert werden, z.B. in Form von *spectral templates*, vgl. KLATT 1989.

¹⁴³ Vgl. MCQUEEN & CUTLER 1997:570: „Evidence from speech normalization suggests that there is an intermediate level of prelexical processing, mediating between the raw acoustic input and the mental lexicon. [...] it is clear that pre-lexical processing entails the transformation of acoustic-phonetic information into more abstract representations.“

¹⁴⁴ MILLER & EIMAS 1995:468; für eine andere Konzeption der Sprachverarbeitung vgl. LAHIRI & MARSLEN-WILSON 1991. Will man statische linguistische Beschreibungen der Struktur von Äußerungen auf Aussagen über dynamische Prozesse der Verarbeitung von Äußerungen beziehen, so braucht man ein Modell der Sprachverarbeitung, das als Bezugspunkt dient. Die Annahme einer

in die erste Kategorie: sie bezieht sich nicht auf die Ebene der lexikalischen Verarbeitung bzw. Worterkennung, sondern auf die Ebene der prälexikalischen Verarbeitung als Rekonstruktion einer phonologischen Repräsentation der zu verarbeitenden Äußerung.

Eine der Fragen, die man im Hinblick auf die Ebene der prälexikalischen Repräsentationen beantworten muss, ist die Frage nach den Einheiten, welche diese Ebene umfasst.¹⁴⁵ Potentielle Einheiten der prälexikalischen Repräsentation sind alle linguistischen Kategorien, die für die Beschreibung der phonologischen Strukturen der Sprache von Bedeutung sind.¹⁴⁶ Ob sich tatsächlich Effekte im Verhalten der Sprecher einer Sprache finden, die Hinweise darauf geben, dass eine bestimmte phonologische Kategorie (z.B. Phoneme oder Silben) auf einer prälexikalischen Ebene der Repräsentation rekonstruiert wird, muss experimentell nachgewiesen werden. Einige experimentelle Ansätze, die darauf hinweisen, dass suprasegmentale Kategorien wie *Silbe und Akzent*, wenn nicht als Einheiten einer prälexikalischen Repräsentation, so zumindest als Einheiten der Segmentation des Signals (s.u. 4.3) gelten müssen, werden im folgenden zusammengefasst. Im Anschluss daran wird in diesem Kapitel gezeigt, dass auch das Merkmal *Sonorität* in eine Konzeption einer Ebene der prälexikalischen Repräsentation in der Sprachwahrnehmung integriert werden muss.

Diese Studie geht davon aus, dass *alle* phonologischen Kategorien und Merkmale, die prinzipiell aus dem akustischen Signal rekonstruiert werden können, für die Sprachverarbeitung genutzt, d.h. entweder auf einer prälexikalischen Ebene repräsentiert werden, oder die Rekonstruktion der linguistischen Struktur einer Äußerung auf andere Weise beeinflussen. Diese Annahme erscheint deshalb plausibel, weil die Wahrnehmung von Sprache einerseits ein erstaunlich robuster Prozess ist, obwohl lautsprachliche Kommunikation meist unter widrigen Bedingungen der Signalproduktion und Signalübertragung¹⁴⁷ erfolgt und weil andererseits das akustische Signal ein höchst komplexes Muster bildet und eine sehr große Menge von Information über Strukturen der

Ebene der prälexikalischen Repräsentation der phonologischen Strukturen einer gehörten Äußerung ist dabei nur eine von verschiedenen Möglichkeiten.

¹⁴⁵ Diese Einheiten kann man auch *units of perception* oder *perceptual units* nennen (vgl. SENDLMEIER 1995). Ebenso kann man die Redeweise von der 'perzeptuellen Realität', der 'mentalenen Realität', oder der 'psychologischen Realität' linguistischer Merkmale und Kategorien auf eine Ebene der prälexikalischen Repräsentation beziehen (für eine Kritik dieser Begriffe vgl. MORAIS, KOLINSKY & NAKAMURA 1996).

¹⁴⁶ Vgl. auch MCQUEEN & CUTLER 1997:570: „It has not been possible, however, to determine which linguistic unit, if any, is constructed at this level of processing. There is support for several different processing units. [...] the form of these representations remains to be determined.“

¹⁴⁷ Faktoren, welche die Wahrnehmung von Sprache wesentlich erschweren, ist z.B. die Tatsache, dass wir viele verschiedene Sprecher mit unterschiedlichen Stimmen verstehen können müssen, meist reden mehrere Sprecher gleichzeitig, dazu kommen ganz verschiedene Störgeräusche aus der Umwelt, Sprecher reden häufig schnell bzw. mit ganz unterschiedlicher Geschwindigkeit, oft mit geringem Artikulationsaufwand, usw.

übermittelten Äußerungen auf höheren Ebenen beinhaltet, z.B. über die morphologische, syntaktische, lexikalische und pragmatische Struktur der Äußerung.

4.2.3 Verarbeitungsstrategien

Kann eine bestimmte sprachliche Äußerung nur auf eine einzige Weise verarbeitet werden, oder ist damit zu rechnen, dass verschiedene Sprecher unter verschiedenen Bedingungen unterschiedliche Verarbeitungsstrategien anwenden? Diese Frage soll im folgenden diskutiert werden. Diese Ausführungen bilden den Hintergrund für die Ausführungen in 4.3.4 zu universalen und sprachspezifischen Verarbeitungsstrategien und werden auch in der Diskussion der experimentellen Ergebnisse der vorliegenden Studie (s.u. 6.3.3) wiederaufgegriffen.

Verschiedene Sprachen unterscheiden sich in dem Inventar von Lauten, das sie benutzen, um Bedeutungen zu unterscheiden, anders ausgedrückt, sie unterscheiden sich in ihrem Phoneminventar. Während z.B. das Deutsche und Englische zwischen einem *stimmlosen* bilabialen Verschlusslaut /p/ und einem *stimmhaften* bilabialen Verschlusslaut /b/ unterscheiden, finden sich im Hindi drei bilabiale Verschlusslaute (LADEFOGED 1993³:145): ein *stimmlos aspirierter* bilabialer Verschlusslaut /p^h/ (wie in dt. *Pappe* – [p^h]appe), ein *stimmlos unaspirierter* bilabialer Verschlusslaut /p/ (wie in dt. *Baum* – [p]aum) und ein *stimmhafter* bilabialer Verschlusslaut /b/ (wie in dt. *Ra[b]e*). Dies bedeutet, dass diese Sprachen jeweils verschiedene Bereiche der universalen Lautskala für das Merkmal [stimmhaft],¹⁴⁸ die sich durch die möglichen Konstellationen der Artikulatoren im Artikulationstrakt und durch die akustischen Entsprechungen dieser Konstellationen ergibt, zu einem Phonem zusammenfassen. Laute, die einem bestimmten Punkt auf dieser Lautskala entsprechen, werden stets als einem dieser Phoneme zugehörig wahrgenommen. Sprecher des Deutschen nehmen bilabiale Verschlusslaute entweder als /p/ oder /b/ wahr, Sprecher des Hindi nehmen bilabiale Verschlusslaute entweder als /p^h/, als /p/, oder als /b/ wahr.¹⁴⁹ Hieraus ergibt sich, dass der Prozess der Rekonstruktion einer prälexikalischen Repräsentation auf der Phonemebene in gewisser Weise sprachspezifisch sein muss: ein und dasselbe akustische Si-

¹⁴⁸ Schränkt man den Blickwinkel auf ein Merkmal wie *Stimmeinsatzzeit* ein, kann man von einer (universalen) Lautskala sprechen, betrachtet man hingegen Konfigurationen aus mehr als einem Merkmal, z.B. Konfigurationen aus Zungenhöhe ([± hoch] und [± tief]), Zungenlage ([± hinten]) und Lippenrundung ([± rund]), durch die sich Vokale kennzeichnen lassen, so ist eher von einem (universalen) Phonemraum zu sprechen.

¹⁴⁹ Dieses Phänomen wird *kategorische Wahrnehmung* genannt, für einen Überblick vgl. KUHLE 1993, HANDEL 1989:266ff.

gnal, z.B. eine CV-Silbe (d.h. eine Konsonant-Vokal-Silbe wie /ba/) mit einem bilabialen Plosiv und einer Stimmeinsatzzeit von 0 ms,¹⁵⁰ wird von Sprechern des Deutschen und des Hindi als Realisierung verschiedener Phonemfolgen wahrgenommen: von ersteren als Phonemfolge /ba/, von letzteren als Phonemfolge /pa/. Aus demselben Signal wird von Sprechern verschiedener Sprachen also eine unterschiedliche symbolische Repräsentation erzeugt.

Nimmt man nun an, dass bei der Sprachwahrnehmung aus dem einkommenden akustischen Signal nicht nur eine Phonemrepräsentation extrahiert wird, sondern auch eine Repräsentation der suprasegmentalen Kategorien und Merkmale, so ergeben sich weit mehr Möglichkeiten der sprachspezifischen Verarbeitung. Unter der Annahme, dass die Erzeugung von Repräsentationen zudem eine zeitliche Dimension hat, d.h. sich in einer zeitlichen Ordnung vollzieht, erhält man eine Vielzahl von Möglichkeiten, aus einem gegebenen akustischen Signal vor dem Hintergrund unterschiedlich strukturierter phonologischer Systeme eine dynamische Repräsentation linguistischer Kategorien und Merkmale zu rekonstruieren.

Betrachtet man die phonologischen Strukturen in den einzelnen Sprachen, so erscheint es auch durchaus sinnvoll, dass sich Sprecher verschiedener Sprachen in der Art und Weise unterscheiden, wie sie aus einem akustischen Signal eine Repräsentation linguistischer Merkmale und Kategorien extrahieren und zwar insbesondere im Hinblick auf suprasegmentale Kategorien wie Silbe und Akzent. So zeichnet sich z.B. das Deutsche dadurch aus, dass Silbengrenzen häufig mit Morphemgrenzen zusammenfallen: das Verb *ver+ankern* wird nicht, nach dem Anlaut-Maximierungs-Prinzip (s.o. 2.3, Anm.), wie in (1a) silbifiziert, sondern wie in (1b), so dass die Silbengrenze der Morphemgrenze zwischen Präfix und Lexem entspricht.¹⁵¹

(1a)	*.ve.r+an.kern.	*[ve.raŋ.kən.]
(1b)	.ver.+an.kern.	[veɐ̯.ʃaŋ.kən.]

Deshalb kann im Deutschen aus der Silbenstruktur einer Äußerung Information über ihre morphologische Struktur abgeleitet werden. Für das Französische hingegen gilt dies nicht, dort wird auch über Morphem- und Wortgrenzen¹⁵² hinweg silbifiziert, z.B. in den Sequenzen (2a) und (2b).

¹⁵⁰ D.h. die Stimmlippenschwingung setzt zeitgleich mit (d.h. mit 0 ms Differenz zu) der Verschlussöffnung ein: in dem Moment, in welchem der mit den Lippen gebildete Verschluss geöffnet wird, beginnen die Stimmlippen zu vibrieren.

¹⁵¹ Vgl. WIESE 1996:65ff. Andere Affixe hingegen erlauben Resilbifizierung über Morphemgrenzen, z.B. in .Ach.t+ung. oder .Leh.r+er.

¹⁵² Jedoch nicht über syntaktische Phrasengrenzen hinweg. Für einen Überblick zur *Liaison* als Resilbifizierungsphänomen im Französischen vgl. DELATTRE 1966.

- (2a) le.[zã].fants
 (2b) peti.[twa].seau

Auf der anderen Seite ist in Sprachen mit festem Wortakzent, wie dem Finnischen, wo die erste Silbe jedes Wortes einen Akzent trägt, oder dem Türkischen, wo meist die letzte Silbe eines Wortes betont ist, der Akzent geeignet, als Indikator für Wortgrenzen in einer Äußerung verarbeitet zu werden.¹⁵³ Für die Sprecher solcher Sprachen stellt daher die Extraktion der metrischen Struktur der Äußerung aus dem akustischen Signal eine mögliche Lösung für das Segmentationsproblem dar, im Gegensatz zu Sprachen mit freiem Wortakzent wie dem Deutschen oder Englischen (aber s.u. 4.3.3). Angesichts der Unterschiede, die sich in den phonologischen Strukturen verschiedener Sprachen finden, ist die Hypothese gerechtfertigt, dass sich die Sprecher dieser Sprachen in den Verarbeitungsstrategien unterscheiden, die sie anwenden, um aus dem akustischen Signal eine symbolische Repräsentation zu rekonstruieren. Weiter unten wird ausgeführt (s.u. 4.3.2), dass sich Unterschiede in den Verarbeitungsstrategien nicht nur beim Vergleich von Sprechern verschiedener Sprachen finden, sondern dass auch andere Faktoren wie die absolute Verarbeitungszeit oder die phonologische Struktur der Äußerung einen Einfluss darauf haben können, ob (auf der Verhaltensebene) die Reaktionszeit in Abhängigkeit von der Silbenstruktur der zu verarbeitenden Äußerung variiert (vgl. SEBASTIÁN-GALLÉS, DUPOUX, SEGUÍ & MEHLER 1992). Im letzten Kapitel dieser Studie (s.u. 6.3.3) wird schließlich gezeigt, dass auch das Geschlecht der Sprecher die Verwendung verschiedener Verarbeitungsstrategien in einer experimentellen Aufgabe bedingen kann.

Was genau ist unter einer Verarbeitungsstrategie zu verstehen? Unterschiedliche Verarbeitungsstrategien entsprechen Unterschieden in der Rekonstruktion einer symbolischen Repräsentation aus einem Signal: Unterschieden in der zeitlichen Dynamik der Repräsentationen und Unterschieden in der kausalen Verknüpfung von Teilrepräsentationen (vgl. hierzu SCHNELLE 1991:120). Dies kann auf der Verhaltensebene z.B. bedeuten, dass in der einen Sprache die Reaktionszeit für eine Testaufgabe mit der Silbenstruktur der zu verarbeitenden Äußerung variiert, in der anderen Sprache aber nicht (vgl. CUTLER, MEHLER, NORRIS & SEGUÍ 1986). Auf der Ebene der Beschreibung dynamischer neuronaler Repräsentation könnte dies bedeuteten, dass die der Silbenstruktur entsprechenden Teilrepräsentation eine andere Dynamik hat oder auf andere Weise

¹⁵³ LAVER 1995:519: „The relatively strong tendency (about 57%) for languages to prefer the trailing edge of words (final or penultimate syllable) for the location of fixed lexical stress is to some extent explainable by suggesting that such placement gives stress a demarcative function, indicating where the word-boundaries are without making the listener expend too much cognitive effort in calculating them.“

mit weiteren Verarbeitungsprozessen kausal verknüpft ist. Variiert auf der Verhaltensebene die Reaktionszeit nicht mit der Silbenstruktur der zu verarbeitenden Äußerungen, so könnte dies auf der neuronalen Ebene z.B. bedeuten, dass die Aktivitätsmuster, die der Silbenstrukturerepräsentation entsprechen, einen sehr geringen Beitrag zur weiteren Dynamik der Aktivitätsmuster liefern.

Die vorangegangenen Ausführungen haben gezeigt, dass aus der Variation der phonologischen Strukturen, die sich in den verschiedenen Sprachen finden, geschlossen werden muss, dass sprachliche Äußerungen auf verschiedene Weise verarbeitet werden können. Eine Variabilität der Verarbeitungsstrategien ist insbesondere für solche Verarbeitungsprozesse zu erwarten, die sich auf die suprasegmentalen Eigenschaften der Sprache beziehen, weil hier die größte Variation zwischen den Sprachen zu finden ist. Diese suprasegmentalen Verarbeitungsstrategien werden möglicherweise erst beim Erwerb einer Sprache mit bestimmten phonologischen Eigenschaften entwickelt (vgl. CUTLER, MEHLER, NORRIS & SEGUÍ 1992, s.u. 4.3.4), während die segmentalen Verarbeitungsstrategien, d.h. die Verarbeitung der distinktiven Eigenschaften von Phonemen, die einkommende akustische Information zwar auf unterschiedlich strukturierte Phonemrepräsentationen abbildet, dies aber möglicherweise unter Verwendung universaler Strategien geschieht. Das Konzept der Verarbeitungsstrategie ist daher von zentraler Bedeutung für die vorliegende Untersuchung, in welcher die funktionale Lateralisierung der Verarbeitung der (suprasegmentalen) Sonoritätskontur untersucht wird und wird weiter unten wieder aufgegriffen (s.u. 4.3.4, 5.4, 6.3.3).

4.2.4 Zwei Annahmen zur Sprachverarbeitung

Abschließend sollen noch einmal die beiden Grundannahmen zur Konzeption einer Ebene der prälexikalischen Repräsentation von zu verarbeitenden sprachlichen Äußerungen formuliert werden, auf denen die folgenden Ausführungen basieren:

- Auf einer prälexikalischen Ebene sind *alle* linguistischen Kategorien und Merkmale repräsentiert, die aus dem akustischen Signal rekonstruierbar sind und die für die Beschreibung der phonologischen Strukturen der Sprache nötig sind.
- Die zeitliche Dynamik der Repräsentationen von linguistischen Kategorien und Merkmalen (Phonemstruktur, Silbenstruktur, Akzentstruktur usw.) auf einer prälexikalischen Ebene ist variabel und hängt (i) von den aktuellen Bedingungen der Sprachverarbeitung (z.B. Stimulustyp, absolute Verarbeitungszeit, Hintergrundgeräusche), (ii) von der phonologischen Struktur einer Sprache ab (z.B. dem

Sprachrhythmus) und (iii) von noch genauer zu spezifizierenden weiteren Faktoren, möglicherweise auch vom Geschlecht¹⁵⁴ der Person, ab.

4.3 Rhythmische Einheiten der Sprache als Segmentationseinheiten

4.3.1 Einleitung

Um in einer sprachlichen Äußerung Wörter zu erkennen, muss der Hörer das Sprachsignal zunächst in Einheiten *segmentieren*, die als potentielle Wörter in Frage kommen; erst dann kann er überprüfen, ob diese Kandidaten tatsächlich als Wörter in seinem Lexikon gespeichert sind.¹⁵⁵ Dieses grundlegende Problem der Sprachwahrnehmung formulieren CUTLER, NORRIS & MCQUEEN (1996) folgendermaßen :

Continuous speech input is the norm; only relatively rarely do we hear isolated spoken word, and even less often do we hear multi-word sequences in which the individual words are separated one from another by, for instance, pauses. [...] Nevertheless, the recognition of speech must involve *segmentation* of utterances into their component words, since only the component word, not the entire utterance, will be represented in the listener's lexical memory. Despite the rarity of clearly marked word boundaries, however, listeners can rapidly and reliably recognise individual words in utterances in their native language. (CUTLER, NORRIS & MCQUEEN 1996:227).

Dieses Problem der Segmentation des Signals stellt sich insbesondere dem Kind, das eine Sprache erst erwirbt: wie kann das Kind erkennen, welche Einheiten eines Sprachsignals als Wörter im zu erwerbenden Lexikon zu speichern sind, die es dann in neuen Signalen wiedererkennen kann? Auf Grund welcher Eigenschaften also kann der Hörer das kontinuierliche Signal in Einheiten segmentieren, die er in einem zweiten Verarbeitungsschritt als Wörter erkennen kann? Ein erster Ansatz besteht darin, das kontinuierliche Signal auf der Basis seiner periodisch sich wiederholenden Eigenschaften zu segmentieren: auf der Basis des Rhythmus der Sprache.¹⁵⁶ Dies ist der Ausgangspunkt für eine Vielzahl von Studien, die in den vergangenen zwei Jahrzehnten zur Funktion des Sprachrhythmus und der damit zusammenhängenden Kategorien Silbe und Akzent für die Sprachverarbeitung durchgeführt worden sind (für einen Überblick vgl. CUTLER & OTAKE (Hrsg.) 1996).

Was ist der Rhythmus der Sprache? LAVER (1995) fasst den Begriff *Rhythmus* folgendermaßen:

¹⁵⁴ Zu Geschlechtsunterschieden in verschiedenen kognitiven Fähigkeiten vgl. KIMURA 1996, HALPERN 1992, KOLB & WHISHAW 1993:190ff.

¹⁵⁵ Für einen Überblick zu Modellen der Worterkennung, die ohne explizite Segmentationsprozedur auskommen, vgl. CUTLER, NORRIS & MCQUEEN 1996:229ff, vgl. auch CUTLER 1995:100ff.

¹⁵⁶ Zu möglichen Funktionen des Sprachrhythmus für den Erwerb von Lexikon und phonologischem System einer Sprache vgl. CUTLER & MEHLER 1993.

Utterances are generally perceived as being spoken with a certain rhythm. The perception of rhythm in speech is predicated on the listener's recognition of a quasi-periodic recurrence in time of a given type of speech unit, such as syllables carrying peaks of prominence (achieved through either or both syllable stress or syllable weight), or syllables themselves. The study of rhythm is a highly controversial area of research, however. (1995:512)

Mit ABERCROMBIE (1967) kann man zwei Arten von Sprachrhythmus unterscheiden: silbenzählenden Rhythmus¹⁵⁷ und akzentzählenden¹⁵⁸ Rhythmus. In Sprachen, die silbenzählenden Rhythmus haben, sind Silben (idealerweise)¹⁵⁹ gleich lang (isochron), so dass die *Silbe* als die sich periodisch wiederholende Einheit gefasst werden kann. Zu den silbenzählenden Sprachen gehört z.B. das Französische, das sich durch ein kleines Inventar an Silbenstrukturen (CV – *bon*, CVC – *choque*, CCVC – *truc*, CCVCC – *Proust*) und festen (wortfinalen bzw. äußerungsfinalen) Akzent auszeichnet. In akzentzählenden Sprachen hingegen sind die Zeitintervalle zwischen zwei betonten Silben (idealerweise)¹⁶⁰ gleich lang, d.h. Folgen von einer *betonten* und allen nachfolgenden *unbetonten* Silben bis zur nächsten betonten Silbe (d.h. Takte oder Füße) bilden die sich periodisch wiederholenden Einheiten. Zu den akzentzählenden Sprachen gehören z.B. das Englische und das Deutsche, die sich durch ein großes Inventar an Silbenstrukturen (z.B. *sprichst*: CCCVCCC), freie Wortbetonung (z.B. /.'Bal.ken./, /.'Bal.kon./) und Reduktion unbetonter Silben (z.B. /,un.b^ə.t^ə./) auszeichnen. Daneben kann man noch eine dritte Art des Rhythmus annehmen, den morenzählenden¹⁶¹ Rhythmus; zu den wenigen Sprachen mit morenzählendem Rhythmus gehören u.a. Japanisch (HARAGUCHI 1996) und Tamil (LAVÉ 1995:529).

Darauf, dass der Sprachrhythmus tatsächlich eine wichtige Funktion für die Sprachwahrnehmung erfüllt, weist die Beobachtung hin, dass bilinguale Sprecher (des Englischen und Französischen¹⁶²), die gelernt haben, Sätze in der einen Sprache in er-

157 Vgl. ABERCROMBIE 1967:97: „As far as is known, every language in the world is spoken with one kind of rhythm or with the other. In the one kind, known as syllable-timed rhythm, the periodic recurrence of movement is supplied by the syllable-producing process: the chest-pulses, and hence the syllables, recur at equal intervals of time – they are isochronous. French, Telugu, Yoruba [...] are syllable-timed languages.“

158 Vgl. ABERCROMBIE 1967:97: „In the other kind, known as stress-timed rhythm, the periodic recurrence of movement is supplied by the stress-producing process: the stress-pulses, and hence the stressed syllables, are isochronous. English, Russian, Arabic [...] are stress-timed languages.“

159 Zu den Faktoren, welche die Länge einer Silbe beeinflussen, gehören u.a. die Struktur der Silbe [komplexe Silben (z.B. frz. *prince* – CCVC) sind länger als CV-Silben (frz. *pin.ceau* – CV.CV)] und die Satzbetonung, die u.a. auf der Variation der Silbendauer beruht (neben der Variation von Tonhöhe und Lautstärke) (vgl. LAVÉ 1995:528).

160 Auf eine betonte Silbe können eine, zwei oder mehr unbetonte Silben folgen (für deutsche Beispiele vgl. WIESE 1996:57), was die absolute Länge des Intervalls zwischen zwei betonten Silben beeinflusst.

161 Als 'Moren' können die Einheiten des Silbengewichts bezeichnet werden. Eine 'leichte' CV-Silbe (/ba/) umfasst demnach eine Mora, eine 'schwere' CVC-Silbe (/bal/) zwei Moren, s.o. 2.3.3, Anm., s.u. 4.3.3.

162 Das Englische hat betonungszählenden, das Französische silbenzählenden Rhythmus.

heblich gesteigertem Tempo (*time-compressed speech*) zu verarbeiten, diese Fähigkeit nicht auf die andere Sprache übertragen können, wenn die andere Sprache einem anderen Rhythmus-Typ zugehört, während dies möglich ist, wenn beide Sprachen demselben Rhythmus-Typ zugehören (z.B. Katalanisch und Spanisch¹⁶³) (MEHLER, SEBASTIÁN, ALTMANN, DUPOUX, CHRISTOPHE & PALLIER 1993).¹⁶⁴ Im folgenden werden einige psycholinguistische Untersuchungen zum Sprachwahrnehmungsverhalten zusammengefasst, die auf eine Funktion des Sprachrhythmus für die Segmentation des Signals in silbenzählenden und betonungszählenden Sprachen hinweisen.

4.3.2 Silbenzählende Sprachen

In silbenzählenden Sprachen bilden Silben die sich periodisch wiederholenden Einheiten im akustischen Signal. Daher ist es eine dem Sprachrhythmus angemessene Strategie, wenn die Sprecher silbenzählender Sprachen das Signal zunächst in Silben segmentieren,¹⁶⁵ bevor sie versuchen, lexikalische (oder andere) Einheiten im Signal wiederzuerkennen.¹⁶⁶

In einer Reihe von Experimenten zur Verarbeitung des Französischen¹⁶⁷ zeigten MEHLER, SEGUÍ & FRAUENFELDER 1981 (vgl. auch MEHLER 1981, SEGUÍ, DUPOUX &

¹⁶³ Beide Sprachen haben silbenzählenden Rhythmus.

¹⁶⁴ Diese Beobachtung muss nicht zwingend auf diese Weise interpretiert werden. Abgesehen davon, dass sich Französisch und Englisch in ihren rhythmischen Eigenschaften unterscheiden, Spanisch und Katalanisch hingegen nicht, sind sich Spanisch und Katalanisch auch in segmental-phonetischer Hinsicht ähnlicher als Englisch und Französisch. Der Effekt könnte daher auch auf dem Faktor *segmental-phonetische Ähnlichkeit* o.ä. beruhen.

¹⁶⁵ Vgl. auch SIEVERS (1871:198): „dass das Ohr des Hörers die zusammenhängende Rede subjectiv in gewisse Theilstücke zerlegt, [...] und dass diese Theilstücke das sind was man als Silben zu bezeichnen pflegt.“

¹⁶⁶ Vgl. hierzu auch LIEBERMAN (1984:186): „it is difficult to see how the perception of speech can occur without a level of neural processing at the syllabic level. [...] Speech perception hypothetically could involve a listener's first segmenting the input signal into a sequence of CV and VC syllable patterns of burst frequencies, formant patterns, noise spectra, and the like and then comparing these patterns against an internal inventory of possible syllable patterns.“ Was LIEBERMAN vorschlägt, ist neben der Segmentation des Signals in Halbsilben oder Diphone (CVs und VCs) eine holistische Erkennung ganzer Silben durch den Abgleich des einkommenden Signals mit gespeicherten *syllable templates*. Unter dieser Sichtweise würde nicht die Silbenstruktur aus der Phonemstruktur einer Äußerung abgeleitet, wie oft vorgeschlagen, sondern umgekehrt wäre die Phonemstruktur gegenüber der Silbenstruktur sekundär: bei der Sprachwahrnehmung würden nicht einzelne Phoneme identifiziert und vielleicht wäre ein Kode aus kleinsten distinktiven Einheiten erst auf der Ebene des Lexikons verfügbar, spielte aber für die Wahrnehmung keine Rolle.

¹⁶⁷ Hier soll kurz angemerkt werden, dass die Resilbifizierung von Lautfolgen über Wortgrenzen hinweg im Französischen weitverbreitet ist [s.o. 4.2.3, (2a-b)]. Daher könnte man gegen die oben gemachte Annahme einwenden, dass die Segmentation in Silben im Französischen keine zuverlässige Strategie für die Worterkennung ist, da im Französischen Silbengrenzen seltener als z.B. im Deutschen oder Englischen mit Wortgrenzen zusammenfallen.

MEHLER 1990), dass die Probanden CV-Zielfolgen¹⁶⁸ (die Silbe *pa*) schneller in CV.X-Wörtern (*.pa.lace.*) als in CVC.X-Wörtern (*.pal.mier.*) erkannten, während sie CVC-Zielfolgen (die Silbe *pal*) schneller in CVC.X-Wörtern (*.pal.mier.*) als in CV.X-Wörtern (*.pa.lace.*) erkannten, d.h. für die abhängige Variable Reaktionszeit findet sich eine Interaktion der beiden Faktoren Silbenstruktur des Stimulus und Typ der Zielfolge.¹⁶⁹ Diese Beobachtung wurde so interpretiert, dass in den Fällen, in denen zur Erkennung der Zielfolge Information über eine Silbengrenze hinweg integriert werden musste (d.h. um *pal* in *.pa.la.ce.* zu erkennen) bzw. in denen eine Silbe in ihre Bestandteile zerlegt werden musste (d.h. um *pa* in *.pal.mier.* zu erkennen), der Prozessor mehr Zeit für die Erkennung braucht, als wenn er ganze Silben als das Zielfolge wiedererkennt (d.h. *pa* in *.pa.lace.* und *pal* in *.pal.mier.*).¹⁷⁰

Dieser Versuch wurde von CUTLER, MEHLER, NORRIS & SEGUÍ (1986) mit englischsprachigen Probanden wiederholt. Das Englische ist eine betonungszählende Sprache und hat, anders als das Französische und andere romanische Sprachen, keine klaren Silbengrenzen.¹⁷¹ Beeinflusst also die Tatsache, dass im Englischen viele Wörter unklare

¹⁶⁸ C steht für einen Konsonanten, V für einen Vokal.

¹⁶⁹ Diese Interaktion von Silbenstruktur des Stimulus und Silbenstruktur der Zielfolge wird im folgenden 'Silbeneffekt' genannt.

¹⁷⁰ Für eine Kritik an der Verwendung von *target-monitoring*-Experimenten zur Untersuchung der 'Einheiten der Wahrnehmung' vgl. MORAIS, KOLINSKY & NAKAMURA 1996. MORAIS ET AL. nehmen an, dass diese Technik nicht automatische Prozesse der Sprachwahrnehmung untersucht, sondern „metalinguistic judgements“ oder „postperceptual introspection“ abfragt. Um Sprachwahrnehmung möglichst ohne Einflüsse bewusster Verarbeitung zu untersuchen, verwenden MORAIS ET AL. ein modifiziertes *dichotic-listening*-Paradigma (zur Technik des dichotischen Hörens s.u. 6.2.3). MORAIS ET AL. präsentieren den Probanden zwei verschiedene Nonwörter gleichzeitig, je eines auf jedes Ohr. Dabei sind die von den Probanden wahrgenommenen Wörter zum Teil Kombinationen aus Teilen beider Stimuluswörter. Der Effekt, dass Bestandteile des einen Wortes in das andere Wort 'wandern', wird als *Migrationseffekt* bezeichnet. Migrationseffekte werden von MORAIS ET AL. folgendermaßen interpretiert: „At least for Portuguese, consonants do have psychological reality at a low level of speech processing, namely the perceptual level. We thus take property migration effects obtained in our dichotic listening situation as reflecting perceptual processing.“ (MORAIS ET AL. 1996:209). Im Portugiesischen finden sich Migrationseffekte häufiger für Phoneme als für Silben oder segmentale Merkmale, im Japanischen häufiger für Moren und Silben als für Phoneme und segmentale Merkmale.

¹⁷¹ Man kann sagen, dass ein Wort keine klaren Silbengrenzen hat, wenn in diesem Wort ein Konsonant dem Auslaut der einen Silbe und dem Anlaut der folgenden Silbe zugeordnet ist, z.B. das /t/ in dt. *Mitte*. Phonologisch kann man diese Ambisilbizität des Konsonanten folgendermaßen begründen (nach GIEGERICH 1992, vgl. auch BLEVINS 1995:230ff):

Das *Prinzip des maximalen Anfangsrandes* (bei GIEGERICH *Syllable Boundary Rule*, 1992:170, *Silbenkontaktgesetz* nach Vennemann 1982:285, s.o. 2.3, Anm.) verlangt, dass dieser Konsonant als Anlaut der zweiten Silbe silbifiziert wird.

Gleichzeitig ist es so, dass betonte Silben im Englischen (und im Deutschen) *schwer* (in anderer Terminologie: *stark*) sein müssen, d.h. aus einem verzweigten Reim bestehen müssen (GIEGERICH 1992:182,146, s.o. 2.3.3). Ist also die erste Silbe eines Wortes betont und besteht aus einem kurzen bzw. ungespannten Vokal wie /i/ in *Mitte*, /e/ in *rette* oder /o/ in *Rotte*, so muss sie über einen Aus-

re Silbengrenzen haben, die Verarbeitungsstrategie, welche die Probanden anwenden? Zeigen auch englische Probanden den gleichen Silbeneffekt wie französische Probanden, oder zeigen sie ihn vielleicht nur bei der Verarbeitung von englischen Wörtern mit klaren Silbengrenzen?

Als Stimuli wurden CVC.C-Wörter (.*bal.co.ny.*) und CV[C]V-Wörter¹⁷² mit ambisilbischen Konsonanten (.*ba[l]an.ce.*) verwendet, als Zielfolgen CV-Silben (*ba*) und CVC-Silben (*bal*). Für englischsprachige Probanden fand sich kein Silbeneffekt, d.h. keine Interaktion von Silbenstruktur des Stimuluswortes und Typ der Zielfolge. Auch als den englischen Probanden das französische Material von MEHLER ET AL. (1981) präsentiert wurde, zeigten sie keinen Silbeneffekt für die Erkennung von CV-Silben und CVC-Silben in französischen CV.C-Wörtern und CVC.C-Wörtern. Als schließlich französischsprachige Probanden mit dem englischen Material getestet wurden, wiesen diese zwar den Silbeneffekt auf, der an französischen Probanden für die Verarbeitung französischer Wörter gezeigt werden konnte, jedoch nur für die Wörter mit eindeutigen Silbengrenzen, nicht für die Testwörter mit ambisilbischem Konsonanten (CV[C]V-Wörter). Hieraus folgerten CUTLER ET AL. (1986), dass Sprecher des Englischen weder bei der Verarbeitung von englischen Wörtern noch von Wörtern einer anderen Sprache eine Segmentationsstrategie anwenden, die Silbengrenzen nutzt, weil es im Englischen viele Wörter mit unklaren Silbengrenzen gibt, während Sprecher des Französischen sowohl bei der Verarbeitung von französischen Wörtern als auch von Wörtern anderer Sprachen eine silbenbasierte Segmentationsstrategie anwenden, soweit dies der phonetisch-phonologischen Struktur der Testwörter entspricht. CUTLER ET AL. (1986) führen also in ihrer Interpretation die unterschiedlichen Ergebnisse für Sprecher des Englischen und des Französischen auf das phonologische System beider Sprachen zurück, einerseits auf den Sprachrhythmus (silbenzählend oder betonungszählend) und andererseits auf die Eindeutigkeit der Silbifizierung (An- oder Abwesenheit von ambisilbischen Konsonanten). Für die Sprecher des Englischen nahmen CUTLER ET AL. (1986) an, dass

lautkonsonanten verfügen, andernfalls ist sie nicht verzweigt, daher nicht schwer und die resultierende Konfiguration verstößt gegen die Akzentregeln des Englischen.

Aus phonologischen Gründen muss man deshalb annehmen, dass in englischen Wörtern mit betontem kurzem bzw. ungespannten Vokal (wie /e/ in 'letter, /i/ in 'pity, /e/ in 'camera) der nachfolgenden Konsonant ambisilbisch sein muss, d.h. sowohl dem Auslaut der ersten als auch dem Anlaut der zweiten Silbe zugeordnet ist.

TREIMAN & DANIS (1988) konnten zeigen, dass Sprecher des Englischen diese phonologisch ambisilbischen Konsonanten auch im psycholinguistischen Experiment nicht eindeutig der einen oder der anderen Silbe zuordnen. Neben Wörtern mit ambisilbischem Konsonanten gibt es im Deutschen und Englischen zahlreiche Wörter mit klaren Silbengrenzen, wie 'can.dy., 'mel.ting., .be 'fore., in denen entweder das Prinzip des maximalen Anfangsrandes keine Anwendung findet, oder in denen der ungespannte Vokal nicht betont ist.

¹⁷² Die Schreibweise [C] indiziert, dass der Konsonant C ambisilbisch ist.

sie anstelle einer silbenbasierten Segmentationsstrategie eine phonembasierte Verarbeitungsstrategie anwenden. Als Hinweis darauf werteten sie die Beobachtung, dass die englischen Probanden CV[C]V-Wörter (*balance*) schneller verarbeiteten als CVC.C-Wörter (*balcony*), d.h. alternierende CV-Folgen schneller als Lautfolgen, die Konsonantencluster beinhalteten.

Für das Niederländische, das ebenso wie das Englische über Wörter mit unklaren Silbengrenzen verfügt,¹⁷³ konnten ZWITSERLOOD, SCHRIEFERS, LAHIRI & VAN DONSELAAR (1993) in einer vergleichbaren Studie dennoch zeigen, dass Sprecher dieser Sprache mit der initialen Silbe kongruente CV- und CVC-Zielfolgen schneller erkennen als inkongruente.¹⁷⁴ Die Abwesenheit eines Silbeneffektes für englischsprachige Probanden kann daher nicht allein durch die Existenz von Wörtern mit unklaren Silbengrenzen im Englischen verursacht sein, wie CUTLER ET AL. (1986) annahmen.

SEBASTIÁN-GALLÉS, DUPOUX, SEGUÍ & MEHLER (1992) demonstrierten anhand von katalanisch- und spanischsprachigen Probanden, dass neben dem Sprachrhythmus (silbenzählend bzw. akzentzählend) und der Eindeutigkeit der Silbifizierung (Existenz ambisilbischer Konsonanten) einer Sprache noch weitere Faktoren einen Einfluss darauf haben, ob Sprecher einen Silbeneffekt zeigen oder nicht. Katalanisch und Spanisch haben wie das Französische eindeutige Silbengrenzen, aber wie das Englische variable Wortbetonung. Beide Sprachen sind wie das Französische silbenzählende Sprachen

Sprecher des *Katalanischen* zeigten in der Untersuchung von SEBASTIÁN-GALLÉS ET AL. (1992) dann einen Silbeneffekt, wenn die erste Silbe des Wortes unbetont war; d.h. im Wort *.vis.'cós*. erkannten die Probanden die kongruente Zielfolge *vis* schneller als die inkongruente Zielfolge *vi*. Für die initialbetonten Wörter hingegen wiesen die Probanden keinen Silbeneffekt auf, d.h. im Wort *'vis.ta*. wurde die mit der ersten Silbe kongruente Zielfolge *vis* ebenso schnell erkannt wie die inkongruente Zielfolge *vi*. Für das *Spanische* fand sich kein Silbeneffekt, wenn die Probanden spontan reagierten, d.h. die Zielfolgen *pa* und *pas* wurden in den Wörtern *'pa.so*. und *.pa.'sión*., *'pas.ta*. und *.pas.'tel*. jeweils gleich schnell erkannt. Wurde die Reaktionszeit der Pro-

¹⁷³ Außerdem findet sich im Niederländischen wie im Englischen Vokalreduktion und freier Akzent. Das Niederländische kann daher (vgl. LAVER 1995:528) als akzentzählende Sprache gelten.

¹⁷⁴ ZWITSERLOOD ET AL. (1993:270) interpretieren diese Ergebnisse jedoch nicht als Evidenz für eine silbenbasierte Segmentationsstrategie, sondern auf die folgende Weise: „prelexical or lexical levels of representation might not be appropriate loci for syllabic information. From where then, might information about syllabic structure derive? One hypothesis about how information about syllables might be implemented is in terms of a parsing routine. In analogy to the syntactic parsing of sentences, the speech input could be mapped onto phonological representations of lexical form. [...] Acoustic cues can aid this process, and knowledge about possible syllables of the language could be implemented in terms of constraints on computation. In such an approach, syllables are not represented units but rather derived or computed entities.“ Zum Modell des Lexikons, das dieser Interpretation der experimentellen Daten zugrundeliegt, vgl. LAHIRI & MARSLÉN-WILSON 1991.

banden hingegen verzögert (indem sie zusätzlich die Bedeutung verarbeiten mussten), fand sich ein Silbeneffekt sowohl für betonte wie unbetonte iniale Silben. Diese Ergebnisse weisen darauf hin, dass erstens die Betonung eines Wortes (im Katalanischen) und zweitens die absolute Reaktionszeit (im Spanischen) einen Einfluss darauf hat, ob Silbengrenzen für die Segmentierung des Signals zur Erkennung von Zielfolgen genutzt werden oder nicht, wobei beide Effekte nur in einer Sprache auftreten, d.h. sprachspezifisch sind.

Diese Ergebnisse wurden von SEBASTIAN-GALLÉS ET AL. (1992) u.a. mit dem Konzept der *akustischen Transparenz* erklärt: Spanisch hat nur fünf Vokale und keine Vokalreduktion; Katalanisch hat acht Vokale und verbreitete Vokalreduktion; Französisch schließlich hat vierzehn Vokale und weitverbreitete Schwa-Reduktion und verfügt außerdem über perzeptuell schwierigere Nasalvokale. Je mehr Vokale eine Sprache hat (die bei der Sprachverarbeitung voneinander unterschieden werden müssen) und je verbreiteter in einer Sprache die Reduktion unbetonter Vokale ist, desto weniger akustische Transparenz besitzt die Sprache und desto schwieriger ist die Verarbeitung sprachlicher Äußerungen allein auf Grund einer phonemischen Verarbeitungsstrategie.

Das Spanische hat unter dieser Hypothese maximale akustische Transparenz und es ist deshalb kein Silbeneffekt zu erwarten, weil eine phonemische Repräsentation ohne große Schwierigkeit aus dem Signal zu rekonstruieren ist. Das Französische ist hingegen maximal intransparent, daher ist die Verarbeitung des einkommenden Signals allein auf der Basis einer phonembasierten Verarbeitungsstrategie sehr schwierig und es ist ein Silbeneffekt zu erwarten.¹⁷⁵

Auf der anderen Seite ist es so, dass unbetonte Silben akustisch weniger transparent sind als betonte, weil sie kürzer und leiser sind als betonte Silben,¹⁷⁶ so dass im Einklang mit den experimentellen Ergebnissen im Katalanischen für die akustisch wenig transparenten unbetonten Silben mit einem Silbeneffekt zu rechnen ist, während sich in betonten, d.h. in akustisch transparenteren Silben, kein Silbeneffekt zeigen sollte, was ebenfalls mit den experimentellen Ergebnissen übereinstimmt. Im Gegensatz zu CUTLER ET AL. (1986), die Unterschiede in den Verarbeitungsstrategien von englisch- und französischsprachigen Probanden auf das phonologische System der Sprachen zurückführen, nehmen SEBASTIÁN-GALLÉS ET AL. (1992) also an, dass die unterschiedli-

¹⁷⁵ Zu bedenken ist jedoch, dass diese Interpretation die Daten zum Englischen nicht integrieren kann: das Englische hat vierzehn Vokale und Vokalreduktion in unbetonten Silben, ist also akustisch wenig transparent, was einen Silbeneffekt erwarten ließe, der sich aber nicht zeigt.

¹⁷⁶ Betonte Silben werden mit mehr artikulatorischem Aufwand hervorgebracht als unbetonte Silben und sind deshalb lauter und länger als unbetonte (LADEFOGED 1993³:113). Darüberhinaus unterscheiden sich betonte Vokale von unbetonten in ihrer akustischen Qualität, weil sie in artikulatorischer Hinsicht deutlicher voneinander differenziert werden als als unbetonte Vokale (LAVER 1995:513).

chen experimentellen Ergebnisse für Sprechergruppen mit unterschiedlichen Muttersprachen und für unterschiedliches Stimulusmaterial eher akustisch-phonetisch zu erklären sind.

4.3.3 Akzentzählende Sprachen

Für silbenzählende Sprachen wie Französisch, Spanisch und Katalanisch konnte, wie oben beschrieben, ein Silbeneffekt demonstriert werden, d.h. eine schnellere Reaktion auf Zielfolgen, die mit der initialen Silben kongruent sind, als auf Zielfolgen, die länger oder kürzer als die initiale Silbe sind. Für akzentzählende Sprachen wie das Englische hingegen zeigte sich kein Silbeneffekt. Was also ist die Segmentationseinheit in akzentzählenden Sprachen, wenn nicht die Silbe? In *silbenzählenden* Sprachen ist die *Silbe* die sich periodisch wiederholende Einheit, in *akzentzählenden* Sprachen ist das Intervall zwischen akzenttragenden Silben die sich periodisch wiederholende Einheit. Die akzenttragenden Silben sind im Englischen stets starke bzw. schwere Silben,¹⁷⁷ d.h. die sich periodisch wiederholenden Einheiten beginnen immer mit starken bzw. schweren Silben. Daher ist es eine dem Sprachrhythmus einer akzentzählenden Sprache angemessene Strategie, wenn die Sprecher das Signal *vor starken Silben* segmentieren, bevor sie versuchen, lexikalische (oder andere) Einheiten im Signal wiederzuerkennen:

An alternative form of the classification hypothesis, applicable to a stress language like English, holds that speech input is classified into feet. [...] Under such a classification system, stored representations would be in foot form, the recognizer would construct a prelexical representation of the signal as a sequence of specific feet, and lexical access could be attempted starting at every foot, that is, at the beginning of every strong syllable. (CUTLER & NORRIS 1992:114)

Im Englischen kann man das Silbengewicht auf die Vokalqualität beziehen, denn in betonten Silben, die immer stark sind, stehen Vollvokale und in unbetonten Silben, die immer schwach sind, stehen reduzierte Vokale:

Strong syllables contain full vowels; the words *eye*, *pill*, *crypt*, and *scrounge* are all strong monosyllables. Weak syllables contain 'reduced' vowels. Usually this is the vowel schwa, as in the second syllable of *scrounges*; but it may also be a very short form of another vowel, as in the second syllable of *pillow* or *cryptic*. (CUTLER & NORRIS 1992:114)

Dies bedeutet, dass für das Englische aus der Qualität des Vokals einer Silbe inferiert werden kann, ob diese Silbe stark oder schwach bzw. betont oder unbetont ist und ob sie also den Beginn eines Taktes oder Fußes bildet oder nicht (GIEGERICH 1992:66ff). Wann immer der Sprecher einen Diphthong oder Vollvokal verarbeitet, kann er daraus

¹⁷⁷ Die Konzepte *stark* und *schwach* bzw. *schwer* und *leicht* beziehen sich auf das Gewicht der Silbe, s.o. 2.2.3.

schließen, dass diese Silbe den Beginn eines Taktes oder Fusses als der rhythmischen Einheit des Englischen bildet und das Signal vor dieser Silbe segmentieren.

Eine akzentbasierte Segmentationsstrategie ist für die Verarbeitung englischer Äußerungen auch aus einem anderen Grunde erfolgversprechend: CUTLER & CARTER 1987¹⁷⁸ zeigten, dass im Vokabular des Englischen dreimal mehr Wörter mit starken Silben als mit schwachen Silben beginnen und dass die mit starken Silben beginnenden Wörter zudem im Mittel doppelt so häufig sind wie Wörter, die mit schwachen Silben beginnen. Wenn Sprecher des Englischen also starke Silben als potentielle Wortanfänge verarbeiten, so erfassen sie einen Großteil der tatsächlichen Wortanfänge und klassifizieren nur einen geringen Teil der nicht wortinitialen Silben falsch.

In einer Reihe von *word-spotting*-Experimenten¹⁷⁹ mit englischen zweisilbigen Nichtwörtern (z.B. *'min.tayf.*, *'min.tef.*), deren erster Teil einem Wort (*mint*) und deren zweiter Teil einem Nichtwort oder Nicht-Morphem (*.tayf.*, *.tef.*) entsprach, konnten CUTLER & NORRIS 1988¹⁸⁰ demonstrieren, dass das Wort *mint* in der Folge *mintef* schneller und mit weniger Fehlern erkannt wird als in der Folge *mintayf*, d.h. das Wort wird schneller und mit weniger Fehlern erkannt, wenn die folgende Silbe schwach ist (CVC: *.tef.*) als wenn die folgende Silbe stark ist (CVVC: *.tayf.*). Diesen Effekt interpretierten CUTLER & NORRIS (1988) folgendermaßen:

- Vor starken Silben wie *.tayf.* in *min.tayf.*, welche potentielle Wortanfänge darstellen, wird das Signal segmentiert, vor schwachen Silben wie *.tef.* in *min.tef.* wird hingegen nicht segmentiert;
- Die Erkennung des Wortes *mint* im derart segmentierten *.min.|.tayf.* erfordert die Integration von Phoneminformation über den Segmentationspunkt hinweg und
- dies führt zu einer Verzögerung der Erkennung und zu einer schlechteren Gesamtleistung für die Erkennung von *mint* in *.min.|.tayf.* im Vergleich zu *.min.tef.*, wo vor der schwachen Silbe *.tef.* keine Segmentation erfolgt ist.

Die Annahme, dass starke Silben als potentielle Wortanfänge behandelt werden und daher das Signal *vor starken Silben* segmentiert wird und ein Worterkennungsversuch gestartet wird, fassen CUTLER & BUTTERFIELD (1992) als *Hypothese der Rhythmischen Segmentation* zusammen:

The rhythmic segmentation hypothesis proposes that listeners processing spoken English operate on the assumption that strong syllables are highly likely to be the initial syllables of lexical words,

¹⁷⁸ Zit. nach CUTLER & NORRIS 1988:114, vgl. auch CUTLER & BUTTERFIELD 1992:219.

¹⁷⁹ Die Probanden hörten kontinuierliche Folgen von zweisilbigen Nicht-Wörtern und mussten darin eingebettete, aber nicht präspezifizierte englische Wörter erkennen. Als abhängige Variablen können sowohl Reaktionszeit als auch Fehlerrate (Zahl der nicht erkannten Wörter) ausgewertet werden.

¹⁸⁰ Für ähnliche Ergebnisse gl. auch MCQUEEN, NORRIS & CUTLER 1994 und NORRIS, MCQUEEN & CUTLER 1995.

whereas weak syllables are most probably not word-initial or, if word-initial, are more likely to be grammatical words. (CUTLER & BUTTERFIELD 1992:232)

Die Daten zur Rolle von Silbengrenzen für die Erkennung von Wörtern (CUTLER & NORRIS 1988) und Konsonant-Vokal-Folgen (CUTLER, MEHLER, NORRIS & SEGUÍ 1986) im Englischen lassen sich wie folgt zusammenfassen: zwar zeigen Sprecher des Englischen keine Sensitivität für Silbengrenzen, wenn sie Konsonant-Vokal-Folgen in Wörtern erkennen sollen (d.h. sie zeigen nicht den Silbeneffekt, der für die Sprecher romanischer Sprachen demonstriert werden konnte), aber sie zeigen eine differentielle Sensitivität für Silbengrenzen in Abhängigkeit vom Gewicht der nachfolgenden Silbe, wenn sie Wörter in kontinuierlichen Nonwort-Folgen erkennen sollen: erstreckt sich das zu erkennende Wort über eine Silbengrenze in der Nonwort-Folge hinweg, so inhibiert diese Silbengrenze die Verarbeitung weniger, wenn die nachfolgende Silbe schwach ist, als wenn die nachfolgende Silbe stark ist. Diese Beobachtung wird so interpretiert, dass der Hörer das Signal vor starken Silben segmentiert (aber nicht vor schwachen Silben), d.h. eine akzent- oder silbengewichtsbasierte Segmentationsstrategie anwendet.

4.3.4 Universale und sprachspezifische Segmentationsstrategien

Sprecher des Französischen verwenden eine silbenbasierte Segmentationsstrategie, Sprecher des Englischen hingegen eine akzentbasierte Segmentationsstrategie. Wie verhalten sich diese beiden Verarbeitungsstrategien zueinander? Sind beide Strategien sprachspezifisch und werden nur dann erworben, wenn die zu erwerbende Sprache bestimmte phonologische Strukturen aufweist, d.h. entweder silbenzählenden oder akzentzählenden Rhythmus, oder ist eine der beiden Strategien allgemeiner, grundlegender, universaler und wird von den Sprechern beider Sprachen beherrscht? Um diese Frage zu beantworten, wurden die beiden oben zitierten Experimentreihen mit bilingualen Sprechern des Englischen und Französischen durchgeführt (CUTLER, MEHLER, NORRIS & SEGUÍ 1992). Die Resultate des ersten Versuchs (Detektion von CV- und CVC-Silben in CV- und CVC-Wörtern) für diese Probandengruppe werden von CUTLER ET AL. (1992) so zusammengefasst:

When a division was made according to 'dominant' language, one of the subgroups [...] performed as English monolinguals do with these materials. The other group, however, [...] performed like French monolinguals when they were listening to French, but like English monolinguals when they were listening to English. [...] Clearly, the French-dominant subjects do use syllabic segmentation [...] But they do not use syllabic segmentation with English words. The English-dominant subjects, however, fail to use syllabic segmentation at all, even though they have, just

like monolingual French speakers, been exposed to French all their lives, and even though its use is highly efficient with French. (CUTLER ET AL. 1992:397)¹⁸¹

Über die Verteilung der silbenbasierten Segmentationsstrategie lässt sich also folgendes sagen: französisch-dominante Probanden wenden diese Strategie auf französisches Material an, aber nicht auf englisches Material. Englisch-dominante Probanden nutzen diese Strategie weder mit französischem noch mit englischem Material. Daraus schließen CUTLER ET AL. (1992), dass die silbenbasierte Segmentation einer zu verarbeitenden Äußerung keine universale, sondern eine sprachspezifische Strategie ist, die nur dann erworben wird, wenn die Muttersprache die entsprechenden phonologischen Eigenschaften besitzt:

we suggest that the explanation for these asymmetries lies in the distinction which we introduced in the introduction, between procedures which are available to all language users and procedures which are available only upon exposure to certain features in the language input. [...] syllabic segmentation is a procedure of the restricted rather than the general kind. (CUTLER ET AL. 1992:397)

Der Versuch zur akzentbasierten Segmentation (*word-spotting* in Folgen von Nonwörtern mit starken und schwachen Silben) konnte, anders als der zur silbenbasierten Segmentation, nur mit englischem Stimulusmaterial durchgeführt werden, da das Französische keinen Unterschied zwischen starken und schwachen Silben macht.¹⁸² Die Ergebnisse dieses Experimentes zeigten, dass englisch-dominante Sprecher auf englisches Material eine akzentbasierte Segmentationsstrategie anwenden, französisch-dominante Sprecher hingegen nicht. Hieraus schließen CUTLER ET AL. (1992), dass auch die akzentbasierte Segmentationsstrategie sprachspezifisch ist, denn wäre sie universal, so würde man erwarten, dass auch französisch-dominante Sprecher sie auf englisches Material anwendeten.

Bei beiden untersuchten Strategien, sowohl bei der silbenbasierten als auch bei der akzentbasierten Segmentation einer Äußerung, handelt es sich um sprachspezifische Verarbeitungsstrategien. Beide Strategien sind nach der Auffassung von CUTLER ET AL. supplementäre Strategien, welche die Verarbeitung von einer bestimmten Art von Sprachmaterial, nämlich solchem mit silbenzählenden oder akzentzählenden Rhythmus, beschleunigen oder stabiler machen, die jedoch nicht unverzichtbar für die Sprachver-

¹⁸¹ CUTLER ET AL. (1986) konnten zeigen, dass französische Sprecher auch mit englischem Material eine silbenbasierte Segmentationsstrategie anwenden (s.o.). Bei dieser Probandengruppe handelte es sich aber nicht um bilinguale Sprecher, sondern um französischsprachige Sprecher ohne Englischkenntnisse, die keine Gelegenheit hatten, eine dem Englischen angemessene Segmentationsstrategie zu entwickeln.

¹⁸² Auch im Französischen gibt es bis zu einem gewissen Grade Silben mit unterschiedlicher struktureller Komplexität, jedoch haben diese Komplexitätsunterschiede keine Funktion für das phonologische System, insbesondere nicht in Bezug auf die Akzentstruktur, da das Französische einen festen Wortakzent hat.

arbeitung sind, da die französisch-dominanten Probanden auch ohne akzentbasierte Segmentationsstrategie englische Äußerungen verarbeiten und die englisch-dominanten Probanden auch ohne silbenbasierte Segmentationsstrategie französische Äußerungen verstehen können:

We speculate that syllabification is only one of a number of possible segmentation routines available to the human language processing device. During language acquisition, speakers adapt their perceptual routines so as to exploit with maximal efficiency the phonological properties of their native language. Effectively, they tend to favor some of the alternative segmentation routines over others, and incorporate them into their characteristic comprehension procedure. [...] Moreover, it appears that speakers of some, possibly all, languages may have more than one routine available to them, and that speech segmentation may involve the application of more than one routine simultaneously. (CUTLER ET AL. 1992:397).

Aus der Annahme, dass silben- und akzentbasierte Strategien sprachspezifisch und supplementär sind, ergibt sich, dass es weitere Verarbeitungsstrategien gibt, von denen einige ebenfalls sprachspezifisch und supplementär sein sollten; andere aber sollten universal sein und von allen Sprechern angewendet werden. Eine solche Hypothese wirft, wie CUTLER ET AL. meinen, weitere Fragen auf, die eine psycholinguistische Theorie der menschlichen Sprachfähigkeit beantworten muss:

On this model, psycholinguistic theory would be called upon to enumerate the full range of possible segmentation routines and specify the conditions under which particular strategies are preferred to others; the general processing model would remain language-universal, with the language-specific variations being predictable from the phonological structure of each language. (CUTLER ET AL. 1992:397)

Das Ziel einer psycholinguistischen Theorie der Sprache besteht nach CUTLER ET AL. also darin, anzugeben, welche Verarbeitungsstrategien überhaupt existieren und unter welchen Bedingungen sie angewendet werden. In Analogie zur Prinzipien- und Parameter-Theorie in der Linguistik ist anzunehmen, dass es universale Verarbeitungsstrategien gibt, die sich auf universale phonologische Strukturen und Prozesse beziehen und sprachspezifische Verarbeitungsstrategien, die sich auf sprachspezifische Strukturen und Prozesse gründen.

Welche weiteren Verarbeitungsstrategien kann es also geben und welche sind die universalen Verarbeitungsstrategien, die von den Sprechern aller Sprachen verwendet werden? Als universal kann wohl die Verarbeitung der distinktiven Merkmale der Sprache, d.h. der Phonemstruktur gelten, da alle Sprachen eine distinktive Phonemebene aufweisen, obwohl, wie oben (s.o. 4.2.3) gezeigt wurde, auch auf dieser Ebene die Art der Abbildung von akustischen Eigenschaften auf die Phonemkategorien einer Sprache von sprachspezifischen Parametersetzungen abhängt. Im Hinblick auf suprasegmentale phonologische Strukturen unterscheiden sich die Einzelsprachen aber weit mehr voneinander als im Hinblick auf die Ebene der distinktiven Merkmale: Alle Sprachen nutzen distinktive Merkmale, wenn auch eine unterschiedliche Auswahl aus der Menge der

möglichen distinktiven Merkmale. Auf der suprasegmentalen Ebene ergibt sich jedoch ein anderes Bild, denn nicht alle Sprachen nutzen Tonhöhenunterschiede zur Differenzierung von Lexemen oder morphosyntaktischen Strukturen, wie dies in Tonsprachen wie Thai der Fall ist. Einige Sprachen haben variablen Wortakzent, der Lexeme unterscheiden kann, z.B. dtsh. 'umfahren und um'fahren, oder engl. 'object und ob'ject. Andere Sprachen haben festen Wortakzent, der Wortgrenzen indizieren kann, z.B. Ungarisch. In manchen Sprachen fallen Silbengrenzen mit Morphem- und Wortgrenzen zusammen, z.B. im Deutschen, in anderen Sprachen wieder nicht, z.B. im Französischen. Die Tatsache, dass eine phonologisch-distinktive Strukturierung der Lauterzeugung auf einer bestimmten Ebene, z.B. auf der Ebene der Tonhöhenvariation, in einer Sprache völlig fehlen kann, sollte auch bedeuten, dass die entsprechende Verarbeitungsstrategie sprachspezifisch ist und nur entwickelt wird, wenn in der kritischen Phase des Spracherwerbs die nötige sprachliche Umgebung vorhanden ist. Daher ist auch zu erwarten, dass die auf prosodische Kategorien wie Silbe und Akzent bezogenen Verarbeitungsstrategien, die in den phonologischen Strukturen verschiedener Sprachen eine unterschiedliche Funktion haben, zu einem hohen Maße sprachspezifisch sind.

Nach der Auffassung von CUTLER ET AL. (1992) sollten sich universale Verarbeitungsstrategien also auf universale Eigenschaften phonologischer Systeme beziehen und sprachspezifische Verarbeitungsstrategien sich aus einzelsprachlichen Variationen ableiten lassen. In Kapitel 2 dieser Arbeit wurde gezeigt, dass die Sonoritätshierarchie und das Sonoritäts-Folge-Prinzip Regularitäten in der Abfolge von Einzellauten in der Silbe zusammenfassen, die sich in nahezu allen Sprachen finden und als universal gelten können. Im Rahmen der Theorie von CUTLER ET AL. könnte man also erwarten, dass alle Sprecher bei der Verarbeitung von Äußerungen ihr Wissen über diese sonoritätsbezogenen phonotaktischen Regularitäten in der Sprache anwenden, um das einkommende Signal zu strukturieren.

4.3.5 Zusammenfassung

Die vorangegangenen Ausführungen haben gezeigt, dass sich im Sprachwahrnehmungsverhalten verschiedener Sprechergruppen Effekte messen lassen, die unter Verwendung der linguistischen Konzepte *Silbe*, *Akzent* und *starke Silbe* erklärt werden können, insbesondere durch die Annahme, dass Sprecher das Signal an *Silbengrenzen* oder *vor starken Silben* segmentieren. Die Segmentation des Signals an Silbengrenzen kann also als eine mögliche Lösung für das Segmentationsproblem der Sprachwahrnehmung gelten (s.o. 4.2.1). Ist auf der Grundlage dieser Beobachtungen auch die Be-

hauptung gerechtfertigt, dass die Sprecher dieser Sprachen Silben, Silbengrenzen, oder das Silbengewicht auf einer prälexikalischen Ebene *repräsentieren*? Wenn bei der Sprachverarbeitung das Signal an einem bestimmten Punkt segmentiert wird, welcher dem Rand einer linguistischen Kategorie entspricht, so ist anzunehmen, dass dieser Punkt als ein Silbenrand oder Taktrand erkannt werden muss, d.h. in gewisser Weise als solcher repräsentiert wird. Aus diesem Grunde werden die oben zusammengefassten experimentellen Ergebnisse als Hinweise darauf gewertet, dass bei der Sprachwahrnehmung auf einer prälexikalischen Ebene eine Repräsentation der *Silbenstruktur* und der *Akzentstruktur* der zu verarbeitenden Äußerung rekonstruiert wird. Wie in den vorangegangenen Ausführungen deutlich geworden ist, ist die Funktion der Silbe in der Sprachwahrnehmung bereits in zahlreichen Studien untersucht worden, zur Rolle der Sonorität in der Sprachwahrnehmung hingegen gibt es bisher keine Studien. Welche Funktion die phonologisch-phonetische Eigenschaft Sonorität also für die Rekonstruktion der Silbenstruktur¹⁸³ und der Akzentstruktur¹⁸⁴ einer Äußerung erfüllen könnte, wird im folgenden Abschnitt erörtert.

4.4 Sonorität in der Sprachwahrnehmung

4.4.1 Einleitung

Soll die Silbenstruktur einer zu verarbeitenden Äußerung rekonstruiert werden, so kann an zwei Punkten der Repräsentation bzw. des Signals angesetzt werden: die Rekonstruktion kann mit den Silbengrenzen, oder mit den Silbennuklei beginnen. In phonologischer Hinsicht haben Silbennuklei maximale Sonorität, Silbengrenzen sind Sonoritätsminima (s.o. 2.3). In akustischer Hinsicht sind Sonoritätsmaxima Signalabschnitte mit maximaler perzeptueller Salienz, denn sie sind lauter, länger und klangvoller (periodischer) als die Signalabschnitte, die den Lauten in den übrigen Strukturpositionen der Silbe entsprechen (s.o. 3.2). Daher ist es denkbar, dass der Hörer bei der Verarbeitung einer sprachlichen Äußerung zunächst die *Sonoritätskontur* einer Äußerung repräsentiert und die maximal salienten *Sonoritätsmaxima* lokalisiert, dass er aus diesen sodann eine Repräsentation von Folgen von Silbennuklei rekonstruiert und schließlich unter Rückgriff auf sein Wissen über die phonotaktischen Regularitäten in der Sprache die Silbenstruktur um die Silbennuklei herum rekonstruiert. Diese Annahme soll im folgenden näher begründet werden.

¹⁸³ Zum Verhältnis von Sonorität und Silbenstruktur s.o. 2.3 und 2.4.

¹⁸⁴ Zum Verhältnis von Sonorität und Akzentstruktur s.o. 3.3.2 und 3.4.

4.4.2 Warum sollte die Sonoritätskontur verarbeitet werden?

Wie in Kapitel 2 hervorgehoben wurde, ist die Silbenstruktur einer Segmentkette eine Funktion ihrer Sonoritätskontur, oder, anders ausgedrückt, es besteht eine eindeutige Beziehung zwischen der Sonoritätskontur einer Segmentkette und ihrer Silbenstruktur. In Kapitel 3 wurde festgestellt, dass es prinzipiell möglich ist, die Sonoritätskontur einer Äußerung aus dem akustischen Signal zu rekonstruieren (s.u. 3.5). Weiter wurde die Annahme gemacht, dass die phonetisch-phonologische Sonorität eines Segmentes als eine Funktion erstens der inhärenten Sonorität der Segmente in einer Segmentkette, zweitens der Silbenstruktur (s.o. 2.5) und drittens der Akzentstruktur dieser Segmentkette konzipiert werden kann (s.o. 3.3.2.2). Unter dieser Annahme kann aus der phonetischen Sonoritätskontur eines akustischen Signals sowohl die Silbenstruktur als auch die Akzentstruktur rekonstruiert werden. Unter einer linguistisch-phonetischen Perspektive ist es also plausibel, dass bei der Sprachwahrnehmung die Sonoritätskontur einer Äußerung verarbeitet wird. Im aktuellen Kapitel wurden experimentelle Ergebnisse zusammengefasst, die so interpretiert wurden, dass die Sprecher bei der Sprachverarbeitung das Signal an Silbengrenzen segmentieren. Dies wurde als Hinweis darauf gewertet, dass Sprecher aus dem zu verarbeitenden Signal eine Silbenstrukturepräsentation rekonstruieren. Daher ist die Annahme auch unter einer psycholinguistischen Perspektive plausibel, dass bei der Sprachverarbeitung die Sonoritätskontur einer Äußerung aus dem akustischen Signal extrahiert wird, um daraus eine phonologische Silbenstrukturepräsentation zu rekonstruieren. Auch unter der Hypothese von CUTLER ET AL. (1992), dass Sprachverarbeitungsstrategien sich auf *universale* und *sprachspezifische phonologische Strukturen* beziehen, sollte angenommen werden, dass die phonetischen Korrelate der Sonoritätskontur einer Äußerung verarbeitet werden, denn das Sonoritäts-Folge-Prinzip (s.o. 2.3), wie CLEMENTS (1990) es ausdrückt,

expresses a strong cross-linguistic tendency, and represents one of the highest-order explanatory principles of modern phonological theory. (CLEMENTS 1990:284).

Silben haben also in nahezu allen Sprachen eine ähnliche Form, die sich mit dem Konzept *Sonorität* beschreiben lässt, dem bestimmbare phonetische Korrelate entsprechen. Die Gründe für dieses universale Formprinzip der Silbe liegen vermutlich in den Bedingungen der Artikulation von Sprache und in den Bedingungen der Sprachwahrnehmung. Die Allgemeingültigkeit dieses Prinzips lässt es jedoch auch als wahrscheinlich erscheinen, dass die Sprecher diese vorsezifizierte Struktur nutzen, um das akustische Signal zu strukturieren, indem sie die Sonoritätskontur verarbeiten und Sonoritätsmaxima und -minima lokalisieren und schließlich die erkannten Phoneme den universalen und einzelsprachlichen phonotaktischen Regeln entsprechend dieser globalen Sonori-

tätsstruktur zuordnen, und eventuell noch bestehende Unsicherheiten in der Segmenterkennung auflösen.

Welche anderen Hinweise gibt es dafür, dass die Annahme gemacht werden muss, dass zu den (universalen oder sprachspezifischen) Sprachverarbeitungsstrategien auch die Wahrnehmungsfunktion *Lokalisierung der Sonoritätsmaxima in einer Äußerung* gehört? In einer Studie zur Wahrnehmung silbischer Konsonanten demonstrierte PRICE (1980) unter Verwendung von Sprachsignalen, die entweder als das englische Wort *plight* (bzw. *prayed*) oder als das englische Wort *polite* (bzw. *parade*) wahrgenommen wurden (s.o. 3.2.2), dass erstens Silbennuklei auch dort wahrgenommen werden können, wo das akustische Signal keinen Hinweis auf die Artikulation eines Vokals gibt und dass zweitens ein und dasselbe phonologische Segment in Abhängigkeit von seiner Dauer, Lautstärke und Klangfülle¹⁸⁵ entweder als Silbennukleus oder als Silbenanlaut wahrgenommen wird (für vergleichbare Daten zum Deutschen vgl. PORTELE & HEUFT 1993, JANNEDY 1994¹⁸⁶). Diese beiden Beobachtungen sind nicht mit der Annahme vereinbar, dass die Silbennuklei einer Äußerung aus ihrer Phonemstruktur erst abgeleitet werden, d.h. aus den phonotaktischen Eigenschaften der Äußerung deduziert werden. Vielmehr muss davon ausgegangen werden, dass es eine Verarbeitungsfunktion gibt, die unabhängig vom segmentalen Gehalt einer Äußerung auf der Basis von phonetisch-akustischen Merkmalen des Signals eine Repräsentation der Silbennuklei der Äußerung erzeugt.

Für diese Annahme spricht auch die Beobachtung von BJELJAC-BABIC, BERTONCINI & MEHLER (1993), die in einer Untersuchung der Sprachwahrnehmungsfähigkeit von Neugeborenen demonstrierten, dass diese zweisilbige von dreisilbigen Äußerungen unterscheiden können; diese Leistung konnten die Neugeborenen unabhängig von der Dauer bzw. Sprechrate der Stimuli erbringen. Hieraus folgerten sie, dass Neugeborene den Rhythmus, die Zahl der rhythmischen 'Schläge', oder die Silben einer Äußerung repräsentieren und zwar lange bevor sie das phonologische System ihrer zukünftigen Muttersprache erworben haben.¹⁸⁷ Auf der Grundlage dieser und anderer Beobachtungen entwickelten MEHLER, DUPOUX, NAZZI & DEHAENE-LAMBERTZ (1996) die Theorie, dass Neugeborene in dieser frühen Phase des Spracherwerbs die Vokale des Signals in Form einer *Time and Intensity Grid Representation* repräsentieren:

¹⁸⁵ *Klangfülle* kann hier als gleichbedeutend mit *Grad der Stimmhaftigkeit* verstanden werden.

¹⁸⁶ JANNEDY, S. (1994): „Rate effects on German unstressed syllables“, *Working Papers in Linguistics, Ohio State University* 44:105-124, zitiert nach BECKMAN (1996).

¹⁸⁷ So beginnen Kinder im Alter zwischen 8 und 12 Monaten, die für ihre Muttersprache spezifische Art von kategorischer Phonemwahrnehmung zu zeigen; für einen Überblick zur Entwicklung der Sprachwahrnehmungsfähigkeit vgl. JUSCZYK 1995.

The TIGRE is a gridlike representation of the vocalic nuclei in the speech signal. All the vowels are represented in it so that the sequential information is preserved. We assume that each vowel receives an index indicating its duration and amplitude. (MEHLER ET AL. 1996:113)

Bereits Neugeborene können verschiedene Sprachen voneinander unterscheiden und zwar auch dann, wenn die Stimuli tiefpassgefiltert sind, so dass im Sprachsignal keine Information über die Identität einzelner Phoneme mehr verfügbar ist. Dies bedeutet, dass sie diese Leistung allein auf Grund prosodischer Eigenschaften wie Intonation, Lautstärke und Dauer erbringen (MEHLER, JUSCZYK, LAMBERTZ, HALSTED, BERTONCINI & AMIEL-TISON 1988, vgl. auch MEHLER & DUPOUX 1990:213ff). Um diese Beobachtung in das Modell zu integrieren, nehmen MEHLER ET AL. (1996) an, dass die Vokale im TIGRE mit zusätzlichen prosodischen Markierungen zu Dauer und Amplitude versehen sind. Das TIGRE-Modell von MEHLER ET AL. enthält also Information auf zwei Ebenen: einerseits Information über die Identität der Vokale, die auf Grund spektraler Information aus dem Signal extrahiert werden kann, andererseits prosodische Information über die Dauer und Lautstärke der Vokale, die dem Rhythmus bzw. der Akzentstruktur der Äußerung entsprechen.

Die alternative Annahme, dass Neugeborene nicht die Vokale, sondern nur die *Sonoritätsmaxima* einer Äußerung und ihre *relative Sonorität* bzw. Akzentuiertheit, d.h. ihre Lautstärke und Dauer relativ zueinander, repräsentieren, ließe dieselben Beobachtungen erwarten, nämlich dass Neugeborene erstens Wörter mit unterschiedlicher Silbenzahl voneinander unterscheiden können und zweitens Sprachen mit unterschiedlichem Sprachrhythmus. Diese Annahme hätte zudem den Vorteil, dass das Format der Repräsentation, die Neugeborene aus dem Signal rekonstruieren, einheitlicher wäre als im TIGRE-Modell, weil nur eine Art von Eigenschaft repräsentiert wird, nämlich die Sonorität und nicht einerseits Phoneminformation und andererseits Information über Amplitude und Dauer. In diesem wichtigen Aspekt unterscheiden sich die Modelle also: das TIGRE-Modell sagt voraus, dass Neugeborene alle oder fast alle Vokale differenzieren können, das Sonoritätsmodell sagt hingegen voraus, dass Neugeborene nur Vokale mit unterschiedlicher Sonorität, d.h. hohe (/i/ und /u/), halbhohe (/e/ und /o/) und tiefe Vokale (/a/) voneinander unterscheiden können.

Zusammenfassend lässt sich sagen, dass die Annahme, dass bei der Wahrnehmung sprachlicher Äußerungen die Sonoritätskontur verarbeitet wird, sowohl unter einer linguistisch-phonetischen als auch unter einer psycholinguistischen Perspektive plausibel ist und den Vorteil hat, dass die Verarbeitung von nur einer einzigen Eigenschaft nötig ist, d.h. der Sonorität, um zwei Arten von phonologischer Repräsentation zu rekonstruieren, d.h. die Silbenstruktur-Repräsentation und die Akzentstruktur-Repräsentation.

4.4.3 Wie könnte die Sonoritätskontur verarbeitet werden?

In Kapitel 3 ist gezeigt worden, dass das phonologische Merkmal Sonorität mehrere akustische Korrelate hat, vor allem die Lautstärke und die Dauer eines Segmentes. Die phonologische Sonoritätskontur ist daher prinzipiell aus dem akustischen Signal rekonstruierbar, wenn auch möglicherweise nur unter Rückgriff auf weitere Parameter wie z.B. die Periodizität bzw. Stimmhaftigkeit eines Signalabschnittes, dem Grad der Dynamik des Signals, oder auf zusätzliche visuelle Information über die Dynamik der Mundbewegungen des Sprechers. Für die Wahrnehmung des Silbenrhythmus bzw. der periodischen Sonoritätsmodulation und für die Lokalisation der Silbennuklei bzw. der Sonoritätsmaxima scheint insbesondere die langsame Amplitudenmodulation im Signal von Bedeutung zu sein; so konnten REMEZ & RUBIN (1990) in einer Studie zur Wahrnehmung unverständlicher sprachähnlicher Signale zeigen, dass die langsame Amplitudenmodulation die Grundlage für die Wahrnehmung der Silbenzahl bildet. PFITZINGER, HEID & BURGER (1996) demonstrierten, dass sich auf der Basis der langsamen Lautstärkemodulation im Signal ein sehr großer Teil der Silbennuklei einer Äußerung automatisch detektieren lässt.

Die Hypothese, dass die Sonoritätskontur auf der Basis der Lautstärkemodulation des Signals und auf der Basis der temporalen Eigenschaften des Signals, d.h. auf der Grundlage des Wechsels von längeren und kürzeren Abschnitten mit konstanten Frequenzeigenschaften bzw. des Wechsels von Segmenten mit längerer und kürzerer Dauer, rekonstruiert wird und zwar in einem zeitlichen Rahmen von Silbendimension, d.h. von 200-400 ms Dauer, fügt sich auch in eine Konzeption der Sprachverarbeitung, die unter hörphysiologischer und computerwissenschaftlicher Perspektive entwickelt wurde und die GREENBERG (1996) folgendermaßen zusammenfasst:

a new theoretical perspective, focusing on time, is beginning to emerge. This framework emphasizes the temporal evolution of coarse spectral patterns as the primary carrier of information within the speech signal, and provides an efficient and effective means of shielding linguistic information against the potentially hostile forces of the natural soundscape such as reverberation and background acoustic interference. [...] this representation provides a principled basis for the segmentation of the speech signal into syllabic units. [...] it is possible, in principle, to gain direct access to the lexicon and grammar through such an auditory analysis of speech. (GREENBERG 1996:1)

Diese Konzeption der Verarbeitung natürlichsprachlicher akustischer Signale stellt nicht mehr die sukzessive Identifikation von einzelnen Phonemen auf der Basis des Frequenzspektrums zu einem Zeitpunkt in den Vordergrund, sondern die Verarbeitung der langsamen Veränderung der Energieverteilung in groben Frequenzbändern, welche Information über die phonologische, die morphologische und die syntaktische Strukturierung der zu verarbeitenden Äußerung trägt. Auf diese Weise ist nicht mehr nur die *bottom-up*-Rekonstruktion der Strukturen auf höheren Ebenen wie Morphologie und

Syntax aus einer Sequenz von identifizierten Phonemen möglich, sondern auch eine unmittelbare Rekonstruktion von Silbengrenzen, Morphemgrenzen, Wortgrenzen und Phrasengrenzen aus dem akustischen Signal (und damit auch die *top-down*-Beeinflussung der Phonemverarbeitung). Eine solche Art der Signalverarbeitung, die sich auf langsame Signalmodulationen stützt, hat im Vergleich zu einer Strategie der sukzessiven Erkennung von einzelnen Phonemen zudem den Vorteil, dass sie, wie GREENBERG ausführt, gegenüber Signalverzerrungen und Störgeräuschen robuster ist.

Dieses Modell der Sprachverarbeitung ist im Rahmen der Forschungen zur *Auditorischen Szenenanalyse* (BREGMAN 1990) zu betrachten, die sich mit der Frage beschäftigt, wie einerseits ein komplexes Signal, das von verschiedenen Schallquellen erzeugt wurde, in verschiedene Informationsströme segregiert wird, die jeweils dem von jeder einzelnen Schallquelle erzeugten Signal entsprechen und wie andererseits die Bestandteile eines einzigen Signals, das aus teilweise abrupten Veränderungen der Energieverteilung in verschiedenen, zum Teil weit auseinanderliegenden Frequenzbereichen bestehen kann, zu einer einzigen Wahrnehmungsgestalt integriert wird. Für die Sprachwahrnehmung ist der erste Aspekt von großer Wichtigkeit, weil unter normalen Kommunikationsbedingungen meist mehr als eine Person redet und sich die Stimmen der Sprecher stets mit Störgeräuschen aus der Umwelt vermischen. Von noch größerer Bedeutung für die Sprachwahrnehmung ist aber vielleicht der zweite Aspekt, weil sich die Energie in einem Sprachsignal über einen weiten Frequenzbereich erstreckt und in stimmhaften Lauten auf mehrere Formanten aufteilt. Eine Äußerung besteht zudem aus einer Folge von sehr verschiedenen Phonemen, was im akustischen Signal schnell alternierenden Sequenzen von Stille (Plosive), Rauschen (Frikative) und Klang (Sonoranten) mit schnell wechselnden Frequenzen entspricht. Alle diese unterschiedlichen Signalabschnitte müssen bei der Sprachwahrnehmung zu einem kohärenten Perzept zusammengefügt werden (vgl. HANDEL 1989:216). Die langsame Amplitudenmodulation im Signal, die auf phonologischer Ebene dem Silbenrhythmus oder der Sonoritätskontur entspricht, könnte hierbei gewissermaßen den 'roten Faden' im Signal bilden, auf dessen Grundlagen nach den 'Prinzipien der guten Gestalt' (BREGMAN 1990) verschiedene Sprachsignale segregiert und die verschiedenen Teile eines einzigen Sprachsignals integriert werden können.

Wie könnte diese langsame Amplitudenmodulation im Sprachsignal auf neuronaler Ebene verarbeitet werden? Im auditorischen Kortex und in den subkortikalen Kernen der Hörbahn finden sich Neuronen, welche phasengekoppelt mit der Modulation des einkommenden Signals feuern, die also jeweils an einem ganz bestimmten Punkt

im Modulationszyklus mit neuronaler Aktivität reagieren.¹⁸⁸ Diese Neuronen reagieren zudem selektiv auf Signale mit einer bestimmten Modulationsfrequenz. Im auditorischen Kortex finden sich dabei vor allem solche Neuronen, die durch langsame Modulationsfrequenzen unter 20-30 Hz aktiviert werden (CREUTZFELDT 1983:207). Abgesehen davon, dass diese Neuronen *modulationsfrequenzspezifisch* reagieren, ist ihre Aktivität auch eine Funktion der Intensität bzw. der *Modulationstiefe* des Signals (CLAREY, BARONE & IMIG 1991:289). Es wäre also möglich, dass im auditorischen Kortex erstens die langsame Amplitudenmodulation des Signals (bzw. der Silbenrhythmus oder die Sonoritätskontur) durch die phasengekoppelte Aktivität *modulationsfrequenzspezifischer* Neuronen repräsentiert wird. Zweitens könnten Neuronen existieren, die phasengekoppelt mit Sonoritätsanstiegen feuern und so Silbenanlaute identifizieren, andere könnten phasengekoppelt mit Sonoritätsmaxima feuern und Nuklei identifizieren usw. Auch Silben von unterschiedlicher struktureller Komplexität, denen unterschiedlich steile Sonoritätsänderungen entsprechen, könnten so unterschieden werden. Drittens könnten *modulationstiefespezifische* Neuronen die relative Sonorität der Sonoritätsmaxima in Relation zu den umgebenden Sonoritätsmaxima detektieren. Einige Neuronen mit einer eher geringen 'besten' Modulationstiefe könnten phasengekoppelt mit allen Sonoritätsmaxima, d.h. mit allen Silben, feuern, andere Neuronen mit einer eher größeren 'besten' Modulationstiefe könnten phasengekoppelt nur mit den akzentuierten Sonoritätsmaxima feuern und auf diese Weise akzentuierte Silben repräsentieren. Die Annahme, dass die dem Silbenrhythmus bzw. der Sonoritätskontur entsprechende Amplitudenmodulation im Signal durch phasengekoppelte Aktivität von Neuronen im zerebralen Kortex detektiert werden kann, wird auch durch die Beobachtung von CREUTZFELDT, OJEMANN & LETTICH (1989a) gestützt, die herausfanden, dass sich im temporalen Kortex der sprachdominanten Hemisphäre Neuronen finden lassen, die im 'Rhythmus der Sprache' feuern. GREENBERG (1996) nimmt daher an, dass bei der Verarbeitung von Sprachsignalen im auditorischen Kortex das niederfrequente Modulationsspektrum berechnet wird, auf dessen Grundlage das Signal in Silben segmentiert wird. Die Sonoritätskontur einer Äußerung bzw. ihre phonetischen Korrelate Dauer und Lautstärke können also prinzipiell aus dem akustischen Signal extrahiert werden und prinzipiell durch die Aktivität von Neuronen mit *modulationsfrequenz- und modulationstiefenspezifischen* Antworteigenschaften repräsentiert werden. Eine Antwort auf die Frage, wie dies genau geschehen könnte, d.h. auf der Grundlage welcher akustischer Eigenschaften des Signals und durch welche funktionalen Eigenschaften von Neuronen

¹⁸⁸ Für einen Überblick vgl. LANGNER 1992, CLAREY, BARONE & IMIG 1991, vgl. auch SCHREINER 1995.

und neuronalen Netzen die Sonoritätskontur einer Äußerung verarbeitet werden könnte, kann zum gegenwärtigen Zeitpunkt nur auf der Basis von Versuchen der Modellierung der auditorischen Wahrnehmung in neuronaler Architektur geliefert werden.

4.5 Zusammenfassung

In diesem Kapitel wurden zunächst experimentelle Beobachtungen zusammengefasst, die zeigen, dass die Kategorie *Silbe* eine notwendige Konstituente einer psycholinguistischen Theorie der Sprachwahrnehmung ist. Im Anschluss daran wurde die Annahme formuliert, dass die phonologisch-phonetische Eigenschaft *Sonorität*, die als Verbindungsglied zwischen segmentaler Struktur und suprasegmentaler Silben- und Akzentstruktur konzipiert wurde (s.u. 2.5, 3.5), zur Strukturierung der wahrgenommenen sprachlichen Äußerungen genutzt wird und dass bei der Verarbeitung eines Sprachsignals insbesondere die maximal salienten Sonoritätsmaxima lokalisiert werden, welche den Silbennuklei entsprechen. Diese Annahme wurde aus phonologisch-phonetischer und aus psycholinguistischer Perspektive begründet. Außerdem wurde in diesem Kapitel hervorgehoben, dass es in der Sprachwahrnehmung universale und supplementäre Verarbeitungsstrategien gibt, deren Verwendung von den phonologischen Eigenschaften der zu verarbeitenden Sprache, von akustisch-phonetischen Eigenschaften des einkommenden Signals sowie von weiteren noch genauer zu spezifizierenden Faktoren abhängt. In diesem Kapitel ist mit der Annahme der Verarbeitungsfunktion *Lokalisierung der Sonoritätsmaxima in einer Äußerung* also die Grundlage geschaffen worden, das phonologisch-phonetische Merkmal Sonorität auf dynamische Aktivierungsprozesse über funktionalen Einheiten im Gehirn zu beziehen. Im nächsten Kapitel kann nunmehr die Frage diskutiert werden, ob die Verarbeitung der Sonoritätskontur einer Äußerung eher mit rechtshemisphärischer oder eher mit linkshemisphärischer Aktivierung korreliert ist.

5 Funktionale Lateralisierung der Sprachwahrnehmung

5.1 Einleitung

In der vorliegenden Studie wird die These vertreten, dass die Verarbeitung der Sonoritätskontur einer Äußerung rechtshemisphärisch lateralisiert ist. Um diese These zu begründen, werden in diesem Kapitel bestehende Erkenntnisse zur funktionalen Lateralisierung der Sprache betrachtet und in einige Theorien zur Hemisphärenspezialisierung eingeordnet. Zunächst (5.2) werden funktionale Lateralitäten in der Sprache und in anderen kognitiven Funktionen zusammengefasst. Im Anschluss daran (5.3) werden einige Theorien zur Hemisphärenspezialisierung für sprachliche und nichtsprachliche kognitive Funktionen skizziert. Da es in dieser Studie um die funktionale Lateralisierung der Wahrnehmung der Sonoritätskontur geht, werden im darauffolgenden Abschnitt (5.4) die bisherigen Erkenntnisse zur Lateralisierung der Wahrnehmung von segmentalen (5.4.2) und suprasegmentalen (5.4.3) Kategorien und Merkmalen genauer betrachtet. Die in diesem Abschnitt dargestellten experimentellen Daten und Beobachtungen dienen auch als Hintergrund für die Konzeption und Interpretation der im nächsten Kapitel (s.u. 6) beschriebenen dichotischen Sprachwahrnehmungsversuche. Im letzten Abschnitt dieses Kapitels (5.5) wird die Hypothese einer rechtshemisphärischen Spezialisierung für die Verarbeitung der Sonoritätskontur einer Äußerung aus den skizzierten Theorien zur Hemisphärenspezialisierung abgeleitet.

5.2 Hemisphärenspezialisierung für sprachliche und nichtsprachliche kognitive Funktionen

5.2.1 Funktionen der linken Hemisphäre

Vor allem was die Kontrolle der Sprachfunktionen betrifft, ist das Gehirn asymmetrisch

organisiert. Nach Schädigungen der linken Hemisphäre durch äußere Einwirkung oder durch Schlaganfälle und Tumoren leiden die Patienten sehr häufig an Sprachstörungen, während Schädigungen der rechten Hemisphäre eher zu anderen kognitiven Defiziten als zu sprachlichen führen. Werden bestimmte Areale der linken Hemisphäre zerstört, so führt dies insbesondere zu gravierenden Defiziten in phonologischen, morphologischen, syntaktischen und lexikalischen Komponenten der Sprachfähigkeit, d.h. zu verschiedenen Formen der Aphasie oder zu Alexie (Verlust der Lesefähigkeit) oder Agraphie (Verlust der Schreibfähigkeit). Läsionen der rechten Hemisphäre führen hingegen zu Defiziten in prosodischen Sprachfunktionen (Aprosodie), d.h. zu produktiven und rezeptiven Defiziten im Hinblick auf die Sprachmelodie und den Sprachrhythmus.

Wahrnehmung und Produktion gesprochener Sprache. Für die Wahrnehmung von ganzen Sätzen, von Wörtern und von einzelnen Lauten lässt sich in dichotischen Studien üblicherweise ein Rechts-Ohr-Vorteil, d.h. eine Überlegenheit der linken Hemisphäre, demonstrieren (für einen Überblick vgl. TARTTER 1988, s.u. 5.4). Auch für die Ausführung von Artikulationsbewegungen lassen sich Asymmetrien in der Amplitude und in der Geschwindigkeit der Mundbewegungen zeigen: die rechte Mundhälfte, die von der linken Hemisphäre kontrolliert wird, wird beim Sprechen weiter geöffnet und schneller bewegt als die linke Mundhälfte (WOLF & GOODALE 1987, HAUSMANN ET AL. 1998).

Wahrnehmung und Produktion geschriebener Sprache. Bei tachistoskopischen¹⁸⁹ Untersuchungen zur Worterkennung hat man festgestellt, dass Wörter, die in der rechten Gesichtsfeldhälfte präsentiert (und die in die linke Hemisphäre projiziert) werden, schneller verarbeitet werden als Wörter, die in der linken Gesichtsfeldhälfte dargeboten (und die in die rechte Hemisphäre projiziert) werden (MISHKIN & FORGAYS 1952, BARTON, GOODGLASS & SHAI 1965). Daraus kann man auf eine Spezialisierung der linken im Vergleich zur rechten Hemisphäre für die Verarbeitung geschriebener Wörter (d.h. für die Lesefähigkeit) schließen. Ein weiteres Argument für die Annahme einer Spezialisierung der linken Hemisphäre für Lesen und Schreiben liefert die Tatsache, dass Alexie (Leseunfähigkeit) ebenso wie Agraphie (Schreibunfähigkeit) bei Rechtshändern nach einer Läsion des Gyrus Angularis der linken, aber nicht der rechten Hemisphäre auftritt (SPRINGER & DEUTSCH 1995:169).

Wahrnehmung und Produktion von Zeichensprache. Eine linkshemisphärische Spezialisierung lässt sich auch für die Produktion und für die Wahrnehmung von Zeichensprache nachweisen. HICKOK, BELLUGI & KLIMA (1996) konnten zeigen, dass

¹⁸⁹ Bei tachistoskopischer Präsentation visueller Reize wird der Stimulus in nur einer Gesichtsfeldhälfte kurz (nicht länger als 100 ms) dargeboten, um zu gewährleisten, dass er nur in eine Hemisphäre, nämlich in die zur Gesichtsfeldhälfte kontralaterale, projiziert wird (s.o. 1.2.1).

Probanden mit linkshemisphärischer Läsion bei der Produktion, der Wahrnehmung und der Wiederholung von sprachlichen Gesten sowie bei Objektbenennungsaufgaben (in Zeichensprache) schlechter abschneiden als rechtshemisphärisch geschädigte Probanden. Dieses Defizit ist ein rein linguistisches Defizit, das nicht auf Störungen der (nichtsprachlichen) Handbewegungskontrolle oder auf ein allgemeines Defizit der räumlichen oder der Bewegungswahrnehmung zurückzuführen ist. DAMASIO ET AL. (1986) konnten an Probanden, die sowohl Lautsprache als auch Zeichensprache beherrschten, eine gleichzeitige Aphasie für Laut- und Zeichensprache nach der Natrium-Amytal-Injektion in die linke Karotisarterie demonstrieren.¹⁹⁰ Der Vorteil für die rechte Gesichtsfeldhälfte (d.h. für die linke Hemisphäre), der sich in tachistoskopischen Studien für die Wahrnehmung von geschriebenen Wörtern findet, konnte auch für Abbildungen von sprachlichen Gesten repliziert werden (GROSSI, SEMENZA, CORAZZA & VOLTERRA 1996).

5.2.2 Funktionen der rechten Hemisphäre

Während die Spezialisierung der linken Hemisphäre für Sprache schon im letzten Jahrhundert erkannt wurde, blieb lange unklar, worin die komplementäre Spezialisierung der rechten Hemisphäre besteht. Intensiv untersucht wurde die Leistungsfähigkeit der rechten Hemisphäre seit den fünfziger Jahren an *Split-Brain*-Patienten (SPRINGER & DEUTSCH 1995:48ff). Bei diesen Patienten wurde das Corpus Callosum durchtrennt, d.h. das Nervenfaserbündel, welches die beiden Hemisphären verbindet. Diese Operation führt dazu, dass (exzitatorische oder inhibitorische) neuronale Aktivität sich nur eingeschränkt über die intakten Fasern der tiefer gelegenen interhemisphärischen Kommissuren (*commissurae colliculi, fornicis* etc.) von der einen in die andere Hemisphäre ausbreiten kann. Daher eignet sich diese Patientengruppe auch besonders für die Untersuchung der kognitiven Fähigkeiten nur einer Hemisphäre, denn im Gegensatz zu *Split-Brain*-Patienten stehen bei Probanden mit intaktem Corpus Callosum beide Hemisphären in ständiger Interaktion, weshalb sich bei letzteren nur *relative* Hemisphärenspezialisierungen in Form von relativen Verarbeitungsvorteilen nachweisen lassen, während bei *Split-Brain*-Patienten die *absoluten* Unterschiede in der Leistungsfähigkeit der beiden Hemisphären deutlicher zutage treten.

Räumlich-visuelle Verarbeitung. In einer Studie mit *Split-Brain*-Patienten zeigten BOGEN & GAZZANIGA (1965), dass diese Probandengruppe eine Vorlage besser mit der

¹⁹⁰ In dieser Untersuchung wurde keine rechtsseitige Injektion vorgenommen, daher lassen sich keine vergleichenden Aussagen machen.

linken Hand, welche durch die rechte Hemisphäre kontrolliert wird, als mit der links-hemisphärisch kontrollierten rechten Hand abzeichnen konnte. LEVY-AGRESTI & SPERRY (1968) fanden heraus, dass *Split-Brain*-Patienten auch das Herausgreifen von vorspezifizierten Objekten ohne visuelle Kontrolle nach (visueller) tachistoskopischer Präsentation des zu greifenden Objektes mit der rechtshemisphärisch kontrollierten linken Hand besser als mit der rechten Hand durchführten. Diese und ähnliche Erkenntnisse führten zu der Annahme, dass die rechte Hemisphäre für räumlich-visuelle Fähigkeiten spezialisiert ist. In diesem Zusammenhang ist noch ein anderer Aspekt zu nennen, der auf eine Spezialisierung der rechten Hemisphäre zurückzuführen ist: das Neglekt-Syndrom. Kontralateraler Neglekt oder Hemineglekt tritt nach Schädigung des Parietallappens der rechten Hemisphäre weitaus häufiger auf als nach entsprechender Schädigung der linken Hemisphäre (KOLB & WHISHAW 1993:406). Bei diesem Syndrom ignoriert der Patient die kontralaterale, d.h. bei rechtshemisphärischer Läsion die linke Hälfte seines Körpers und die linke Hälfte seiner Umwelt in allen sensorischen Modalitäten. Eine linksseitige Läsion hingegen führt in der Regel nicht zu kontralateralem Neglekt.

Emotion und Affekt. Ein weiterer Aspekt, für den eine Spezialisierung der rechten Hemisphäre angenommen wird, ist die Kontrolle von Emotion und Affekt.¹⁹¹ So findet sich eine Überlegenheit der rechten Hemisphäre für die Wahrnehmung von Emotionen in Gesichtsausdrücken (BRYDEN 1982), ebenso erscheint die von der rechten Hemisphäre kontrollierte linke Gesichtshälfte als emotional ausdrucksvoller als die rechte Gesichtshälfte (PIZZAMIGLIO, CALTAGIRONE & ZOCCOLATTI 1989). Daneben finden sich auch Unterschiede im emotionalen Verhalten nach rechtshemisphärischer und linkshemisphärischer Läsion: Patienten mit linkshemisphärischer Läsion neigen eher zu Depression („catastrophic reaction“), Patienten mit rechtshemisphärischer Läsion eher zu Indifferenz oder Euphorie (GAINOTTI 1989). Nicht zu vergessen ist schließlich die Spezialisierung der rechten Hemisphäre für die Verarbeitung und für die Produktion der emotionalen Prosodie in der Sprache (ROSS 1981, ROSS, THOMSON & YENKOSKY 1997) und für die Verarbeitung von Wörtern mit emotionaler Bedeutung (BOROD, BLOOM & HAYWOOD 1997).

5.2.3 Sprachfunktionen der rechten Hemisphäre

Während die rechte Hemisphäre von Patienten mit durchtrenntem Corpus Callosum

¹⁹¹ Für einen Überblick vgl. BOROD 1992, DAVIDSON 1995.

nicht in der Lage ist, die Kontrolle über die Sprachproduktion zu übernehmen, kann sie dennoch Wörter und Sätze verstehen (ZAIDEL 1978,¹⁹² zit. nach SPRINGER & DEUTSCH 1995:61). Auch nach linksseitiger Injektion von Natrium-Amytal, d.h. bei einer Betäubung der linken Hemisphäre, bricht das Sprachverständnis nicht völlig zusammen (WADA & RASMUSSEN 1960, MACGLONE 1984). Abgesehen von der Tatsache, dass die rechte Hemisphäre Sprache *verstehen* kann, ist sie für die Kontrolle *prosodischer*, *pragmatischer* und *nichtpropositionaler* Sprachfunktionen zuständig.¹⁹³

Prosodie. Während linkshemisphärische Läsionen aphasische Störungen verursachen (DAMASIO 1992), kann eine rechtshemisphärische Läsion zu einer Aprosodie führen (ROSS 1981, ROSS ET AL. 1997), d.h. zur Unfähigkeit, Äußerungen mit angemessenem Rhythmus (Betonung) und angemessener Melodie (Intonation) zu produzieren, bzw. die Bedeutung des Rhythmus und der Melodie einer Äußerung richtig zu interpretieren. Dieses Defizit kann sich auf die emotionale, pragmatische, syntaktische oder lexikalische Funktion von Rhythmus und Melodie erstrecken. Die rechte Hemisphäre ist jedoch nicht für die Wahrnehmung und Produktion all dieser Funktionen spezialisiert, sondern zumindest syntaktische und lexikalische Funktionen der Prosodie scheinen eher von der linken Hemisphäre kontrolliert zu werden (s.u. 5.3.1).

Pragmatische Sprachfunktionen. Patienten mit rechtshemisphärischer Schädigung haben im Vergleich zu linksseitig geschädigten Patienten Schwierigkeiten mit den pragmatischen Aspekten der Sprache, d.h. mit der kontextabhängigen Interpretation von Äußerungen im Hinblick auf die innertextliche und die außersprachliche Situiertheit dieser Äußerungen und zwar sowohl bei der Sprachverarbeitung wie auch bei der Sprachproduktion. Hierzu gehört:

- die *übertragene Bedeutung* von Wörtern oder Sätzen in indirekten Aufforderungen¹⁹⁴, in ironischer Redeweise¹⁹⁵, in metaphorischen Ausdrücken¹⁹⁶ und in Sprichwörtern¹⁹⁷ (BROWNELL 1988, MOLLOY, BROWNELL & GARDNER 1990);
- die *Korreferenz* von verschiedenen Substantiven (*das Auto – der Wagen*) und Pronomina (*der Wagen – er*) im Text, die auf Grund grammatischer Regeln oder

¹⁹² ZAIDEL, E.: „Auditory language comprehension in the right hemisphere following cerebral commissurotomy and hemispherectomy: a comparison with child language and aphasia“, in: CARAMAZZA, A. & ZURIF, E.B. (Hrsg.): *Language Acquisition and Language Breakdown*. Baltimore: Johns Hopkins University Press.

¹⁹³ Für Überblicksdarstellungen zur Sprachfähigkeit der rechten Hemisphäre vgl. CODE 1987, CHIARELLO (Hrsg.) 1988, JOANETTE, GOULET & HANNEQUIN 1990, VAN LANCKER (Hrsg.) 1997, BEEMAN & CHIARELLO (Hrsg.) 1997.

¹⁹⁴ Bsp.: *Weißt Du, wie spät es ist?* als Aufforderung, die Uhrzeit zu nennen.

¹⁹⁵ Bsp.: *Das ist ja eine schöne Bescherung.*

¹⁹⁶ Bsp.: *am Fuße des Berges.*

¹⁹⁷ Bsp.: *Morgenstund hat Gold im Mund.*

kontextueller Hinweise rekonstruiert werden kann (LEONARD, WATERS & CAPLAN 1997);

- die *Textgrammatik*, d.h. die Anordnung der Informationseinheiten in einem mündlichen oder schriftlichen Text, so dass ein thematischer Zusammenhang der Teiläußerungen hergestellt und die relevante Information genannt wird, aber die irrelevante Information ausgespart bleibt, sowie die sprachliche Markierung der Beziehungen zwischen den Informationseinheiten des Textes: die Herstellung von Textkohäsion durch Konjunktionen, Präpositionen und den Satzbau (JOANETTE & GOULET 1990, ALBYN-DAVIS, O'NEIL-PIROZZI & COON 1997).

In einer neurolinguistischen Untersuchung pragmatischer Sprachfunktionen muss der Proband stets mehr als einen einzigen Satz verarbeiten, denn er muss einen Satz im situativen Kontext oder gleich ganze Texte verarbeiten. Daher stellt sich hier mehr als bei der Untersuchung von z.B. Phonologie und Syntax die Frage, welche Art von Verhalten untersucht werden soll: Spontansprache, Nacherzählung einer Bildergeschichte oder Beantwortung von Fragen zu einem Text? Zudem muss das Problem gelöst werden, wie ein pragmatisches Defizit in der Sprachproduktion definiert werden soll: mit Hilfe von quantifizierenden Kriterien wie 'Zahl der benutzten Pronominalausdrücke oder Konjunktionen', 'Verhältnis der Anzahl der Nebensätze zur Anzahl der Hauptsätze', oder eher mit nur schwer zu quantifizierenden Kriterien der Art 'Verständlichkeit des produzierten Textes'?¹⁹⁸ Ein drittes Problem liegt darin, dass sich die Verarbeitung von pragmatischer Information nicht oder nur schwer mit den Methoden der lateralisierten Reizdarbietung (Tachistoskopie, dichotisches Hören) untersuchen lässt, weil diese Methoden auf sehr kurze Präsentations- und Reaktionszeiten angewiesen sind, um zu gewährleisten, dass die Reaktion erfolgt, bevor Information über den Stimulus in die andere Hemisphäre gelangt, während die Einheiten, die eine pragmatische Dimension aufweisen, größer als einzelne Wörter oder Sätze sind und deshalb auch längere Präsentations- und Verarbeitungszeiten nötig machen. Trotz zahlreicher Untersuchungen bleibt noch genau zu klären, welche sprachlichen und kommunikativen Fähigkeiten dem Bereich 'pragmatische Sprachfähigkeiten' zuzuordnen sind und in Bezug auf welche dieser Fähigkeiten sich Defizite für bestimmte Patientengruppen nachweisen lassen.

*Nicht-propositionale Sprache.*¹⁹⁹ Bei vielen Aphasikern bleiben trotz ausgedehnter Schädigung der linken Hemisphäre lexikalische und nicht-lexikalische Sprachautomatismen oder rekurrente Äußerungen erhalten (CODE 1997, CODE 1987:63ff). Hierunter fallen z.B. idiomatische Wendungen wie *Have a nice day*, *Good to see you*, Flü-

¹⁹⁸ Für eine Diskussion dieser Fragen vgl. STEMMER & JOANETTE 1997.

¹⁹⁹ 'Nichtpropositionale Sprache' soll hier verstanden werden als 'nicht auf eine Proposition bezogene Äußerungen', d.h. nicht-referentielle und nicht-prädizierende Äußerungen.

che wie *Sacre nom de Dieu*²⁰⁰ oder nichtlexikalische Automatismen wie [tã].²⁰¹ CODE (1997) vertritt daher die These, dass die rechte Hemisphäre auch im intakten Gehirn die Produktion nicht-propositionaler Äußerungen kontrolliert,²⁰² während die linke Hemisphäre für die Produktion der propositionalen Sprache zuständig ist. Hierzu passt auch die Beobachtung, dass Patienten mit linkshemisphärischer Schädigung und weitreichenden sprachlichen Defiziten dennoch in der Lage sein können, Liedtexte singend fehlerfrei zu reproduzieren (CODE 1987:93). Charakteristisch für Sprachautomatismen und rekurrente Äußerungen ist eine intakte Silbenstruktur bzw. Phonotaktik, was auf die Sprachproduktion von hirngeschädigten Patienten allgemein zutrifft (DRESSLER 1985, zit. nach STARK & STARK 1990). Diese Beobachtung kann so interpretiert werden, dass eine Repräsentation der *phonologischen Sonoritätsbeziehungen* zwischen einzelnen Segmenten oder Oberklassen auch nach linkshemisphärischer Schädigung erhalten bleibt (BUCKINGHAM 1990). Hierzu ist jedoch zu bemerken, dass die Abwesenheit phonotaktischer Fehler auch auf der Intaktheit der (möglicherweise bilateral repräsentierten) Artikulationsroutinen von Silbendimension²⁰³ beruhen kann und nicht unbedingt nur auf intakte Sonoritätsbeziehungen oder Silbifizierungsroutinen auf phonologischer Ebene zurückzuführen ist.²⁰⁴ Abgesehen von der intakten Silbenstruktur lässt sich hinzufügen, dass Sprachautomatismen und rekurrente Äußerungen häufig eine reduzierte artikulatorische Komplexität aufweisen, insofern das Verhältnis Vokale / Konsonanten erhöht ist (d.h. mehr Vokale, weniger Konsonanten) und sich unter den Konsonanten mehr Plosivkonsonanten und Nasale und weniger Frikative finden (CODE 1997).

200 Äußerung eines Patienten von Aubertin, zit. nach CODE 1997:44.

201 Rekurrente Äußerung von Leborgne, einem Patienten von Paul Broca, zit. nach CODE 1997:44.

202 Dies wird plausibler, wenn man hinzufügt, dass rekurrente Äußerungen in der normalen Sprache häufig emotionalen Gehalt (z.B. Flüche) oder pragmatische Funktion (z.B. Verlegenheitsäußerungen, Gliederungssignale) haben.

203 Zu der Annahme, dass der Prozess der Produktion von Sprache auch einen Verarbeitungsschritt der Projektion phonologisch-phonetischer Repräsentationen auf Artikulationsroutinen von Silbendimension umfasst, vgl. LEVELT & WHEELDON 1994.

204 Zu der Frage, ob sich aus der Beobachtung, dass die Äußerungen von Patienten mit linkshemisphärischer Schädigung eine intakte Silbifizierung aufweisen, Aussagen über die Repräsentation von Silbifizierungsroutinen in der rechten oder linken Hemisphäre ableiten lassen, vgl. CODE & BALL 1994, SUSSMANN 1984.

5.3 Theorien zur Hemisphärenspezialisierung

5.3.1 Einleitung

Wie lassen sich die oben beschriebenen Beobachtungen zur funktionalen Lateralisierung von sprachlichen und nichtsprachlichen kognitiven Funktionen zusammenfassen und auf ein allgemeines Organisationsprinzip des Gehirns zurückführen? Eine Annahme, die aus diesen Beobachtungen abgeleitet werden könnte, geht davon aus, dass die linke Hemisphäre für die Kontrolle des Sprachverhaltens und die rechte Hemisphäre für die Kontrolle von visuospatialen Fähigkeiten und emotionalem Verhalten spezialisiert ist. Diese Annahme impliziert, dass die eine Hemisphäre über ein bestimmtes funktionales System, z.B. über die Sprache, verfügt, welches die andere Hemisphäre nicht besitzt.

Funktionale Lateralisierung kann jedoch auch unter einem anderen Gesichtspunkt betrachtet werden. Es könnte auch die These formuliert werden, dass komplexe kognitive Fähigkeiten (wie Sprache, Raumwahrnehmung und emotionales Erleben) auf weniger komplexen Fähigkeiten basieren. Unter dieser These könnte erwartet werden, dass die Lateralisierung komplexer Fähigkeiten auf die Lateralisierung weniger komplexer Fähigkeiten zurückgeht und dass es möglicherweise elementare Verarbeitungsfunktionen gibt, in denen sich die beiden Hemisphären unterscheiden. Diese Annahme wird durch die Beobachtung unterstützt, dass auch einige Tierarten Hemisphärenspezialisierungen aufweisen, die der linkshemisphärischen Spezialisierung für Sprache beim Menschen vergleichbar sind (für einen Überblick vgl. HELLIGE 1993), z.B. eine Spezialisierung der linken Hemisphäre für die Gesangsproduktion bei Vögeln (NOTTEBOHM 1984²⁰⁵), für die Vokalisationsdiskrimination bei Makaken (HEFFNER & HEFFNER 1984²⁰⁶) und für die Diskrimination von Formanttransitionen²⁰⁷ bei Ratten (FITCH, BROWN, O'CONNOR & TALLAL 1993). Die Annahme, dass sich die Hemisphären in ihren Verarbeitungsprinzipien unterscheiden, bietet zudem einen theoretischen Rahmen, in welchen Thesen und Beobachtungen zur funktionalen Lateralisierung einzelner Komponenten der Sprachfähigkeit eingeordnet werden können, wie z.B. auch die These einer rechtshemisphärischen Spezialisierung für die Verarbeitung der Sonoritätskontur

205 NOTTEBOHM, F. (1984): „Learning, forgetting, and brain repair“, in: Geschwind, N. & Galaburda, A.M. (Hrsg.): *Cerebral Dominance. The Biological Foundation*. Cambridge, Mass.: Harvard University Press, zitiert nach Springer & Deutsch 1995:239.

206 HEFFNER, H.E. & HEFFNER, R.S. (1984): „Temporal lobe lesions and perception of species specific vocalization by macaques“, in: *Science* 226, 75-76, zit. nach Kolb & Wishaw 1993:207.

207 Formanttransitionen sind die Frequenzänderungen, die bei der Produktion einer Konsonant-Vokal-Silbe entstehen, wenn der Vokaltrakt von der Konfiguration, mit welcher der konsonantische Verschluss gebildet wird, in die vokalische Konfiguration übergeht.

einer Äußerung. Aus diesem Grunde sollen im folgenden einige Theorien zu den Unterschieden in den Verarbeitungsprinzipien der Hemisphären im Hinblick auf die Sprachfähigkeit skizziert werden.

5.3.2 Hypothese einer ‘funktionalen’ Lateralisierung

Unerklärt bleibt im Rahmen der Dichotomie *verbale vs. visuospatiale Fähigkeiten* die Tatsache, dass sich für einige prosodische und pragmatische Sprachfunktionen eine Spezialisierung der rechten Hemisphäre findet. Um diese Daten zu integrieren, ist die Hypothese einer ‘funktionalen Lateralisierung’ der Sprachfähigkeit (ZURIF & MENDELSON 1972, EMMOREY 1987, SHIPLEY-BROWN ET AL. 1988, MOEN 1993, PERKINS ET AL. 1996) formuliert worden. Linguistisch-funktionale Aspekte der Sprachfähigkeit sind dieser Hypothese zufolge linkshemisphärisch und linguistisch-nichtfunktionale Aspekte rechtshemisphärisch lateralisiert. Linguistisch-funktional wird dabei als mehr oder weniger gleichbedeutend mit *bedeutungsunterscheidend* oder *distinktiv* verstanden. Der linken Hemisphäre wird unter einer solchen Sichtweise eine spezifisch linguistische Art der Informationsverarbeitung zugesprochen, über welche die rechte Hemisphäre nicht verfügt.²⁰⁸

Diese Hypothese lässt erwarten, dass die Verarbeitung einer phonetischen Eigenschaft wie Tonhöhe, die in der einen Sprache Wörter voneinander unterscheidet, in der anderen aber nicht, bei Sprechern der einen Sprache eher linkshemisphärisch lateralisiert und bei Sprechern der anderen Sprache eher rechtshemisphärisch lateralisiert ist. Ein solches Ergebnis fand sich z.B. für die Verarbeitung von Wortpaaren des Thai, einer Tonsprache, die sich nur in der Tonhöhenkontur unterscheiden (VAN LANCKER & FROMKIN 1972, 1978). Sprecher des Thai zeigten für die Diskrimination dieser Wörter eine linkshemisphärische Spezialisierung, Sprecher des Englischen aber wiesen für die Diskrimination keinen Ohr-Vorteil auf. Die Hypothese einer funktionalen Lateralisierung erweist sich im Kontext der Lateralisierung prosodischer Sprachfunktionen als nützlich, um die zum Teil sehr inkonsistenten Ergebnisse zur Verarbeitung von lexikalischer, syntaktischer und pragmatischer Prosodie in eine Theorie zu integrieren und soll daher später wiederaufgegriffen werden (s.u. 5.4.4). Zu bedenken ist jedoch auch, dass diese Theorie nur eine Bedeutung im Hinblick auf das Sprachsystem, aber keine Relevanz für eine allgemeine Theorie der Hemisphärenspezialisierung im Hinblick auf andere kognitive Fähigkeiten besitzt.

²⁰⁸ Daher könnte man diese Hypothese auch als eine Version der Annahme einer absoluten Funktionsspezialisierung betrachten.

5.3.3 Die linke Hemisphäre ist spezialisiert für die Verarbeitung von akustischer Information in schneller Folge

Für die Wahrnehmung von Konsonant-Vokal-Silben, die sich nur im initialen Konsonanten unterscheiden (*ba, da, ga*), findet sich in dichotischen Identifikations-Experimenten ein sehr stabiler Rechts-Ohr-Vorteil (ASBJØRNSSEN & BRYDEN 1996, HUGDAHL 1995, JÄNCKE 1993, STUDDERT-KENNEDY & SHANKWEILER 1970), der auf eine linkshemisphärische Spezialisierung für die Identifikation von Konsonanten bzw. für die Verarbeitung der relevanten akustischen Merkmale des Signals (d.h. der Formanttransition, die als Konsonant wahrgenommen wird) schließen lässt. SCHWARTZ & TALLAL (1980) konnten demonstrieren, dass ein solcher Rechts-Ohr-Vorteil, der sich für Stimuli mit Formanttransitionen von 40 ms Dauer messen lässt, signifikant geringer ist für dieselben Stimuli, wenn die Formanttransition sich über eine Dauer von 80 ms erstreckt, also halb so schnell verläuft. Hieraus schlossen SCHWARTZ & TALLAL (1980), dass der Rechts-Ohr-Vorteil für Sprache von der Präsenz schneller Veränderungen im akustischen Signal abhängt und dass die linke Hemisphäre spezialisiert ist für die Verarbeitung von schnell aufeinanderfolgenden akustischen Ereignissen (TALLAL, MILLER & FITCH 1993).

Diese These wird auch durch andere Erkenntnisse gestützt. TALLAL & NEWCOMBE (1978) konnten zeigen, dass Patienten mit linkshemisphärischer Läsion im Vergleich zu einer Kontrollgruppe ein Defizit für die Verarbeitung²⁰⁹ von nicht-sprachlichen auditorischen Stimuli in schneller Folge aufweisen, rechtshemisphärisch geschädigte Patienten jedoch nicht. ROBIN, TRANEL & DAMASIO (1990) fanden heraus, dass linkshemisphärisch geschädigte Patienten (aber nicht rechtshemisphärisch geschädigte) bei auditiven Verarbeitungstests, welche eine feine zeitliche Auflösung erfordern,²¹⁰ schlechter abschnitten, rechtshemisphärisch geschädigte Patienten (nicht aber linkshemisphärisch geschädigte) hingegen bei Verarbeitungstests, welche die Verarbeitung spektraler Information erforderten.²¹¹ Ebenso wurde festgestellt, dass Kinder mit *specific language impairment* (SLI),²¹² d.h. mit sprachlichen Defiziten, aber durch-

209 Aufgaben: Diskrimination und Identifikation von Zwei-Ton-Folgen mit sehr kurzem Inter-Stimulus-Intervall.

210 Aufgaben: Lücken-Detektion, Rhythmusidentifikation.

211 Aufgaben: Diskrimination der Frequenz zweier Töne und Tonhöhen-Zuordnung.

212 TALLAL ET AL. (1993:28) fassten darunter die folgende Patientengruppe: „LI refers to children who were developing normally in every respect, but failed to develop language at the expected rate. Criteria for inclusion as an LI subject in our research studies began with exclusion of all children with sensory hearing loss, general mental retardation, paralysis or lack of sensation in the oral musculature, or frank neurological or psychiatric disorders. [...] potential subjects had to demonstrate a non-verbal performance IQ of 85 or above, and a significant discrepancy between both their chronologi-

schnittlicher Begabung in anderen kognitiven Fähigkeiten, gleichermaßen Defizite in der Diskrimination von in schneller Folge präsentierten Tonsequenzen wie in der Diskrimination von schnellen natürlichsprachlichen CV-Formanttransitionen zeigen (TALLAL, MILLER & FITCH 1993). Gleichzeitig zeigen Kinder mit SLI nicht die übliche anatomische Asymmetrie²¹³ im linken parietalen (Planum Temporale) und rechten frontalen Bereich (Gyri des Operculum) (JERNIGAN, HESSELINK & TALLAL 1991,²¹⁴ zit. nach TALLAL, MILLER & FITCH 1993:43). Werden diesen Kindern aber dieselben Stimuli verlangsamt dargeboten, unterscheidet sich die Diskriminationsleistung der sprachgestörten Kinder nicht von der Kontrollgruppe (MERZENICH ET AL. 1996).

Diese Beobachtungen sprechen für die Annahme, dass die linke Hemisphäre spezialisiert ist für die Verarbeitung von schnell aufeinanderfolgenden akustischen Ereignissen. Aber was ist der komplementäre Teil dieser Hypothese? Tentativ könnte man annehmen, dass die rechte Hemisphäre spezialisiert ist für die Verarbeitung von sich langsam ändernden akustischen Signalen oder für die Verarbeitung von gleichzeitig einkommender akustischer Information, d.h. für die Verarbeitung des Spektrums oder „Timbre“ eines Reizes zu einem Zeitpunkt.

5.3.4 Globale und lokale Informationsverarbeitung

Präsentiert man links- bzw. rechtshemisphärisch geschädigten Patienten hierarchisch strukturierte Stimuli, z.B. eine Abbildung des Buchstaben *H*, der aus vielen Buchstaben *Z* gebildet wird, so zeigt sich, dass die Patientengruppe mit rechtshemisphärischer Läsion beim Abzeichnen der Abbildung eher lokale Information reproduziert, d.h. viele *Z*s zeichnet, diese aber nicht korrekt zu einem *H* anordnet, während die Patientengruppe mit linkshemisphärischer Schädigung eher die globale Information reproduziert, aber nicht die lokale, d.h. ein großes *H* zeichnet, aber die *Z*s, aus denen es besteht, weglässt

cal and mental age in receptive and/or expressive language developments.“ Für einen Überblick zum SLI vgl. auch FLETCHER & INGHAM 1995.

²¹³ Die im Zusammenhang mit der Sprache wichtigste anatomische Asymmetrie der Hemisphären betrifft das Planum Temporale, d.h. das posterior vom auditorischen Kortex gelegene Gebiet, in welchem Teile des Wernicke'schen Sprachzentrums liegen (ZILLES & REHKÄMPER 1994:226). Das Planum Temporale der sprachdominanten linken Hemisphäre ist bei einer Mehrheit der Menschen größer als das Planum Temporale der rechten Hemisphäre (GESCHWIND & LEVITSKY 1968). Andererseits ist der primäre auditorische Kortex, der in den Heschl'schen Querwindungen liegt, auf der rechten Seite größer als auf der linken, ebenso wie die Oberfläche des frontalen Operculums (des Broca'schen Zentrums), während aber die in den Sulci befindliche Fläche links größer ist (KOLB & WHISHAW 1993:164).

²¹⁴ JERNIGAN, T., HESSELINK, J. & TALLAL, P. (1991): „Cerebral structure on magnetic resonance imaging in language-learning impaired children“, *Archives of Neurology* 48, 539-545.

(DELIS, ROBERTSON & EFRON 1986). Diese Beobachtung kann so interpretiert werden, dass die linke Hemisphäre für die Verarbeitung von lokaler Information spezialisiert ist und die rechte Hemisphäre für die Verarbeitung von globaler Information.

Die hierarchische Struktur akustischer Reize, d.h. ihre Komposition aus Struktureinheiten auf verschiedenen Ebenen, die sich global bzw. lokal zueinander verhalten, ist zunächst weniger offensichtlich als die Schichtung visueller Reize in globale und lokale Merkmale. Ein Grund hierfür ist die Tatsache, dass die hierarchische Struktur akustischer Reize sich erst in der Zeit entfaltet, während die hierarchische Struktur visueller Signale unmittelbar gegeben ist. Besonders deutlich wird die Schichtung akustischer Stimuli dort, wo die *Struktur der Teile*, nicht die *Funktion als Ganzes*, in den Vordergrund tritt, nämlich in der Musik und in der Sprache.

Im Hinblick auf die Sprachfähigkeit könnte das Konzept *Verarbeitung von globalen Merkmalen* beispielsweise auf die Verarbeitung (i) der Gestalt von Texten bezogen werden, d.h. auf *pragmatische* Sprachfunktionen, (ii) auf die Gestalt von akustischen Signalen, die ganzen Wörtern oder Sätzen entsprechen, d.h. auf *prosodische* Sprachfunktionen, oder z.B. (iii) auf die Gestalt von Phonemfolgen, d.h. auf *sonoritätsbezogene* Sprachfunktionen. Das Konzept *Verarbeitung von lokalen Merkmalen* hingegen könnte z.B. auf die Verarbeitung der (i) Bestandteile von Texten bezogen werden, d.h. auf die *syntaktische* und *lexikalische* Verarbeitung, (ii) auf die Bestandteile von Wörtern, d.h. auf die Wahrnehmung von *Phonemen*, oder beispielsweise (iii) auf die Bestandteile von Einzellauten, d.h. die Wahrnehmung von *distinktiven Merkmalen*.

Problematisch bleibt jedoch die Relativität der Begriffe *global* und *lokal*, deren Bedeutung nur im Kontext einer experimentellen Aufgabe klar definiert werden kann. Bei der Sprachverarbeitung unter natürlichen Bedingungen hingegen wird Information auf vielen verschiedenen Ebenen der Hierarchie gleichzeitig verarbeitet und *jede dieser Ebenen verhält sich global zu der einen Ebene und lokal zu einer anderen*. Daher ist diese Theorie zur Hemisphärenspezialisierung auf die funktionale Lateralisierung der Sprache nur bedingt anwendbar.

5.3.5 Die Ortsfrequenz-Hypothese

In einem gewissen Zusammenhang mit der Global-Lokal-Hypothese steht die Annahme, dass, zumindest für die visuelle Modalität, die rechte Hemisphäre spezialisiert ist für die Verarbeitung von Stimuluskomponenten mit niedrigen Ortsfrequenzen und die linke Hemisphäre für die Verarbeitung von Stimuluskomponenten mit hohen Ortsfrequenzen (vgl. HELBIGE 1993:79ff, SPRINGER & DEUTSCH 1995:93). IVRY & LEBBY

(1993) konnten zeigen, dass ein vergleichbarer Effekt auch in der auditorischen Verarbeitung existiert (vgl. auch IVRY & LEBBY 1997): für die Diskrimination eines Stimulus mit einer höheren Frequenz als die durchschnittliche Frequenz der Vergleichsstimuli fanden sie einen Rechts-Ohr-Vorteil bzw. einen Verarbeitungsvorteil für die linke Hemisphäre, für die Diskrimination eines Stimulus mit einer niedrigeren Frequenz hingegen fanden sie einen Links-Ohr-Vorteil bzw. einen Verarbeitungsvorteil für die rechte Hemisphäre. Ebenso wie die Annahme einer Spezialisierung der linken Hemisphäre für die Verarbeitung von akustischen Ereignissen in schneller Folge ist die Ortsfrequenz-Hypothese zwar auf Erkenntnisse zur Lateralisierung der Sprachwahrnehmung und der nichtsprachlichen auditiven Verarbeitung anwendbar, berührt aber nicht die Lateralisierung der Sprachfunktionen auf höheren Ebenen, wie z.B. eine mögliche rechtshemisphärische Spezialisierung für pragmatische Sprachfunktionen.

5.3.6 Zusammenfassung

Die drei letztgenannten Theorien, die annehmen, dass die linke Hemisphäre spezialisiert ist für (i) die Verarbeitung von akustischer Information in schneller Folge, (ii) für die lokale Informationsverarbeitung oder (iii) für die Verarbeitung von Stimuli mit hohen Ortsfrequenzen, lassen sich im Hinblick auf die auditorische und die Sprachwahrnehmung wie folgt zusammenfassen: Die linke Hemisphäre scheint für die Verarbeitung von akustischen Ereignissen in schneller Folge spezialisiert, d.h. für die Verarbeitung von Stimulusanteilen mit hoher Frequenz, bzw. für die Verarbeitung von lokalen Merkmalen des Gesamtsignals, während die rechte Hemisphäre möglicherweise für die Verarbeitung von langsameren Änderungen eines akustischen Ereignisses spezialisiert ist, d.h. für die Verarbeitung niedrigerer Frequenzen, bzw. für die Verarbeitung von globalen Merkmalen des Gesamtsignals.

In den vorangegangenen Ausführungen wurden Beobachtungen zur funktionalen Lateralisierung sprachlicher und nichtsprachlicher kognitiver Funktionen zusammengefasst. Da es in dieser Studie um die Sprachwahrnehmung geht, wird im nächsten Abschnitt der Blickwinkel weiter eingeschränkt und es werden die bisherigen Erkenntnisse zur funktionalen Lateralisierung der Wahrnehmung von Sprache betrachtet. Im Anschluss daran wird die Diskussion zu den Theorien der Hemisphärenspezialisierung wieder aufgenommen und es wird gezeigt, dass sowohl die Annahme einer 'funktionalen' Lateralisierung der Sprache als auch die Annahme einer Spezialisierung der linken Hemisphäre für die Verarbeitung von akustischer Information in schneller Folge geeignet ist, die Beobachtungen zu Lateralitäten in der Sprachwahrnehmung zu erfassen. Aus

diesen beiden Annahmen wird schließlich auch die Hypothese dieser Arbeit abgeleitet, dass sich für die Wahrnehmung der Sonorität eine rechtshemisphärische Spezialisierung findet.

5.4 Lateralitätseffekte in der Sprachwahrnehmung

5.4.1 Einleitung

Ein großer Teil der Erkenntnisse zur funktionalen Lateralisierung der Sprachwahrnehmung sind mit der Methode des dichotischen Hörens gewonnen worden (für einen Überblick vgl. HUGDAHL (Hrsg.) 1988). Beim dichotischen Hören werden dem Probanden zwei verschiedene Sprachsignale gleichzeitig präsentiert, je eines auf jedem Ohr (vgl. Abbildung 6-1). Da die Hörbahn gekreuzt ist und unter beidohriger Stimulation die Signale vom rechten Ohr nur über die kontralaterale Verbindung in die linke Hemisphäre und vom linken Ohr nur in die rechte Hemisphäre projiziert werden (KIMURA 1961, MILNER, TAYLOR & SPERRY 1968), lässt sich aus einem Verarbeitungsvorteil für das eine oder das andere Ohr, d.h. aus einer schnelleren oder korrekteren Verarbeitung der linksohrig präsentierten Reize im Vergleich zu den rechtsohrig präsentierten Reizen oder umgekehrt, auf die Spezialisierung einer Hemisphäre im Vergleich zur anderen schließen.

Die Methode des dichotischen Hörens geht auf BROADBENT (1954)²¹⁵ zurück, der, um Aufmerksamkeitsprozesse in der auditorischen Verarbeitung zu untersuchen, den Probanden zwei verschiedene akustische Stimuli (Zahlwörter) gleichzeitig darbot, je ein Wort auf jedes Ohr, so dass das Sprachverarbeitungssystem mehr Information erhält, als es verarbeiten kann. Sind die zwei dargebotenen Reize nicht sehr ähnlich, wie z.B. die Zahlwörter *eins* und *fünf*, so werden beide Wörter wahrgenommen (BRYDEN 1988). Wenn beide Wörter sehr ähnlich sind, wie z.B. *Bank* und *lang*, nimmt der Proband oft nur ein einziges Wort wahr. Dabei kann ein Reiz den anderen *dominieren*, so dass nur das Wort *Bank*, aber nicht das gleichzeitig präsentierte Wort *lang* wahrgenommen wird (WEXLER & HALWES 1983),²¹⁶ oder Bestandteile beider Wörter fusio-

²¹⁵ BROADBENT, D.E.: „The role of auditory localization in attention and memory“, in: *Journal of Experimental Psychology* 47, 191-196, zit. nach CODE 1987:25.

²¹⁶ Dieses Phänomen machen sich viele dichotische Tests zunutze, besonders jene, die CV-Silben verwenden, da die Dominanz des einen Perzepts über das andere dazu führt, dass die Verarbeitungsunterschiede zwischen den beiden Ohren bzw. Hemisphären besonders deutlich hervortreten (dieses Prinzip wurde erstmals im „fused dichotic test“ von WEXLER & HALWES 1983 genutzt). Das Phänomen Stimulusdominanz ist aber gleichzeitig ein Störfaktor in dichotischen Tests und zwar dann, wenn die Dominanz des einen Perzepts über das andere nicht durch die Unterschiede in der Verar-

nieren und werden als ein einziges Wort wahrgenommen (*Überblendung*), so dass z.B. *Bank* und *lang* als *blank* wahrgenommen wird (MORAIS, KOLINSKY & NAKAMURA 1996).

Die für die Lateralitätsforschung interessante Frage besteht darin, welche Stimuluskomponenten effizienter verarbeitet werden: die dem linken Ohr präsentierten oder die dem rechten Ohr dargebotenen. Im allgemeinen findet sich ein Rechts-Ohr-Vorteil (im folgenden REA, *right ear advantage*), d.h. eine Spezialisierung der linken Hemisphäre, für die Verarbeitung von sprachlicher akustischer Information (KIMURA 1961) und ein Links-Ohr-Vorteil (im folgenden LEA, *left ear advantage*) für die Verarbeitung von nicht-sprachlicher akustischer Information wie Musik (KIMURA 1964) oder Umweltgeräuschen. Hängt dieser REA für Sprache von der linguistischen Funktion der Reize ab oder von der Präsenz bestimmter phonetischer oder akustischer Eigenschaften im Sprachsignal? Um diese Frage zu beantworten, werden im folgenden experimentelle Daten zur Lateralisierung der Verarbeitung von segmentalen und suprasegmentalen Kategorien und Merkmalen der Sprache zusammengefasst, die in dichotischen Studien und in klinischen Studien mit links- und rechtshemisphärisch geschädigten Patienten erhoben wurden.

5.4.2 Lateralitätseffekte in der Verarbeitung von segmentalen Merkmalen

SHANKWEILER & STUDDERT-KENNEDY (1967) und STUDDERT-KENNEDY & SHANKWEILER (1971) konnten einen REA für die Verarbeitung der initialen und finalen Plosiv-Konsonanten (/p/, /t/, /k/, /b/, /d/, /g/) von sinnlosen Konsonant-Vokal-Silben (wie *ba*, *da*, *ga*, *pa*, *ta*, *ka*) nachweisen,²¹⁷ fanden jedoch keinen Ohr-Vorteil für die Wahrnehmung von Vokalen. Ohr-Vorteile für die Wahrnehmung von Vokalen lassen sich nur dann zeigen, wenn die experimentelle Aufgabe in der einen oder anderen Weise manipuliert wird: wenn die Teststimuli von unterschiedlichen Vokaltrakten erzeugt wurden (DARWIN 1971), wenn sie von zwei unterschiedlich hohen Stimmen gesprochen wurden (HAGGARD & PARKINSON 1971), wenn sie in weißem Rauschen präsentiert wurden (WHITE & HOUSE 1973) oder wenn die Vokal-Erkennungs-Aufgabe in

beitungseffizienz beider Hemisphären verursacht wird, sondern durch phonologische oder phonetische Eigenschaften der Stimuli, was dazu führt, dass der Proband unabhängig vom Präsentationsohr stets nur das eine, aber nie das andere Wort wahrnimmt.

²¹⁷ Vgl. auch ASBJØRNSEN & BRYDEN 1996, ASBJØRNSEN & HUGDAHL 1995, JÄNCKE 1992a, JÄNCKE 1992b, JÄNCKE 1993, JÄNCKE & LUPPEN 1993, HUGDAHL 1995, WEXLER & HALWES 1983.

eine Konsonanten-Erkennungs-Aufgabe eingebettet wurde (SPELLACY & BLUMSTEIN 1970).

Außer für Plosiv-Konsonanten konnte ein REA auch für Liquid-Konsonanten (/l/, /r/) (CUTTING 1974), für Frikative²¹⁸ (/s/, /z/, /f/, /v/) (DARWIN 1971) und für Gleitlaute (/w/, /j/)²¹⁹ (HAGGARD & PARKINSON 1971) demonstriert werden. HAYDEN (1979) wies außerdem nach, dass die Größe des REA mit der Lautklasse des untersuchten Kontrastes signifikant variiert.²²⁰ Der größte Ohr-Vorteil lässt sich für Plosiv-Konsonanten zeigen, der kleinste für Frikative und für Gleitlaute. Basieren die unterschiedlichen Lateralitätseffekte für verschiedene Konsonantenklassen auf der Tatsache, dass diese Laute durch verschiedene Arten von akustischer Information kodiert werden? Um diese Frage beantworten zu können, wird im folgenden kurz zusammengefasst, auf Grund welcher akustischer Eigenschaften die verschiedenen Konsonantenklassen identifiziert werden.²²¹ Plosivkonsonanten können nur auf Grund der Formanttransition identifiziert werden, welche resultiert, wenn der Vokaltrakt vom konsonantischen Verschluss in die vokalische Öffnung übergeht;²²² ähnliches gilt für Nasale und Liquide, diese werden aber zusätzlich durch die Formantstruktur, d.h. durch die relative Frequenz der Formanten²²³ kodiert; Frikative können sowohl auf Grund des Frequenzbereiches der Friktion, d.h. des Rauschens, welches durch die Engebildung im Vokaltrakt bei der Artikulation von Frikativkonsonanten entsteht, als auch auf Grund des Verlaufs der Formanttransitionen identifiziert werden; Gleitlaute schließlich werden durch die relative Frequenz der Formanten und ihre sehr langsame Transition kodiert. HAYDENS Ergebnisse stützen daher die Annahme, dass der REA für Sprachlaute auf der Verarbeitung von Formanttransitionen beruht: je wesentlicher die Verarbeitung der Formanttransitionen für die Lauterkennung und je schneller der Verlauf der Transitionen, desto

218 Aber nur, wenn die Frikative akustisch durch Formanttransitionen kodiert waren, nicht aber, wenn sie durch den Frequenzbereich der Friktion kodiert waren. Zu den akustischen *cues* für die Phonemerkennung vgl. auch LADEFOGED 1993³:203ff, PICKETT 1985²:169ff, zu derartigen *trading relations* zwischen zwei oder mehr akustischen *cues* vgl. REPP 1982.

219 Die Tatsache, dass sich für Gleitlaute, aber nicht für Vokale ein REA demonstrieren lässt, kann als weiterer Hinweis darauf gelten, dass sich Gleitlaute von ihren vokalischen Entsprechungen auf für die Sprachwahrnehmung relevante Weise unterscheiden, vgl. hierzu die Diskussion in Kapitel 2.

220 HAYDEN (1979) konnte die folgenden signifikanten Unterschiede zeigen: (i) der REA für Frikative ist kleiner als der für Affrikata, der wiederum kleiner als der für Liquide ist und (ii) der REA für Gleitlaute ist kleiner als der für Nasale und dieser kleiner als der für Plosive. Ein signifikanter REA konnte nur für Plosive und Nasale demonstriert werden.

221 Vgl. LADEFOGED 1993³:203ff, PICKETT 1985²:169ff.

222 Die Bewegung vom konsonantischen Verschluss zur vokalischen Öffnung geht mit einer Änderung der Form des Vokaltraktes, d.h. seiner Resonanzeigenschaften einher. Diese Änderung der Resonanzeigenschaften wiederum führt zu einer Änderung (Transition) der Frequenz der Formanten.

223 Formanten sind die Frequenzbereiche des Spektrums, in denen sich Signalanteile mit besonders hoher Energie konzentrieren.

größer der REA. Diese Interpretation ist konsistent mit den Ergebnissen von SCHWARZ & TALLAL (1980), die nachwiesen, dass der REA für CV-Silben von der Geschwindigkeit der Formanttransition abhängt (s.o. Abschnitt 5.3.3).

Mit dieser Annahme lässt sich auch erklären, warum unter gewissen Bedingungen ein REA für Vokale nachweisbar ist: nämlich dann, wenn die Vokale nicht allein auf Grund der Verarbeitung der statischen Formantstruktur, d.h. des Spektrums zu einem Zeitpunkt, identifiziert werden können: Erstens dann, wenn die Stimuli von verschiedenen Sprechern, d.h. durch unterschiedliche Resonanzräume, erzeugt wurden, so dass z.B. der Vokal /a/ in verschiedenen Stimuli durch ganz unterschiedliche spektrale Muster repräsentiert wird, die nur in der relativen Lage der Formanten zueinander übereinstimmen, die sich aber in der absoluten Frequenz der Formanten unterscheiden; und zweitens wenn der Stimulus durch Rauschen überdeckt wird, so dass die Formantstruktur des Signals weniger prominent ist. In diesen Fällen muss der Sprecher zusätzliche akustische *cues* auswerten, d.h. zusätzliche oder andere Verarbeitungsstrategien anwenden, um den Vokal zu identifizieren. Der REA für die Vokalwahrnehmung wird unter diesen experimentellen Bedingungen dadurch hervorgerufen, dass auch die dynamischen Formanttransitionen verarbeitet werden, die beim Übergang von der konsonantischen zur vokalischen Vokaltraktkonfiguration erzeugt werden und die aus diesem Grund stets sowohl für die konsonantische wie auch für die vokalische Konfiguration charakteristisch sind.

SPELLACY & BLUMSTEIN (1970) demonstrierten einen REA für die Erkennung von Vokalen, indem sie die Vokal-Identifikations-Aufgabe in eine Konsonanten-Identifikations-Aufgabe²²⁴ einbetteten. Während (Plosiv-)Konsonanten nur auf Grund der Formanttransitionen zu identifizieren sind, können die Vokale sowohl auf Grund der Lage der Formanten als auch auf Grund der Formanttransitionen erkannt werden. Unter diesen experimentellen Bedingungen zeigt sich für die Vokalerkennung, ebenso wie für die Konsonantenerkennung, ein REA. Dieser REA lässt sich durch die Annahme erklären, dass die Probanden die Konsonanten-Identifikations-Strategie generalisierend auf alle Stimuli anwenden, auf die sie anwendbar ist, also auch auf die Vokale. In derselben Studie konnten SPELLACY & BLUMSTEIN zeigen, dass die Einbettung des Phonem-Erkennungs-Tests in einen Erkennungs-Test für nichtsprachliche Stimuli (Melodien, Umweltgeräusche) zu einem LEA für die Erkennung der Vokale²²⁵ führt, während der REA für die Erkennung der Konsonanten²²⁶ nicht beeinflusst wird. Das Fehlen eines Ohr-Vorteils für Vokale unter normalen Bedingungen der Sprachverar-

224 Als Stimuli wurden stimmhafte und stimmlose Plosivkonsonanten verwendet.

225 Vokale können auf Grund der spektralen Zusammensetzung des Signals identifiziert werden.

226 Konsonanten können nur durch die Formanttransitionen kodiert werden

beitung im Gegensatz zu einem nachweisbaren REA oder LEA unter bestimmten experimentellen Bedingungen kann so interpretiert werden, dass Vokale unter den normalen Bedingungen auf der Basis zweier verschiedener akustischer Dimensionen identifiziert werden, deren Verarbeitung komplementär lateralisiert ist, während experimentelle Bedingungen zu einer größeren Gewichtung der einen oder der anderen akustischen Dimension im Verarbeitungsprozess führen können. Die Ergebnisse von SPELLACY & BLUMSTEIN (1970) weisen auch darauf hin, dass die Möglichkeit, Reize auf der Basis von verschiedenen Eigenschaften zu identifizieren,²²⁷ d.h. verschiedene *Verarbeitungstrategien* anzuwenden, zu unterschiedlichen Lateralitätseffekten für dieselben Sprachreize führen kann.

Die vorangegangenen Ausführungen lassen sich folgendermaßen zusammenfassen: Erstens basiert der REA für Sprache auf einer Spezialisierung der linken Hemisphäre für die Verarbeitung von Formanttransitionen, oder allgemeiner formuliert, für die Verarbeitung von akustischen Ereignissen in schneller Folge und zweitens lässt sich der REA für Sprache auch für sinnlose Konsonant-Vokal-Silben und nicht nur für bedeutungsvolle Lautfolgen demonstrieren. Drittens können die experimentellen Bedingungen die Verarbeitungsstrategien beeinflussen, welche von den Probanden angewendet werden, was zu verschiedenen Lateralitätseffekten für identische Stimuli führen kann.

5.4.3 Lateralitätseffekte in der Verarbeitung von suprasegmentalen Merkmalen

Zum Bereich der suprasegmentalen Eigenschaften der Sprache gehören die drei Phänomene Intonation, Akzent und Ton. Auch das Merkmal Sonorität gehört, wie im zweiten Kapitel gezeigt wurde, zu den suprasegmentalen Merkmalen (s.o. 2.4.5, 2.5).²²⁸ Während jedoch die funktionale Lateralisierung von Intonation, Akzent und Ton in zahlreichen Studien untersucht wurde, hat man sich bislang nicht mit der Frage beschäftigt, ob die Verarbeitung der Sonoritätskontur einer Äußerung lateralisiert ist. Bevor im nächsten Abschnitt dieses Kapitels (s.u. 5.5) eine Hypothese zur funktionalen Lateralisierung der Sonoritätskontur formuliert wird, soll daher zunächst die funktionale Lateralisierung der drei anderen suprasegmentalen Eigenschaften Intonation, Akzent

²²⁷ Sprachliche Reize entsprechen sehr komplexen akustischen Mustern mit einer Struktur und einer Bedeutung auf verschiedenen Ebenen und können stets auf der Basis unterschiedlicher Eigenschaften identifiziert werden.

²²⁸ Genauer gesagt wurde Sonorität als Verbindungsglied zwischen suprasegmentaler und segmentaler Ebene gekennzeichnet.

und Ton betrachtet werden.

In der neuropsychologischen und neurolinguistischen Literatur wird der Gegenstandsbereich der suprasegmentalen Eigenschaften häufig mit dem Begriff *Prosodie* belegt, wobei meist zwischen *emotionaler* und *linguistischer* Prosodie unterschieden wird. Nähert man sich diesem Phänomenbereich von der phonetischen Seite, so kann man ihn in Anlehnung an LAVER (1995) mit den Begriffen *Prosodie*²²⁹ und *Metrik*²³⁰ beschreiben, wobei Prosodie die auf Tonhöhe und Lautstärke bezogenen akustischen Eigenschaften und Metrik die auf Rhythmus oder Prominenz bezogenen Eigenschaften bezeichnet. Nähert man sich dem Bereich schließlich von der linguistischen Seite, so kann man funktionale Kategorien der suprasegmentalen Organisation der Sprache verwenden, wie eben *Satzintonation*, *Satzakzent*, *Wortakzent* oder *Ton*. Die folgenden Ausführungen verwenden die linguistische Begrifflichkeit.

5.4.3.1 Satzprosodie

Sätze können sich in ihrer Intonation, d.h. in ihrem Tonhöhenverlauf, oder in ihrer Betonung, d.h. in der relativen Hervorhebung einzelner Wörter oder Phrasen, unterscheiden. Wenn sich zwei Sätze in linguistisch-funktionaler Hinsicht unterscheiden, z.B. ein Aussagesatz und ein Fragesatz, so differieren diese zwei Sätze meist sowohl auf der Ebene der Intonation als auch auf der Ebene der Satzbetonung. *Satzintonation* und *Satzbetonung* werden daher im folgenden auch unter dem Begriff *Satzprosodie* zusammengefasst. In den folgenden Ausführungen werden zunächst verschiedene linguistisch-funktionale Kontraste differenziert, die in den hier zitierten Studien zur Satzprosodie verwendet wurden. Danach werden einige dichotische und klinische Studien zur Lateralisierung der Verarbeitung der Satzprosodie zusammengefasst. Auf akustisch-phonetischer Seite entsprechen den verschiedenen funktionalen Kategorien der Satzprosodie unterschiedliche, sehr komplexe Muster der Variation von Tonhöhe, Lautstärke und Segment- und Silbendauer im Zeitverlauf (vgl. CRUTTENDEN 1986).

(1) Modale Intonation²³¹

229 Unter 'Prosodie' versteht LAVER (1995:450): „the suprasegmental texture of speech [...] the patterns of prosodic organization of pitch and loudness, [...] the variations of melody and sonority in individual utterances.“

230 Unter 'Metrik' versteht LAVER (1995:511): „the grouping, accentuation and rhythm of the delivery of its component parts [...] the metrical structure of speech in terms of the relationships between its syllables, stress and rhythm.“; „The metrical organization of speech [...] integrates facets of all four basic dimensions of speech – quality, duration, loudness and pitch.“

231 Auf der Basis der in der hier zitierten Literatur verwendeten Stimulussätze werden an dieser Stelle vier verschiedene funktionale Kontraste differenziert. Davon werden einige eher durch den reinen

Du hast also das ganze grüne Buch gelesen!	deklarativ
Du hast also das ganze grüne Buch gelesen?	interrogativ
Bis morgen wird das grüne Buch gelesen!	imperativ

Diese Art der Satzprosodie hat eine pragmatische Dimension, insofern man *deklarativ*, *interrogativ* und *imperativ* auch als Sprechaktkategorien auffassen kann. Angesichts der rechtshemisphärischen Spezialisierung für pragmatische Sprachfunktionen ist daher für die Verarbeitung dieser Stimuli eine rechtshemisphärische Beteiligung zu erwarten.

(2) *Emphatische Betonung*

Wer hat das grüne Buch gelesen? – Ich habe das grüne Buch gelesen.
 Was hast Du zuletzt gelesen? – Ich habe *das grüne Buch* gelesen.
 Hast Du das grüne Buch mal in die Hand genommen? – Ja, ich habe das grüne Buch *gelesen*.

(3) *Satzkonstituentenstrukturbezogene Intonation und Betonung*²³²

The boy said, the girl is fat.
 The boy, said the girl, is fat.

The woman said, I see the paper, boy.
 The woman said, I see the paperboy.

Zu dieser Art von Satzprosodie ist zu bemerken, dass zwischen zwei Sätzen, die sich auf diese Weise unterscheiden, nicht nur *prosodische* Unterschiede bestehen, sondern vor allem *syntaktische* Unterschiede. Die Verarbeitung solcher Sätze erfordert daher zusätzlich zu ihrer segmentalphonologischen und prosodischen Verarbeitung auch die Rekonstruktion ihrer morphosyntaktischen Struktur. Aus diesem Grund ist zu erwarten, dass die Verarbeitung dieser Stimuli auch eine links-hemisphärische Beteiligung erfordert.

(4) *Pragmatische Intonation – Fortsetzung*

Sie hat das grüne Buch gelesen (– das hat sie mir eben erzählt).
 Sie hat das grüne Buch gelesen (und auch gleich verzettelt).

Diese Art der Satzprosodie trägt Information über die Einbettung eines Satzes in einen Text und wird daher auch als *pragmatische Betonung* bezeichnet. Auch hier ist wegen der textpragmatischen Dimension des funktionalen Kontrastes eine rechtshemisphärische Beteiligung zu erwarten.

BLUMSTEIN & COOPER (1974) konnten in einer dichotischen Studie mit Stimuli vom Typ (1) ein LEA für die Verarbeitung der Satzprosodie in (unverständlichen) tiefpass-

Tonhöhenverlauf, d.h. die Intonation, kodiert, z.B. der modale Kontrast zwischen Deklarativsätzen und Interrogativsätzen. Andere werden eher durch die unterschiedliche Betonung differenziert, d.h. durch Lautstärke-, Dauer- und Tonhöhenvariation in einzelnen Phrasen, wie z.B. der emphatische Kontrast.

²³² Diese Art von Stimuli wird nur von PERKINS, BARAN & GANDOUR (1996) verwendet.

gefilterten²³³ natürlichsprachlichen Sätzen und in sinnlosen, aber ungefilterten Sätzen nachweisen. Auch SHIPLEY-BROWN ET AL. (1988) fanden ein LEA für satzprosodische Kontraste vom Typ (1) und (4). Weniger konsistent sind die Daten, die in klinischen Studien gewonnen wurden. WEINTRAUB, MESULAM & KRAMER (1981) zeigten ein Defizit bei rechtshemisphärisch geschädigten Patienten (im folgenden RHDs, *right hemisphere damaged subjects*) im Vergleich zu einer Kontrollgruppe²³⁴ mit Sätzen vom Typ (1) und (2). BAUM, KELSCH DANILOFF, DANILOFF & LEWIS (1982) hingegen fanden ein Defizit²³⁵ bei linkshemisphärisch geschädigten Patienten (im folgenden LHDs, *left hemisphere damaged subjects*) für die Diskrimination von emphatischer Betonung vom Typ (2). HEILMAN, BOWERS, SPEEDIE & COSLETT (1984) kamen zu dem Ergebnis, dass für die Identifikation von emotionaler Prosodie RHDs schlechter als LHDs und als Kontrollsubjekte abschneiden, für die Identifikation der Satzprosodie vom Typ (1) in tiefpassgefilterten Sätzen sowohl RHDs als auch LHDs schlechter als die Kontrollgruppe abschneiden und sich RHDs und LHDs nicht voneinander unterscheiden. COHEN, BRANCH & HYND (1994) fanden ein Defizit für RHDs²³⁶ im Vergleich zu LHDs und einer Kontrollgruppe für die Verarbeitung von Intonationskontrasten vom Typ (1). PERKINS, BARAN & GANDOUR (1996) konnten unterschiedliche Ergebnisse für Identifikationsaufgaben und Diskriminationsaufgaben zeigen. In einer Identifikationsaufgabe mit Stimuli vom Typ (1) zeigten LHDs ein Defizit im Vergleich zu einer Kontrollgruppe, RHDs unterschieden sich hingegen nicht von der Kontrollgruppe.²³⁷ In einer Diskriminationsaufgabe mit denselben Stimuli zeigten sich keine Unterschiede zwischen den Probandengruppen; wurden die Stimuli tiefpassgefiltert, so waren die RHDs (anders als die LHDs) schlechter als die Kontrollgruppe.²³⁸ PERKINS, BARAN & GANDOUR (1996) untersuchten die Verarbeitung der Satzprosodie mit Stimuli vom Typ (3) und demonstrierten ein Defizit für LHDs (nicht aber für die RHDs) im Vergleich zu einer Kontrollgruppe.²³⁹ In einer umfangreichen Studie zur Wahrnehmung von emotionaler Prosodie und modaler Satzprosodie vom Typ (1)²⁴⁰ schließlich kamen PELL & BAUM (1997a, b) mit verschiedenen Stimuli (sinnvolle Sätze, sinnlose Sätze aus Pseu-

233 In tiefpassgefilterten Sprachsignalen (*cutoff frequency*: typischerweise um 500 Hz) sind nur die tieferen Frequenzanteile des ursprünglichen Signals erhalten, welche als *cues* für die suprasegmentale Struktur der Äußerung verarbeitet werden; die höheren Frequenzen, die zur Erkennung einzelner Phoneme nötig sind, fehlen hingegen.

234 Die RHDs wurden aber nicht mit einer linkshemisphärisch geschädigten Kontrollgruppe verglichen.

235 Im Vergleich zur Kontrollgruppe, RHDs wurden nicht getestet.

236 Bei den Probanden handelte es sich um Kinder.

237 Jedoch unterschieden sich die RHDs nicht von den LHDs.

238 Auch hierbei unterschieden sich die RHDs nicht von den LHDs.

239 RHDs unterschieden sich nicht von den LHDs.

240 Interrogative vs. deklarative vs. imperative Betonung.

dowörtern, tiefpassgefilterte Sätze) zu dem Ergebnis, dass in einer *Diskriminationsaufgabe* weder RHDs noch LHDs schlechter als die Kontrollgruppe abschnitten,²⁴¹ und dass in einer *Identifikationsaufgabe* sowohl RHDs als auch LHDs ein Defizit gegenüber der Kontrollgruppe aufwiesen, sich aber nicht voneinander unterschieden. Wie lassen sich diese Daten interpretieren?

- Viele der Studien können ein selektives Defizit der einen Patientengruppe, aber nicht der anderen Patientengruppe gegenüber der Kontrollgruppe zeigen, kommen aber dennoch zu dem Ergebnis, dass sich RHDs und LHDs nicht signifikant voneinander unterscheiden. Diese Beobachtung lässt sich so interpretieren, dass die Komponenten der komplexen Fähigkeit *Verarbeitung der Satzprosodie* auf beide Hemisphären verteilt sind, so dass sowohl LHDs als auch RHDs beeinträchtigt sind. Eine solche Interpretation wird auch unterstützt durch die Daten von VAN LANCKER & SIDTIS (1992), die in einer Studie zur Lateralisierung der Verarbeitung von emotionaler Prosodie kein selektives Defizit von RHDs oder LHDs nachweisen konnten, die aber zeigten, dass RHDs zur Identifikation der emotionalen Prosodie eher die Tonhöhenkontur nutzen, LHDs hingegen eher die temporalen Charakteristika des Signals.
- Einige Studien kommen unter Verwendung identischer Stimuli zu unterschiedlichen Ergebnissen für Diskriminations- und Identifikationsaufgaben: in Identifikationsaufgaben zeigt sich eher ein Defizit der LHDs, in Diskriminationsaufgaben häufig überhaupt kein selektives Defizit einer Probandengruppe. Diese Ergebnisse weisen darauf hin, dass die Probanden in Identifikations- und Diskriminationsaufgaben verschiedene *Verarbeitungsstrategien* anwenden und dass die für die *Identifikation*, d.h. für die (sprachliche oder nichtsprachliche) *Kategorisierung* oder *Bezeichnung* von Sprachstimuli, erforderlichen Verarbeitungsschritte mehr als die für die bloße Diskrimination nötigen von der linken Hemisphäre abhängen.
- Für tiefpassgefilterte Stimuli zeigt sich häufig ein selektives Defizit der RHDs bzw. ein LEA. Der Grund hierfür könnte darin liegen, dass die tiefpassgefilterten Sätze nicht lexikalisch, syntaktisch und segmental-phonologisch verarbeitet werden, so dass die linkshemisphärische Beteiligung, die sich bei verständlichen Sätzen stets findet und die eine rechtshemisphärische Spezialisierung für die prosodische Verarbeitung überdecken kann,²⁴² wegfällt.

²⁴¹ LHDs und RHDs sind langsamer als Kontrollsubjekte, machen aber nicht mehr Fehler.

²⁴² Zwei der hier zitierten Studien (BLUMSTEIN & COOPER 1974, PERKINS, BARAN & GANDOUR 1996) finden keine Hemisphärenspezialisierung mit ungefilterten Stimuli, aber eine rechtshemisphärische Spezialisierung mit gefilterten Stimuli, ebenso die Studie von BEHRENS (1985) zum Wortakzent (s.u.).

- Für die Verarbeitung der (i) modalen Intonation [Typ (1)] eines Satzes finden sich vier Studien, die auf eine rechtshemisphärische Spezialisierung hinweisen und jeweils eine Studie, welche die Annahme einer linkshemisphärischen Spezialisierung bzw. einer bilateralen Repräsentation stützt. (ii) Für die Verarbeitung der emphatischen Betonung [Typ (2)] finden sich zwei Studien, die auf rechtshemisphärische und eine Studie, die auf linkshemisphärische Spezialisierung hinweist. (iii) Für die Verarbeitung der konstituentenstrukturbezogenen Intonation und Betonung [Typ (3)] findet sich eine Studie, welche die Annahme einer linkshemisphärischen Lateralisierung dieser Fähigkeit unterstützt. (iv) Für die Verarbeitung pragmatischer Intonation [Typ (4)] weist eine Studie auf rechtshemisphärische Spezialisierung hin.

Die vorangegangenen Ausführungen haben gezeigt, dass erstens verschiedene Arten der Satzprosodie unterschieden werden müssen, für die sich unterschiedliche Lateralitätseffekte demonstrieren lassen und dass zweitens die Verwendung von gefilterten oder ungefilterten Signalen und von Diskriminations- oder Identifikationsaufgaben einen Einfluss auf die experimentellen Resultate haben kann. Für die Verarbeitung von modaler Satzprosodie [Typ (1)] überwiegt die Zahl der Studien, welche die Annahme einer rechtshemisphärischen Spezialisierung stützen, zur Lateralisierung der Verarbeitung der übrigen Arten der Satzprosodie hingegen sind weitere Untersuchungen erforderlich, bevor Aussagen von allgemeinerer Gültigkeit gemacht werden können.

5.4.3.2 Wortakzent

Die Wörter in den folgenden Wortpaaren unterscheiden sich nur in der Betonung, d.h. im Wortakzent:

jemanden *um*fahren – ein Hindernis *um*fahren
*hot*dog – hot *dog*
*black*board – black *board*
*ligh*tousekeeping – light *house*keeping
object (Substantiv) – *object* (Verb)
transport (Substantiv) – *transport* (Verb)

Die akustischen *cues* für den Wortakzent sind Tonhöhe, Lautstärke und Silbendauer: betonte Silben in einem Wort sind relativ zu unbetonten lauter und länger und weisen eine höhere Grundfrequenz auf (vgl. LAYER 1995:512). Alle hier zitierten Studien zur Verarbeitung des Wortakzentes verwendeten Sprachstimuli dieser Art. In einer dichotischen Studie zur Lateralisierung der Verarbeitung des Wortakzents fand BEHRENS (1985) einen REA für die Verarbeitung ungefilterter Wörter, einen LEA für dieselben Wörter nach Tiefpassfilterung und keinen Ohr-Vorteil für ungefilterte Pseudowörter. BLUMSTEIN & GOODGLASS (1972) kamen zu dem Ergebnis, dass sich RHDs und LHDs im Hinblick auf diese Fähigkeit nicht voneinander unterscheiden, dass aber beide Gruppen ein Defizit im Vergleich zur Kontrollgruppe aufweisen. WEINTRAUB, MESULAM & KRAMER (1981) zeigten ein Defizit von RHDs im Vergleich zur Kontrollgruppe, untersuchten aber keine LHDs. BAUM, KELSCH, DANILOFF, DANILOFF & LEWIS (1982) fanden ein Defizit bei LHDs im Vergleich zur Kontrollgruppe.²⁴³ EMMOREY (1987) demonstrierte ein selektives Defizit für LHDs (im Gegensatz zu RHDs) im Vergleich zur Kontrollgruppe.

Die Daten zur Lateralisierung der Verarbeitung des Wortakzentes sind ähnlich inkonsistent wie die zur Intonationsverarbeitung; will man die Daten aber zusammenfassen, so scheinen sie eher auf eine linkshemisphärische Spezialisierung für die Verarbeitung des Wortakzentes hinzuweisen. Zu bedenken ist jedoch, dass die Verarbeitung einiger dieser Stimuli in besonderem Maße die Rekonstruktion ihrer morphosyntaktischen Struktur erfordert, denn die Phrasen *blackboard* und *black board* unterscheiden sich nicht nur in ihrer Bedeutung. Bei diesen Formen handelt es sich nicht um verschiedene Lexeme, sondern um zwei Ausdrücke mit unterschiedlicher morphosyntaktischer Struktur: *blackboard* ist ein Kompositum, d.h. ein aus zwei Lexemen zusammengesetztes komplexes Lexem mit einer nur eingeschränkt kompositionalen Bedeutung, und

²⁴³ RHDs wurden nicht getestet.

black board ist eine Folge von zwei Lexemen, die aus einem Substantiv und einem Adjektiv besteht.²⁴⁴ Sollen die Stimuli identifiziert werden (z.B. indem der Proband auf die Abbildung einer Tafel oder auf die Abbildung eines schwarzen Tablett zeigt), so muss erstens das Akzentmuster wahrgenommen und zweitens die morphosyntaktische Struktur der Stimuli rekonstruiert werden. Daher ist zu erwarten, dass bei der Verwendung solcher Testwörter in Studien zur Lateralisierung der Verarbeitung des Wortakzentes eine eventuell bestehende rechtshemisphärische Spezialisierung für die Verarbeitung der phonetischen *cues* für das Akzentmuster durch eine mögliche linkshemisphärische Spezialisierung für die morphosyntaktische Verarbeitung überdeckt wird. Wie für die Satzprosodie gilt auch für den Wortakzent, dass sich bei der Verwendung tiefpassgefilterter, unverständlicher Stimuli eher eine rechtshemisphärische Spezialisierung demonstrieren lässt als unter Verwendung verständlicher Stimuli.

5.4.3.3 Lexikalischer Ton

In manchen Sprachen gibt es Wortpaare, die sich nur in der Tonhöhe oder im Tonhöhenverlauf unterscheiden. Solche Sprachen heißen Tonsprachen. Zu den Tonsprachen zählen z.B. Thai, Norwegisch und Chinesisch.²⁴⁵ MOEN (1993) konnte in einer dichotischen Studie zum Norwegischen einen REA für die Verarbeitung des Tonkontrastes demonstrieren, MOEN & SUNDET (1996) fanden in einer klinischen Studie mit norwegischen Probanden ein selektives Defizit für LHDs im Vergleich zu RHDs und Kontrollsubjekten für diesen Kontrast. In zwei Studien zum Thai konnten VAN LANCKER & FROMKIN (1973, 1978) einen REA für die Wahrnehmung des Tonkontrastes zeigen,²⁴⁶ ebenso fanden GANDOUR & DARDARANANDA (1983) hierfür ein selektives Defizit für thaisprachige LHDs. Anders als die Daten zur Verarbeitung von Satzprosodie und Wortakzent sind diese Ergebnisse konsistent und weisen auf eine linkshemisphärische Spezialisierung für die Verarbeitung des lexikalischen Tons hin.

5.4.3.4 Emotionale Prosodie

Die emotionale Prosodie von Sätzen wird durch dieselben akustischen Dimensionen kodiert wie die verschiedenen Arten der Satzprosodie, unterscheidet sich von dieser aber dadurch, dass die Eigenschaften des akustischen Signals nicht auf linguistische

²⁴⁴ Diese eingeschränkte Vergleichbarkeit der beiden Elemente eines Stimuluspaares findet sich nicht bei den Substantiv-Verb-Paaren (*object – object*).

²⁴⁵ Unter den südostasiatischen und afrikanischen Sprachen finden sich viele Tonsprachen.

²⁴⁶ Dies gilt jedoch nur für Sprecher des Thai, nicht für amerikanische Probanden, die kein Thai verstanden.

Kategorien bezogen werden, sondern auf emotionale Kategorien wie ‘traurig’, ‘wütend’ usw. Für die Wahrnehmung emotionaler Prosodie konnte in mehreren dichotischen Studien (HAGGARD & PARKINSON 1971, LEY & BRYDEN 1982, HERRERO & HILLIX 1990) ein LEA demonstriert werden; auch ROSS, THOMSON & YENKOSKY (1997) kommen in einem Überblick zur Lateralisierung der emotionalen Prosodie zu dem Ergebnis, dass die Mehrzahl der Studien die Annahme einer rechtshemisphärischen Spezialisierung für die Wahrnehmung und Produktion emotionaler Prosodie stützt.

5.4.4 Zusammenfassung

Wie die vorangegangenen Ausführungen gezeigt haben, findet sich für die Wahrnehmung von *Konsonanten* in vielen Studien eine linkshemisphärische Spezialisierung, während sich für die Wahrnehmung von *Vokalen* kein Ohr-Vorteil zeigen lässt. Andere Untersuchungen zeigen, dass die Verarbeitung von lexikalischem *Ton* linkshemisphärisch lateralisiert ist, während die Ergebnisse zu *Wortakzent* und *Satzprosodie* eher inkonsistent sind. Für die Verarbeitungs des Wortakzents weisen die zitierten Studien auf eine stärkere linkshemisphärische Spezialisierung hin, was jedoch auch im Kontext der Tatsache zu sehen ist, dass die verwendeten Stimuli häufig nicht in lexikalischer, sondern in morphosyntaktischer Hinsicht kontrastieren. Für die Verarbeitung der Satzprosodie weisen die hier referierten Untersuchungen eher auf eine rechtshemisphärische Beteiligung hin. Eine linkshemisphärische Spezialisierung lässt sich im Allgemeinen nicht nur für *bedeutungsvolle* Sprachstimuli, sondern auch für die Verarbeitung von *bedeutungslosen* Konsonant-Vokal-Silben zeigen. Dennoch scheint es so, dass verständliche Sprachstimuli eher als unverständliche Sprachstimuli zu einem Verarbeitungsvorteil für die linke Hemisphäre führen.

5.5 Lateralitätseffekte in der Sprachwahrnehmung und Theorien zur Hemisphärenspezialisierung

Wie können diese Ergebnisse zur Lateralisierung der Verarbeitung von segmentalen und suprasegmentalen phonologischen Merkmalen in der Sprache in eine allgemeinere Theorie zur Hemisphärenspezialisierung eingeordnet werden? Zwei der oben skizzierten Theorien, die Hypothese einer ‘funktionalen’ Lateralisierung und die Hypothese einer linkshemisphärischen Verarbeitung von akustischer Information in schneller Folge, scheinen hier besonders interessant.

Zunächst betrachten wir die Ergebnisse vor dem Hintergrund der Hypothese einer ‘funktionalen’ Lateralisierung. Die Annahme, dass die Verarbeitung von Eigenschaften, die eine distinktive, d.h. bedeutungsunterscheidende, Funktion im Sprachsystem haben, linkshemisphärisch lateralisiert ist, ermöglicht eine angemessene Zusammenfassung der Daten: Die Verarbeitung von informationsreichen²⁴⁷ Konsonanten, die Verarbeitung des bedeutungsunterscheidenden Wortakzents und des bedeutungsunterscheidenden lexikalischen Tons, ebenso die Verarbeitung der satzkonstituentenstrukturbezogenen Prosodie [Typ (3)], die als distinktiv im Hinblick auf syntaktische Strukturen gelten kann, ist linkshemisphärisch lateralisiert. Für die Verarbeitung der wenig informationstragenden, d.h. wenig distinktiven Vokale findet sich hingegen kein Ohr-Vorteil und für die Verarbeitung der nicht-distinktiven modalen, emphatischen und pragmatischen Satzprosodie [Typ (1), (2), (4)] eher eine rechtshemisphärische Spezialisierung. In diesem Zusammenhang muss auch die Beobachtung erwähnt werden, dass eine Tiefpassfilterung der Stimuli, die sie unverständlich macht und damit die ‘linguistische Funktionalität’ auf der phonologischen, der morphosyntaktischen und der lexikalischen Ebene beseitigt, die Wahrscheinlichkeit erhöht, dass sich eine rechtshemisphärische Spezialisierung zeigt.

Die Hypothese, dass die linke Hemisphäre spezialisiert ist für die Verarbeitung von akustischen Ereignissen in schneller Folge, ist besonders geeignet, um die Daten zu Lateralitätseffekten für die Verarbeitung von segmentalen Merkmalen zusammenzufassen (s.o. 5.4.2). Auch die Daten zur Verarbeitung der suprasegmentalen Merkmale passen mehr oder weniger zu dieser Hypothese. (i) Der lexikalische Ton ist eine Eigenschaft einzelner Silben²⁴⁸ von 300-400 ms Dauer und entspricht in akustischer Hinsicht dem Tonhöhenverlauf²⁴⁹ in diesem Zeitbereich. (ii) Der Wortakzent ist eine Eigenschaft

²⁴⁷ Dass Konsonanten in den meisten Sprachen mehr Information tragen als Vokale, ergibt sich aus der Tatsache, dass sich in den meisten Sprachen mehr Konsonanten finden als Vokale. Der höhere Informationsgehalt der Konsonanten lässt sich so interpretieren, dass die Konsonanten eher eine bedeutungsunterscheidende Funktion besitzen, während Vokale eher die Funktion der Gruppierung der distinktiven Konsonanten um einen Silbenmittelpunkt herum erfüllen.

²⁴⁸ Neben silbenbasierten Tonsystemen gibt es auch wortbasierte Tonsysteme, z.B. im Schwedischen und Norwegischen. Auch in wortbasierten Tonsystemen finden sich schnelle Veränderungen der Tonhöhe im Zeitbereich einer Silbe von 300-400 ms. Der Unterschied zwischen beiden Tonsystemen besteht darin, dass mögliche Tonhöhenverläufe in wortbasierten Tonsystemen für ganze Wörter phonologisch vorsepezifiziert sind, in silbenbasierten Tonsystemen hingegen für einzelne Silben (LAVER 1995:463).

²⁴⁹ Dies gilt nur für Konturtöne. Daneben finden sich Registertöne, die einer bestimmten Tonhöhe relativ zu einer Basistonhöhe entsprechen, die also nicht dynamisch sind. Diese Tatsache würde im Rahmen der oben diskutierten Dichotomie erwarten lassen, dass die Verarbeitung von Register- und Kontur-tönen unterschiedlich lateralisiert ist. Thai, das von VAN LANCKER & FROMKIN (1973, 1978) und GANDOUR & DARDARANANDA (1983) untersucht wurde, gehört zu den Konturtonsprachen (LAVER 1995:467).

von Wörtern und entspricht einer Veränderung von Tonhöhe, Lautstärke und Dauer im Zeitbereich der Dauer eines Wortes. (iii) Die Satzprosodie ist eine Eigenschaft von Sätzen und entspricht einer Veränderung von Tonhöhe, Lautstärke und Dauer im Zeitbereich der Dauer eines Satzes. Der linkshemisphärisch lateralisierte lexikalische Ton entspricht also einer schnellen Änderung akustischer Eigenschaften, der Wortakzent einer langsameren Änderung und die eher rechtshemisphärisch lateralisierte Satzprosodie einer sehr langsamen Änderung der akustischen Eigenschaften.

Beide Hypothesen zur Hemisphärenspezialisierung eignen sich gleichermaßen zur Zusammenfassung der oben zitierten Beobachtungen zur Lateralisierung der Verarbeitung von segmentalen und suprasegmentalen Kategorien und Merkmalen in der Sprache. Die erste Hypothese konzipiert die Grundlage der Lateralisierung dabei als eine Funktion der Art und Weise der kognitiven Verarbeitung des Sprachreizes, d.h. seiner linguistischen Funktionalität, die zweite Hypothese als eine Funktion der akustischen Verarbeitung des Sprachsignals in Abhängigkeit von der Präsenz bestimmter akustischer Eigenschaften im Sprachsignal. Wahrscheinlich ist jedoch, dass im Experiment beobachtete Lateralitätseffekte für eine Komponente der Sprachfähigkeit immer sowohl eine Funktion der akustisch-phonetischen Eigenschaften des Sprachsignals als auch seiner linguistischen Funktionalität sind.

Die oben zusammengefassten Daten zu funktionalen Asymmetrien in der Verarbeitung von segmentaler und suprasegmentaler Information in der Sprache weisen auch darauf hin, dass *Verarbeitungsstrategien* einen starken Einfluss auf Lateralitätseffekte haben können (s.o. 5.4.2). Dies ist einerseits für die Wahrnehmung von Vokalen gezeigt worden, die auf Grund verschiedener akustischer *cues* verarbeitet werden können. Veranlasst man den Probanden durch eine Änderung der experimentellen Bedingungen, andere akustische *cues* zu nutzen, als unter den natürlichen Bedingungen der Sprachverarbeitung verwendet werden, so ändert sich auch der Ohr-Vorteil. Andererseits wurde in den vorangegangenen Ausführungen hervorgehoben, dass die Verwendung von Diskriminations- im Gegensatz zu Identifikationsaufgaben und von ungefilterten im Gegensatz zu tiefpassgefilterten Stimuli zu unterschiedlichen Ergebnissen führen kann. Mit *unverständlichen* tiefpassgefilterten Stimuli lässt sich eher als mit verständlichen ungefilterten Stimuli eine *rechtshemisphärische Spezialisierung* nachweisen und in Identifikationsaufgaben, die ein *Verstehen* des Reizes erfordern, lässt sich eher als in Diskriminationsaufgaben eine *linkshemisphärische Spezialisierung* zeigen.

Auch in diesem Zusammenhang scheinen wieder die zwei Faktoren *akustische Basis* und *linguistische Funktion* relevant zu sein, denn ein Einfluss der Verarbeitungsstrategie auf im Experiment beobachtete Lateralitätseffekte lässt sich sowohl im Hinblick auf die Verarbeitung der *akustischen Basis* als auch im Hinblick auf die Verar-

beitung der *Bedeutung* des Stimulus demonstrieren. Dies kann als weiteres Argument dafür gelten, dass die Theorie einer 'Funktionalen' Lateralisierung und die Hypothese einer linkshemisphärischen Spezialisierung für die Verarbeitung von akustischen Ereignissen in schneller Folge gleichermaßen Gültigkeit besitzen. Aus dieser Annahme folgt schließlich, dass die Hemisphärenspezialisierung, die sich für eine komplexe Sprachverarbeitungsaufgabe demonstrieren lässt, eine Funktion sowohl der Hemisphärenspezialisierung für die zugrundeliegende auditive Verarbeitung als auch eine Funktion der linguistisch-funktionalen Verarbeitung des Reizes ist.

5.6 Ist die Verarbeitung der Sonoritätskontur einer Äußerung lateralisiert?

Die bisherigen Ausführungen haben gezeigt, dass sich die Erkenntnisse zu funktionalen Lateralitäten in der Sprachwahrnehmung mittels zweier Theorien gleichermaßen zusammenfassen lassen. Welche Voraussage ergibt sich aus diesen Theorien für eine mögliche funktionale Lateralisierung der Sonoritätskontur? Im folgenden wird gezeigt, dass aus diesen beiden Theorien geschlossen werden muss, dass für die Verarbeitung der Sonoritätskontur einer Äußerung eine rechtshemisphärische Spezialisierung zu erwarten ist.

Hypothese einer 'Funktionalen' Lateralisierung. In Kapitel 2 (s.o. 2.4) wurde darauf hingewiesen, dass sich in den meisten Sprachen keine zwei Wörter finden, die sich nur in der Sonoritätskontur bzw. in der Silbenstruktur, aber nicht in den segmentalen Eigenschaften unterscheiden (zu Ausnahmen s.o. 2.4.1, Anm.). Die Sonoritätskontur bzw. die Silbenstruktur erfüllt also im phonologischen System der meisten Sprachen *keine distinktive* (d.h. bedeutungsunterscheidende) *Funktion*. In Kapitel 3 (s.o. 3.2.2) wurde festgestellt, dass sich auch keine zwei Segmente finden lassen, die sich nur in einem Oberklassenmerkmal und sonst in keinem weiteren segmentalen Merkmal unterscheiden. Die Oberklassenmerkmale, auf die sich das phonologische Merkmal Sonorität zurückführen lässt, sind also ebenfalls *keine distinktiven Merkmale*. Eine Ausnahme hiervon bildet das Merkmal Silbizität: dieses ist distinktiv, insofern sich der Vokal /i/ und der Gleitlaut /j/ nur in diesem und sonst in keinem weiteren Merkmal unterscheiden. Die Distribution von Vokalen und Gleitlauten ist jedoch durch die Silbenstruktur einer Äußerung, diese wiederum durch die segmentale Sonorität der darin vorkommenden Segmente bestimmt (s.o. 2.4.4). Aus diesem Grunde kann auch das Merkmal [silbisch] keine bedeutungsunterscheidende Funktion erfüllen und muss daher als *nichtdistinktives, prosodisch zu definierendes phonologisches Merkmal* gelten.

Die Sonoritätskontur einer Äußerung hat also keine bedeutungsunterscheidende Funktion. Eher ist es so, dass dem Sonoritäts-Folge-Prinzip entsprechende Segmentketten, also solche, die einen zyklischen, quasiperiodischen Anstieg und Abstieg der Sonorität aufweisen, eine bestimmte, für die Sprachproduktion und für die Sprachwahrnehmung vorteilhafte Form oder Gestalt aufweisen. Auf artikulatorischer Seite entspricht der Sonoritätszyklus einem Zyklus aus einer graduellen Öffnungsbewegung und einer anschließenden graduellen Schließbewegung, was unter kinetischem Aspekt die effizienteste Art der Sprachproduktion ist. Auf perzeptiv-phonetischer Seite entspricht der Sonoritätszyklus einer langsamen Änderung der Parameter Lautstärke, Dauer und Periodizität, auf Grund derer im Signal Silbennuklei und Silbengrenzen lokalisiert werden können (s.o. 4.3.3). Daher ist es so, dass die Verarbeitung der Sonoritätskontur einer Äußerung eine strukturierende Funktion für die Sprachwahrnehmung erfüllt und möglicherweise zu den universalen Sprachverarbeitungsstrategien gehört (s.o. 4.3.4), aber keine *bedeutungsunterscheidende* oder *strukturunterscheidende*²⁵⁰ Funktion besitzt, allgemein gesagt, keine linguistisch-funktionale Relevanz für die Sprachwahrnehmung hat. Daher ist unter dieser Hypothese für die Verarbeitung der Sonoritätskontur einer Äußerung eine rechtshemisphärische Spezialisierung zu erwarten.

Hypothese einer linkshemisphärischen Spezialisierung für die Verarbeitung von akustischen Ereignissen in schneller Folge. Der phonologischen Sonoritätskontur entspricht in phonetischer Hinsicht eine langsame, quasiperiodische Änderung der Lautstärke des Signals sowie eine langsame Änderung der Parameter Dauer und Periodizität, d.h. ein zyklischer Wechsel von länger dauernden und kürzer dauernden stetigen Signalabschnitten und von periodischen und aperiodischen Signalabschnitten (s.o. 3). Ein Zyklus hat dabei die zeitliche Dimension einer Silbe, d.h. eine Dauer von 200-400 Millisekunden.²⁵¹ Dies entspricht einer Modulationsfrequenz von 2,5-5 Hz. In Kapitel 4 (s.o. 4.4.3) wurde herausgestellt, dass die Wahrnehmung des Silbenrhythmus in einer Äußerung bzw. die Wahrnehmung der Zahl der Silbennuklei auf der Verarbeitung der langsamen Amplitudenmodulation im Signal basiert. Das phonologische Merkmal Sonorität wird also nicht auf der Grundlage von akustischen Ereignissen in schneller Folge oder auf der Grundlage von sich schnell ändernder akustischer Information identifiziert, wie z.B. die Plosivkonsonanten, für die sich der stabilste Rechts-Ohr-Vorteil nachweisen lässt, sondern auf der Grundlage der sich sehr langsam ändernden akustischen Eigenschaften Lautstärke, Dauer und Periodizität. Deshalb ist auch unter dieser Hypothe-

²⁵⁰ *Strukturunterscheidend* bedeutet hier: *distinktiv* im Hinblick auf die syntaktische oder morphologische *Struktur*, wie z.B. die Satzintonation und Satzbetonung syntaktisch distinktiv sein kann (s.o. 5.4.3.1), oder die Wortbetonung morphologisch distinktiv (s.o.5.4.3.2).

²⁵¹ Vgl. GREENBERG 1996. LENNEBERG (1972) gibt eine mittlere Silbendauer von 360 ms an.

se für die Verarbeitung der Sonoritätskontur einer Äußerung eine rechtshemisphärische Spezialisierung zu erwarten.

Beide Theorien unterstützen also die Annahme einer rechtshemisphärischen Repräsentation der Verarbeitung der Sonoritätskontur einer Äußerung. Auch die oben erwähnte Beobachtung, dass nach linkshemisphärischer Schädigung die sprachlichen Äußerungen der Patienten eine intakte Phonotaktik aufweisen (s.o. 5.2.3), ist mit der Annahme einer rechtshemisphärischen Spezialisierung für die Verarbeitung des Sonoritätsverlaufes vereinbar, denn hieraus könnte gefolgert werden, dass diese Patienten über eine intakte Repräsentation der segmentalen Sonorität bzw. der Sonoritätshierarchie und des Sonoritäts-Folge-Prinzips verfügen, die als Basis für die Silbifizierung und artikulatorische Implementation der bei der Sprachproduktion aktivierten Lautfolgen dienen könnten. Dies könnte so interpretiert werden, dass die *Produktion der Sonoritätskontur* einer Äußerung nach linkshemisphärischer Schädigung intakt ist, d.h. rechtshemisphärisch lateralisiert sein könnte und analog hierzu könnte angenommen werden, dass auch die *Wahrnehmung der Sonoritätskontur* einer Äußerung rechtshemisphärisch lateralisiert sein könnte.

Schließlich ist diese Annahme auch unter der Hypothese einer *globalen Verarbeitung* von Reizen durch die rechte Hemisphäre und einer *lokalen Informationsverarbeitung* durch die linke Hemisphäre plausibel (s.o. 5.3.4). Für das visuelle System konnte gezeigt werden, dass die Verarbeitung der *Gestalt* eines visuellen Reizes rechtshemisphärisch lateralisiert ist und die Verarbeitung der *Details* eines Reizes linkshemisphärisch. Bezogen auf die Wahrnehmung von Sprache unter natürlichen Bedingungen können die einzelnen Phoneme, d.h. die segmentale Organisation, als lokale Merkmale oder Details von sprachlichen Äußerungen gelten und die suprasegmentale Organisation, d.h. die Intonation und Betonung, als globale Merkmale von sprachlichen Äußerungen. Weiter oben wurde hervorgehoben (s.o. 2.4.2, 3.4), dass sich die *Oberklassenmerkmale* bzw. die segmentale *Sonorität* mit CHOMSKY & HALLE (1968:302) auf das „Skelett der Sprache“ oder die *Gestalt* der Sprache beziehen lassen. Sonorität wurde als dasjenige Merkmal der Sprache gekennzeichnet (s.o. 2.4.5), welches die Verbindung zwischen segmentaler Struktur und suprasegmentaler Silbenstruktur herstellt und welches die Integration einer Folge von Segmenten in eine übergeordnete Struktur, d.h. in Silben, ermöglicht. Wenn also im Hinblick auf die Sprachwahrnehmung, bzw. auf die Erkennung eines Wortes unter experimentellen Bedingungen eine globale und eine lokale Art der Informationsverarbeitung unterschieden werden sollen, so ist anzunehmen, dass die Sonoritätskontur einer Äußerung als globales Merkmal verarbeitet wird und daher eine rechtshemisphärische Spezialisierung für diese Wahrnehmungsfunktion zu erwarten ist.

Auch unter der Hypothese einer Spezialisierung der rechten Hemisphäre für die Verarbeitung der Signalanteile mit *tiefer Frequenz* und der linken Hemisphäre für die Verarbeitung der Signalanteile mit *hoher Frequenz* (IVRY & LEBBY 1993, 1997) ist eine rechtshemisphärische Spezialisierung für die Verarbeitung der Sonoritätskontur zu erwarten, denn diese korreliert mit den sehr langsamen Amplitudenmodulationen im Signal. Wie in Kapitel 3 (s.o. 3.2.2) ausgeführt wurde, gibt auch die Stimmeinsatzzeit eines Konsonanten einen Hinweis darauf, ob ein Sonoritätsmaximum folgt oder nicht. Die Stimmeinsatzzeit eines Konsonanten wird (s.o. 3.2.2, Anm.) auf der Grundlage der zeitlichen Differenz zwischen der Verschlussöffnung und dem Einsatz der Stimmlippen-schwingung, d.h. dem Einsetzen einer periodischen Grundschwingung im Signal, verarbeitet. Da es die Anwesenheit oder Abwesenheit der Grundfrequenz F_0 des Signals ist, die als Hinweis auf die Stimmhaftigkeit des Konsonanten verarbeitet wird, nehmen IVRY & LEBBY (1997) an, dass die Verarbeitung der Stimmhaftigkeit eine rechtshemisphärische Beteiligung erfordert (vgl. auch SIMOS, MOLFESE & BRENDEN 1997). Für die Verarbeitung der Lautstärk modulation und der Stimmeinsatzzeit als phonetische Korrelate der phonologischen Sonoritätskontur ist daher eine Spezialisierung der rechten Hemisphäre zu erwarten.

Mit einer Theorie der differentiellen Aktivierung semantischer Information in den beiden Hemisphären, wie sie von CHIARELLO (1997)²⁵² vertreten wird, ist die Annahme einer rechtshemisphärischen Spezialisierung für die Verarbeitung von segmentaler Sonorität und Lautklassenmerkmalen ebenfalls vereinbar, wenn man die Aktivierung von lexikalischen Konzepten, d.h. von Wortbedeutungen und von phonologischen Konzepten, d.h. von Merkmalen wie [nasal], [vokalisches], [betont] etc., als analog zur Aktivierung von Kategorien wie Phonem /a/, Silbe /ba/, Takt ['σσσ] etc. betrachtet. CHIARELLO (1997) nimmt an, dass bei der Verarbeitung eines Wortes in der rechten Hemisphäre sehr viele ähnliche Bedeutungen unselektiv aktiviert werden, während in der linken Hemisphäre nur die in den gegebenen Kontext passenden Bedeutungen selektiv aktiviert werden. Auf die phonologische Ebene übertragen könnte dies bedeuten, dass bei der Verarbeitung eines Signalabschnittes in einem Sprachsignal in der linken Hemisphäre nur das am besten in den Kontext und zum gegebenen akustischen Muster passende Phonem aktiviert wird, während in der rechten Hemisphäre alle ungefähr passenden Phoneme, d.h. eine ganze *Oberklasse* von Phonemen bzw. alle Phoneme mit gleichem *Sonoritätswert* aktiviert werden. Wird eine ganze Sequenz von Lauten verarbeitet, so könnte es so sein, dass in der rechten Hemisphäre eine Folge von *Oberklassen*

²⁵² Vgl. auch RICHARDS & CHIARELLO 1997. Für ähnliche Ergebnisse aus dem Bereich der Verarbeitung von Wörtern im Satzzusammenhang vgl. FAUST 1997.

aktiviert wird, d.h. eine Folge von *Sonoritätswerten* oder eine *Sonoritätskontur* und in der linken Hemisphäre eine Folge von *distinktiven Phonemen*. Unter dieser Hypothese würde also bei der Wahrnehmung von sprachlichen Äußerungen in der linken Hemisphäre eine *Repräsentation der distinktiven Merkmale* erzeugt, die aus dem Signal rekonstruiert werden können, in der rechten Hemisphäre hingegen eine *Repräsentation der nichtdistinktiven Oberlassenmerkmale*, für die sich Hinweise im Signal finden, z.B. eine Formantstruktur [+ sonorant], oder Friktion ([– sonorant]). Anders formuliert, eine *Repräsentation der Sonoritätskontur* einer Äußerung würde dieser Theorie zufolge in der rechten Hemisphäre rekonstruiert.

Die Erkenntnisse zur differentiellen Aktivierung semantischer Information in beiden Hemisphären bilden auch den Ausgangspunkt für die Annahme (BEEMAN & CHIARELLO 1997), dass die linke Hemisphäre sprachliche Information ‘code-spezifisch’ und modular verarbeitet, die rechte Hemisphäre hingegen ‘code-unspezifisch’ und nicht-modular. Das Konzept der ‘code-spezifischen’ Informationsverarbeitung (CHIARELLO 1997) bezieht sich in diesem Zusammenhang auf die Annahme, dass die linke Hemisphäre eine ganz bestimmte Art sprachlicher Information (z.B. lexikalische Information) auf der Basis einer begrenzten Menge von (in diesem Fall lexikalischen) Verarbeitungsroutinen auf eine bestimmte Kategorie von Bedeutungen (d.h. auf lexikalische Bedeutungen) abbildet, während die rechte Hemisphäre eher verschiedene Arten von Information (z.B. sprachliche *und* bildliche Information) mit denselben Verarbeitungsroutinen unspezifisch auf verschiedene Arten von Bedeutungen (d.h. auf lexikalische, emotionale und piktoriale Bedeutungen) abbilden kann. Nimmt man nun an, dass die Verarbeitung der Sonoritätskontur einer Äußerung nicht nur als Grundlage der Rekonstruktion einer phonologischen Silbenstrukturepräsentation, sondern auch einer phonologischen Akzentstrukturepräsentation dient (s.o. 3.4 und 4.4.2), so könnte man dies als Hinweis darauf werten, dass die Sonoritätskontur eher ‘code-unspezifisch’ und nicht-modular verarbeitet wird und daher, der Hypothese von CHIARELLO und BEEMAN zufolge, eher in der rechten Hemisphäre verarbeitet werden sollte.

Zusammenfassend lässt sich sagen, dass für die Verarbeitung der Sonoritätskontur einer Äußerung auf der Grundlage der in diesem Kapitel zusammengefassten experimentellen Daten und vor dem Hintergrund aller hier skizzierten Theorien zur Hemisphärenspezialisierung eine rechtshemisphärische Lateralisierung zu erwarten ist.

5.7 Zusammenfassung

In diesem Kapitel wurde zunächst über Studien zur Spezialisierung der linken und der rechten Hemisphäre für sprachliche und nichtsprachliche kognitive Funktionen berichtet. Im Anschluss daran wurden einige Theorien zur Hemisphärenspezialisierung skizziert, die, wie sich zeigte, mehr oder weniger geeignet sind, Beobachtungen zur Hemisphärenspezialisierung für Sprachfunktionen zusammenzufassen. Sodann wurden Ergebnisse aus Studien zur funktionalen Lateralisierung der Verarbeitung von phonemischer und prosodischer Information in der Sprachwahrnehmung referiert. Hierbei wurde insbesondere hervorgehoben, dass unterschiedliche experimentelle Bedingungen in dichotischen Sprachverarbeitungs-Experimenten zu unterschiedlichen Ohr-Vorteilen führen können.

Als Faktoren, welche die experimentellen Ergebnisse beeinflussen können, müssen erstens solche Bedingungen genannt werden, welche die Aufmerksamkeit der Probanden auf eine bestimmte Art von akustischer Information lenken, z.B. die Erhöhung der 'perzeptuellen Schwierigkeit' des Signals (TARTTER 1988). Als zweiter Faktor sind die Bedingungen zu nennen, welche die Bedeutungshaltigkeit der Reize beeinflussen, z.B. eine Tiefpassfilterung des Signals und als dritter Faktor solche Bedingungen, welche die kognitiven Strategien der Probanden beeinflussen, z.B. die Verwendung einer Diskriminations- oder einer Identifikationsaufgabe. Diese Beobachtungen zum Einfluss experimenteller Bedingungen auf die Ohr-Vorteile bzw. Hemisphärenvorteile, die im Experiment beobachtet werden können, werden im folgenden Kapitel wiederaufgegriffen.

Vor dem Hintergrund bisheriger Beobachtungen zur funktionalen Lateralisierung der Wahrnehmung segmentaler und suprasegmentaler linguistischer Kategorien und Merkmale wurde, gestützt auf zwei Theorien zur Hemisphärenspezialisierung, schließlich die These formuliert, dass die Verarbeitung der Sonoritätskontur einer Äußerung rechtshemisphärisch lateralisiert ist. Diese These wurde in zwei dichotischen Sprachwahrnehmungsversuchen experimentell überprüft, die im nächsten Kapitel beschrieben werden.

6 Lateralitätseffekte in der Verarbeitung der Sonoritätskontur einer Äußerung

6.1 Einleitung

Die These dieser Studie, dass die Verarbeitung der Sonoritätskontur einer Äußerung rechtshemisphärisch lateralisiert ist (s.u. 5.5), wurde in zwei verschiedenen Sprachwahrnehmungsexperimenten mit dichotischer Reizpräsentation getestet. Wie im vorangegangenen Kapitel hervorgehoben wurde, können Ohr-Vorteile für die Sprachwahrnehmung sowohl für bedeutungsvolle als auch für bedeutungslose Reize demonstriert werden (s.o. 4.5.2). Die Bedeutungshaltigkeit eines Reizes kann aber den Ohr-Vorteil beeinflussen, insofern mit bedeutungsvollen Reizen eher als mit bedeutungslosen Reizen Rechts-Ohr-Vorteile nachgewiesen werden können (s.o. 5.4.4, 5.4.3.1). Aus diesem Grunde wurde das Merkmal ‘Sonorität’ in den beiden Experimenten auf unterschiedliche Weise operationalisiert: in Experiment 1 wurden natürlichsprachliche *bedeutungsvolle deutsche Wörter* als Stimuli verwendet, in Experiment 2 wurden natürlichsprachliche *bedeutungslose Silben* verwendet, die aus ein- bis fünfsilbigen Iterationen der bedeutungslosen Silbe ‘na’ bestanden.

In Experiment 1 bestand die Aufgabe der Probanden darin, die präsentierten Wörter zu *identifizieren*, in Experiment 2 bestand die Testaufgabe darin, eine dreisilbige Zielfolge im präsentierten dichotischen Stimuluspaar zu *detektieren*. Die Verarbeitungskomponente, die in diesen beiden Experimenten untersucht wurde, war die Wahrnehmungsfunktion *Lokalisierung der Sonoritätsmaxima in einer Äußerung*. In beiden Experimenten unterschieden sich die verwendeten Reize in der *Zahl der Sonoritätsmaxima*, aber nur minimal in jeder anderen Hinsicht. Die Wörter in Experiment 1 unterschieden sich in der Zahl der Sonoritätsmaxima, d.h. in der Zahl der Silben, aber nur

minimal in den Segmenten, aus denen das Wort bestand. Ebenso unterschieden sich die Reize in Experiment 2 nur in der Zahl der Sonoritätsmaxima, d.h. in der Anzahl der Iterationen der Silbe /na/, die einzelnen iterierten Silben unterschieden sich jedoch nicht voneinander.

Im folgenden werden zunächst die beiden Experimenten gemeinsamen methodischen Grundlagen erläutert (6.2), danach folgt eine getrennte Beschreibung und Diskussion der beiden Experimente (6.3. und 6.4). Die Hypothese, die in den beiden Experimenten getestet wurde, war die Annahme, dass sich für die Lokalisierung der Sonoritätsmaxima in einer sprachlichen Äußerung ein Links-Ohr-Vorteil, d.h. eine rechtshemisphärische Spezialisierung findet. Die Voraussage für Experiment 1 war, dass sich für die Wahrnehmung von n-silbigen gegenüber n+1-silbigen Wörtern ein Links-Ohr-Vorteil findet, die Voraussage für Experiment 2 war, dass sich für die Detektion einer dreisilbigen sinnlosen Silbenfolge bei gleichzeitiger Präsentation von ein-, zwei- vier- und fünfsilbigen Distraktorfolgen ein Links-Ohr-Vorteil findet.

6.2 Allgemeines zur Methode

6.2.1 Probanden

An den beiden Experimenten nahmen 48 rechtshändige²⁵³ Probanden teil, davon 24 Frauen und 24 Männer. Die Händigkeit der Versuchspersonen wurde mit einer modifizierten Version des *Edinburgh Handedness Inventory* (Oldfield 1971) überprüft. Rechtshändigkeit bzw. Linkshändigkeit wird in diesem Test durch ein Testergebnis auf einer Skala von +100 bzw. -100 abgebildet. Der mittlere Lateralitätsquotient der männlichen Probanden lag bei +89 (Wertebereich +60 bis +100), der mittlere Lateralitätsquotient der weiblichen Probanden bei +87 (Wertebereich +60 bis +100).

Ein Teil der Versuchspersonen wurde aus dem Versuchspersonenpool der psychologischen Fakultät rekrutiert, ein anderer Teil der Versuchspersonen setzte sich aus Studenten und Mitarbeitern des Sprachwissenschaftlichen Instituts zusammen. Das Alter der männlichen Probanden betrug im Mittel 30 Jahre (Wertebereich 21-37 Jahre), das Alter der weiblichen Probanden im Mittel 26 Jahre (Wertebereich 19-37 Jahre).

²⁵³ Während bei 96% der Rechtshänder der größere Teil der Sprachfunktionen in der linken Hemisphäre repräsentiert ist, ist dies nur bei 70% der Linkshänder der Fall. 15% der Linkshänder haben bilaterale Repräsentation der Sprache und 15% eine rechtshemisphärische Repräsentation der Sprache (RASMUSSEN, T. & MILNER, B. (1977): „The role of early left brain in determining lateralization of cerebral speech function“, in: *Annals of the New York Academy of Sciences* 299, 355-369, zitiert nach KOLB & WHISHAW 1993:173.

Keiner der Probanden berichtete von einer bekannten Beeinträchtigung des Hörvermögens. Zwei Personen wurden nicht zur Teilnahme zugelassen, weil sie angaben, auf einem Ohr weniger gut als auf dem anderen zu hören.

6.2.2 Verfahren

Alle Versuchspersonen nahmen an Experiment 1 und an Experiment 2 teil. Beide Experimente wurden jeweils in einer einzigen Sitzung durchgeführt und jedes Experiment dauerte ca. 15 Minuten. Die Hälfte der Versuchspersonen nahm zuerst an Experiment 1 und dann an Experiment 2 teil, die andere Hälfte der Versuchspersonen wurde in umgekehrter Reihenfolge getestet (zuerst Experiment 2, dann Experiment 1, vgl. Tabelle 6-1). Die Versuchspersonen wurden einzeln in einem ruhigen Raum getestet. Die gesamte Versuchsdurchführung wurde durch einen Computer kontrolliert. Die Versuchsanweisungen wurden in schriftlicher Form auf dem Computermonitor präsentiert. Der Versuchsleiter saß den Probanden gegenüber, auf der Rückseite des Bildschirms. Bei Rückfragen wurden zusätzliche Erklärungen durch den Versuchsleiter geliefert. Bevor die Versuchspersonen mit dem Experiment begannen, wurde ihnen der Händigkeitsfragebogen präsentiert.

Tabelle 6-1. Versuchspersonengruppen für Experiment 1 und 2 im Überblick

Geschlecht	Testreihenfolge	Kopfhörerposition ²⁵⁴	beginnende Hand in Experiment 2			
Männer	1 – 2	ABAB(AB) ²⁵⁵	links	3		
			rechts	3	6	
		BABA(BA)	links	3	6	
			rechts	3	12	
		2 – 1	ABAB(AB)	links	3	12
				rechts	3	6
	BABA(BA)	links	3	6		
		rechts	3	24		
	Frauen	1 – 2	ABAB(AB)	links	3	
				rechts	3	6
		BABA(BA)	links	3	6	
			rechts	3	12	
2 – 1		ABAB(AB)	links	3	12	
			rechts	3	6	
BABA(BA)	links	3	6			
	rechts	3	48			

²⁵⁴ Kopfhörerposition A bedeutet: der rechte Kanal des Kopfhörers liegt am rechten Ohr, Kopfhörerposition B bedeutet: der rechte Kanal des Kopfhörers liegt am linken Ohr.

²⁵⁵ Experiment 1 besteht aus 4 Blöcken (d.h. ABAB), Experiment 2 aus 6 Blöcken (d.h. ABABAB).

6.2.3 Dichotisches Hören

Bei beiden Experimenten handelte es sich um Sprachwahrnehmungsexperimente mit *dichotischer Stimuluspräsentation*, d.h. in beiden Experimenten wurden dem Probanden zwei verschiedene Reize zeitgleich dargeboten, je ein Reiz auf jedem Ohr (vgl. Abbildung 6-1).²⁵⁶ Da die Hörbahn gekreuzt ist und bei beidohriger Reizpräsentation die Information nur über die kontralateralen auditorischen Bahnen in den auditorischen Kortex weitergeleitet wird, führt die dichotische Darbietung von akustischen Reizen dazu, dass der rechtsohrig präsentierte Reiz nur in die linke Hemisphäre und der linksohrig präsentierte Reiz nur in die rechte Hemisphäre projiziert wird (KIMURA 1961, MILNER, TAYLOR & SPERRY 1968). Die Methode des dichotischen Hörens basiert auf der Annahme, dass sich eine Hemisphärenspezialisierung für die Verarbeitung der präsentierten Reize darin zeigt, dass die kontralateral zur spezialisierten Hemisphäre präsentierten Reize effizienter verarbeitet werden als die ipsilateral zur spezialisierten Hemisphäre dargebotenen (KIMURA 1961). *Effizientere Verarbeitung* kann dabei bedeuten, dass die Probanden im Mittel auf den Reiz, welcher dem einen Ohr dargeboten wird, schneller und / oder mit weniger Fehlern reagieren als auf den Reiz, welcher dem anderen Ohr dargeboten wird. Eine effizientere Reizverarbeitung der einen im Vergleich zur anderen Hemisphäre kann sich jedoch auch auf eine andere Weise zeigen: durch *Stimulusdominanz*. Verwendet man in einem dichotischen Test sehr ähnliche Stimuli, so nimmt der Proband nicht zwei verschiedene Stimuli, sondern nur noch einen einzigen Reiz wahr: ein Reiz *dominiert* über den anderen (WEXLER & HALWES 1983) Eine Hemisphärenspezialisierung für die Verarbeitung einer Art von Stimuli zeigt sich in einem solchen dichotischen Test dadurch, dass die Probanden im Mittel häufiger den Reiz wahrnehmen, der dem einen Ohr dargeboten wurde, als den Reiz, der dem anderen Ohr dargeboten wurde.

Dieses Phänomen der Stimulusdominanz wird in Experiment 1 genutzt, in welchem sehr ähnliche Wörter präsentiert werden. Auf Grund der großen Ähnlichkeit der Reize ist zu erwarten, dass die Probanden in jedem Testdurchgang nur ein einziges Wort wahrnehmen und zwar entweder das dem rechten Ohr präsentierte Wort, oder das dem linken Ohr präsentierte Wort. Eine häufigere Wahrnehmung des linksohrig präsentierten Reizes gegenüber dem rechtsohrig präsentierten Reiz wird in Experiment 1 als Hinweis auf eine Spezialisierung der kontralateralen, d.h. der rechten Hemisphäre für die Verarbeitungsfunktion *Lokalisierung der Sonoritätsmaxima* gewertet. In Experiment 2 hingegen werden weniger ähnliche Stimuli verwendet, für die keine Stimulus-

²⁵⁶ Für einen Überblick zum dichotischen Hören vgl. HUGDAHL (Hrsg.) 1988, vgl. auch die Ausführungen in 6.2.3.

dominanz zu erwarten ist. Als Hinweis auf eine eventuell bestehende Hemisphärenspezialisierung für die Verarbeitungsfunktion *Lokalisierung der Sonoritätsmaxima* wird in Experiment 2 daher die Reaktionszeit und die Fehlerrate für jedes Ohr bzw. jede Hemisphäre ausgewertet.

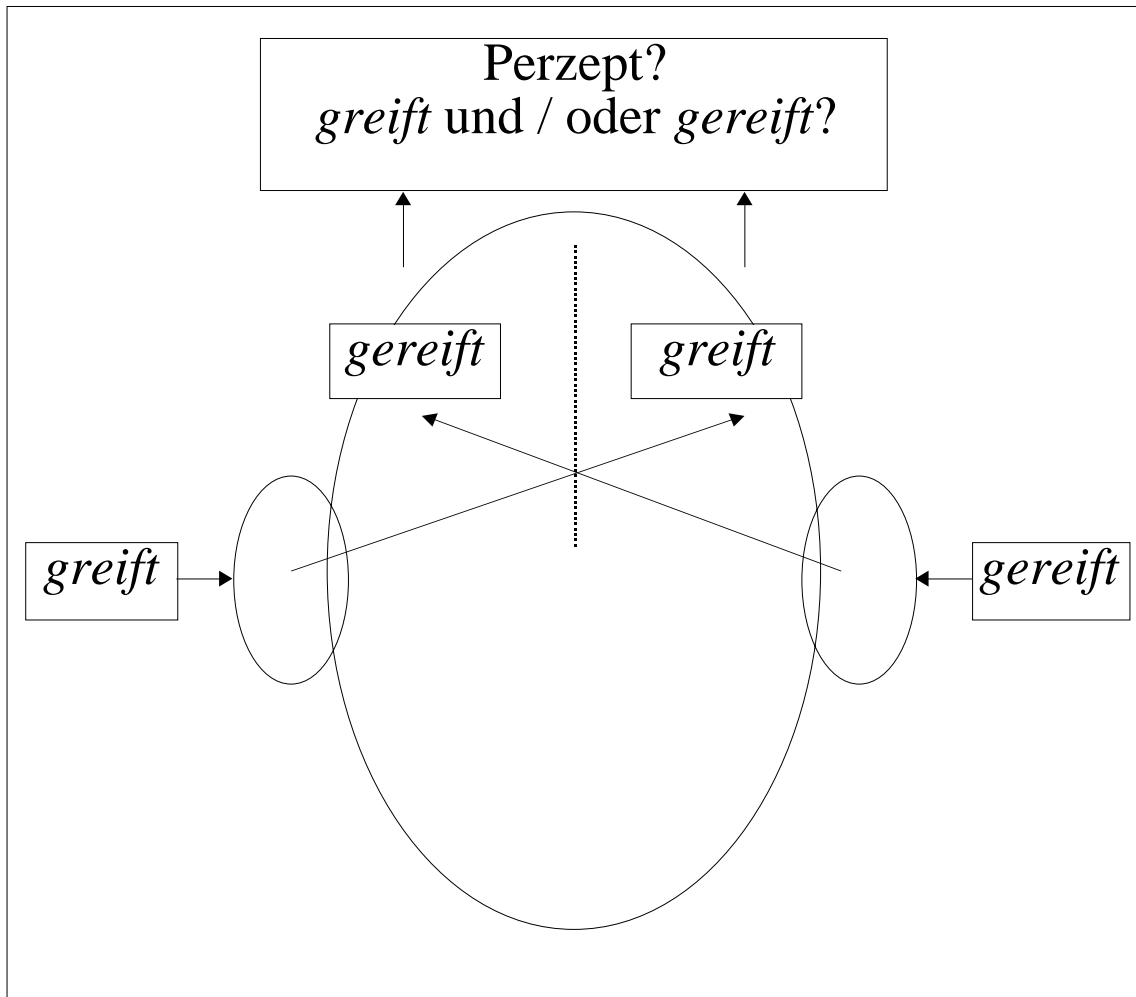


Abbildung 6-1. Dichotisches Hören.

6.3 Experiment 1

6.3.1 Methode

6.3.1.1 Stimuluswörter

Die Versuchsstimuli für das Experiment 1 bestanden aus 8 deutschen Wortpaaren. Die Elemente eines jeden Paares unterschieden sich in der Silbenzahl, aber nur minimal in den Segmenten, aus denen sie bestanden: das $n+1$ -silbige Wort bestand aus denselben

Segmenten wie das *n*-silbige Wort und aus einem zusätzlichen Schwa, d.h. einem reduzierten Zentralvokal wie in *breit*, nach dem ersten Konsonanten (vgl. Tabelle 6-2). Die zusätzliche Silbe im *n+1*-silbigen Wort gegenüber dem *n*-silbigen Wort entspricht einem zusätzlichen Sonoritätsmaximum (vgl. Abbildung 6-1). Abgesehen von der Anzahl der Sonoritätsmaxima unterschieden sich beide Wörter in der Silbenstruktur. Während der Liquidkonsonant (/r/ bzw. /l/) im *n*-silbigen Wort dem Anlaut der ersten Silbe zugeordnet war, gehörte er im *n+1*-silbigen Wort zum Anlaut der zweiten Silbe (vgl. Abbildung 6-2).

Tabelle 6-2. Phonologische Struktur der 8 Stimuluswortpaare.

Wort-paar	Wort 1		Wort 2		Segmentale Struktur				
	n-silbig einsilbig oder zweisilbig		n+1-silbig zweisilbig oder dreisilbig		C ₁ Obstruent /b/ oder /g/	Schwa oder ø	C ₂ Liquid /r/ oder /l/	VV oder V: oder V	C ₃ oder C ₃ C ₄
1	' <i>breit</i>	'σ	<i>be</i> 'reit	σ 'σ	b	ə	r	aɪ	t
2	' <i>braten</i>	'σσ	<i>be</i> 'raten	σ 'σσ				a:	tn
3	' <i>bleibt</i>	'σ	<i>be</i> 'leibt	σ 'σ			l	aɪ	bt
4	' <i>blieben</i>	'σσ	<i>be</i> 'lieben	σ 'σσ				i:	bn
5	' <i>greift</i>	'σ	<i>ge</i> 'reift	σ 'σ	g	ə	r	aɪ	ft
6	' <i>grast</i>	'σ	<i>ge</i> 'rast	σ 'σ				a:	st
7	' <i>gleiten</i>	'σσ	<i>ge</i> 'leiten	σ 'σσ			l	aɪ	tn
8	' <i>glitten</i>	'σσ	<i>ge</i> 'litten	σ 'σσ				ɪ	tn

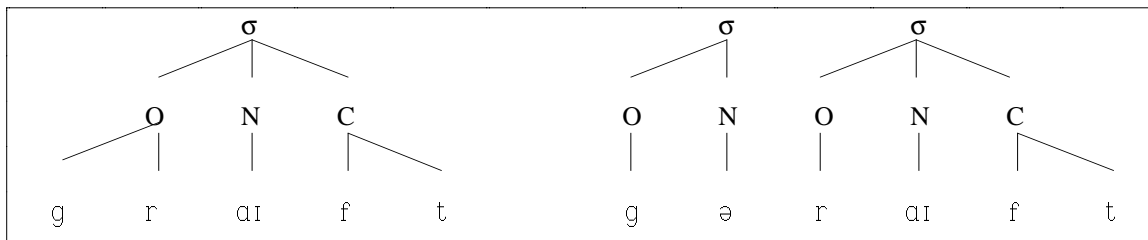


Abbildung 6-2. Silbenstruktur der Stimuli, exemplarisch dargestellt am Wortpaar 5: *greift* – *gereift*.

5			•				•		•		
4											
3			•						•		
2											
1	•			•		•				•	•
	g	r	aɪ	f	t	g	ə	r	aɪ	f	t

Abbildung 6-3. Sonoritätskontur der Stimuluswortpaare, exemplarisch dargestellt am Wortpaar 5: *greift* – *gereift*.²⁵⁷

6.3.1.2 Herstellung des Stimulusmaterials

Für die Aufnahme des Stimulusmaterials wurden die 16 Testwörter in denselben neutralen Satz eingefügt (*Sag jetzt ___ kurz und knapp*). Aus diesen 16 Sätzen, welche die Testwörter enthielten, wurden vier Listen mit je 16 Sätzen (d.h. mit insgesamt 64 Sätzen) in randomisierter Reihenfolge erzeugt. Die 4 Listen wurden von einer Sprecherin ohne phonetische Ausbildung (J.M.) nacheinander laut vorgelesen und mit einem Kassettenrekorder (Philipps AZ 8052) und einem Mikofon (Sennheiser MD 735) in einem schalltoten Raum aufgenommen. Dann wurden die 64 aufgenommenen Sätze mit einem Computer (Power Macintosh 6100) digitalisiert (16 bit, sampling rate 44.1 kHz) und auf der Festplatte abgespeichert. Diese Audio-Dateien wurden mit kommerzieller Audio-Software (*SoundEdit 16* und *Signalize 3.1*) auf demselben Computer weiterbearbeitet.

Aus den Sätzen wurden die 4 verschiedenen Versionen der 16 Testwörter ausgeschnitten und einzeln abgespeichert. Dann wurden für jedes Wortpaar aus den vier Versionen von Wort 1 und den vier Versionen von Wort 2 die zwei am besten zueinander passenden Versionen (hinsichtlich der Gesamtlänge und des Intonationsverlaufs) von Wort 1 und Wort 2 ausgewählt.

Die Länge der beiden ausgewählten Versionen für jedes Wortpaar wurde von der Versuchsleiterin manuell angeglichen, indem in den Signalen einzelne oder mehrere F0-Perioden zwischen zwei Nulldurchgängen ausgeschnitten bzw. einzelne F0-Perioden iteriert wurden. Die Dauer des resultierenden Signals wurde immer wieder an Breitbandspektogrammen beider Wörter kontrolliert. Für die Synchronisierung der beiden Signale wurde die Differenz der Länge beider Wörter berechnet: das *n*-silbige Wort

²⁵⁷ Zugrundegelegt wird die fünfstufige Sonoritätshierarchie von CLEMENTS (1990), s.u. 2.4.4. CLEMENTS (1990) unterscheidet Obstruenten, Nasale, Liquide und Vokale, die fünfte Stufe entspricht silbischen Vokalen, d.h. Vokalen.

wurde um die Hälfte der Differenz verlängert und das $n+1$ -silbige Wort um die Hälfte der Länge verkürzt.

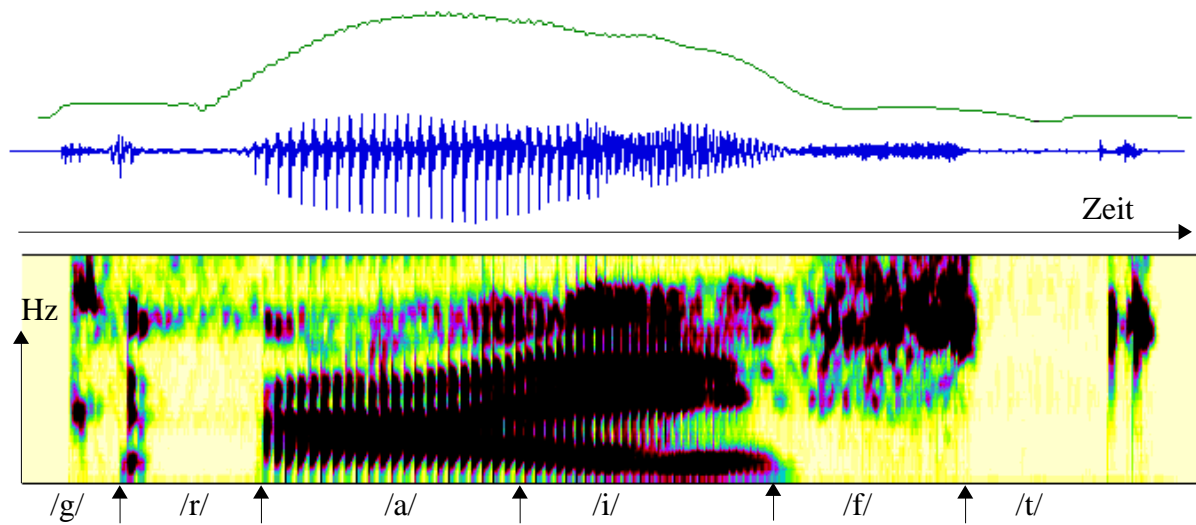


Abbildung 6-4. Breitbandspektrogramm, akustisches Signal und Einhüllende des Stimuluswortes 'greift'.

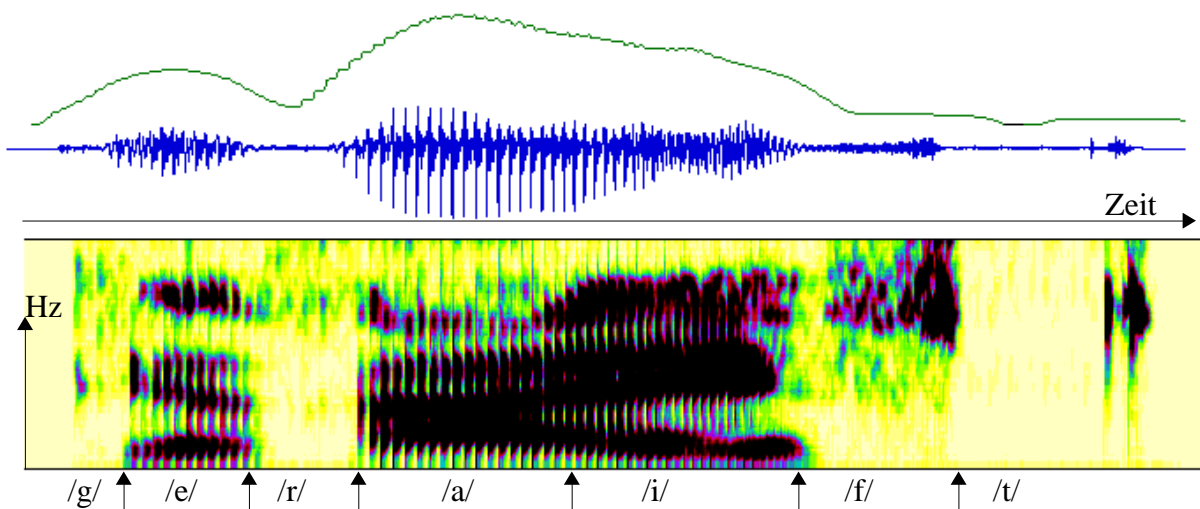


Abbildung 6-5. Breitbandspektrogramm, akustisches Signal und Einhüllende des Stimuluswortes 'gereift'.

Der Wortanfang,²⁵⁸ das Ende des betonten Vokals und alle nachfolgenden Segmente wurden anhand von Breitbandspektrogrammen manuell synchronisiert. Die nötige Verkürzung bzw. Verlängerung des Signals wurde möglichst gleichmäßig auf das Schwa

²⁵⁸ Als Öffnung des bilabialen (/b/) bzw. velaren (/g/) Verschlusses des initialen Konsonanten. Dies entspricht in den meisten Fällen dem Einsatz der Aspiration (da im Dt. auch die stimmhaften Konsonanten, hier /b/ und /g/, wortinitial als entstimmte Varianten realisiert werden, vgl. KOHLER 1977), aber in einigen Fällen dem Stimmeinsatz.

(/ə/, nur bei den $n+1$ -silbigen Wörtern), auf den Liquid /r/ bzw. /l/ und den betonten Vokal /a/ bzw. /i/ verteilt und jeweils in der Mitte der Segmente (beim Diphthong /ai/ in der Mitte des ersten und des zweiten Segmentes) vorgenommen. Die mittlere Dauer der 8 resultierenden Wortpaare war 503 ms (Wertebereich von 414 ms bis 608 ms). Vor und nach jedem Wort wurden 50 ms Stille eingefügt. Die Intensität der beiden Wörter eines Wortpaares wurde danach manuell angeglichen. Für jedes Wortpaar wurden zwei Stereo-Dateien hergestellt: (i) n -silbiges Wort auf dem linken Kanal und $n+1$ -silbiges Wort auf dem rechten Kanal und (ii) $n+1$ -silbiges Wort auf dem linken Kanal und n -silbiges Wort auf dem rechten Kanal.

6.3.1.3 Versuchsaufbau

Die Versuchsdurchführung wurde durch einen Computer (Power Macintosh 6100) kontrolliert (COHEN, MACWHINNEY, FLATT & PROVOST 1993), der die Reihenfolge der Stimuluspräsentation für jeden Block neu randomisierte und für jeden Durchgang die experimentelle Bedingung, die Reaktionswahl und die Reaktionszeit des Probanden aufzeichnete. Die dichotischen Stimuli wurden über ein Paar mit dem Computer verbundenen Stereo-Kopfhörer (Sennheiser HD 475) mit einer Intensität von ca. 65 dB dargeboten.²⁵⁹ Die Versuchsanweisungen wurden den Probanden vor dem Experiment in schriftlicher Form auf dem Computermonitor präsentiert. Jede Anweisung wurde solange auf dem Bildschirm gezeigt, bis der Proband durch Tastendruck die nächste Anweisung (oder den Start des Experiments) abrief. Die Probanden wussten nicht, dass sie in jedem Durchgang gleichzeitig zwei verschiedene Wörter hören würden; sie erhielten lediglich die Anweisung, durch Tastendruck mitzuteilen, welche von zwei Antwortalternativen sie gehört hatten. Da die beiden Testwörter sehr ähnlich waren, d.h. gleich lang und weitgehend aus denselben Segmenten bestanden, wurde erwartet, dass in den meisten Fällen beide Reize zu einem einzigen Perzept fusionieren bzw. ein Reiz den anderen dominieren würde und der Proband in den meisten Durchgängen nur ein einziges Wort wahrnehmen würde. Zusätzlich wurde angenommen, dass in den Fällen, in denen der Proband zwei verschiedene Reize wahrnahm und sich dennoch für die eine oder andere Antwortalternative entscheiden musste, sich die Antwortwahl gleichmäßig auf Wort 1 und Wort 2 bzw. auf das linksohrig präsentierte und das rechtsohrig präsentierte Wort verteilen würde.

²⁵⁹ Die Intensität wurde mit einem 1/4 Zoll Bruel & Kjaer Pegelmesser mit Messverstärker 1608 gemessen.

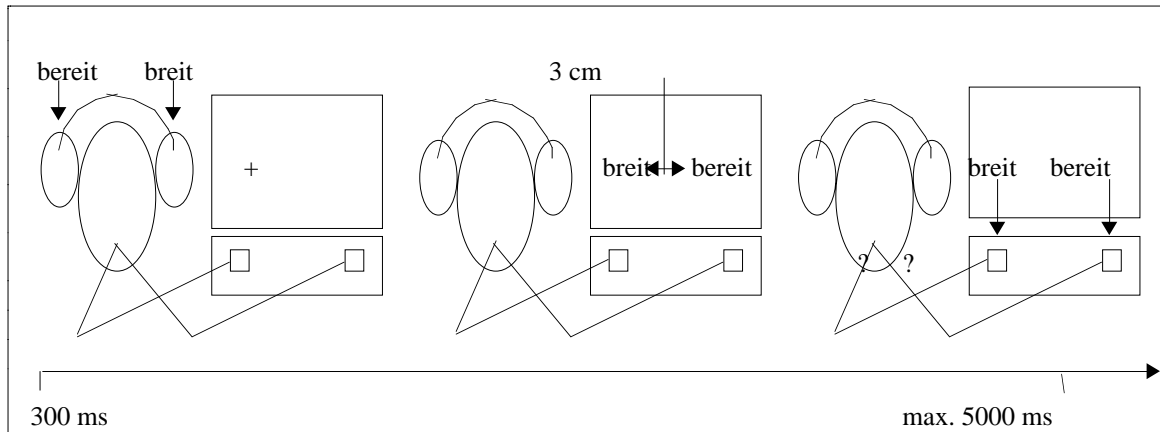


Abbildung 6-6. Versuchsaufbau in Experiment 1.

Um die Aufmerksamkeit der Probanden auf beide Ohren zu richten, begann jeder Durchgang mit einem diotisch (d.h. beiden Ohren) präsentierten 400 Hz-Ton von 300 ms Dauer. Danach wurde zunächst ein dichotisches Wortpaar (akustisch) präsentiert (50 ms Stille + Wortpaar + 50 ms Stille), dann wurden die beiden Wörter, die dichotisch dargeboten worden waren, in geschriebener Form am Computerbildschirm (visuell) präsentiert, je eines in jeder Bildschirmhälfte. Der letzte Buchstabe des linken Wortes und der erste Buchstabe des rechten Wortes waren durch einen Abstand von 3 cm voneinander getrennt. Die Probanden waren angewiesen, mit der rechten Hand die rechte Taste zu drücken, wenn sie das Wort gehört hatten, welches nach der akustischen Präsentation in der rechten Bildschirmhälfte erschien und mit der linken Hand die linke Taste zu drücken, wenn sie das Wort gehört hatten, welches in der linken Bildschirmhälfte erschien. Die beiden Antwortalternativen blieben auf dem Bildschirm, bis der Proband eine der beiden Tasten drückte, aber es wurden nur Reaktionen innerhalb der ersten 5000 ms nach dem Beginn der Darbietung der Reaktionswahl ausgewertet. Die Messung der Reaktionszeit begann mit dem Einsetzen der Darbietung der Reaktionswahl. Für jeden Durchgang wurde nur eine einzige Reaktion aufgezeichnet. Zwei aufeinanderfolgende Durchgänge wurden durch ein Inter-Trial-Intervall von 300 ms getrennt (vgl. Abbildung 6-6 und Abbildung 6-7).

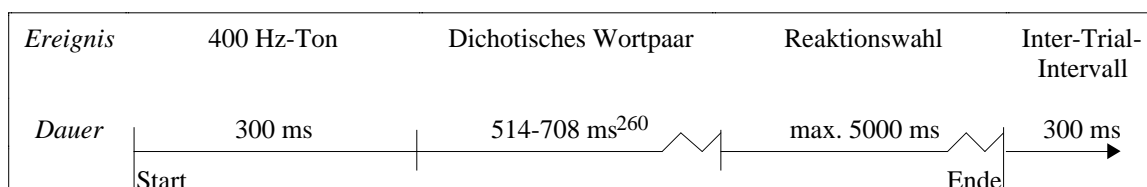


Abbildung 6-7. Ereignisse und Zeitverlauf in einem Durchgang in Experiment 1.

²⁶⁰ 50 ms Stille + Wortpaar (Dauer 414-608 ms) + 50 ms Stille.

Um eine Beeinflussung der Ergebnisse durch Worteffekte (Präferenz für Wort 1 oder Wort 2) und durch Reaktionswahleffekte (Präferenz für die linke oder rechte Bildschirmhälfte bzw. Hand) zu kontrollieren, wurde jedes der acht dichotischen Wortpaare in jedem Block unter vier Bedingungen präsentiert (siehe Tabelle 6-3).

Tabelle 6-3. Präsentationsbedingungen für jedes Wortpaar.

Bedingung	akustischer Stimulus		visuelle Reaktionswahlpräsentation	
	linkes Ohr	rechtes Ohr	linke Hälfte	rechte Hälfte
1	Wort 1	Wort 2	Wort 1	Wort 2
2	Wort 1	Wort 2	Wort 2	Wort 1
3	Wort 2	Wort 1	Wort 1	Wort 2
4	Wort 2	Wort 1	Wort 2	Wort 1

Jeder Block bestand aus 32 Durchgängen (8 Wortpaare \times 4 Bedingungen). Das Experiment bestand aus 4 Blöcken, d.h. jede Versuchsperson wurde in 128 Durchgängen getestet (4 Blöcke \times 32 Durchgänge). Vor den ausgewerteten 128 Durchgängen gab es 16 randomisierte Übungsdurchgänge, die nicht in die Auswertung eingingen.

Die Position des Kopfhörers, d.h. die Lage des linken und rechten Kanals relativ zum linken und rechten Ohr des Probanden, wurde nach jedem Block gewechselt (d.h. nach 32 Durchgängen, ohne Berücksichtigung der Übungsdurchgänge), um eventuell bestehende Unterschiede zwischen beiden Kanälen des Kopfhörers in der Signalübertragung zu kontrollieren. Die eine Hälfte der Versuchspersonen begann mit Kopfhörerposition A (rechter Kanal – rechtes Ohr), die andere Hälfte mit der Kopfhörerposition B (rechter Kanal – linkes Ohr).

Vor dem Experiment durchlief der Proband einen Priming-Block, bestehend aus 64 Durchgängen, um Einflüsse von Wortfrequenz und Wortvertrautheit auf die Reaktionswahl des Probanden im Experiment zu minimieren. In diesem Priming-Block wurden 64 Wörter in randomisierter Reihenfolge in schriftlicher Form auf dem Bildschirm präsentiert. In jedem Durchgang wurde nur je ein Wort in der Mitte des Bildschirms gezeigt (gleiche Schrift und Schriftgröße wie im Experiment). Der Proband wurde angewiesen, das präsentierte Wort zu lesen und dann eine Taste zu drücken, um das nächste Wort abzurufen. Jedes Priming-Wort wurde solange auf dem Bildschirm gezeigt, bis der Proband eine Taste drückte.

Die Liste der 64 Priming-Wörter bestand erstens aus den 8 Stimuluswortpaaren, d.h. aus 16 einzelnen Wörtern. Zu jedem der 16 Stimuluswörter wurde ein semantisch verwandtes Wort (z.B. *schmal* zu *breit*, +16 Wörter) und ein morphologisch verwandtes

Wort (z.B. *Breite* zu *breit*, +16 Wörter) hinzugefügt. Schließlich wurden 8 weitere Wortpaare präsentiert (d.h. +16 Wörter), deren Elemente sich phonologisch ähnlich zueinander verhielten wie die Stimuluswortpaare, die aber nicht als Testwörter verwandt wurden (z.B. *gleich* und *gelaicht*).

Die Voraussage, die in diesem Experiment getestet wurde, war die Annahme, dass der linksöhrig präsentierte Reiz häufiger über den rechtsöhrig präsentierten Reiz *dominieren* (s.o. 6.2.3.) würde als umgekehrt und daher die Probanden im Mittel häufiger das linksöhrig präsentierte Wort als das rechtsöhrig präsentierte Wort wahrnehmen würden.

6.3.2 Ergebnisse

Für jeden Probanden wurden für die 128 Testdurchgänge die Anzahl der Reaktionen ('Anzahl der Reaktionen') und die mittlere Reaktionszeit ('Reaktionszeit') für jedes Ohr berechnet.²⁶¹ Nur Reaktion innerhalb von 5 Sekunden nach dem Beginn der Anzeige der Reaktionswahl wurden berücksichtigt.²⁶² Alle Werte wurden auch für jedes Wortpaar getrennt aufgezeichnet (vgl. Tabelle 6-4). Außerdem wurden die mittleren Reaktionszeiten unabhängig vom Präsentationsöhr für jedes Wortpaar und für jedes einzelne Wort berechnet (vgl. Tabelle 6-5). Schließlich wurde auch berechnet, wie häufig die Probanden unabhängig vom Präsentationsöhr auf Wort 1 bzw. auf Wort 2 reagierten. Alle Werte wurden auch für jede Geschlechtergruppe getrennt aufgezeichnet.

Wie aus Tabelle 6-1 ersichtlich ist, zeigte sich für die Gesamtpopulation keine häufigere Reaktion auf das linksöhrig präsentierte Wort als auf das rechtsöhrig präsentierte Wort. Die Gruppe der männlichen Probanden hingegen reagierte im Mittel häufiger auf das linksöhrig präsentierte Wort als auf das rechtsöhrig präsentierte, während die Gruppe der weiblichen Probanden im Mittel häufiger auf das rechtsöhrig präsentierte Wort als auf das linksöhrig präsentierte Wort reagierte (vgl. auch Abbildung 6-8). Die Standardabweichung für die Gruppe der weiblichen Probanden war nahezu doppelt so hoch wie die Standardabweichung für die Gruppe der männlichen Probanden.

²⁶¹ Da der Proband sich in jedem Durchgang für eines der beiden präsentierten Wörter entscheiden musste und Mehrfachantworten nicht möglich waren, wurde für jeden Durchgang nur eine einzige Reaktionswahl und Reaktionszeit aufgezeichnet: entweder 'Reaktion auf das Wort, das dem linken Ohr präsentiert wurde', oder 'Reaktion auf das Wort, das dem rechten Ohr präsentiert wurde'. Dies bedeutete auch, dass die 'Trefferquote' der Probanden immer bei 100% lag, da keine Fehler gemacht werden konnten und die Probanden in jedem Durchgang reagieren mussten.

²⁶² Weniger als 0.5 % der Fälle (30 von 6144 Fällen) fielen dadurch aus der Auswertung.

Tabelle 6-4. Mittelwerte und Standardabweichungen der Variablen 'Anzahl der Reaktionen' und 'Reaktionszeit' in Abhängigkeit vom Präsentationsohr für die Gesamtpopulation, für die männliche und weibliche Probandengruppe und für die Gesamtpopulation nach Wortpaar (1 bis 8).

	Mittelwerte (Mittelw.) und Standardabweichungen (S.D.)							
	'Anzahl der Reaktionen'				'Reaktionszeit' in ms			
	links		rechts		links		rechts	
	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.
Gesamt	63.58	3.30	63.79	4.18	914.91	272.44	893.90	260.90
Männer	64.96	2.33	62.58	2.75	919.20	300.13	904.09	278.51
Frauen	62.21	3.59	65.00	5.01	910.62	248.11	883.72	247.62
Gesamt 1	8.04	1.64	7.81	1.73	925.82	316.66	887.21	275.86
Gesamt 2	7.73	1.28	8.13	1.33	908.82	340.90	880.12	335.98
Gesamt 3	8.15	1.07	7.81	1.14	920.24	318.80	891.25	318.15
Gesamt 4	7.73	1.25	8.15	1.34	959.58	357.33	899.67	312.94
Gesamt 5	8.02	1.38	7.92	1.41	916.76	356.55	855.92	257.55
Gesamt 6	7.71	1.22	8.23	1.21	908.65	281.05	956.27	350.19
Gesamt 7	8.08	1.25	7.90	1.28	945.65	317.46	924.38	295.43
Gesamt 8	8.12	0.98	7.85	0.92	814.75	269.99	853.68	278.99

Tabelle 6-5. Mittelwerte (Mittelw.) und Standardabweichungen (S.D.) der Variable Reaktionszeit für alle Wortpaare und nach Wortpaaren getrennt, jeweils für das n-silbige Wort (Wort 1) und das n+1-silbige Wort (Wort 2).

	Mittelwerte und Standardabweichungen der Variable 'Reaktionszeit' in ms					
	Gesamt		Wort 1 (n-silbig)		Wort 2 (n+1-silbig)	
	Mittelw.	S.D.	Mittelw.	S.D.	Mittelw.	S.D.
Gesamt	967.92	433.93	1053.74	500.44	897.25	355.81
Gesamt 1	947.27	354.97	979.58	385.47	916.33	324.26
Gesamt 2	945.85	402.35	1027.33	461.37	879.65	337.83
Gesamt 3	950.79	409.89	1031.31	514.68	877.43	268.71
Gesamt 4	1002.21	436.78	1138.87	561.80	922.49	324.82
Gesamt 5	916.08	337.90	932.26	291.59	901.31	377.90
Gesamt 6	1062.29	581.41	1164.15	582.87	973.43	571.48
Gesamt 7	1003.95	412.32	1106.74	478.47	901.15	305.51
Gesamt 8	910.60	496.20	1094.19	739.15	807.33	234.62

Für die Gesamtpopulation und für beide Geschlechtergruppen fand sich eine im Mittel schnellere Reaktion auf rechtsohrig präsentierte Wörter als auf linksohrig präsentierte Wörter (vgl. Tabelle 6-4). Wie aus Tabelle 6-5 ersichtlich ist, variierten die Reaktions-

zeiten in Abhängigkeit vom Wortpaar und in Abhängigkeit vom wahrgenommenen Wort. Die Reaktionszeiten für Reaktionen auf das n+1-silbige Wort (Wort 2) waren für alle Wortpaare kürzer als die Reaktionszeiten für Reaktionen auf das n-silbige Wort (Wort 1). Darüberhinaus nahmen die Probanden beider Geschlechtergruppen häufiger das n+1-silbige (Wort 2) als das n-silbige Wort (Wort 1) wahr (vgl. Tabelle 6-6).

Tabelle 6-6. Mittelwerte und Standardabweichungen der Variable 'Anzahl der Reaktionen' in Abhängigkeit vom gehörten Wort (Wort 1 oder Wort 2) für die Gesamtpopulation und nach Geschlechtergruppen getrennt.

	Mittelwerte und Standardabweichungen der Variablen 'Anzahl der Reaktionen'			
	Wort 1 (n-silbig)		Wort 2 (n+1-silbig)	
	Mittelwert	S.D.	Mittelwert	S.D.
Gesamt	40,62	23,29	87,60	24,09
Männer	39,92	22,86	87,63	22,97
Frauen	41,35	24,22	87,58	25,66

Die Ergebnisse für die Variablen 'Anzahl der Reaktionen' und 'Reaktionszeit' wurden einer dreifaktoriellen Varianzanalyse unterzogen, mit der Variable 'Geschlecht' (männlich, weiblich) als Gruppenvariable und den Variablen 'Ohr' (links, rechts) und 'Wortpaar' (Wortpaar 1 bis 8) als *between-subjects*-Faktoren.

Anzahl der Reaktionen: Die Haupteffekte für 'Geschlecht' und 'Ohr' waren nicht signifikant ('Geschlecht': $F(1,46) = 0.19$, n.s.; 'Ohr': $F(1,46) = 0.05$, n.s.), wohl aber die Zweifach-Interaktion für 'Geschlecht' und 'Ohr' ($F(1,46) = 7.30$, $p < 0.01$). Die Zweifach-Interaktion für 'Ohr' und 'Wortpaar' und die Dreifach-Interaktion für 'Geschlecht', 'Ohr' und 'Wortpaar' waren nicht signifikant (alle F-Werte(1,46) < 0.91 , n.s.) (vgl. Tabelle 6-7).

Zusätzlich wurden zwei separate einfaktorielle Varianzanalysen für jede Geschlechtergruppe durchgeführt. Während sich für die Gruppe der männlichen Versuchspersonen ein signifikanter Effekt für den Faktor 'Ohr' zeigte ($F(1,23) = 5.69$, $p < 0.025$), konnte für die Gruppe der weiblichen Probanden kein signifikanter Unterschied zwischen beiden Ohren demonstriert werden ($F(1,23) = 2.92$, n.s.).

Reaktionszeit: Eine dreifaktorielle Varianzanalyse ergab keine signifikanten Haupteffekte oder Interaktionen (alle F-Werte(1,46) < 2.77 , n.s.), abgesehen von einem signifikanten Haupteffekt für den Faktor 'Wortpaar' ($F(7,322) = 2.66$, $p < 0.05$) (vgl. Tabelle 6-8).

Tabelle 6-7. Zusammenfassung der ANOVA für die Variable 'Anzahl der Reaktionen'.

Dreifaktorielle Varianzanalyse			
Abhängige Variable: 'Anzahl der Reaktionen'			
Haupteffekte	Geschlecht	$F(1,46) = 0.19$	n.s.
	Ohr	$F(1,46) = 0.05$	n.s.
	Wortpaar	$F(7,322) = 1.89$	n.s.
Zweifach-Interaktionen	Geschlecht \times Ohr	$F(1,46) = 7.30$	$p < 0.01$
	Geschlecht \times Wortpaar	$F(7,322) = 1.58$	n.s.
	Ohr \times Wortpaar	$F(7,322) = 0.91$	n.s.
Dreifach-Interaktion	Geschlecht \times Ohr \times Wortpaar	$F(7,322) = 0.48$	n.s.

Tabelle 6-8. Zusammenfassung der ANOVA für die Variable 'Reaktionszeit'.

Dreifaktorielle Varianzanalyse			
Abhängige Variable: 'Reaktionszeit'			
Haupteffekte	Geschlecht	$F(1,46) = 0.04$	n.s.
	Ohr	$F(1,46) = 2.77$	n.s.
	Wortpaar	$F(7,322) = 2.66$	$p < 0.011$
Zweifach-Interaktionen	Geschlecht \times Ohr	$F(1,46) = 0.27$	n.s.
	Geschlecht \times Wortpaar	$F(7,322) = 0.63$	n.s.
	Ohr \times Wortpaar	$F(7,322) = 1.65$	n.s.
Dreifach-Interaktion	Geschlecht \times Ohr \times Wortpaar	$F(7,322) = 0.84$	n.s.

Schließlich wurden zwei zweifaktorielle Varianzanalysen für die Variablen 'Reaktionszeit' und 'Anzahl der Reaktionen' mit der Gruppenvariable 'Geschlecht' und der Variablen 'gehörtes Wort' (Wort 1, Wort 2) als *between-subject*-Faktor durchgeführt. Für die Variable 'Reaktionszeit' ergab diese Varianzanalyse keinen signifikanten Geschlechtseffekt und keine signifikante Interaktion (beide F-Werte(1,45) < 0.22, n.s.), aber einen Haupteffekt für den Faktor 'gehörtes Wort' ($F(1,45) = 44,62$, $p < 0.001$). Auch für die Variable 'Anzahl der Reaktionen' zeigte sich kein Geschlechtseffekt und keine Interaktion (alle F-Werte(1,45) < 0.17, n.s.), im Gegensatz zu einem Haupteffekt für den Faktor 'gehörtes Wort' ($F(1,45) = 13.74$, $p < 0.001$).

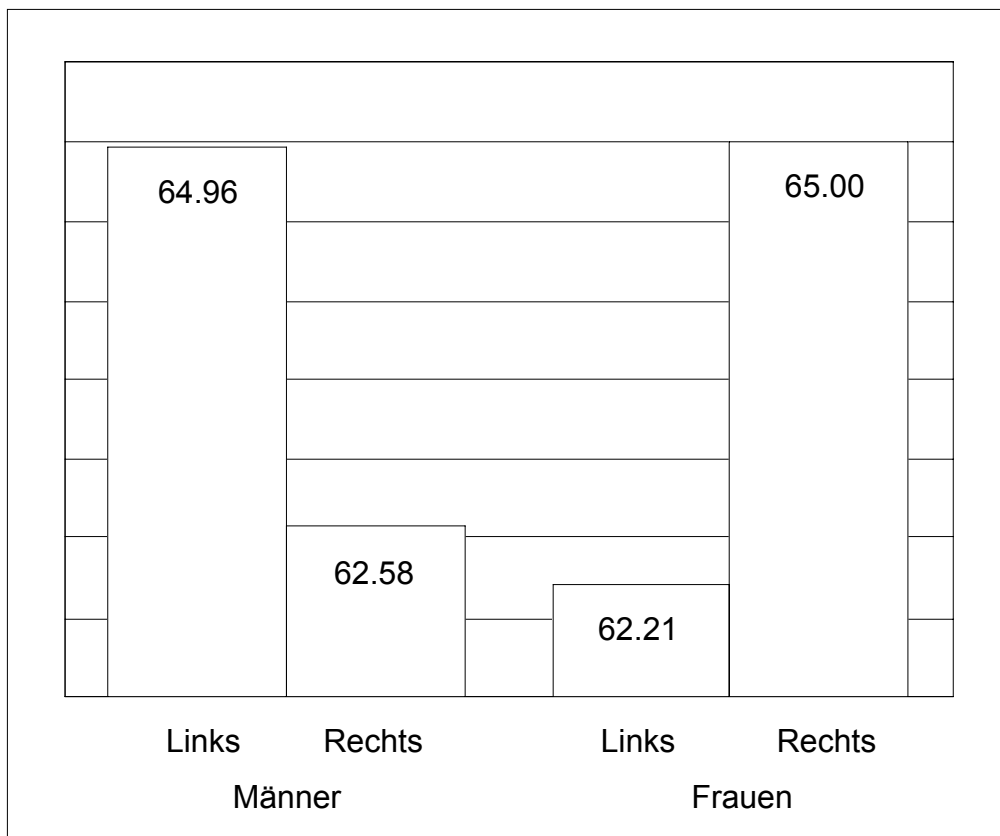


Abbildung 6-8. Mittelwerte der Variablen 'Treffer' je Ohr nach Geschlechtergruppen in absoluten Zahlen,

6.3.3 Diskussion

In diesem Experiment konnte ein Geschlechtsunterschied in der Lateralität für die Verarbeitungsfunktion 'Lokalisierung der Sonoritätsmaxima in einer Äußerung' nachgewiesen werden, mit einem Links-Ohr-Vorteil für die Gruppe der männlichen Probanden und keinem Ohr-Vorteil für die Gruppe der weiblichen Probanden. Ein Geschlechtsunterschied in der Lateralität bzw. ein Ohr-Vorteil für die Gruppe der männlichen Probanden konnte nur für die Variable 'Anzahl der Reaktionen' nachgewiesen werden, aber nicht für die Variable 'Reaktionszeit'. Daneben zeigte sich, dass die Reaktionszeiten in Abhängigkeit vom Faktor 'Wortpaar' und vom gehörten Wort (Wort 1 oder Wort 2) variierten. Die Reaktionszeiten auf n+1-silbige Wörter waren dabei kürzer als die Reaktionszeiten auf n-silbige Wörter. Dieser schnelleren Reaktion auf n+1-silbige Wörter entsprach eine allgemeine Präferenz für die Wahrnehmung n+1-silbiger Wörter gegenüber n-silbigen Wörtern, d.h. Wort 2 dominierte häufiger über Wort 1 als umgekehrt.

Der Geschlechtsunterschied in der Lateralität, der sich in diesem Experiment zeigte, weist auf eine rechtshemishärische Spezialisierung für die Verarbeitungsfunktion 'Lokalisierung der Sonoritätsmaxima' für die Gruppe der männlichen Probanden hin, im Gegensatz zu einer eher bilateralen Repräsentation für die Gruppe der weiblichen Probanden. Dieses Ergebnis unterstützt die Hypothese einer stärkeren Lateralisierung für rezeptive und produktive Sprachfunktionen bei Männern im Vergleich zu Frauen (vgl. MCGLONE 1977, 1980, BRYDEN 1982, KIMURA 1983, HAUSMANN ET AL. 1998, WOLF & GOODALE 1987, für einen Forschungsüberblick vgl. HISCOCK ET AL. 1994).²⁶³ Auch mit bildgebenden Verfahren konnte in Studien zur phonologischen Sprachverarbeitung eine stärkere Lateralisierung der Aktivitätserhöhung bei Männern gegenüber einer eher bilateralen Aktivitätserhöhung bei Frauen nachgewiesen werden (PUGH ET AL. 1994, SHAYWITZ ET AL. 1995). Die Ergebnisse dieses Experimentes sind also zunächst so zu bewerten, dass die Verarbeitungsstrategien, mittels welcher die hier verwendete Versuchsaufgabe ausgeführt wurde, bei Männern eher rechtshemisphärisch lateralisiert sind und bei Frauen eher bilateral repräsentiert. Dieser Geschlechtsunterschied kann auf zweierlei Weise interpretiert werden. Einerseits könnte angenommen werden, dass Männer und Frauen dieselben Verarbeitungsstrategien angewendet haben, diese jedoch unterschiedlich lateralisiert sind. Andererseits wäre es möglich, dass Männer und Frauen die Versuchsaufgabe auf der Grundlage unterschiedlicher Verarbeitungsstrategien durchgeführt haben. Diese zweite Möglichkeit soll im folgenden näher betrachtet werden.

Zunächst muss darauf hingewiesen werden, dass die in diesem Experiment verwendeten Stimuli sich *nicht nur* in der Zahl der Sonoritätsmaxima unterscheiden, sondern dass das n+1-silbige Wort gegenüber dem n-silbigen Wort über einen zusätzlichen Vokal verfügt, d.h. über einen sehr kurzen reduzierten Zentralvokal (vgl. hierzu auch Abbildung 6-4 und Abbildung 6-5). In Kapitel 2 (s.o. 2.4) wurde darauf hingewiesen, dass abgesehen von quasihomophonen Lautfolgen (s.o. 3.2.2), in denen sich ein Segmentpaar nur im Merkmal [silbisch] unterscheidet (z.B. dt. *kann* und *kannen* in einer bestimmten Aussprachevariante),²⁶⁴ es im Deutschen und auch in anderen Sprachen keine Lautfolgen gibt, die sich nur in der Zahl der Silbennuklei bzw. in der Zahl der Sonoritätsmaxima unterscheiden, die aber in segmentaler Hinsicht nicht differieren. Die Ursache hierfür liegt darin, dass die Relation zwischen segmentalem Sonoritätswert und

²⁶³ Für die Annahme, dass Frauen für die *nichtsprachliche* auditive Verarbeitung stärker lateralisiert sind als Männer vgl. BOUCHER & BRYDEN 1997, PIAZZA 1980.

²⁶⁴ Für Minimalpaare wie *kann* und *kannen*, *sein* und *seinen* hatte sich in Vorversuchen gezeigt, dass diese nicht mehr stabil differenziert werden können, wenn zum Zwecke der dichotischen Präsentation ihre Länge angeglichen wird. Daher wurden im hier beschriebenen Versuch Wortpaare vom Typ *breit* und *bereit* verwendet.

Silbenstrukturposition in einer Lautfolge eindeutig ist, weil eine gegebene Lautfolge nur auf eine einzige Weise silbifiziert werden kann,²⁶⁵ d.h. so, dass das sonoranteste Segment dem Silbennukleus zugeordnet wird (s.o. 2.3). Diese Tatsache bedeutet für die Sprachwahrnehmung (s.o. 4.4), dass das phonologisch-phonetische Merkmal Sonorität als Grundlage für die Segmentation eines Signals in Silben und als Grundlage für die Lokalisierung der Silbengipfel dienen kann. Auf der anderen Seite führt diese Tatsache dazu, dass der phonologisch-phonetische Sonoritätsverlauf für experimentelle Zwecke nur schwer isoliert werden kann.

Bezogen auf das hier durchgeführte Experiment bedeutet dies, dass die beiden Wörter vom Typ *greift* und *gereift* nicht nur dadurch unterschieden werden können, dass auf einer suprasegmentalen Ebene die Sonoritätsmaxima lokalisiert werden, sondern auch dadurch, dass auf einer segmentalen Ebene die Vokale identifiziert werden: Das Wort *greift* besteht aus dem Diphthong /ai/, das Wort *gereift* hingegen aus einem zusätzlichen Schwa und dem Diphthong /ai/. Es wäre also möglich, dass bei der Erkennung der Wörter in diesem Experiment zwei Verarbeitungsstrategien miteinander konkurrieren. Die erste Verarbeitungsstrategie ist die Lokalisierung der Sonoritätsmaxima, die zweite Verarbeitungsstrategie ist die Identifikation der Vokale. Wie in Kapitel 5 gezeigt wurde, können für die Identifikation von Vokalen nur dann Ohr-Vorteile nachgewiesen werden, wenn die experimentellen Bedingungen so variiert werden, dass die Probanden nicht das Frequenzspektrum zu einem Zeitpunkt, sondern die Frequenzänderungen in der Zeit zur Identifikation der Vokale nutzen (DARWIN 1971, HAGGARD & PARKINSON 1971, WHITE & HOUSE 1973, SPELLACY & BLUMSTEIN 1970). Die Verarbeitungsfunktion 'Identifikation der Vokale' ist also unter normalen Verarbeitungsbedingungen nicht lateralisiert.

Da erstens in diesem Experiment für die Gruppe der weiblichen Probanden kein Ohr-Vorteil demonstriert werden konnte, zweitens die Stimuli auch auf der Grundlage der Vokale identifiziert werden können und drittens die Verarbeitung von Vokalen nicht lateralisiert ist, könnte angenommen werden, dass die weiblichen Probanden keine sonoritätsbasierte, sondern eine vokalbasierte Verarbeitungsstrategie angewendet haben. Die Ergebnisse dieses Experimentes, mit einem Geschlechtsunterschied in der Lateralität und einer rechtshemisphärischen Spezialisierung bei Männern im Gegensatz zu eher bilateraler Aktivierung bei Frauen könnten demnach so interpretiert werden, dass Männer eine andere *Verarbeitungsstrategie* angewendet haben als Frauen: während die

²⁶⁵ Abgesehen von Aussprachevarianten wie *edles*: /e.[d]les./ und /e[t].les./, *Handlung*: /.han[t].lung./ – /.han.[d]lung./, vgl. hierzu WIESE 1996 (Punkte bezeichnen Silbengrenzen).

Männer die Stimuli auf der Grundlage der Sonoritätskontur verarbeitet haben, haben die Frauen die Stimuli auf der Basis der Vokale identifiziert.

In Kapitel 4 wurde gezeigt, dass die Anwendung unterschiedlicher Sprachverarbeitungsstrategien von verschiedenen Faktoren abhängt: einerseits von akustisch-phonetischen Faktoren wie der Akzentuiertheit der initialen Silbe oder der akustischen Transparenz des Vokalsystems, andererseits von phonologischen Faktoren wie dem phonologischen System der Muttersprache (s.o. 4.3). Insbesondere haben sich für die Sprecher verschiedener Sprachen unterschiedliche Strategieeffekte messen lassen (CUTLER ET AL. 1992, s.o. 4.3.4). Ausgehend von den Ergebnissen dieses Experimentes könnte also die Hypothese formuliert werden, dass auch das Geschlecht eines Sprechers (bzw. die damit zusammenhängenden kognitiven und biologischen Unterschiede) ein Faktor ist, der die Verwendung unterschiedlicher Verarbeitungsstrategien in der Sprachwahrnehmung beeinflusst. Diese Annahme erscheint angesichts der Tatsache plausibel, dass sich für Frauen in sprachbezogenen Experimenten nicht nur andere Lateralitätseffekte zeigen lassen (vgl. BRYDEN 1982, HAUSMANN ET AL. 1998, WOLF & GOODALE 1987, KOLB & WHISHAW 1993:191ff, SPRINGER & DEUTSCH 1995:211ff), sondern dass sich für Männer und Frauen auch qualitative Unterschiede im Sprachverhalten finden (vgl. KIMURA 1996, HALPERN 1992, KOLB & WHISHAW 1993:196ff).

Legt man die in Kapitel 4 formulierte Annahme zugrunde, dass Verarbeitungsstrategien in der Sprachwahrnehmung auf Unterschiede in der funktionalen Dynamik von phonologischen Repräsentationen auf verschiedenen Ebenen der Sprache bezogen werden können (s.o. 4.2.3), so ist die Aussage, dass Frauen eine vokalbezogene Verarbeitungsstrategie anwenden und Männer eine sonoritätsbezogene Verarbeitungsstrategie, nicht gleichbedeutend mit der Aussage, dass Frauen nur Vokale repräsentieren, aber nicht den Sonoritätsverlauf und Männer nur den Sonoritätsverlauf, aber nicht die Vokale. Vielmehr ist diese Aussage so zu verstehen, dass beide Probandengruppe beide Arten von Information nutzen, um das Sprachsignal zu segmentieren und um invariante Einheiten zu erkennen, dass sich aber Männer und Frauen in der funktionalen Dynamik und in der funktionalen Gewichtung der Repräsentationen der Vokale und der Repräsentation der Sonoritätskontur unterscheiden. Diese Unterschiede könnten dazu führen, dass die eine Probandengruppe die Versuchsaufgabe eher auf der Basis der einen Art von Repräsentation durchführt, weil diese stabiler oder schneller verfügbar ist, die andere Probandengruppe jedoch auf der Basis einer anderen Art von Repräsentation. Die Beantwortung der Frage, ob sich der in diesem Experiment demonstrierte Geschlechtsunterschied in der Lateralität tatsächlich, wie hier angenommen, auf Unterschiede in den Verarbeitungsstrategien beider Probandengruppen zurückführen lässt, erfordert weitere experimentelle Studien. In diesem Zusammenhang könnte untersucht werden,

welchen Einfluss eine akustische Manipulation der Stimuli auf den Geschlechtsunterschied in der Lateralität hätte. Eine Tiefpassfilterung der Reize (s.o. 5.4.3.1, Anm.) würde die Verständlichkeit der Wörter beseitigen und somit auch ihre Identifikation oder Diskrimination auf der Grundlage der Vokale unmöglich machen. Unter der Annahme, dass die weiblichen Probanden in diesem Experiment eine vokalbasierte Verarbeitungsstrategie angewendet haben und deshalb anders als die männlichen Probanden keinen Ohr-Vorteil aufwiesen, wäre daher zu erwarten, dass sich nach einer Tiefpassfilterung der Reize kein Unterschied in der Lateralität für die Diskrimination von *breit* und *bereit* zeigen sollte.

Wie die vorangegangenen Ausführungen gezeigt haben, können die in diesem Experiment verwendeten Wörter auf der Basis von mindestens zwei verschiedenen Verarbeitungsstrategien unterschieden werden. Wie ist vor diesem Hintergrund also die für die Gruppe der männlichen Probanden demonstrierte rechtshemisphärische Spezialisierung zu bewerten? Wie die Ausführungen in Kapitel 5 gezeigt haben, lässt sich für die die Erkennung von Konsonanten kein Links-Ohr-Vorteil nachweisen (CUTTING 1974, DARWIN 1971, HAYDEN 1979, JÄNCKE 1992a, b, JÄNCKE 1993), für die Identifikation von Vokalen lässt sich nur dann ein Links-Ohr-Vorteil zeigen, wenn die Vokal-Erkennungs-Aufgabe in eine Erkennungsaufgabe für nichtsprachliche Reize eingebettet wird (SPELLACY & BLUMSTEIN 1970). Der hier demonstrierte Links-Ohr-Vorteil für die männlichen Probanden lässt sich also nicht auf eine phonembasierte Verarbeitungsstrategie zurückführen, da die Probanden eine sprachbezogene Aufgabe durchführen mussten, d.h. die Identifikation von bedeutungsvollen deutschen Wörtern.

Gegen die Annahme, dass der für die männlichen Probanden demonstrierte Links-Ohr-Vorteile auf eine rechtshemisphärische Spezialisierung der Sonoritätskontur hinweist, könnte man weiter einwenden, dass sich n-silbige und n+1-silbige Wörter darin unterscheiden, dass die n-silbigen Wörter mit einer betonten Silbe beginnen und die n+1-silbigen Wörter mit einer unbetonten. Es wäre also denkbar, dass die beiden Wörter auch auf der Basis ihrer Akzentstruktur unterschieden werden können, auf der Grundlage der Detektion einer initialen betonten Silbe im Fall der n-silbigen Wörter und einer inialen unbetonten Silbe im Falle der n+1-silbigen Wörter.²⁶⁶ Wie die Ausführungen in Kapitel 5 gezeigt haben, wurde für die Verarbeitung des Akzentmusters einer Äußerung häufiger ein Rechts-Ohr-Vorteil als ein Links-Ohr-Vorteil demonstriert

²⁶⁶ Wie auch im Englischen sind betonte Silben im Deutschen immer *starke* Silben (s.o. 2.2.3, 4.3.3), d.h. eine betonte Silbe lässt sich daran erkennen, dass sie einen langen Vokal wie /a:/ und /i:/ oder einen Diphthong wie /ai/ aufweist. Ein reduzierter Vokal wie Schwa weist hingegen immer auf eine unbetonte Silbe hin. Daher könnte man sagen, dass eine akzentbasierte Verarbeitungsstrategie in diesem Experiment dieselbe Art von Information verarbeitet wie eine vokalbasierte Verarbeitungsstrategie.

(BEHRENS 1985, EMMOREY 1987, BAUM ET AL. 1982). Daher ist die Annahme, dass die Stimuli auf der Basis der Akzentstruktur und nicht auf der Basis der Sonoritätskonkur identifiziert wurden, wenig wahrscheinlich.

Wie bereits bemerkt wurde, fanden sich für die Variable 'Reaktionszeit' keine Ohr-Vorteile. Jedoch zeigte sich, dass die Reaktionszeiten in Abhängigkeit vom Wort, welches wahrgenommen wurde, variierten. Diese Varianz zwischen den einzelnen Wörtern ist darauf zurückzuführen, dass Faktoren wie Wortfrequenz, Wortart und Abstraktheit oder Konkretheit der Bedeutung bei der Auswahl der Stimuli nicht berücksichtigt wurden. Da in diesem Experiment ein phonologischer Kontrast untersucht wurde, der sich im Deutschen in nur sehr wenigen Wortpaaren findet, d.h. ein Unterschied in der Zahl der Sonoritätsmaxima bei möglichst identischer segmentaler Struktur, war es nicht möglich, diese Faktoren zusätzlich zu kontrollieren. Das Problem der Reaktionszeitenvarianz könnte durch die Verwendung von deutschen Pseudo-Wörtern gelöst werden, welche dieselben phonologischen Kontraste aufweisen, die jedoch keine Häufigkeitsdimension, Wortart- oder Konkretheitsdimension besitzen.²⁶⁷

Schließlich zeigte sich in diesem Experiment eine schnellere Reaktion der Probanden auf den Reiz, wenn sie das n+1-silbige Wort wahrgenommen hatten, als wenn sie das n-silbige Wort wahrgenommen hatten. Unter der Annahme, dass die zusätzliche Silbe in n+1-silbigen Wörtern gegenüber n-silbigen Wörtern (bei genau gleicher absoluter Länge) zu mehr phonologischer Komplexität führt, die mehr Verarbeitungsaufwand für die Worterkennung erfordern sollte, ist dieses Ergebnis erstaunlich. Andererseits ist zu bemerken, dass das n-silbige Wort (z.B. /gr-ei-ft./)²⁶⁸ mit einer CCV-Silbe mit komplexem Anlaut beginnt, das n+1-silbige Wort (z.B. /g-e.r-ei-ft./) jedoch mit zwei CV-Silben mit nur einem Anlaut-Konsonanten. Daher könnte es auch so sein, dass nicht die Komplexität des Gesamtwortes der Faktor ist, welcher den Verarbeitungsaufwand (gemessen als Verarbeitungszeit) bestimmt, sondern die Komplexität der initialen Silbe.

Außerdem nahmen die Probanden das n+1-silbige Wort nicht nur schneller wahr als das n-silbige Wort, sondern auch häufiger. Dies bedeutet, dass das n+1-silbige Wort in vielen Fällen unabhängig vom Präsentationsohr über das n-silbige Wort dominierte. Dieser Effekt ist wahrscheinlich auf akustisch-phonetische Faktoren zurückzuführen (vgl. Abbildung 6-4 und Abbildung 6-5), z.B. auf die Tatsache, dass das n+1-silbige Wort, das ein zusätzliches Sonoritätsmaximum aufweist, am Wortanfang deutlich lauter ist als das n-silbige Wort, was dazu führen könnte, dass die Aufmerksamkeit des Pro-

²⁶⁷ Diese Pseudo-Wörter würden aber auch nur phonologisch, nicht semantisch verarbeitet.

²⁶⁸ Punkte indizieren Silbengrenzen, der Bindestrich trennt Anlaut, Nukleus und Coda voneinander.

banden auf das lautere Wort gerichtet wird. Dieser Effekt lässt sich jedoch nicht vermeiden, wenn man die Verarbeitung zweier Reize untersuchen will, von denen einer zu einem früheren Zeitpunkt als der andere ein Sonoritätsmaximum aufweist, die sich also in der Zahl der Sonoritätsmaxima unterscheiden. Die Dominanz des Reizes mit dem zusätzlichen Sonoritätsmaximum hat im hier beschriebenen Experiment vermutlich dazu geführt, dass die gemessenen Ohr-Unterschiede für die Variable ‘Anzahl der Reaktionen’ kleiner waren, als sie es bei einem geringeren Ausmaße von akustisch-phonetisch determinierter Stimulusdominanz hätten sein können.

Zusammenfassend lässt sich sagen, dass dieses Experiment die Hypothese, dass die Verarbeitungsfunktion ‘Lokalisierung der Sonoritätsmaxima’ rechtshemisphärisch lateralisiert ist, bestätigt hat. Der Geschlechtsunterschied in der Lateralität, der sich in diesem Experiment fand, wurde durch die Annahme erklärt, dass die männlichen Probanden eine rechtshemisphärisch lateralisierte sonoritätsbasierte Verarbeitungsstrategie angewendet haben, die weiblichen Probanden hingegen eine bilateral repräsentierte Vokalidentifikationsstrategie. Die Beantwortung der Frage, ob der Geschlechtsunterschied in der Lateralität auf Unterschiede in den Verarbeitungsstrategien zurückzuführen ist, erfordert weitere experimentelle Untersuchungen. Eine zweite Frage, die in weiterführenden Studien zu untersuchen ist, bezieht sich darauf, ob der hier demonstrierte Ohr-Vorteil auf der kognitiven bzw. linguistischen Verarbeitung der Stimuli basiert, oder ob er auf Lateralitäten in der Verarbeitung der zugrundeliegenden akustischen Eigenschaften des Sprachsignals zurückzuführen ist. Weiter oben (s.o. 5.4.4) wurde hervorgehoben, dass aus verschiedenen Gründen davon auszugehen ist, dass funktionale Lateralitäten für komplexe sprachliche Leistungen stets eine Funktion sowohl der Verarbeitung der zugrundeliegenden akustischen Eigenschaften des Signals als auch der spezifisch kognitiven oder linguistischen Komponente dieser Funktion ist. Unter diesem Aspekt sind auch die Ergebnisse des zweiten Experimentes von Interesse, in welchem statt bedeutungsvoller deutscher Wörter sinnlose Silbenfolgen verwendet wurde. Dieses Experiment wird im nächsten Abschnitt beschrieben.

6.4 Experiment 2

6.4.1 Methode

6.4.1.1 Stimuli

Während in Experiment 1 die Lateralisierung der Verarbeitungsfunktion ‘Lokalisierung der Sonoritätsmaxima in einer Äußerung’ mit *bedeutungsvollen Wörtern* untersucht wurde, dienten in Experiment 2 *bedeutungslose Silbenfolgen* als Stimuli. In noch einem weiteren Punkt unterschied sich Experiment 2 von Experiment 1: in Experiment 1 bestand die Aufgabe der Probanden darin, einen wahrgenommenen Sprachreiz als deutsches Wort zu identifizieren, in Experiment 2 hingegen sollten die Probanden möglichst schnell eine Zielfolge mit einer bestimmten Zahl von Sonoritätsmaxima detektieren.

Die Versuchsstimuli für das Experiment 2 bestanden aus dichotischen Paaren von ein- bis fünfsilbigen Iterationen der Silbe /na/, d.h. aus dichotischen Paaren der Folgen /na/ (‘einsilbig’), /na-na/ (‘zweilsilbig’), /na-na-na/ (‘dreisilbig’), /na-na-na-na/ (‘viersilbig’) und /na-na-na-na-na/ (‘fünfsilbig’). Als Zielfolge wurde die dreisilbige Folge /na-na-na/ ausgewählt, die übrigen ein- und zweisilbigen bzw. vier- und fünfsilbigen Folgen dienten als Distraktoren. In einem Vorversuch waren zunächst die dreisilbige Folge /na-na-na/ als Zielfolge und die zweisilbige (/na-na/) und die viersilbige (/na-na-na-na/) Folge als Distraktoren verwendet worden. Hierbei zeigte sich, dass die Zielfolgen-Detektionsrate nahe 100% lag, so dass eventuell bestehende Ohr-Vorteile zumindest in Bezug auf die Variable ‘Detektionsrate’ nur schwer nachzuweisen gewesen wären. Aus diesem Grunde wurde das Stimulus-Inventar um die einsilbige Distraktor-Folge /na/ und die fünfsilbige Distraktor-Folge /na-na-na-na-na/ ergänzt. Die Stimuli für das Experiment 2 wurden mit derselben technischen Ausrüstung aufgenommen und weiterbearbeitet und von derselben Sprecherin gesprochen wie die Stimuli für das Experiment 1.

In diesem Experiment wurden fünf verschiedene Silbenfolgen präsentiert: einsilbige, zweisilbige, dreisilbige, viersilbige und fünfsilbige Silbenfolgen. Da erstens die Stimuli dichotisch präsentiert wurden und aus diesem Grunde genau gleich lang sein mussten und da sie zweitens die durchschnittliche Länge natürlichsprachiger Silben nicht wesentlich überschreiten oder unterschreiten sollten,²⁶⁹ war es nötig, die Stimulussilbenfolgen in drei Gruppen mit unterschiedlicher Gesamtlänge einzuteilen. Jede der

²⁶⁹ Diese Bedingung wurde gestellt, weil in diesem Experiment eine Verarbeitungsfunktion untersucht werden soll, von der angenommen wurde, dass sie zur Menge der Verarbeitungsstrategien für natürlichsprachliche Äußerungen gehört. Wenn sich der dargebotene Reiz in akustisch-phonetischer Hinsicht von den üblicherweise verarbeiteten natürlichsprachlichen Äußerungen unterscheidet, so ist zu erwarten, dass andere, möglicherweise nicht sprachverarbeitungsspezifische Strategien angewendet werden.

drei Gruppen musste dabei auch die dreisilbige Zielfolge /na-na-na/ enthalten. Um also zu gewährleisten, dass jedes einzelne /na/ ungefähr im Längenbereich natürlichsprachlicher Silben liegt (ca. 200-400ms) und jede Längengruppe die dreisilbige Zielfolge enthält, wurden dichotische Stimuluspaare in drei unterschiedlichen Längen (600 ms, 800 ms und 1000 ms) hergestellt (vgl. Tabelle 6-9). Für die Länge 1 (600 ms) wurden die einsilbige, zweisilbige und dreisilbige Folge (d.h. /na/, /na-na/ und /na-na-na/), für Länge 2 (800 ms) die zweisilbige, dreisilbige und viersilbige Folge (d.h. /na-na/, /na-na-na/ und /na-na-na-na/) und für Länge 3 die dreisilbige, viersilbige und fünfsilbige Folge (d.h. /na-na-na/, /na-a-na-na/ und /na-na-na-na-na/) untereinander gepaart. Identische Folgen wurden nicht miteinander gepaart, z.B. nicht /na-na/ mit /na-na/ (vgl. Tabelle 6-10).

Da alle Silbenfolgen trotz unterschiedlicher Silbenzahl dieselbe Gesamtlänge haben mussten, war es nötig, einzelne /na/-Einheiten mit unterschiedlicher Länge zu iterieren (vgl. Tabelle 6-9), um z.B. eine dreisilbige und eine fünfsilbige /na/-Folge mit identischer Gesamtlänge zu erhalten. Darüberhinaus mussten dreisilbige /na/-Folgen in drei Längen, d.h. 600 ms, 800 ms und 1000 ms hergestellt werden, da die dreisilbige Zielfolge in jeder Stimuluslängengruppe vorkommen musste. Die zweisilbigen und die viersilbigen /na/-Folgen mussten in je zwei Stimuluslängen hergestellt werden.

Tabelle 6-9. Länge der einzelnen Stimuli und Stimuluskomponenten

Gesamtlänge der Stimuli	Stimuli	Länge jedes einzelnen /na/
600 ms	<i>na</i>	600 ms
	<i>na-na</i>	300 ms
	<i>na-na-na</i>	200 ms
800 ms	<i>na-na</i>	400 ms
	<i>na-na-na</i>	266 ms
	<i>na-na-na-na</i>	200 ms
1000 ms	<i>na-na-na</i>	333 ms
	<i>na-na-na-na</i>	250 ms
	<i>na-na-na-na-na</i>	200 ms

Alle /na/-Folgen wurden hergestellt, indem ausgehend von einem einzigen /na/ sieben verschiedene /na/-Versionen mit einer Länge von 600, 400, 333, 300, 266, 250 und 200 Millisekunden mit dem Programm SoundEdit 16 für Macintosh erzeugt wurden, aus denen durch Iteration jeweils eines /na/ mit der entsprechenden Länge die neun verschiedenen Versuchsstimuli hergestellt wurden, je drei für jede Länge (vgl. Tabelle 6-9). Zwischen zwei /na/-Einheiten wurden dabei keine Pausen eingefügt. Am Anfang

und am Ende jedes Stimulus wurden 50 ms Stille in die Audio-Datei eingefügt. Die resultierenden Stimuli hatten eine Gesamtlänge von 700 (600+100), 900 (800+100) und 1100 (1000+100) ms.

Tabelle 6-10. Dichotische Stimuluszusammenstellung für Experiment 2.

	links \ rechts			
Länge 1– 600ms		<i>na</i>	<i>na-na</i>	<i>na-na-na</i>
	<i>na</i>	–	X	X
	<i>na-na</i>	X	–	X
Länge 2– 800ms		<i>na-na</i>	<i>na-na-na</i>	<i>na-na-na-na</i>
	<i>na-na</i>	–	X	X
	<i>na-na-na</i>	X	–	X
Länge 3–1000ms		<i>na-na-na</i>	<i>na-na-na-na</i>	<i>na-na-na-na-na</i>
	<i>na-na-na</i>	–	X	X
	<i>na-na-na-na</i>	X	–	X
	<i>na-na-na-na-na</i>	X	X	–

Aus der oben beschriebenen Art der Zusammenstellung der dichotischen Paare (vgl. Tabelle 6-10) ergibt sich die folgende Stimuluszusammenstellung:

- Es gibt für jede Länge 6 *verschiedene* dichotische Paare.
- Bei 3 Längen gibt es insgesamt 18 *verschiedene* dichotische Paare.
- Für jede Länge gibt es 4 *Zielfolgen*,²⁷⁰ d.h. dichotische Paare, welche die Folge /na-na-na/ enthalten und auf deren Präsentation der Proband durch Tastendruck reagieren soll (schattierte Felder), sowie 2 *Distraкторfolgen* (weiße Felder).
- Bei 3 Längen gibt es insgesamt 12 Zielfolgen.
- Von 18 verschiedenen dichotischen Paaren insgesamt sind 12 Paare Zielfolgen und 6 Paare Distraktorfolgen.

Um also eine Zielfolgen-Wahrscheinlichkeit von 50% zu erreichen (so dass bei jeder einzelnen Präsentation die Wahrscheinlichkeit, dass es sich um eine Zielfolge handelt, genauso groß ist wie die Wahrscheinlichkeit, dass es sich um eine Distraktorfolge handelt), werden die 6 Distraktorfolgen doppelt so häufig dargeboten wie die 12 Zielfolgen.

²⁷⁰ Mit *Zielfolge* ist hier gemeint: ein dichotisches Paar, welches die dreisilbige Zielfolge /na-na-na/ enthält. Mit *Distraкторfolge* ist hier gemeint: ein dichotisches Paar, welches die dreisilbige Zielfolge /na-na-na/ *nicht* enthält.

Tabelle 6-11. Darbietungshäufigkeit der Stimuli in Experiment 2.

	links \ rechts				
Länge 1– 600ms		<i>na</i>	–	12	6
		<i>na-na</i>	12	–	6
		<i>na-na-na</i>	6	6	–
Länge 2– 800ms		<i>na-na</i>	–	6	12
		<i>na-na-na</i>	6	–	6
		<i>na-na-na-na</i>	12	6	–
Länge 3–1000ms		<i>na-na-na</i>	–	6	6
		<i>na-na-na-na</i>	6	–	12
		<i>na-na-na-na-na</i>	6	12	–

Ein Block besteht daher aus den folgenden Stimulusdarbietungen:

- Aus 12 Zielfolgen-Darbietungen, je 4 pro Länge, davon je 2 auf jedes Ohr.
- Aus $2 \times 6 = 12$ Distraktorfolgen-Darbietungen, je 4 pro Länge.
- Aus 24 Durchgängen insgesamt.

Das Experiment besteht aus sechs Blöcken, also aus den folgenden Stimulusdarbietungen insgesamt (vgl. auch Tabelle 6-11):

- Aus 72 Zielfolgen-Darbietungen, je 24 pro Länge, davon je 12 auf jedes Ohr (schattierte Felder; Stimuli, in welchen die Zielfolge dem rechten Ohr präsentiert wird, sind schwach schattiert, solche, in welchen Zielfolgen dem linken Ohr präsentiert wird, stark schattiert),
- und aus 72 Distraktorfolgen-Darbietungen (weiße Felder), je 24 pro Länge, je 12 pro Ohr,
- d.h. aus 144 Durchgängen insgesamt.

6.4.1.2 Versuchsaufbau

Der Versuch wurde unter denselben Bedingungen durchgeführt wie Experiment 1. Der Computer zeichnete für alle Durchgänge, in denen die Versuchsperson die Reaktionstaste drückte, die experimentelle Bedingung und die Reaktionszeit auf. Die Versuchsanweisungen wurden den Probanden in schriftlicher Form auf dem Computermonitor präsentiert. Jede Anweisung wurde solange auf dem Bildschirm gezeigt, bis der Proband durch Tastendruck die nächste Anweisung oder den Start des Experiments abrief. Die Aufgabe des Probanden bestand in diesem Experiment darin, immer dann (mit einer Hand) die Reaktionstaste zu drücken, wenn er die Folge *na-na-na* hörte (d.h. die ‘dreisilbige’ Folge), unabhängig davon, auf welchem Ohr er diese Folge hörte und welche Folge er auf dem anderen Ohr hörte (vgl. auch Abbildung 6-9).

Das Experiment bestand aus 6 Blöcken zu je 24 Durchgängen, d.h. aus 144 Durchgängen insgesamt. Um den Probanden mit der Zielfolge vertraut zu machen, wurde ihm vor Beginn des Experimentes die 'dreisilbige' Zielfolge *na-na-na* in den drei Längen je dreimal (d.h. insgesamt neunmal) in randomisierter Reihenfolge diotisch (d.h. auf beiden Ohren dasselbe Signal, ohne Distraktor) präsentiert. Danach folgten 16 randomisierte Übungsdurchgänge, die nicht in die Auswertung eingingen.

Die Hälfte der Probanden begann das Experiment mit der rechten Hand, die andere Hälfte mit der linken Hand. Nach der Hälfte der Durchgänge (d.h. nach 3 Blöcken bzw. 72 Durchgängen) wurde die reagierende Hand gewechselt, d.h. die Probanden, die mit der rechten Hand begonnen hatten, fuhrten nach der ersten Hälfte des Experimentes mit der linken Hand fort und umgekehrt. Wie in Experiment 1 begann auch hier die Hälfte der Probanden mit der Kopfhörerposition A und die andere Hälfte mit der Kopfhörerposition B (für einen Überblick über die Probandengruppen vgl. Tabelle 6-1). Nach jedem Block wurde die Kopfhörerposition gewechselt.

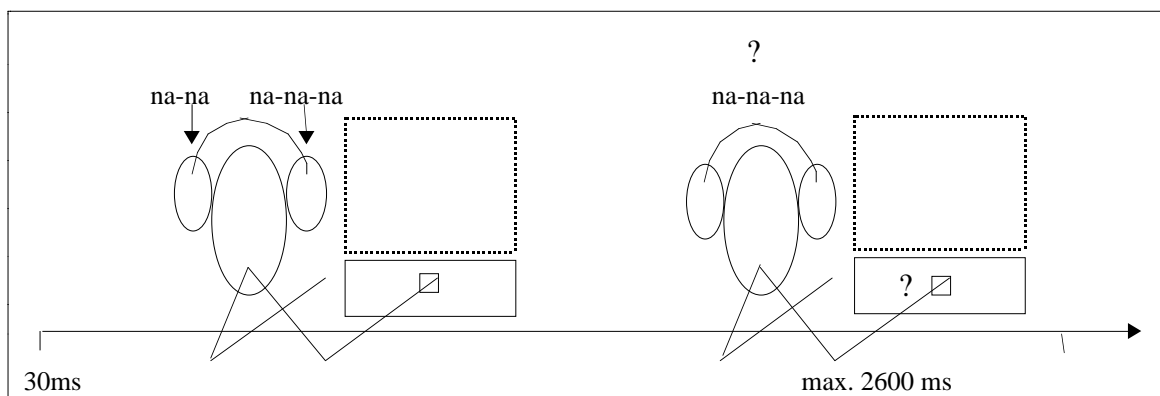


Abbildung 6-9. Versuchsaufbau.

Um die Aufmerksamkeit der Probanden auf beide Ohren zu richten, begann jeder Durchgang mit einem diotisch präsentierten 400 Hz-Ton von 300 ms Dauer. Danach wurde das dichotische Stimuluspaar präsentiert, d.h. auf dem linken Ohr eine /na/-Folge und auf dem rechten Ohr eine andere /na/-Folge. Reaktionen der Probanden innerhalb von 1500 ms nach Ende des Stimulus wurden als Antworten aufgezeichnet und gingen in die Auswertung ein. Für jeden Durchgang wurde nur eine einzige Reaktion des Probanden aufgezeichnet.

In diesem Experiment wurde die These getestet, dass die Probanden dreisilbige Zielfolgen, welche dem linken Ohr präsentiert wurden, schneller und mit weniger Fehlern detektieren würden als dreisilbige Zielfolgen, welche dem rechten Ohr präsentiert wurden.

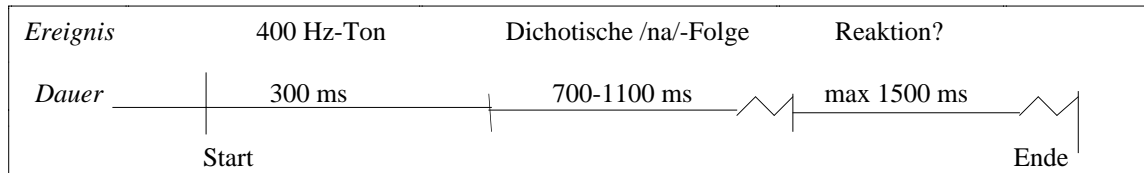


Abbildung 6-10. Ereignisse und Zeitverlauf in einem Durchgang in Experiment 2

6.4.2 Ergebnisse

Für jede Versuchsperson wurden für die 144 Testdurchgänge die Anzahl der Reaktionen ('Treffer') auf die präsentierten Zielfolgen und die mittlere Reaktionszeit ('Reaktionszeit') für das linke und das rechte Ohr berechnet. Nur Reaktion innerhalb von 1500 ms nach Stimulusende wurden berücksichtigt. Die Reaktionszeit wurde ab dem Stimulusende gemessen. Daneben wurden die Anzahl der Reaktionen auf Distraktorfolgen ('false alarms') und die mittlere Reaktionszeit der Reaktionen auf Distraktorfolgen ('Reaktionszeit false alarms') berechnet. Alle Werte wurden auch für jede Stimuluslänge getrennt aufgezeichnet.

Wie aus Tabelle 6-12 zu ersehen ist, unterschieden sich die Werte der Variable 'Treffer' weder für die Gesamtpopulation noch für die Geschlechtergruppen in Abhängigkeit vom Ohr, welchem die Zielfolge präsentiert wurde. Für die Gruppe der weiblichen Probanden fand sich eine höhere Gesamttrefferrate als für die Gruppe der männlichen Probanden. Außerdem variierte die Trefferrate mit der Stimuluslänge; die höchsten Werte für die Variable 'Treffer' fanden sich für die Stimuluslänge 3. Schließlich variierten auch die Unterschiede zwischen den Werten für 'Treffer' zwischen den Ohren mit der Stimuluslänge: für beide Geschlechtergruppe fanden sich für die Länge 1 und 3 mehr Reaktionen auf linksohrig präsentierte als auf rechtsohrig präsentierte Zielfolgen, für die Länge 2 hingegen mehr Reaktionen auf rechtsohrig präsentierte als auf linksohrig präsentierte Zielfolgen.

Für die Variable 'Reaktionszeit' zeigte sich für die Gesamtpopulation und für die Gruppe der weiblichen Probanden eine schnellere Reaktion auf rechtsohrig präsentierte Zielfolgen als auf linksohrig präsentierte Zielfolgen (vgl. Tabelle 6-13). Die Reaktionszeit variierte für beide Geschlechtergruppen mit der Stimuluslänge. Die kürzeste Reaktionszeit fand sich für die Stimuluslänge 2, die Reaktionszeiten für die beiden anderen Längen unterschieden sich kaum.

Tabelle 6-12. Mittelwerte und Standardabweichungen (S.D.) der Variable 'Treffer' in absoluten Zahlen für die Gesamtpopulation und für jede Geschlechtergruppe, auch nach Stimuluslänge und Zielfolgenpräsentationsohr getrennt.

		'Treffer'					
		linkes Ohr		rechtes Ohr		Gesamt	
		Mittelwert	S.D.	Mittelwert	S.D.	Mittelwert	S.D.
Gesamt	alle Längen	20.77	5.76	21.00	5.56	41.77	10.48
	1	6.31	3.48	6.15	3.48	13.29	6.00
	2	6.21	2.09	6.73	2.04	12.94	3.61
	3	8.25	2.88	8.12	2.96	16.38	5.58
Männer	alle Längen	19.38	4.96	19.50	5.43	38.88	9.58
	1	5.5	4.04	5.21	3.82	10.71	6.78
	2	6.04	1.81	6.63	2.12	12.66	3.58
	3	7.83	2.88	7.67	2.94	15.50	5.61
Frauen	alle Längen	22.17	6.25	22.50	5.38	44.67	10.72
	1	7.13	2.66	7.08	2.87	14.21	5.19
	2	6.37	2.37	6.83	1.99	13.21	3.69
	3	8.67	2.87	8.58	2.96	17.25	5.52

Tabelle 6-13. Mittelwerte und Standardabweichungen (S.D.) der Variable 'Reaktionszeit' in ms für die Gesamtpopulation und für jede Geschlechtergruppe, auch nach Stimuluslänge und Zielfolgenpräsentationsohr getrennt.

		'Reaktionszeit' in ms					
		linkes Ohr		rechtes Ohr		Gesamt	
		Mittelwert	S.D.	Mittelwert	S.D.	Mittelwert	S.D.
Gesamt	1	1091.60	219.20	1058.49	181.83	1082.99	217.70
	2	920.66	192.86	930.24	156.37	953.41	210.92
	3	1035.44	169.04	1006.19	167.69	1054.12	225.39
Männer	1	1088.70	270.39	1096.58	210.79	1096.85	241.95
	2	899.04	202.36	930.39	162.79	955.19	236.08
	3	1069.41	193.14	1041.79	168.68	1099.21	215.77
Frauen	1	1094.02	171.53	1025.37	149.27	1072.15	198.32
	2	943.23	184.17	930.08	153.19	951.61	183.69
	3	1001.48	136.67	970.58	162.52	1010.45	227.56

Die Anzahl der Fehlerreaktionen auf Distraktorfolgen ('false alarms') variierte schließlich auch mit der Stimuluslänge; die höchste Fehlerrate fand sich für beide Geschlech-

tergruppen für die Länge 2, die niedrigste Fehlerrate für die Länge 1 (vgl. Tabelle 6-14).

Tabelle 6-14. Mittelwerte und Standardabweichungen (S.D.) der Variablen 'false alarms' und 'Reaktionszeit false alarms' in ms für die Gesamtpopulation und für jede Geschlechtergruppe, nach Stimuluslänge getrennt.

		'false alarms'		'Reaktionszeit false alarms' in ms	
		Mittelwert	S.D.	Mittelwert	S.D.
Gesamt	1	1.23	4.35	1144.26	327.43
	2	3.94	4.30	1023.20	272.66
	3	2.44	3.26	1151.06	327.20
Männer	1	1.00	4.48	1152.89	327.51
	2	3.88	5.04	1057.43	319.56
	3	2.38	3.19	1234.95	264.91
Frauen	1	1.46	4.30	1141.02	349.95
	2	4.00	3.53	988.96	219.59
	3	2.50	3.40	1077.05	365.52

Die Ergebnisse für die Variablen 'Treffer' und 'Reaktionszeit' wurden einer dreifaktoriellen Varianzanalyse unterzogen, mit der Variablen 'Geschlecht' (männlich, weiblich) als Gruppenvariable und den Variablen 'Ohr' (links, rechts) und 'Stimuluslänge' (Länge 1 bis 3) als *between-subjects*-Faktoren.

Treffer: Die Haupteffekte für 'Geschlecht' und 'Ohr' waren nicht signifikant (alle F-Werte(1,46) < 3.89, n.s.), wohl aber der Haupteffekt für den Faktor 'Länge' (F(1,46) = 8.46, $p < 0.001$). Es fanden sich keine signifikanten Zweifach- oder Dreifach-Interaktionen (vgl. Tabelle 6-15).

Tabelle 6-15. Zusammenfassung der ANOVA für die Variable 'Treffer'.

Dreifaktorielle Varianzanalyse			
Abhängige Variable: 'Treffer'			
Haupteffekte	Geschlecht	F(1,46) = 3.89	$p < 0.055$
	Ohr	F(1,46) = 0.13	n.s.
	Länge	F(2,92) = 8.46	$p < 0.001$
Zweifach-Interaktionen	Geschlecht × Ohr	F(1,46) = 0.03	n.s.
	Geschlecht × Länge	F(2,92) = 1.03	n.s.
	Ohr × Länge	F(2,92) = 2.78	$p < 0.067$
Dreifach-Interaktion	Geschlecht × Ohr × Länge	F(2,92) = 0.16	n.s.

Reaktionszeit: Auch hier fand sich kein signifikanter Haupteffekt für ‘Geschlecht’ und ‘Ohr’ (alle F-Werte(1,46) < 0.7, n.s.), im Gegensatz zu einem signifikanten Haupteffekt für den Faktor ‘Länge’ (F(1,46) = 12.78, $p < 0.001$). Signifikante Interaktionen konnten nicht demonstriert werden (vgl. Tabelle 6-16).

Tabelle 6-16. Zusammenfassung der ANOVA für die Variable ‘Reaktionszeit’.

Dreifaktorielle Varianzanalyse			
Abhängige Variable: ‘Reaktionszeit’			
Haupteffekte	Geschlecht	F(1,38) = 0.41	n.s.
	Ohr	F(1,38) = 0.70	n.s.
	Länge	F(2,76) = 12.78	$p < 0.001$
Zweifach-Interaktionen	Geschlecht × Ohr	F(1,38) = 2.60	n.s.
	Geschlecht × Länge	F(2,76) = 1.79	n.s.
	Ohr × Länge	F(2,76) = 0.72	n.s.
Dreifach-Interaktion	Geschlecht × Ohr × Länge	F(2,76) = 1.39	n.s.

Zusätzlich wurden die Ergebnisse für die Variablen ‘*false alarms*’ und ‘Reaktionszeit *false alarms*’ einer zweifaktoriellen Varianzanalyse unterzogen, mit der Variable ‘Geschlecht’ (männlich, weiblich) als Gruppenvariable und der Variablen ‘Stimuluslänge’ (Länge 1 bis 3) als *between-subjects*-Faktor (für die Mittelwerte vgl. Tabelle 6-14).

False alarms: Es zeigte sich ein Haupteffekt für ‘Länge’ (F(1,46) = 5.47, $p < 0.01$), aber kein Haupteffekt für ‘Geschlecht’. Die Interaktion von ‘Geschlecht’ und ‘Länge’ war nicht signifikant (alle F-Werte(1,46) < 0.12, n.s.).

Reaktionszeit false alarms: Es fanden sich keine signifikanten Haupteffekte oder Interaktionen (alle F-Werte(1,46) < 0.55, n.s.).

6.4.3 Diskussion

In diesem Experiment konnte für keine der Probandengruppen ein signifikanter Ohr-Vorteil für die untersuchte Verarbeitungsfunktion nachgewiesen werden. Die Hypothese, dass sich für die Detektion der dreisilbigen Silbenfolge ein Links-Ohr-Vorteil in Form einer höheren Trefferrate und einer kürzeren Reaktionszeit für linksohrig präsentierte Zielfolgen finden würde, wurde daher durch die Daten nicht bestätigt.

Es zeigte sich, dass sowohl die Gesamtrefferrate als auch die Reaktionszeit und schließlich auch die Fehlerrate (*false alarms*) in Abhängigkeit von der Stimuluslänge (Länge 1 bis 3) variierte. Dabei fand sich für die Stimuli der Länge 2 (zwei-, drei- und

viersilbige Folgen) im Mittel eine niedrigere Trefferrate, eine kürzere Reaktionszeit und eine höhere Fehlerrate als für die Stimuli der beiden anderen Längen. Die höchste Trefferrate fand sich für die Stimuli der Länge 3 (drei-, vier- und fünfsilbige Folgen). Auch die Trefferrate je Ohr variierte in Abhängigkeit von der Stimuluslänge, dieser Effekt wurde jedoch nicht signifikant. Dabei zeigte sich für die Längen 1 und 3 eine im Mittel höhere Trefferrate für linksohrig präsentierte Zielfolgen, für die Länge 2 jedoch eine im Mittel höhere Trefferrate für rechtsohrig präsentierte Zielfolgen.

Die unterschiedlichen Ergebnisse, die sich für die drei Stimuluslängen für die Variablen ‘Anzahl der Reaktionen’, ‘Reaktionszeit’ und ‘Anzahl der false alarms’ zeigten, weisen darauf hin, dass die drei Stimuluslängen, d.h. die Länge 1 mit den ein-, zwei- und dreisilbigen, die Länge 2 mit den zwei-, drei- und viersilbigen und die Länge 3 mit den drei-, vier- und fünfsilbigen Folgen nicht auf dieselbe Weise verarbeitet werden. Diese drei Stimulusgruppen unterscheiden sich nicht nur in der absoluten Länge, sondern vor allem in der ‘phonologischen’ Komplexität²⁷¹ der dichotischen Paare und sind daher möglicherweise nicht gleichermaßen geeignet, um die Verarbeitungsfunktion ‘Lokalisierung der Sonoritätsmaxima in einer Äußerung’ zu operationalisieren. Eine mögliche Interpretation der Tatsache, dass die Hypothese einer Hemisphärenspezialisierung für die getestete Verarbeitungsfunktion durch die Daten nicht bestätigt werden konnte, ist daher, dass die Detektion der Zielfolgen der Längen 1 und 3 einerseits und der Länge 2 andererseits verschiedene Verarbeitungsstrategien erfordert, die möglicherweise komplementär lateralisiert sind, so dass sich für das Gesamtergebnis kein Ohr-Vorteil zeigt. In jedem Fall muss darauf hingewiesen werden, dass die Präsentation von ein- bis fünfsilbigen Silbenfolgen zu einer so großen Varianz in den Daten geführt hat, dass sich abgesehen von einem Längeneffekt keine signifikanten Ergebnisse zeigen ließen. Die Längenunterschiede in den Stimuli wurden, wie oben erwähnt, deshalb eingeführt, weil die Präsentation von nur zwei-, drei- und viersilbigen Silbenfolgen zu einer Gesamtrefferrate von nahezu 100% geführt hatte. Eine andere und möglicherweise bessere Strategie, diesem Deckeneffekt entgegenzuwirken, hätte darin bestehen können, nicht ein Zielfolgen-monitoring-Paradigma zu verwenden, sondern ein Ohr-monitoring-Paradigma (vgl. HARSHMAN & LUNDY 1988), so dass die Probanden nicht nur die Zielfolge hätten detektieren müssen, sondern zusätzlich das Ohr hätten identifizieren müssen, auf welchem sie die Zielfolge hörten. Ein solches Versuchsparadigma hätte möglicherweise eher zu Ohr-Unterschieden in der Genauigkeit der Reaktionen geführt.

²⁷¹ Insofern eine fünfsilbige Folge wie /na-na-na-na-na/ ‘phonologisch’ komplexer ist als eine einsilbige Folge wie /na/.

Die Daten dieses Experimentes weisen für die Gruppe der weiblichen Probanden eine höhere Gesamttrefferrate auf als für die Gruppe der männlichen Probanden, dieser Unterschied wurde jedoch nicht signifikant. Diese Beobachtung könnte so interpretiert werden, dass die weiblichen Probanden in diesem Experiment andere Strategien anwendeten als die männlichen Probanden. Da sich für die beiden Geschlechtergruppen aber keine weiteren Unterschiede, etwa in der Reaktionszeit, finden, lässt sich diese Annahme nicht näher präzisieren.

Zusammenfassend lässt sich sagen, dass die Konzeption dieses Versuches möglicherweise dazu geführt hat, dass die Probanden unter den verschiedenen experimentellen Bedingungen unterschiedliche Verarbeitungsstrategien anwendeten. Außerdem muss die Möglichkeit in Betracht gezogen werden, dass die hier verwendeten sinnlosen Silbenfolgen nicht als Sprachreize verarbeitet wurden, sondern eher als monotonrhythmische Folgen derselben sprachähnlichen Vokalisation. Diese Möglichkeit einer nichtsprachlichen Verarbeitung könnte eine Erklärung dafür liefern, warum in diesem Experiment, anders als im ersten Experiment, kein Ohr-Vorteil nachgewiesen werden konnte. Gleichzeitig muss darauf hingewiesen werden, dass ein Zielfolgen-monitoring-Paradigma möglicherweise ungeeignet ist, um eventuell bestehende Ohr-Unterschiede in der Verarbeitungseffizienz eines Reizes zu messen und stattdessen zum Nachweis einer funktionalen Lateralisierung der Sprachwahrnehmung ein Ohr-monitoring-Paradigma verwendet werden sollte. Aus diesen Gründen muss angenommen werden, dass die Konzeption des hier beschriebenen Versuchs nicht geeignet war, die funktionale Lateralisierung der Verarbeitungsfunktion 'Lokalisierung der Sonoritätsmaxima' zu untersuchen.

6.5 Zusammenfassung

In diesem Kapitel wurden zwei Sprachwahrnehmungsexperimente mit dichotischer Stimuluspräsentation beschrieben. In beiden Experimenten wurde die Hypothese getestet, dass sich für die Lokalisation der Sonoritätsmaxima in einer sprachlichen Äußerung eine rechtshemisphärische Spezialisierung findet. Im ersten Experiment konnte ein Geschlechtsunterschied in der Lateralität für die Wahrnehmung der Sonoritätsmaxima demonstriert werden, mit einer rechtshemisphärischen Spezialisierung für die Gruppe der männlichen Probanden und einer eher bilateralen Repräsentation für die Gruppe der weiblichen Probanden. Dieser Geschlechtsunterschied wurde durch die Annahme erklärt, dass männliche und weibliche Probanden unterschiedliche Verarbeitungsstrategien anwendeten. Die Hemisphärenspezialisierung, die sich für die männlichen Proban-

den zeigen ließ, ging in die erwartete Richtung und kann als Hinweis darauf interpretiert werden, dass die Verarbeitung der Sonoritätskontur in einer Äußerung rechtshemisphärisch lateralisiert ist.

Im zweiten Experiment konnte für keine Probandengruppe ein Ohr-Vorteil demonstriert werden. Dieses Ergebnis wurde darauf zurückgeführt, dass die Konzeption des Experimentes nicht geeignet war, um die funktionale Lateralisierung der Verarbeitung der Sonoritätskontur zu untersuchen. Ebenso wurde darauf hingewiesen, dass die unterschiedlichen Ergebnisse beider Experimente dadurch zustande gekommen sein könnten, dass im ersten Versuch *bedeutungsvolle* deutsche Wörter, im zweiten Versuch hingegen *sinnlose* Silbenfolgen verwendet wurden. Ein Ohr-Vorteil für die Verarbeitung der Sonoritätskontur kann möglicherweise nur mit bedeutungsvollen Wörtern nachgewiesen werden, die nicht nur auf einer phonologischen, sondern auch auf einer semantischen Ebene verarbeitet werden. Mit bedeutungslosen Silbenfolgen lässt sich eine funktionale Lateralisierung für die Verarbeitung der Sonoritätskontur vielleicht überhaupt nicht nachweisen.²⁷² In diesem Zusammenhang ist auch von Bedeutung, dass im ersten Experiment eine Identifikationsaufgabe verwendet wurde, im zweiten Experiment hingegen eine Zielfolgen-Detektionsaufgabe. Wie im vorangegangenen Kapitel hervorgehoben wurde (s.o. 5.4.4), können die beiden Faktoren ‘Verständlichkeit der Stimuli’ und ‘Art der Testaufgabe’ die funktionalen Lateralitäten für eine ganze Reihe von segmentalen und suprasegmentalen Sprachverarbeitungsfunktionen beeinflussen.

Zusammenfassend lässt sich sagen, dass die Daten des ersten Experimentes darauf hinweisen, dass die Verarbeitung der Sonoritätskontur einer Äußerung rechtshemisphärisch lateralisiert ist. Damit bestätigen sie die Hypothese dieser Arbeit. Weitere Untersuchungen müssen klären, ob sich eine rechtshemisphärische Spezialisierung für die Verarbeitung der Sonorität auch unter anderen experimentellen Bedingungen demonstrieren lässt, z.B. unter Verwendung von bedeutungslosen oder unverständlichen Wörtern. Auch die Annahme, dass dem hier demonstrierten Geschlechtsunterschied in der Lateralität Unterschiede in den angewendeten Verarbeitungsstrategien zugrunde liegen, muss in weiteren Studien näher untersucht werden. Schließlich muss auch die Frage gestellt werden, ob die Lateralitätseffekte für die Testaufgabe, die hier demonstriert wurden, eher auf die linguistisch-funktionale, d.h. auf die phonologische und semantische, Verarbeitung der Reize, oder eher auf Aspekte der akustisch-phonetischen Verarbeitung zurückzuführen sind.

²⁷² Vgl. hierzu auch die Beobachtung von GALLÉS ET AL. (1992), dass sich ein Silbeneffekt im Spanischen nur dann zeigen ließ, wenn die Stimuluswörter zusätzlich semantisch verarbeitet werden mussten (s.o. 4.3.2).

7 Schluss

7.1 Zusammenfassung

In dieser Arbeit wurde die Frage untersucht, ob die Verarbeitung der Sonoritätskontur einer Äußerung lateralisiert ist. Die Hypothese, dass sich für die Lokalisierung der Sonoritätsmaxima in einem Wort eine rechtshemisphärische Spezialisierung findet, konnte durch die Ergebnisse eines dichotischen Sprachwahrnehmungsversuchs bestätigt werden.

Die vorliegende Untersuchung ist im Zusammenhang von Forschungen zur neuronalen Realisierung der menschlichen Sprachfähigkeit zu sehen. Daher wurde im ersten Kapitel zunächst ein Überblick über die Methoden gegeben, mit denen die neuronalen Grundlagen der menschlichen Sprachfähigkeit untersucht werden können. Die Vorgehensweise des hier beschriebenen Projektes bestand darin, ausgehend von einer Beschreibungseinheit der sprachwissenschaftlichen Theorie, d.h. dem phonologischen Merkmal Sonorität, funktionale Einheiten im Gehirn zu identifizieren, deren Aktivität mit der Verarbeitung dieses phonologischen Merkmals korreliert ist, indem die Frage untersucht wurde, ob die Verarbeitung der Sonorität lateralisiert ist.

Hierzu war es im zweiten Kapitel nötig, zunächst den phonologischen Status der Kategorie Silbe und des Merkmals Sonorität zu klären und insbesondere zu zeigen, dass die Silbe eine elementare Kategorie der phonologischen Theorie ist und dass das Merkmal Sonorität nicht auf die segmentalen Merkmale der phonologischen Theorie zurückzuführen ist, sondern das Verbindungsglied zwischen der segmentalen Ebene und der suprasegmentalen Ebene in der Sprache bildet. Als phonetische Korrelate der phonologischen Sonorität wurden im dritten Kapitel die akustischen Parameter Lautstärke, Dauer und Stimmhaftigkeit identifiziert. In diesem Kapitel wurde auch gezeigt, dass diese phonetischen Korrelate geeignet sind, um aus dem akustischen Signal eine Silbenstruktur und eine Akzentstruktur zu rekonstruieren. Im vierten Kapitel wurde die Frage diskutiert, welche Funktion die Kategorie Silbe und das Merkmal Sonorität für

die Sprachwahrnehmung erfüllen könnten. Hier wurde die Annahme formuliert, dass die Verarbeitung der phonetischen Korrelate der Sonoritätskontur einer Äußerung und insbesondere die Lokalisierung der Sonoritätsmaxima und der Sonoritätsminima, zur Strukturierung des zu verarbeitenden Signals dienen. Diese Annahme wurde unter phonetisch-phonologischer unter unter psycholinguistischer Perspektive begründet. Die Ausführungen im vierten Kapitel hatten die Funktion, das phonetisch-phonologische Merkmal Sonorität in ein psycholinguistisches Modell der Sprachwahrnehmung zu integrieren. Damit wurde die Voraussetzung dafür geschaffen, um die statische phonologische Repräsentation auf das beobachtbare Verhalten menschlicher Sprecher und schließlich auf dynamische Aktivierungsprozesse von funktionalen Einheiten im Gehirn beziehen zu können.

Im fünften und sechsten Kapitel dieser Arbeit wurde die Frage gestellt, ob die Verarbeitung der Sonoritätskontur einer Äußerung lateralisiert ist. Zunächst wurden im fünften Kapitel einige Studien zur funktionalen Lateralisierung der Sprache referiert, insbesondere Studien zur Lateralisierung der Sprachwahrnehmung. Vor dem Hintergrund dieser Studien wurde die Hypothese formuliert und begründet, dass für die Verarbeitung der Sonoritätskontur eine rechtshemisphärische Spezialisierung zu erwarten ist. Im sechsten Kapitel wurden zwei dichotische Sprachwahrnehmungsversuche beschrieben, in denen die Hypothese einer rechtshemisphärischen Spezialisierung für die Verarbeitung der Sonoritätskontur einer Äußerung getestet wurde. Die Ergebnisse des ersten Experimentes zeigten einen Geschlechtsunterschied in der Lateralität für die Verarbeitung der Sonoritätskontur, mit einer rechtshemisphärischen Spezialisierung für die Gruppe der männlichen Probanden und einer eher bilateralen Repräsentation der in der Testaufgabe aktivierten funktionalen Systeme für die Gruppe der weiblichen Probanden. Der Geschlechtsunterschied in der Lateralität, der in diesem Experiment gezeigt werden konnte, wurde durch die Annahme erklärt, dass sich Männer und Frauen in den Verarbeitungsstrategien unterschieden, die sie bei der Durchführung der Testaufgabe anwendeten. Insbesondere wurde angenommen, dass die männlichen Probanden eine rechtshemisphärisch lateralisierte sonoritätsbasierte Verarbeitungsstrategie verwendeten, während die weiblichen Probanden eine bilateral repräsentierte vokalbasierte Verarbeitungsstrategie zur Identifikation der Stimuli nutzten. Im zweiten Experiment konnte weder ein Ohr-Vorteil für eine der Probandengruppen noch ein Geschlechtsunterschied in der Lateralität gezeigt werden. Dieses Ergebnis wurde so interpretiert, dass einerseits das erste und das zweite Experiment nicht dieselbe Wahrnehmungsfunktion operationalisierten und dass andererseits der Versuchsaufbau im zweiten Experiment nicht geeignet war, eine funktionale Lateralisierung der Sprachverarbeitung nachzuweisen.

Den Ausgangspunkt der vorliegenden Studie bildete das phonologische Merkmal Sonorität. Dieses Merkmal ist, wie in Kapitel 3 ausgeführt wurde, in verschiedenen Studien auch unter akustischer und artikulatorischer Perspektive untersucht worden. In psycholinguistischen Modellen zur Sprachwahrnehmung und in neuropsychologischen Modellen zur funktionalen Lateralisierung der Sprache wurde das Merkmal Sonorität bisher nicht berücksichtigt. In dieser Arbeit ist gezeigt worden, dass die Eigenschaft Sonorität nicht nur eine elementare Einheit der phonologischen und phonetischen Beschreibung der Sprache ist, sondern darüberhinaus erstens eine noch zu präzisierende Funktion für die Sprachwahrnehmung erfüllt und zweitens eine Rolle für solche Sprachwahrnehmungsfunktionen spielt, für die sich eine funktionale Lateralisierung zeigen lässt.

Zusammenfassend lässt sich sagen, dass die Ergebnisse dieser Untersuchung darauf hinweisen, dass die Verarbeitung der Sonoritätskontur einer Äußerung rechtshemisphärisch lateralisiert ist. Damit ist neben der emotionalen Intonation einer Äußerung und neben ihrer Satzprosodie eine weitere Eigenschaft von lautsprachlichen Strukturen identifiziert worden, für deren Verarbeitung sich eine rechtshemisphärische Spezialisierung zeigen lässt.

7.2 Ausblick

Die Ergebnisse dieser Untersuchung werfen eine Reihe von Fragen auf. Vergleicht man die beiden Experimente miteinander, so stellt sich insbesondere die Frage, ob sich eine rechtshemisphärische Spezialisierung für die Verarbeitung der Sonoritätskontur nur mit *bedeutungsvollen* Wörtern zeigen lässt, oder ob der Links-Ohr-Vorteil auch mit *bedeutungslosen* Wörtern und sinnlosen Silben nachweisbar wäre. Interessant wäre es auch herauszufinden, ob der für die männlichen Probanden beobachtete Ohr-Vorteil eine linguistisch-funktionale, d.h. lexikalische oder phonologische Verarbeitung der Reize als Wörter des Deutschen voraussetzt, oder ob er auf der akustisch-phonetischen Verarbeitung der Signale basiert. Ein erster Ansatz, diese Frage genauer zu untersuchen, könnte darin bestehen, die drei akustischen Parameter, welche als Korrelate des phonologischen Sonoritätsverlaufs identifiziert wurden, isoliert zu betrachten, ähnlich wie dies von PRICE (1980) für die (beidohrige) Wahrnehmung der Silbizität durchgeführt wurde. So könnten Sprachsignale hergestellt werden, die sich nur im Parameter 'Lautstärke des Liquids' unterscheiden und die entweder als das einsilbige oder als das zweisilbige Wort wahrgenommen werden. Auf dieselbe Weise könnte man für die beiden anderen Parameter 'Dauer des Liquids' und 'Stimmeinsatzzeit des Plosivkonsonanten'

verfahren. Erhielte man für die drei verschiedenen phonetischen Korrelate differierende Ohr-Vorteile, so wäre dies ein erster Hinweis darauf, dass die beobachteten Ohr-Vorteile nicht nur eine Funktion der phonologischen bzw. semantischen Verarbeitung der Wörter sind, sondern auch der akustisch-phonetischen Verarbeitung der Parameter Lautstärke, Dauer und Stimmhaftigkeit. Schließlich könnte untersucht werden, ob sich Ohr-Vorteile auch für nichtsprachliche Reize demonstrieren lassen, die sich in ähnlichen akustischen Eigenschaften unterscheiden wie die hier verwendeten Sprachsignale, z.B. in der langsamen Amplitudenmodulation, die aber keine weiteren Eigenschaften natürlichsprachlicher akustischer Signale aufweisen.

In dieser Studie wurde die *Wahrnehmung* des Merkmals Sonorität untersucht. Eine zweite Frage, die sich in diesem Zusammenhang stellt, ist die Frage nach der funktionalen Lateralisierung der *Produktion* der Sonoritätskontur einer Äußerung. Im zweiten und dritten Kapitel dieser Arbeit wurde die Sonoritätskontur einer Äußerung auf die grobe artikulatorische Gestalt der Sprache, auf die Auslenkung, die Geschwindigkeit und die Koordination der artikulatorischen Gesten bezogen. RICHARDS & CHIARELLO (1997) vertreten die Auffassung, dass die rechte Hemisphäre eine wichtige, jedoch noch genauer zu spezifizierende Funktion für die Planung und Ausführung von sprachlichen und nichtsprachlichen intentionalen Bewegungen erfüllt. Zudem ließen sich auch für die Sprachproduktion Geschlechtsunterschiede in der Lateralität demonstrieren (HAUSMANN ET AL. 1998, WOLF & GOODALE 1987). Daher wäre es von großem Interesse, die Frage zu untersuchen, ob sich für die *artikulatorische Differenzierung* von n-silbigen und n+1-silbigen Wörtern wie *sein* und *sein(e)n* oder *breit* und *b(e)reit* vergleichbare Hemisphärenvorteile zeigen lassen und möglicherweise auch ein vergleichbarer Geschlechtsunterschied, wie für die dichotische Wahrnehmung von Minimalpaaren wie *breit* und *bereit*. In einer Studie zur *Produktion* der Sonoritätskontur wären schließlich auch Erkenntnisse zu Defiziten in der Silbenstruktur und Phonetik in den Äußerungen linkshemisphärisch und rechtshemisphärisch geschädigter Patienten zu berücksichtigen (CODE 1997).

Betrachtet man das allgemeinere Forschungsprogramm, das im ersten Kapitel dieser Arbeit skizziert wurde, so ist offensichtlich, dass in dieser Studie nur ein einziger Aspekt der neuronalen Grundlagen des phonologischen Merkmals Sonorität untersucht wurde, d.h. die funktionale Lateralisierung der Wahrnehmung dieses Merkmals. Zunächst könnte man fragen, ob sich ähnliche Hemisphärenunterschiede auch in klinischen Studien mit hirngeschädigten Probanden zeigen lassen. Die Untersuchung von Probanden mit fokaler Hirnläsion wäre zudem geeignet, um die neuronalen Grundlagen der Verarbeitung der Sonoritätskontur einer Äußerung über die funktionale Lateralisierung hinaus in enger umrissenen funktionalen Systemen des Gehirns zu lokalisieren. In

klinischen Studien könnte ebenso die Frage untersucht werden, für welche der drei akustischen Parameter der Sonorität sich in Wahrnehmung und Produktion Hemisphärenvorteile demonstrieren lassen. So wäre es denkbar, dass rechtshemisphärisch geschädigte Patienten bei der Produktion deutscher Wörter wie *breit* und *b(e)reit* oder *sein* und *sein(e)n* zwar die *Dauer* des Liquids bzw. des Nasalkonsonanten variieren, um beide Wörter zu differenzieren, aber nicht die *Lautstärke* bzw. *Lautstärkemodulation*, während sich für Patienten mit linkshemisphärischer Schädigung möglicherweise das umgekehrte Muster zeigen ließe.

Besonders interessant wäre auch die Untersuchung der Verarbeitung der Sonoritätskontur mit elektrophysiologischen Methoden wie EEG und MEG. Mit diesen Methoden könnten nicht nur die bei der Verarbeitung der Sonoritätskontur aktivierten funktionalen Einheiten genauer lokalisiert werden, sondern es ließen sich auch Aussagen über den Zeitverlauf der Verarbeitung der Sonorität ableiten. Wenn zudem vergleichbare Daten zum Zeitverlauf der Verarbeitung der Phoneme und der Silben in einer Äußerung zur Verfügung stehen würden, so könnten Aussagen über den relativen Zeitverlauf beider Arten der Verarbeitung mit Aussagen über Strategieunterschiede in Sprachwahrnehmungsversuchen korreliert werden, wie sie von CUTLER ET AL. 1986, CUTLER ET AL. 1992 und GALLÉS ET AL. 1992 (s.o. 4.3) nachgewiesen wurden. Auf diese Weise könnte auch die Annahme (s.o. 4.2.3) überprüft werden, dass den Strategieunterschieden, die sich für das Sprachverhalten beobachten lassen, Unterschiede in der zeitlichen Dynamik der zugrundeliegenden neuronalen Repräsentationen entsprechen. Eine solche Untersuchung wäre gleichfalls geeignet, die hier gemachte Annahme zu überprüfen, dass dem Geschlechtsunterschied in der Lateralität für die Verarbeitung von Minimalpaaren wie *breit* und *bereit* ein Unterschied in den angewendeten Verarbeitungsstrategien entspricht.

Am Ende muss darauf hingewiesen werden, dass nicht nur die Frage nach den neuronalen Grundlagen der Sonorität weitere Untersuchungen erfordert, sondern dass auch die Funktion der Sonorität für die Sprachwahrnehmung genauer erforscht werden muss. Mögliche Funktionen der Sonorität in der Sprachverarbeitung wurden bereits im vierten Kapitel dieser Arbeit skizziert. Eine Integration des Merkmals Sonorität in eine psycholinguistische Theorie der Sprachfähigkeit setzt jedoch eine weitergehende Untersuchung von sonoritätsbasierten, silbenbasierten und phonembasierten Sprachverarbeitungsstrategien voraus.

Ein letzter Punkt, der weitere Untersuchungen erfordert, ist die Frage nach dem Verhältnis von Geschlechtsunterschieden in der Lateralität und eventuell bestehenden Unterschieden in den Sprachverarbeitungsstrategien, die beide Geschlechtergruppen anwenden. Während die Annahme als bestätigt gelten kann, dass sich die Sprecher ver-

schiedener Sprachen darin unterscheiden, welche Verarbeitungsstrategien sie anwenden (CUTLER ET AL. 1992, s.o. 4.3), herrscht bislang keine Klarheit darüber, wie Geschlechtsunterschiede in den sprachlichen Fähigkeiten und in der funktionalen Lateralisierung der Sprache zu bewerten sind.

8 Literaturverzeichnis

- ABERCROMBIE, D. (1967): *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- ALBYN-DAVIS, G., O'NEILL-PIROZZI, T.M. & COON, M. (1997): „Referential cohesion and logical coherence of narration after right hemisphere stroke“, *Brain and Language* 56, 183-210.
- ASBJØRNSEN, A.E. & BRYDEN, M.P. (1996): „Biased attention and the fused dichotic words test“, *Neuropsychologia* 34, 407-411.
- ASBJØRNSEN, A.E. & HUGDAHL, K. (1995): „Attentional effects in dichotic listening“, *Brain & Language* 49, 189-201.
- BARTON, M.I., GOODGLASS, H. & SHAI, A. (1965): „Differential recognition of tachistoscopically presented English and hebrew words in right and left visual fields“, *Perceptual and Motor Skills* 21, 431-437.
- BAUM, S.R., KELSCH DANILOFF, B., DANILOFF, R. & LEWIS, J. (1982): „Sentence comprehension by Broca,s aphasics: effects of some suprasegmental variables“, *Brain & Language* 17, 261-272.
- BECKMAN, M.E. (1996): „When is a syllable not a syllable?“, in: Otake, T. & Cutler, A. (Hrsg.): *Phonological Structure and Language Processing*. Berlin: de Gruyter, 95-123.
- BECKMAN, M.E., EDWARDS, J. & FLETCHER, J. (1992): „Prosodic structure and tempo in a sonority model of articulatory dynamics“, in: Docherty, G.J. & Ladd, D.R. (Hrsg.): *Papers in Laboratory Phonology II. Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, 68-86.
- BEEMAN, M. & CHIARELLO, C. (1997): „Concluding remarks: Getting the whole story right“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 377-389.
- BEEMAN, M. & CHIARELLO, C. (Hrsg.) (1997): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum.
- BEHRENS, S. J. (1985): „The perception of stress and lateralization of prosody“, *Brain & Language* 26, 332-348.
- BERG, T. (1992): „Umriss einer psycholinguistischen Theorie der Silbe“, in: Eisenberg, P. (Hrsg.) (1992): *Silbenphonologie des Deutschen*. Tübingen: Narr, 45-99.
- BINDER, J.R., FROST, J.A., HAMMEKE, T.A., COX, R.W., RAO, S.M. & PRIETO, T. (1997): „Human brain language areas identified by functional magnetic resonance imaging“, *The Journal of Neuroscience* 17, 353-362.

- BJELJAC-BABIC, R., BERTONCINI, J. & MEHLER, J. (1993): „How do 4-days old infants categorize multisyllabic utterances?“, in: *Developmental Psychology* 29, 711-721.
- BLEVINS, J. (1995): „The syllable in phonological theory“, in: Goldsmith, J.A. (Hrsg.): *The Handbook of Phonological Theory* (Blackwell Handbooks in Linguistics). Cambridge, Mass.: Blackwell, 206-244.
- BLUMSTEIN, S. & COOPER, W.E. (1974): „Hemispheric processing of intonation contours“, *Cortex* 10, 146-158.
- BLUMSTEIN, S.E. (1995): „The neurobiology of language“, in: Miller, J.L. & Eimas, P.D. (Hrsg.): *Speech, Language and Communication* (Handbook of Perception and Cognition). San Diego usw.: Academic Press, 339-370.
- BOGEN, J. & GAZZANIGA, M. (1965): „Cerebral commissurotomy in man: minor hemisphere dominance for certain visuospatial functions“, *Journal of Neurosurgery* 23, 394-399.
- BOROD, C. (1992): „Interhemispheric and intrahemispheric control of emotion: a focus on unilateral brain damage“, *Journal of Consulting and Clinical Psychology* 60, 339-348.
- BOROD, J.C., BLOOM, R.L. & SANTSCHI-HAYWOOD, C. (1997): „Verbal aspects of emotional communication“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 285-308.
- BOUCHER, R. & BRYDEN, M.P. (1997): „Laterality effects in the processing of melody and timbre“, *Neuropsychologia* 35, 1467-1473.
- BREGMAN, A.S. (1990): *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Mass., London: MIT Press.
- BROWNELL, H. (1988): „Appreciation of metaphoric and connotative word meaning by brain-damaged patients“, in: Chiarello, C. (Hrsg.): *Right Hemisphere Contributions to Lexical Semantics*. Berlin usw.: Springer, 19-31.
- BRYDEN, M.P. (1982): *Laterality: Functional Asymmetry in the Intact Brain*. New York: Academic Press.
- BRYDEN, M.P. (1988): „An overview of the dichotic listening procedure and its relation to cerebral organisation“, in: *Handbook of Dichotic Listening*. Hrsg. von K. Hugdahl. Chichester: Wiley, 1-41.
- BUCKINGHAM, H.W. JR. (1990): „Principle of sonority, doublet creation and the check off monitor“, in: Nespoulous, J.L. & Villiard, P. (Hrsg.): *Morphology, Phonology and Aphasia*. New York usw.: Springer, 193-205.
- CAPLAN, C. (1994): „Language and the brain“, in: Gernsbacher, M.A. (Hrsg.): *Handbook of Psycholinguistics*. San Diego usw.: Academic Press, 1023-1053.
- CAPLAN, D. (1995): „The cognitive neuroscience of syntactic processing“, in: Gazzaniga, M. (Hrsg.): *The Cognitive Neurosciences* (A Bradford Book). Cambridge, Mass., London: The MIT Press, 971-879.
- CHIARELLO, C. (1997): „On codes of meaning and the meaning of codes: semantic access and retrieval within and between the hemispheres“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 141-160.
- CHIARELLO, C. (Hrsg.) (1988): *Right Hemisphere Contributions to Lexical Semantics*. Berlin usw.: Springer.
- CHOMSKY, N. & HALLE, M. (1968): *The Sound Pattern of English*. New York: Harper

and Row.

- CHOMSKY, N. & LASNIK, H. (1995): „The Theory of Principles and Parameters“, in: Chomsky, N.: *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, Mass., London: MIT Press, 13-127.
- CHOMSKY, N. (1988): *Language and Problems of Knowledge*. Cambridge, Mass.: MIT Press.
- CLAREY, J.C., BARONE, P. & IMIG, T.J. (1992): „Physiology of the thalamus and cortex“, in: Popper, A.N. & Fay, R.R. (Hrsg.): *The Mammalian Auditory Pathway: Neurophysiology*. New York, usw.: Springer, 232-334.
- CLEMENTS, G.N. & KEYSER, S.J. (1983): *CV Phonology: A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press.
- CLEMENTS, G.N. (1990): „The role of the sonority cycle in core syllabification“, in: Kingston, J. & Beckman, M. (Hrsg.): *Papers in Laboratory Phonology 1: Between the grammar and physics of speech*. Cambridge: Cambridge University Press.
- CODE, C. (1987): *Language, Aphasia and the Right Hemisphere*. Chichester usw.: Wiley.
- CODE, C. (1997): „Can the right hemisphere speak?“, in: *Brain and Language* 57, 38-59.
- CODE, C. & BALL, C. (1994): „Syllabification of aphasic recurring utterances: contributions of sonority theory“, *Journal of Neurolinguistics* 8, 257-265.
- COHEN, J., MACWHINNEY, B., FLATT, M. & PROVOST, J. (1993): „Psycope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers“, *Behavior Research Methods, Instruments, and Computers* 25, 257-271.
- COHEN, M.J., BRANCH, W.B. & HYND, G.W. (1994): „Receptive prosody in children with left of right hemisphere dysfunction“, *Brain & Language* 47, 171-181.
- COLES, M.G.H. & RUGG, M.D. (1995): „Event-related brain potentials: an introduction“, in: Rugg, M.D. & Coles, M.G.H. (Hrsg.): *Electrophysiology of Mind. Event-related potentials and cognition* (Oxford Psychology Series 25). Oxford usw.: Oxford University Press, 1-26.
- COSLETT, H.B. & SAFFRAN, E.M. (1997): „Reading and the right hemisphere: evidence from acquired dyslexia“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 105-132.
- CREUTZFELDT, O., OJEMANN, G. & LETTICH, E. (1989a): „Neuronal activity in the human temporal lobe. I. Responses to speech“, *Experimental Brain Research* 77, 451-475.
- CREUTZFELDT, O., OJEMANN, G. & LETTICH, E. (1989b): „Neuronal activity in the human temporal lobe. II. Responses to the subjects own voice“, *Experimental Brain Research* 77, 476-489.
- CREUTZFELDT, O.D. (1983): *Cortex Cerebri*. Berlin usw.: Springer.
- CROSSON, B. (1992): *Subcortical Functions in Language and Memory*. New York, London: Guilford.
- CRUTTENDEN, A. (1986): *Intonation*. Cambridge: Cambridge University Press.
- CRYSTAL, T. & HOUSE, A. (1982): „Segmental durations in connected speech signals:

- preliminary results“, *Journal of the Acoustical Society of America* 72, 705-716.
- CUTLER, A. & BUTTERFIELD, S. (1992): „Rhythmic cues to speech segmentation: Evidence from juncture misperceptions“, *Journal of Memory and Language* 31, 218-236.
- CUTLER, A. & CARTER, D.M. (1987): „The predominance of strong initial syllables in the English vocabulary“, *Computer Speech and Language* 2, 133-142.
- CUTLER, A. & MEHLER, J. (1993): „The periodicity bias“, *Journal of Phonetics* 21, 103-108.
- CUTLER, A. & OTAKE, T. (1994): „Mora or phoneme? Further evidence for language-specific listening“, *Journal of Memory and Language* 33, 824-844.
- CUTLER, A. (1995): „Spoken word recognition and production“, in: Miller, J.L. & Eimas, P.D. (Hrsg.): *Speech, Language and Communication* (Handbook of Perception and Cognition). San Diego usw.: Academic Press, 97-136.
- CUTLER, A., MEHLER, J., NORRIS, D. & SEGUÍ, J. (1986): „The syllable's differing role in the segmentation of French and English“, *Journal of Memory and Language* 25, 385-400.
- CUTLER, A., MEHLER, J., NORRIS, D. & SEGUÍ, J. (1992): „The monolingual nature of speech segmentation by bilinguals“, *Cognitive Psychology* 24, 381-410.
- CUTLER, A., NORRIS, D. & MCQUEEN, J. (1996): „Lexical access in continuous speech: Language-specific realisations of a universal model“, in: Otake, T. & Cutler, A. (Hrsg.): *Phonological Structure and Language Processing*. Berlin: de Gruyter, 227-242.
- CUTLER, A., NORRIS, D. (1988): „The role of strong syllables in segmentation for lexical access“, *Journal of Experimental Psychology: Human Perception and Performance* 14, 113-121.
- DAMASIO, A.R. (1988): „Concluding remarks: neuroscience and cognitive science in the study of language and the brain“, in: Plum, F. (Hrsg.): *Language, Communication and the Brain*. New York: Raven, 275-282.
- DAMASIO, A.R. (1992): „Aphasia“, *The New England Journal of Medicine* 326, 531-538.
- DAMASIO, A.R., BELLUGI, U., DAMASIO, H., POIZNER, H. & VAN GLIDER, J. (1986): „Sign language aphasia during left-hemisphere Amytal injection“, *Nature* 322, 363-365.
- DARWIN, C.J. (1971): „Ear differences in the recall of fricatives and vowels“, *Quarterly Journal of Experimental Psychology* 23, 46-62.
- DAUER, R.M. (1983): „Stress-timing and syllable timing reanalyzed“, *Journal of Phonetics* 11, 51-62.
- DELATTRE, P. (1966): *Studies in French and Comparative Phonetics*. The Hague: Mouton.
- DELIS, D.C., ROBERTSON, L.C. & EFRON, R. (1986): "Hemispheric specialization of memory for visual hierarchical stimuli", *Neuropsychologia* 24, 205-214.
- DÉMONET, J.-F., CHOLLET, F., RAMSAY, S., CARDEBAT, D., NESPOULOUS, J.-L., WISE, R., RASCOL, A. & FRACKOWIAK, R. (1992): „The anatomy of phonological and semantic processing in normal subjects“, *Brain* 115, 1753-1768.
- DOGIL, G. & LÜSCHÜTZKY, H.C. (1989): *Some notes on sonority and segmental strength* (Phonetica Saraviensia 10). Saarbrücken.

- DRESSLER, W. (1985): *Morphonology*. Ann Arbor: Karoma.
- DURAND, JACQUES (1990): *Generative and Non-linear Phonology*. London: Longman.
- EISENBERG, P. (Hrsg.) (1992): *Silbenphonologie des Deutschen*. Tübingen: Narr.
- EMMOREY, K. (1987): „The neurological substrates for prosodic aspects of speech“, *Brain & Language* 30, 305-320.
- FADIGA, L. & GALLESE, V. (1997): „Action representation and language in the brain“, in: *Theoretical Linguistics* 23, 267-280.
- FARNETANI, E. (1997): „Coarticulation and connected speech processes“, in: Hardcastle, W.J. & Laver, J. (Hrsg.): *The Handbook of Phonetic Sciences* (Blackwell Handbooks in Linguistics). Cambridge, Mass., Oxford: Blackwell, 371-404.
- FAUST, M. (1997): „Obtaining evidence of language comprehension from sentence priming“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 161-186.
- FEDIO, P., AUGUST, A., PATRONAS, N, SATO, S. & KUFTA, C. (1997): „Semantic, phonological, and perceptual changes following left and right intracarotid injection“, *Brain and Cognition* 33, 98-117.
- FITCH, R.H., BROWN, C.P., O'CONNOR, K. & TALLAL, P. (1993): „Functional lateralization for auditory temporal processing in male and female rats“, *Behavioral Neurosciences* 107, 844-850.
- FLETCHER, P. & INGHAM, R. (1995): „Grammatical Impairment“, in: Fletcher, P. & MacWhinney, B. (Hrsg.): *The Handbook of Child Language*. Oxford: Blackwell, 603-622.
- FOKES, J. & BOND, Z.S. (1993): „The elusive / illusive syllable“, in: *Phonetica* 50, 102-123.
- FOWLER, C.A. (1995): „Speech production“, in: Miller, J.L. & Eimas, P.D. (Hrsg.): *Speech, Language and Communication* (Handbook of Perception and Cognition). San Diego usw.: Academic Press, 30-59.
- FRAZIER, L. (1987): „Structure in auditory word recognition“, in: *Cognition* 25, 157-187.
- FRAZIER, L. (1995): „Issues of representation in psycholinguistics“, in: Miller, J.L. & Eimas, P.D. (Hrsg.): *Speech, Language and Communication* (Handbook of Perception and Cognition). San Diego usw.: Academic Press, 1-27.
- FRY, D. B. (1979): *The Physics of Speech*. Cambridge: Cambridge University Press 1979.
- GAINOTTI, G. (1989): „Disorders of emotion and affect in patients with unilateral brain damage“, in: *Handbook of Neuropsychology*. Hrsg.v. Boller, F. & Grafman, J., Vol. 3. Amsterdam usw.: Elsevier, 345-361.
- GANDOUR, J. & DARDARANANDA, R. (1983): „Identification of tonal contrasts in Thai aphasic patients“, *Brain and Language* 18, 98-114.
- GARMAN, M. (1990): *Psycholinguistics* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- GERNSBACHER, M.A. (Hrsg.) (1993): *Handbook of Psycholinguistics*. San Diego usw.: Academic Press.
- GESCHWIND, N. & LEVITSKY, W. (1968): „Left-right asymmetries in temporal speech region“, *Science* 161-186-187.

- GIEGERICH, H.J. (1992): *English Phonology*. Cambridge: Cambridge University Press.
- GOLDSMITH, J. & LARSON, G. (1990): „Local modelling and syllabification“, in: Deaton, K., Noske, M. & Ziolkowski, M. (Hrsg.): *Papers from the 26th Annual Regional Meeting of the Chicago Linguistics Society*. Part 2. Chicago: Chicago Linguistic Society, 129-141
- GOLDSMITH, J. (1989): *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
- GOLDSMITH, J. (1993): „Harmonic phonology“, in: Goldsmith, J. (Hrsg.): *The Last Phonological Rule* (Studies in Contemporary Linguistics). Chicago, London: The University of Chicago Press, 21-60.
- GRAMMONT, M. (1939): *Traité de Phonétique*. Paris: Delagrave.
- GREENBERG, JOSEPH (1978): „Some generalizations concerning initial and final consonant clusters“, in: Greenberg, J. (Hrsg.): *Universals of Human Language*. Bd. 2: Phonology. Stanford: Stanford University Press, 243-280.
- GREENBERG, S. (1988): „The ear as a speech analyzer“, *Journal of Phonetics* 16, 139-150.
- GREENBERG, S. (1996): „Understanding speech understanding: Towards a unified theory of speech perception“, in: *Proceedings of the ESCA Workshop on the Auditory Basis of Speech Perception*. Keele University, July 15-July 19, 1996. Hrsg. von Ainsworth, W.A. & Greenberg, S. European Speech Communication Association, 1-8.
- GROSSI, G., SEMENZA, C., CORAZZA, S. & VOLTERRA, V. (1996): „Hemispheric specialization for sign language“, *Neuropsychologia* 34, 737-740.
- HAGGARD, M.P. & PARKINSON, A.M. (1971): „Stimulus and task factors in the perceptual lateralization of speech signals“, *Quarterly Journal of Experimental Psychology* 23, 168-177.
- HALL, T.A. (1992): *Syllable Structure and Syllable Related Processes in German*. Tübingen: Niemeyer.
- HALPERN, D. F. (1992): *Sex Differences in Cognitive Abilities*. New York: Erlbaum.
- HANDEL, S. (1989): *Listening. An Introduction to the Perception of Auditory Events* (A Bradford Book). Cambridge, Mass., London: The MIT Press.
- HARAGUCHI, S. (1996): „Syllable, mora and accent“, in: Otake, T. & Cutler, A. (Hrsg.): *Phonological Structure and Language Processing*. Berlin: de Gruyter, 77-94.
- HARLEY, T.A. (1995): *The Psychology of Language*. Hove: Erlbaum.
- HARSHMAN, R.A. & LUNDY, M.E. (1988): „Can dichotic listening measure ‘degree of laterality’?“, in: *Handbook of Dichotic Listening*. Hrsg. von K. Hugdahl. Chichester: Wiley, 283-321.
- HAUSMANN, M., BEHRENDT-KÖRBITZ, S., KAUTZ, H., LAMM, C., RADELT, F. & GÜNTÜRKÜN, G. (1998): „Sex differences in oral asymmetries during word repetition“, in: *Neuropsychologia*.
- HAYDEN, M.E., KIRSTEN, E. & SINGH, S. (1979): „Role of distinctive features in dichotic presentation of 21 English consonants“, *Journal of the Acoustical Society of America* 65, 1039-1046.
- HAYES, BRUCE (1994): *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- HEIKE, G. (1992): „Zur Phonetik der Silbe“, in: Eisenberg, P. (Hrsg.) (1992): *Silbenphonologie des Deutschen*. Tübingen: Narr, 1-44.

- HEIL, P., RAJAN, R. & IRVINE, D.R.F. (1992): „Sensitivity of neurons in cat primary cortex to tones and frequency-modulated stimuli. I. Effects of variation of stimulus parameters“, in: *Hearing Research* 63, 108-134.
- HEILMAN, K.M., BOWERS, D., SPEEDIE, L. & COSLETT, H.B. (1984): „Comprehension of affective and nonaffective prosody“, *Neurology* 34, 917-921.
- HELLIGE, J.B. (1993): *Hemispheric Asymmetry: What's Right and What's Left* (Perspectives in Cognitive Neuroscience). Cambridge, Mass.: Harvard University Press.
- HERRERO, J.V. & HILLIX, W.A. (1990): „Hemispheric performance in detecting prosody: a competitive dichotic listening task“, *Perceptual and Motor Skills* 71, 479-486.
- HISCOCK, M., INCH, R., JACEK, C., HISCOCK-KALIL, C. & KALIL, K. (1994): „Is there a sex difference in human laterality? I. An exhaustive survey of auditory laterality studies from six neuropsychology journals“, *Journal of Clinical and Experimental Neuropsychology* 16, 423-435.
- HISCOCK, M., ISRAELIAN, M., INCH, R., JACEK, C. & HISCOCK-KALIL, C. (1995): „Is there a sex difference in human laterality? II. An exhaustive survey of visual laterality studies from six neuropsychology journals“, *Journal of Clinical and Experimental Neuropsychology* 17, 590-610.
- HUGDAHL, K. (1995): „Dichotic listening: probing temporal lobe functional integrity“, in: *Brain Asymmetry* (A Bradford Book). Hrsg. v. Davidson, R.J. & Hugdahl, K. Cambridge, Mass.: MIT Press, 123-156.
- HUGDAHL, K. (1995): „Dichotic listening: probing temporal lobe functional integrity“, in: Davidson, R.J. (Hrsg.): *Brain Asymmetry*. Cambridge, Mass.: MIT Press 1995, 123-156.
- HYMAN, LARRY (1985): *A Theory of Phonological Weight*. Dordrecht: Foris.
- IVRY, P. & LEBBY, R. (1993): „Hemispheric differences in auditory perception are similar to those found in visual perception“, *Psychological Science* 4, 41-45.
- IVRY, R. & LEBBY, P.C. (1997): „The neurology of consonant perception: specialized module or distributed processors“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 3-26.
- JAKOBSON, R. & HALLE, M. (1956): *Fundamentals of Language*. The Hague: Mouton.
- JÄNCKE, L. & LUPPEN, A. (1993): „Dichotischer Konsonant-Vokal-Recall-Test“, *Sprache Stimme Gehör* 17, 54-58.
- JÄNCKE, L. (1992): „A German dichotic consonant vowel recall and monitoring test. Reliability and comparison of both procedures“, *Zeitschrift für Neuropsychologie* 3.
- JÄNCKE, L. (1992): „Stimulus-dependent ear asymmetry in a dichotic monitoring task“, *Perceptual and Motor Skills* 75, 691-695.
- JÄNCKE, L. (1993): „Do ear advantage scores obtained in a consonant-vowel recall test vary with respect to the required response condition?“, *Neuropsychologia* 31, 499-501.
- JÄNCKE, L. (1994): „Hemispheric priming affects right-ear advantage in dichotic listening“, *International Journal of Neuroscience* 74, 71-77.
- JÄNCKE, L., STEINMETZ, H. & VOLKMANN, J. (1992): „Dichotic listening: what does it measure?“, *Neuropsychologia* 30, 941-950.

- JESPERSEN, O. (1904): *Lehrbuch der Phonetik*. Autorisierte Übersetzung von Hermann Davidsen. Leipzig und Berlin: Teubner.
- JOANETTE, Y. & GOULET, P. (1990): „Narrative discourse in right-brain-damaged right-handers“, in: Joannette, Y. & Brownell, H. (Hrsg.): *Discourse Ability and Brain Damage. Theoretical and Empirical Perspectives*. New York usw.: Springer, 131-153.
- JOANETTE, Y., GOULET, P. & HANNEQUIN, D. (1990): *Right Hemisphere and Verbal Communication*. New York usw.: Springer.
- JONES-GOTMAN, M., ROULEAU, I. & SNYDER, P. J. (Hrsg.): *Clinical and Research Contributions of the Intracarotid Amobarbital Procedure to Neuropsychology* (Brain and Cognition Special Issue). *Brain and Cognition* 33 (1997).
- JUSZCYK, P.W. (1995): „Language acquisition: speech sounds and the beginning of phonology“, in: Miller, J.L. & Eimas, P.D. (Hrsg.): *Speech, Language and Communication* (Handbook of Perception and Cognition). San Diego usw.: Academic Press, 263-295.
- KAHN, D. (1980): *Syllable-based Generalization in English Phonology* (Doctoral Dissertation, MIT 1976). New York: Garland.
- KEATING, P. (1983): „Comments on the jaw and syllable structure“, in: *Journal of Phonetics* 11, 401-406.
- KENSTOWICZ, M. (1994): *Phonology in Generative Grammar* (Blackwell Textbooks in Linguistics). Cambridge, Mass, Oxford: Blackwell.
- KIMURA, D. (1961): „Cerebral dominance and the perception of verbal stimuli“, *Canadian Journal of Psychology* 15, 166-171.
- KIMURA, D. (1964): „Left-right differences in the perception of melodies“, *Quarterly Journal of Experimental Psychology* 16, 355-358.
- KIMURA, D. (1967): „Functional asymmetry of the brain in dichotic listening“, *Cortex* 3, 163-168.
- KIMURA, D. (1983): „Sex differences in cerebral organization for speech and praxic functions“, *Canadian Journal of Psychology* 37, 19-35.
- KIMURA, D. (1992): *Neuromotor Mechanisms in Human Communication* (Oxford Psychology Series 20). New York usw.: Oxford University Press.
- KIMURA, D. (1996): „Sex, sexual orientation and sex hormones influence human cognitive function“, *Current Opinion in Neurobiology* 6, 259-263.
- KISCHKA, U., WALLECH, C.-W. & WOLF, G. (Hrsg): *Methoden der Hirnforschung. Eine Einführung*. Heidelberg, Berlin: Spektrum Akademischer Verlag 1997.
- KLATT, D.H. (1989): „Review of selected models of speech perception“, in: Marslen-Wilson, W.D. (Hrsg.): *Lexical Representation and Process*. Cambridge, Mass.: MIT Press, 169-226.
- KOHLER, K. (1986): „Invariance and variability in speech timing: from utterance to segment“, in: Perkell, J.S. & Klatt, D. (Hrsg.): *Invariance and Variability in Speech Processes*. Hillsdale, N.J.: Erlbaum, 268-289.
- KOHLER, K.J. (1977): *Einführung in die Phonetik des Deutschen* (Grundlagen der Germanistik). Berlin: Erich Schmidt.
- KOLB, B. & WHISHAW, I.Q. (1993): *Neuropsychologie*. Heidelberg, Berlin, Oxford: Spektrum (Originalausgabe: *Fundamentals of Human Neuropsychology*, Third Edition. New York, Oxford: Freeman 1990).

- KUHL, P. (1993): „Developmental speech perception: Implications for models of language impairment“, in: *Annals of the New York Academy of Sciences* 628, 248-263.
- LADEFOGED, PETER (1993³): *A Course in Phonetics*. Fort Worth usw.: Harcourt Brace College Publishers.
- LAHIRI, A. & MARSLER-WILSON, W.D. (1991): „The mental representation of lexical form: a phonological approach to the lexicon“, *Cognition* 38, 245-294.
- LAKE, D.A. & BRYDEN, M.P. (1976): „Handedness and sex differences in hemispheric asymmetry“, *Brain & Language* 3, 266-282.
- LANGNER, G. (1992): „Periodicity coding in the auditory system“, in: *Hearing Research* 60, 115-142.
- LARSON, G. (1990): „Local computational networks and the distribution of segments in the Spanish syllable“, in: Deaton, K., Noske, M. & Ziolkowski, M. (Hrsg.): *Papers from the 26th Annual Regional Meeting of the Chicago Linguistics Society*. Part 2. Chicago: Chicago Linguistic Society, 257-272.
- LAVER, J. (1995): *Principles of Phonetics* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- LENNEBERG, E.H. (1972): *Biologische Grundlagen der Sprache*. Frankfurt: Suhrkamp (Originalausgabe: *Biological Foundations of Language*. New York: Wiley 1967).
- LEONARD, C.L., WATERS, G.S & CAPLAN, D. (1997): „The use of contextual information related to general world knowledge by right brain damaged subjects“, *Brain and Language* 57, 343-359.
- LEUNINGER, H. (1989): *Neurolinguistik*. Opladen: Westdeutscher Verlag.
- LEUNINGER, H. (1996): *Reden ist Schweigen, Silber ist Gold*. München: dtv.
- LEVELT, W.J.M. & WHEELDON, L. (1994): „Do speakers have access to a mental syllabary“, *Cognition* 50, 239-269.
- LEVELT, W.J.M. (1989): *Speaking. From Articulation to Intention* (ACL-MIT Press Series in Natural Language Processing. A Bradford Book). Cambridge, Mass., London: The MIT Press.
- LEVIN, J. (1985): *A Metrical Theory of the Syllable*. Doctoral Dissertation, MIT.
- LEVITT, A., HEALEY, A.F. & FENDRICH, D.W. (1991): „Syllable-internal structure and the sonority hierarchy: differential evidence from lexical decision, naming, and reading“, *Journal of Psycholinguistic Research* 20, 377-363.
- LEVY-AGRESTI, J. & SPERRY, R.W. (1968): „Differential perceptual capacities in major and minor hemispheres“, *Proceedings of the National Academy of Science, U.S.A.* 61, 115.
- LEY, R.G. & BRYDEN, M.P. (1984): „A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content“, *Brain & Cognition* 1, 3-9.
- LIEBERMAN, P. (1984): *The Biology and Evolution of Language*. Cambridge, Mass.: Harvard University Press.
- LINDBLOM, B. (1983): „Economy of speech gestures“, in: MacNeilage, P.F. (Hrsg.): *The Production of Speech*. New York usw.: Springer, 217-245.
- MCGLONE, J. (1977): „Sex differences in the cerebral organization of verbal functions in patients with unilateral brain lesions“, *Brain* 100, 775-793.

- MCGLONE, J. (1978): „Sex differences in functional brain asymmetry“, *Cortex* 14, 122-128.
- MCGLONE, J. (1980): „Sex differences in human brain organization: a critical survey“, *Behavioral and Brain Sciences* 3, 215-227.
- MCGLONE, J. (1984): „Speech comprehension after unilateral injection of sodium amytal“, *Brain and Language* 22, 150-157.
- MCQUEEN, J. & CUTLER, A. (1997): „Cognitive Processes in Speech Perception“, in: Hardcastle, W.J. & Laver, J. (Hrsg.): *The Handbook of Phonetic Sciences* (Blackwell Handbooks in Linguistics). Cambridge, Mass., Oxford: Blackwell, 566-585.
- MCQUEEN, J.M., NORRIS, D.G. & CUTLER, A. (1994): „Competition in spoken word recognition: Spotting words in other words“, *Journal of Experimental Psychology: Language, Memory and Cognition* 20, 621-638.
- MEHLER, J. & DUPOUX, E. (1990): *Naître Humain*. Paris: Odile Jacob.
- MEHLER, J. & SEGUÍ, J. & FRAUENFELDER, U. (1981): „The role of the syllable in language acquisition and perception“, in: Myers, T., Laver, J. & Anderson, J. (Hrsg.): *The Cognitive Representation of Speech*. Amsterdam: North Holland, 295-305.
- MEHLER, J., BERTONCINI, J., DUPOUX, E. & PALLIER, C. (1996): „The role of supra-segmentals in speech perception and acquisition“, in: Otake, T. & Cutler, A. (Hrsg.): *Phonological Structure and Language Processing*. Berlin: de Gruyter, 145-169.
- MEHLER, J., DUPOUX, E., NAZZI, T. & DEHAENE-LAMBERTZ, G. (1995): „Coping with linguistic diversity: The infant's viewpoint“, in: Morgan, J.L & Demuth, K. (Hrsg.): *From Signal to Syntax*. Hillsdale, N.J.: Erlbaum, 101-116.
- MEHLER, J., JUSZYK, P.W., LAMBERTZ, G., HALSTED, G., BERTONCINI, J. & AMIELTISON, C. (1988): „A precursor of language acquisition in young infants“, in: *Cognition* 29, 143-187.
- MEHLER, J., SEBASTIÁN, N., ALTMANN, G., DUPOUX, E., CHRISTOPHE, A. & PALLIER, C. (1993): „Understanding compressed sentences: The role of rhythm and meaning“, in: *Annals of the New York Academy of Sciences* 628, 272-282.
- MERZENICH, M.M., JENKINS, W.M., JOHNSTON, P., SCHREINER, C., MILLER, S.L. & TALLAL, P. (1996): „Temporal processing deficits of language-learning impaired children ameliorated by training“, *Science* 271, 77-81.
- MILLER, J.L. & EIMAS, P.D. (1995): „Speech perception: From signal to word“, *Annual Review of Psychology* 46, 467-492.
- MILNER, B., TAYLOR, L. & SPERRY, R.W. (1968): „Lateralized suppression of dichotically presented digits after commissural section in man“, *Science* 161, 184-186.
- MISHKIN, M. & FORGAYS, D.G. (1952): „Word recognition as a function of retinal locus“, *Journal of Experimental Psychology* 43, 43-48.
- MOEN, I. & SUNDET, K. (1996): „Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage“, *Brain and Language* 53, 267-281.
- MOEN, I. (1993): „Functional lateralization of the perception of Norwegian word-tone – evidence from a dichotic listening experiment“, *Brain & Language* 44, 400-413.
- MOLLOY, R., BROWNELL, H. & GARDNER, H. (1990): „Discourse comprehension by right-hemisphere stroke patients: deficits of prediction and revision“, in: Joannette,

- Y. & Brownell, H. (Hrsg.): *Discourse Ability and Brain Damage. Theoretical and Empirical Perspectives*. New York usw.: Springer, 113-130.
- MONDOR, T.A. & BRYDEN, M.P. (1991): „The influence of attention on the dichotic REA“, *Neuropsychologia* 1991, 1179-1190.
- MORAIS, J., KOLINSKY, R. & NAKAMURA, M. (1996): „The psychological reality of speech units in Japanese“, in: Otake, T. & Cutler, A. (Hrsg.): *Phonological Structure and Language Processing*. Berlin: de Gruyter, 203-216.
- NELKEN, I., PRUT, Y., VADDIA, E. & ABELES, M. (1994): „Population responses to multifrequency sounds in the cat auditory cortex“, in: *Hearing Research* 72, 223-236.
- NORRIS, D.G., MCQUEEN, J.M. & CUTLER, A. (1995): „Competition and segmentation in spoken word recognition“, *Journal of Experimental Psychology: Language, Memory and Cognition* 21, 1209-1228.
- NOSKE, R. (1992): *A Theory of Syllabification and Segmental Alternation*. Tübingen: Niemeyer.
- NYGAARD, L.C. & PISONI, D.B. (1995): „Speech perception: New directions in theory and research“, in: Miller, J.L. & Eimas, P.D. (Hrsg.): *Speech, Language and Communication* (Handbook of Perception and Cognition). San Diego usw.: Academic Press, 63-96.
- OHALA, J. & KAWASAKI, H. (1984): „Prosodic phonology and phonetics“, *Phonology Yearbook* 1, 87-92.
- OJEMANN, G.A. (1988): „Effects of cortical and subcortical stimulation on human language“, in: Plum, F. (Hrsg.): *Language, Communication and the Brain*. New York: Raven, 101-115.
- OJEMANN, G.A. (1989): „Cortical organization of language“, in: *The Journal of Neuroscience* 11, 2281-2287.
- OLDFIELD, R.C. (1971): „The assessment and analysis of handedness: the Edinburgh inventory“, *Neuropsychologia* 9, 97-113.
- OSTERHOUT, L. & HOLCOMB, P.J. (1995): „Event-related potentials and language comprehension“, in: Rugg, M.D. & Coles, M.G.H. (Hrsg.): *Electrophysiology of Mind. Event-related potentials and cognition* (Oxford Psychology Series 25). Oxford usw.: Oxford University Press, 171-215.
- OTAKE, T., HATANO, G., CUTLER, A. & MEHLER, J. (1993): „Mora or syllable? Speech segmentation in Japanese“, *Journal of Memory and Language* 32, 258-278.
- PAPANICOLAOU, A.C. (1995): „An introduction to magnetoencephalography with some applications“, *Brain and Cognition* 27, 331-352.
- PELL, M.D. & BAUM, S.R. (1997a): „The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults“, *Brain and Language* 57, pp. 80-99.
- PELL, M.D. & BAUM, S.R. (1997b): „Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody“, *Brain & Language* 57, 195-214.
- PENFIELD, W. & ROBERTS, L. (1959): *Speech and Brain Mechanisms*. Princeton: Princeton University Press.
- PERKINS, J.M., BARAN, J.A. & GANDOUR, J. (1996): „Hemispheric specialization in processing intonation contours“, *Aphasiology* 10, 343-362.
- PFITZINGER, H., BURGER, S. & HEID, S. (1996): „Syllable detection in read and spon-

- taneous speech“, in: *Proceedings ICSLP*, Philadelphia, Vol. 2, 1261-1264.
- PIAZZA, D.M. (1980): „The influence of sex and handedness in the hemispheric specialization of verbal and nonverbal tasks“, *Neuropsychologia* 18, 163-176.
- PICKETT, J.M. (1985²): *The Sounds of Speech Communication* (Perspectives in Audio-logy). Baltimore: University Park Press.
- PIERREHUMBERT, J. & NAIR, R. (1995): „Word games and syllable structure“, *Language and Speech* 38, 77-114.
- PIKE, K. (1943): *Phonetics*. Ann Arbor: The University of Michigan Press.
- PINEL, J.P.J. (1997): *Biopsychologie. Eine Einführung*. Heidelberg, Berlin: Spektrum Akademischer Verlag.
- PIZZAMIGLIO, L., CALTAGIRONE, C. & ZOCCOLATTI, P. (1989): „Facial expression of emotion“, in: Boller, F. & Grafman, J. (Hrsg.): *Handbook of Neuropsychology*. Vol. 3. Amsterdam usw.: Elsevier, 383-401.
- POEPEL, D. (1996): „A critical review of PET studies of phonological processing“, *Brain and Language* 55, 317-351.
- POMPINO-MARSCHALL, B. (1995): *Einführung in die Phonetik*. Berlin, New York: de Gruyter.
- POPPER, A.N. & FAY, R.R. (Hrsg.) (1991): *The Mammalian Auditory Pathway: Neurophysiology* (Springer Handbook of Auditory Research). New York usw.: Springer.
- POPPER, A.N. & FAY, R.R. (Hrsg.) (1992): *The Mammalian Auditory Pathway: Neurophysiology*. New York, usw.: Springer.
- PORTELE, T. & HEUFT, B. (1994): „‘Seinen’ oder nicht ‘seinen’ - Prosodische Markierung der Silbenkerne reduzierter Ensilben“, *Fortschritte der Akustik DAGA*, 1313-1316.
- POSNER, M.I. & RAICHLER, M.E. (1996): *Bilder des Geistes*. Heidelberg: Spektrum.
- PRICE, P. J. (1980): „Sonority and syllabicity: acoustic correlates of perception“, *Phonetica* 37, 327-343.
- PRINCE, A. & SMOLENSKY, P. (1993): *Optimality Theory*. Rutgers Center for Cognitive Science.
- PUGH K.R., SHAYWITZ B.A., SHAYWITZ S.E., CONSTABLE R.T., SKUDLARSKI P., FULBRIGHT R.K., BRONEN R.A., SHANKWEILER D.P., KATZ L., FLETCHER J.M., GORE J.C. (1996): „Cerebral organization of component processes in reading“, *Brain* 119, 1221-1238.
- RAICHLER, M.E. (1994): „Images of the mind: studies with modern imaging techniques“, *Annual Review of Psychology* 54, 133-156.
- REMEZ, R. & RUBIN, P.E. (1990): „On the perception of speech from time-varying acoustic information: contributions of amplitude variation“, *Perception and Psychophysics* 48, 313-325.
- REPP, B.H. (1982): „Phonetic trading relations and context effects: new experimental evidence for a speech mode of perception“, *Psychological Bulletin* 92, 81-110.
- RICHARDS, L. & CHIARELLO, C. (1997): „Activation without selection: parallel right hemisphere roles in language and intentional movement?“, *Brain and Language* 57, 151-178.
- RIZZOLATTI, G. & ARBIB, M.A. (1998): „Language within our grasp“, in: *Trends in*

Neuroscience 21.

- RIZZOLATTI, G., FADIGA, L., GALLESE, V. & FOGASSI, L. (1996): „Premotor cortex and the recognition of motor actions“, in: *Cognitive Brain Research* 3, 131-141.
- ROBIN, D.A., TRANEL, D. & DAMASIO, H. (1990): „Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions“, *Brain and Language* 39, 539-555.
- ROCA, I. (1994): *Generative Phonology* (Linguistic Theory Guides). London, New York: Routledge.
- ROCH LECOURE, A., LHERMITTE, F. & BRYANS, B. (1983): *Aphasiology*. London: Baillière Tindall.
- RODE C., WAGNER M. & GÜNTÜRKÜN O. (1995): „Menstrual cycle affects functional cerebral asymmetries“, *Neuropsychologia* 33, 855-65.
- ROSS, E.D. (1981): „The aprosodias: functional-anatomic organization of the affective components of language in the right hemisphere“, *Archives of Neurology* 38, 561-569.
- ROSS, E.D., THOMPSON, R.D. & YENKOSKY, J. (1997): „Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions“, *Brain and Language* 56, 27-54.
- SAUSSURE, F. DE (1972): *Cours de Linguistique Générale*. Hrsg. v. C. Bally und A. Séchehaye. Kritische Ausgabe von Tullio de Mauro. Paris: Payot (Bibliothèque Scientifique Payot).
- SCHNELLE, H. (1991): *Die Natur der Sprache. Die Dynamik der Prozesse des Sprechens und Verstehens* (Grundlagen der Kommunikation und Kognition). Berlin, New York: de Gruyter.
- SCHREINER, C.E. (1995): „Order and disorder in auditory cortical maps“, in: *Current Opinion in Neurobiology* 5, 498-496.
- SCHWARTZ, J. & TALLAL, P. (1980): „Rate of acoustic change may underlie hemispheric specialization for speech perception“, *Science* 207, 1380-1381.
- SCHWARZ, A., KISCHKA, U. & RIHS, F. (1997): „Funktionelle bildgebende Verfahren“, in: Kischka, U., Wallesch, C.-W. & Wolf, G. (Hrsg.): *Methoden der Hirnforschung. Eine Einführung*. Heidelberg, Berlin: Spektrum Akademischer Verlag, 295-318.
- SEBASTIÁN-GALLÉS, N., DUPOUX, E., SEGUÍ, J. & MEHLER, J. (1992): „Contrasting syllabic effects in Catalan and Spanish“, *Journal of Memory and Language* 31, 19-32.
- SEGUI, J., DUPOUX, E. & MEHLER, J. (1990): „The role of the syllable in speech segmentation, phoneme identification and lexical access“, in: Altman, G.T.M. (Hrsg.): *Cognitive Models of Speech Processing* (ACL-MIT Press Series in Natural Language Processing. A Bradford Book). Cambridge, Mass.: MIT Press, 263-280.
- SELKIRK, E. (1982): „The syllable“, in: Hulst, H.v.d. & Smith, N. (Hrsg.): *The Structure of Phonological Representations*. Part II. Dordrecht: Foris, 337-383.
- SELKIRK, E. (1984): „On the major class features and syllable theory“, in: Aronoff, M. & Oehrle, R.T. (Hrsg.): *Language Sound Structure. Studies in Phonology presented to Morris Halle by his Teachers and Students*. Cambridge, Mass.: MIT Press.
- SENDLMEIER, W.F. (1995): „Feature, phoneme, syllable or word: How is speech men-

- tally represented“, *Phonetica* 52, 131-143.
- SHANKWEILER, D. & STUDDERT-KENNEDY, M. (1967): „Identification of consonants and vowels presented to left and right ears“, *Quarterly Journal of Experimental Psychology* 19, 59-63.
- SHANNON, R. V., ZENG, F.-G., KAMATH, V., WYGONSKI, J. & EKELID, M. (1995): „Speech recognition with primarily temporal cues“, *Science* 279, 303-304.
- SHAYWITZ, B.A., SHAYWITZ, S.E., PUGH, K.R., CONSTABLE, R.T., SKUDLARSKI, P., FULBRIGHT, R.K., BRONEN, R.A., FLETCHER, J.M., SHANKWEILER, D.P. & KATZ, L. ET AL. (1995): „Sex differences in the functional organization of the brain for language“, *Nature* 373, 607-609.
- SHIPLEY-BROWN, F., DINGWALL, W.O., BERLIN, C.I., YENI-KOMSHIAN, G. & GORDON-SALANT, S. (1988): „Hemispheric processing of affective and linguistics intonation contours in normal subjects“, *Brain and Language* 33, 16-26.
- SIEVERS, E. (1901⁵): *Grundzüge der Phonetik*. Leipzig: Breitkopf und Härtel.
- SILVERMAN, K.E. & PIERREHUMBERT, J.B. (1990): „The timing of prenuclear high accents in English“, in: Kingston, J. & Beckman, M.E. (Hrsg.): *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press 72-106.
- SIMOS, P.S., MOLFESE, D.L. & BRENDEN, R.A. (1997): Behavioral and electrophysiological indices of voicing-cue discrimination: laterality patterns and development“, *Brain and Language* 57, 122-150.
- SPELLACY, F. & BLUMSTEIN, S. (1970): „The influence of language set on ear preference in phoneme recognition“, *Cortex* 6, 430-439.
- SPENCER, A. (1996): *Phonology*. Cambridge, Mass., Oxford: Blackwell.
- SPRINGER, S.P. & DEUTSCH, G. (1995): *Linkes – rechtes Gehirn*. 3. Auflage. (Originalausgabe: *Left Brain, Right Brain* (A Series of Books in Psychology). New York: Freeman) Heidelberg usw.: Spektrum Akademischer Verlag.
- STARK, K.H. & STARK, J.A. (1990): „Syllable structure in Wernicke’s aphasia“, in: Ne-spoulous, J.L. & Villiard, P. (Hrsg.): *Morphology, Phonology and Aphasia*. New York usw.: Springer, 213-234.
- STEMMER, B. & JOANETTE, Y. (1997): „The interpretation of narrative discourse of brain-damaged individuals within the framework of a multilevel discourse model“, in: Beeman, M. & Chiarello, C. (Hrsg.): *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah, N.J.: Erlbaum, 329-348.
- STERIADE, D. (1982): *Greek Prosodies and the Nature of Syllabification*. Doctoral Dissertation, MIT.
- STUDDERT-KENNEDY, M. & SHANKWEILER, D. (1970): „Hemispheric specialization for speech perception“, *Journal of the Acoustical Society of America* 48, 579-594.
- SUSSMAN, H. (1984): „A neuronal model for syllable representation“, *Brain and Language* 22, 167-177.
- TALLAL, P. & NEWCOMBE, F. (1978): „Impairment of auditory perception and language comprehension in dysphasia“, *Brain and Language* 5, 13-24.
- TALLAL, P., MILLER, S. & FITCH, R. H. (1993): „Neurobiological basis for speech: a case for the preeminence of temporal processing“, *Annals of the New York Academy of Science* 682, 27-47.

- TALLAL, P., MILLER, S.L., BEDI, G. BYMA, G., WANG, X., NAGARAJAN, S.S., SCHREINER, C. & JENKINS, W. (1996): „Language comprehension in language-learning impaired children improved with acoustically modified speech“, *Science* 271, 81-84.
- TARTTER, C. (1988): „Acoustic and phonetic feature effects in dichotic listening“, in: *Handbook of Dichotic Listening*. Hrsg. von K. Hugdahl. Chichester: Wiley, 283-321.
- TREIMAN, R. & DANIS, C. (1988): „Syllabification of intervocalic consonants“, *Journal of Memory and Language* 27, 87-104.
- TREIMAN, R. & KESSLER, B. (1995): „In defense of an onset-rime syllable structure for English“, *Language and Speech* 38, 127-142.
- TREIMAN, R. (1983): „The structure of spoken syllables: Evidence from novel word games“, *Cognition* 15, 49-74.
- TREIMAN, R., STRAUB, K. & LAVERY, P. (1994): „Syllabification of bisyllabic non-words: evidence from short-term memory errors“, *Language and Speech* 37, 45-60.
- TRUBETZKOY, N.S. (1977⁶): *Grundzüge der Phonetik*. Göttingen: Vandenhoeck & Ruprecht.
- VALLET, E.M., KREUTZER, M.L. & RICHARD, J.P. (1992): „Syllable phonology and song segmentation: testing their salience in female canaries“, *Behaviour* 121, 155-167.
- VAN LANCKER, D. & FROMKIN, V.A. (1973): „Hemispheric specialization for pitch and ‘tone’: Evidence from Thai“, *Journal of Phonetics* 1, 101-109.
- VAN LANCKER, D. & FROMKIN, V.A. (1978): „Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers“, *Journal of Phonetics* 6, 19-23.
- VAN LANCKER, D. & SIDTIS, J.J. (1992): „The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: all errors are not created equal“, *Journal of Speech and Hearing Research* 35, 963-970.
- VAN LANCKER, D. (Hrsg.) (1997): *Current Studies of Right Hemisphere Function. Brain & Language* 57, Special Issue.
- VENNEMANN, T. (1982): „Zur Silbenstruktur der deutschen Standardsprache“, in: Vennemann, T. (Hrsg.): *Silben, Segmente, Akzente*. Tübingen: Niemeyer.
- VENNEMANN, T. (1986): *Neuere Entwicklungen in der Phonologie*. Berlin, New York, Amsterdam: Mouton de Gruyter.
- VENNEMANN, T. (1988): *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin usw.: de Gruyter.
- VENNEMANN, T. (1988): *Preference Laws for Syllable Structure*. Berlin, New York, Amsterdam: Mouton de Gruyter.
- WADA, J. & RASMUSSEN, T. (1960): „Intracarotid injection of sodium amytal for the lateralization of speech dominance“, *Journal of Neurosurgery* 17, 266-282.
- WALLESCH, C.-W. & DEUSCHL, G. (1997): „Elektrophysiologie am Menschen“, in: Kischka, U., Wallesch, C.-W. & Wolf, G. (Hrsg.): *Methoden der Hirnforschung. Eine Einführung*. Heidelberg, Berlin: Spektrum Akademischer Verlag, 167-185.
- WEBSTER, D.B., POPPER, A.N. & FAY, R.R. (Hrsg.) (1992): *The Mammalian Auditory Pathway: Neuroanatomy* (Springer Handbook of Auditory Research). New York

usw.: Springer.

- WEINTRAUB, S., MESULAM, M. & KRAMER, L. (1981): „Disturbances in prosody: a right hemisphere contribution to language“, *Archives of Neurology* 38, 742-744.
- WEISS, M.A. & HOUSE, A.S. (1973): „Perception of dichotically presented vowels“, *Journal of the Acoustical Society of America* 53, 51-58.
- WEXLER, B.E. & HALWES, T. (1983): „Increasing the power of dichotic methods: the fused dichotic words test“, *Neuropsychologia* 21, 59-66.
- WIESE, R. (1988): *Silbische und Lexikalische Phonologie. Studien zum Chinesischen und zum Deutschen*. Tübingen: Niemeyer.
- WIESE, R. (1996): *The Phonology of German*. Oxford: Clarendon.
- WILLIAMS, S.M. (1987): „Stimulus-independence of dichotic testing for perceptual asymmetry: sentences versus unrelated words“, *Perceptual and Motor Skills* 65, 97-98.
- WITELSON, S.F. (1976): „Sex and the single hemisphere: specialization of the right hemisphere for spatial processing“, *Science* 193, 425-427.
- WOLF, M.E. & GOODALE, M.A. (1987): „Oral asymmetries during verbal and non-verbal movements of the mouth“, *Neuropsychologia* 25, 375-396.
- ZATORRE, R. J., MEYER, E., GJEDDE, A. & EVANS, A. C. (1996): „PET studies of phonetic processing and speech: review, replication and reanalysis“, *Cerebral Cortex* 6, 21-30.
- ZATORRE, R.J., EVANS, A.C., MEYER, E. & GJEDDE, A. (1992): „Lateralization of phonetic processing and pitch discrimination in speech processing“, *Science* 256, 846-849.
- ZEC, D. (1995): „Sonority Constraints on Syllable Structure“, *Phonology* 12, 85-129.
- ZENNER, H.-P. (1994): *Hören. Physiologie, Biochemie, Zell- und Neurobiologie*. Stuttgart, New York: Thieme.
- ZILLES, K. & REHKÄMPER, G. (1994): *Funktionelle Neuroanatomie*. Berlin usw.: Springer.
- ZURIF, E. & MENDELSON, M. (1972): „Hemispheric asymmetries for the perception of speech sounds: the influence of intonation and structure“, *Perception and Psychophysics* 1, 329-332.
- ZWITSERLOOD, P., SCHRIEFERS, H., LAHIRI, A. & DONSELAAR, W. VAN (1993): „The role of syllables in the perception of spoken Dutch“, *Journal of Experimental Psychology: Learning, Memory and Cognition* 19, 1-12.

Curriculum Vitae

Judith Meinschaefer

Geburtsdatum 27.06.1969 in Herne

Gegenwärtige Position

seit 1998 Wissenschaftliche Mitarbeiterin, Fachbereich Sprachwissenschaft, Universität Konstanz.

Ausbildung

1998 Promotion (Dr.phil.), Allgemeine Sprachwissenschaft, Ruhr-Universität Bochum.

1995-1997 Graduiertenkolleg „Kognition, Gehirn und Neuronale Netze“ (KoGNet), Ruhr-Universität Bochum.

1995 M.A., Romanische Sprachwissenschaft, Ruhr-Universität Bochum.

1990-1991 Studium der Allgemeinen Sprachwissenschaft, Université François Rabelais, Tours, Frankreich.

1988-1994 Magisterstudium der Romanischen Philologie, Hispanistik, der Allgemeinen Sprachwissenschaft, Philosophie und Neueren Geschichte, Ruhr-Universität Bochum.

Stipendien

1998 Promotionsstipendium, Wilhelm-und-Günter-Esser-Stiftung

1998 EG-Teilstipendium, Forschungspraktikum in Italien

1995-1997 DFG-Promotionsstipendium

1990-1991 DAAD-Stipendium, Studium in Frankreich

1990 Sprachkursstipendium, Spanisches Aussenministerium

Anstellungen

seit 1998 Wissenschaftliche Mitarbeiterin, Universität Konstanz

1998 Wissenschaftliche Hilfskraft, Ruhr-Universität Bochum

1992-1995 Studentische Hilfskraft, Ruhr-Universität Bochum