

Deutsche Phonetik

Eine Einführung

Koloman Brenner
Balázs Huszka
Csaba Werk-Marinkás



BÖLCSÉSZ
KONZORCIUM



Magyarország célba ér



2006

Kiadta a Bölcsész Konzorcium

A Konzorcium tagjai:

- Eötvös Loránd Tudományegyetem
- Pécsi Tudományegyetem
- Szegedi Tudományegyetem
- Debreceni Egyetem
- Pázmány Péter Katolikus Egyetem
- Berzsenyi Dániel Főiskola
- Eszterházy Károly Főiskola
- Károli Gáspár Református Egyetem
- Miskolci Egyetem
- Nyíregyházi Főiskola
- Pannon Egyetem
- Kodolányi János Főiskola
- Szent István Egyetem

Szerkesztette: Brenner Koloman

Műszaki szerkesztő: Szabó János

Szakmai lektor: Mády Katalin

Nyelvi lektor: Isabelle von Zitzowitz

A kötet megjelenése az Európai Unió támogatásával,
a Nemzeti Fejlesztési Terv keretében valósult meg:

A felsőoktatás szerkezeti és tartalmi fejlesztése – HEFOP-3.3.1-P-2004-09-0134/1.0

ISBN 963 9704 83 0

© Bölcsész Konzorcium. Minden jog fenntartva!

Bölcsész Konzorcium HEFOP Iroda

H-1088 Budapest, Múzeum krt. 4/A.

tel.: (+36 1) 485-5200/5772 – dekanbtk@ludens.elte.hu

Koloman Brenner – Balázs Huszka – Csaba Werk-Marinkás

Deutsche Phonetik – eine Einführung

Budapest – Veszprém

2006

Inhaltsverzeichnis

0. Einleitung	
1. Sprache und Sprechen	
2. Die deutsche Sprache in ihrer gesprochenen Form	
3. Die Phonetik als interdisziplinäre Wissenschaft	
3.1. Forschungsgegenstand der Phonetik.....	
3.2. Die wichtigsten Bereiche der Phonetik.....	
3.3. Kurze Geschichte der Phonetik.....	
4. Neurophysiologische Vorgänge der Sprechschallerzeugung	
4.1. Eine grobe Skizze des Nervensystems.....	
4.2. Zellstrukturen im Nervensystem.....	
4.3. LULLIES' Drei-Ebenen-Modell der Artikulationssteuerung	
5. Materiell-energetische Vorgänge während der Artikulation	
5.1. Biologische Grundlagen der Sprechproduktion.....	
5.1.1. Der Atmungsorganismus	
5.1.2. Das Stimmorgan	
5.1.3. Das Ansatzrohr	
5.2. Vokale.....	
5.2.1. Klassifizierung der Vokale nach primären artikulatorischen Merkmalen	
5.2.1.1. Zungenlage.....	
5.2.1.2. Zungenhöhe.....	
5.2.1.3 Lippentätigkeit	
5.2.2. Ein Aspekt der sekundären Modifikation: die Nasalierung.....	
5.3. Allgemeine Tendenzen bei der Realisierung der Vokale in der deutschen Standardvarietät.....	
5.3.1 Stellungsbedingte Variation	

5.3.2	Quantitätstendenzen bei den Vokalen	
5.4.	Die Vokale der deutschen Standardvarietät	
5.4.1.	Monophthonge (Einzelvokale)	
5.4.2.	Reduktionsvokale	
5.4.3.	Diphthonge	
5.5.	Das Vokalsystem der deutschen Standardvarietät auf Grund der artikulatorischen Merkmale	
5.6.	Konsonanten	
5.6.1.	Klassifizierung der Konsonanten nach primären artikulatorischen Merkmalen	
5.6.1.1.	Artikulationsstelle	
5.6.1.2.	Artikulationsmodus	
5.6.1.3.	Stimmbeteiligung	
5.6.2.	Aspekte der sekundären Modifikationen	
5.7.	Allgemeine Tendenzen bei der Realisierung der Konsonanten in der deutschen Standardvarietät	
5.8.	Die Konsonanten der deutschen Standardvarietät	
5.8.1.	Verschlusslaute	
5.8.1.1.	Plosive	
5.8.1.2.	Nasale	
5.8.1.3.	Vibranten	
5.8.2.	Engelaute	
5.8.2.1.	Frikative	
5.8.2.2.	Approximanten	
5.8.2.3.	Laterale	
5.8.3.	Affrikaten	
5.9.	Die Koartikulation und die Assimilation	
6.	Sprechakustik	
6.1.	Akustische Grundlagen	
6.1.1.	Die sogenannte „stumme Schallform“	
6.1.2.	Der Explosionsschall	
6.1.3.	Der Frik(a)tionsschall	
6.1.4.	Der (quasi)periodische Schall - der Klang	
6.2.	Korrelation zwischen artikulatorischen und akustischen Vorgängen	

7. Wahrnehmung und Perzeption	
7.1. Funktionelle (Neuro-) Anatomie des auditorischen Systems	
7.2. Psychoakustik des Gehörs. Wahrnehmungsgrößen	
7.3. Die Stufen oder Ebenen des Sprachverstehens	
7.3.1. Die auditive Ebene	
7.3.2. Die phonetische Ebene	
7.3.3. Die phonologische Ebene	
7.4. Perzeptionsmodelle und -theorien.....	
7.4.1. Analyse mit Synthese	
7.4.2. Die Motor-Theorie.....	
7.4.3. Die kategoriale Theorie	
7.4.4. Die Quantaltheorie.....	
7.4.5. Die Kohorten-Theorie.....	
8. Literatur	
Abbildungsverzeichnis	
Das Zeicheninventar der IPA	

0. Einleitung

Das vorliegende Werk möchte etwas prinzipiell Unmögliches schaffen: Es möchte für das neue BA-Studium der Fachrichtung Germanistik in Ungarn (und in anderen Ländern, *horribile dictu* auch in Ländern, in denen Deutsch als Staatssprache gilt...) ein Lehrbuch zur Verfügung stellen, das sowohl den neusten Stand der Phonetik als Wissenschaft repräsentiert, als auch den wissbegierigen Studentinnen und Studenten ein leicht „verdauliches“ Werk anbietet, das sie nicht vom Fach der Phonetik (bei Gott nicht besonders beliebt bei den meisten Germanistik-Studenten) schreiend weglaufen lässt. Das Grundkonzept dieses Lehrbuchs entstand nach vielen hitzigen Debatten und den notwendigen Kompromissen, die aus der oben angeführten „aussichtslosen“ Lage, bzw. aus den doch unterschiedlichen Auffassungen, die die Autoren vertraten, resultierten. Die Phonetik ist ein tiefer Brunnen und wir versuchten – ohne dabei schwindlig zu werden – die mannigfaltigen Bereiche so zusammenzufassen, dass die neusten und manchmal sehr detaillierten Darstellungen den Studenten zeigen: Die Phonetik ist nicht einfach ein Fach, das uns zeigt wie ich Deutsch zu sprechen habe, sondern eine interdisziplinäre und sehr moderne Wissenschaft, die es verdient, im Mittelpunkt des Studiums zu stehen.

Abbildungen, lateinische Ausdrücke und anatomische Details mögen das Lehrbuch an einigen Stellen auf den ersten Blick kompliziert, unserer bescheidenen Hoffnung nach aber mit der Zeit auch immer interessanter. Das Wunder der menschlichen Kommunikation mit Hilfe gesprochener Sprache wird mit seinen biologischen Grundlagen (auch auf der neurophonetischen Ebene), mit ihren physischen, akustischen Erscheinungen und zu guter Letzt in seinem perceptiven Rahmen dargestellt. Damit auch die deutsche Sprache ihren ge-

bührenden Platz bekommt, wird einerseits die Vielfalt und die Variation in der gesprochenen Form derselben berücksichtigt, aber auch die Lauttypen einer – zwar eher als ein Idealbild, als eine sprecherbezogene Realität existierenden – Standardvarietät des Deutschen dargestellt, mit Hinweisen auf wichtige Varianten. Dies sollte den Germanistikstudenten in Ungarn – sowohl im Bereich Deutsch als Fremdsprache (DaF), als auch im Bereich Deutsch als Minderheitensprache (DaM) – als Ausgangspunkt angeboten werden, da es wünschenswert wäre, dass gerade in einem Land mit einer so langen deutschen Tradition (und zwar mittel- und oberdeutsch geprägt von der dialektalen Warte her) nicht die fremden (nördlich geprägten) Normen und Standardvorstellungen verbreitet werden.

Wir gingen davon aus, dass die Zielgruppe dieses Lehrbuchs sehr heterogen sein wird, da viele angehende Studenten – dank der neuen Möglichkeiten – über ausgezeichnete Deutschkenntnisse verfügen, andere wiederum verfügen über ein eher schulisch geprägtes Sprachvermögen. Aus diesem Grund war es schwer zu entscheiden, welches sprachliche Niveau aber auch – wie bereits angedeutet – welches wissenschaftliche Niveau wir unseren Lesern zumuten wollen. Dieser Spagat war nicht einfach und konnte deshalb vielleicht nicht immer optimal gelöst werden. Die notwendige Einfachheit ist zwar vorhanden, allerdings wollten wir die im deutschen Kulturkreis vorhandene Tradition bezüglich eines wissenschaftlichen Stils auch nicht zur Gänze ablegen. Unser Ziel war es, dem Leser (dem Studenten ebenso wie der Lehrkraft) etwas anzubieten, was teilweise auch eine Vorbereitung für das spätere MA-Studium bietet. In diesem Sinne wünschen wir angenehme (wenn auch manchmal arbeitsame) Stunden mit unserer „Deutschen Phonetik – eine Einführung“.

Die Verfasser

1. Sprache und Sprechen

Die menschliche Sprache und ihr „Ursprung“ ist ein Phänomen, welches in den Wissenschaften recht unterschiedlich definiert, gedeutet und beschrieben wird. Die Fragestellungen bezüglich der Sprache gehören zu den grundsätzlichen in der Philosophie. Insbesondere in der Erkenntnistheorie werden sprachliche Bezüge immer wieder behandelt. Wollen wir das Wunder der menschlichen Sprache verstehen und mit Hilfe der Phonetik aus einem besonderen Blickwinkel analysieren, so muss am Anfang unserer Überlegungen die Frage nach der Entstehung der menschlichen Sprache stehen. In den Schöpfungsmythen der unterschiedlichen Kulturkreise wird diese Frage meist auf religiöse und / oder kultische Art beantwortet. Die Bedeutung dieser Frage für den Menschen wird daran deutlich. Schon bei den so genannten Naturvölkern gab es beispielsweise Amulette, die die Lungen und die Luftröhre abbildeten (vgl. WÄNGLER 1983: 22). Die Frage nach der Fähigkeit des Sprechens und dessen anatomische Grundlagen beschäftigte die Menschen also offensichtlich von der ersten Stunde an.

Im alten Ägypten – wie der griechische Geschichtsschreiber HERODOT schreibt – ließ der Pharao PSAMMETICH I. (7. Jahrhundert v. Chr.) folgenden Versuch¹ durchführen: Auf der Suche nach einer „Ursprache“ ließ man zwei Knaben in einer Hütte ohne menschliche Kommunikation aufwachsen. Das erste von ihnen gesprochene Wort, so die Idee, sollte Hinweise auf die Existenz einer „Ursprache“ geben. Das Experiment scheiterte: Die Knaben riefen „Bekos“ (phrygisch für „Brot“). Auch aus Indien gibt es aus dieser frühen Zeit aus der Sicht der Phonetik und der Sprachwissenschaft Interessantes

1 Ziel des Versuchs war es zu beweisen, dass die Ägypter „von Natur aus“ das älteste Kulturvolk auf Erden wären.

zu berichten. In der sog. Rig-Veda (die Veden waren kultische Bücher der alten indischen Sanskrit sprechenden Völker) wird darüber berichtet, wie die Götter die Rede schufen und wie sie verwendet werden sollte. Im alten Griechenland erreichte die Sprechpflege durch die Hochkultur in der Dramatik ein immens hohes Niveau.

Wie bereits oben angeführt, bedeutete die Frage nach dem Ursprung und dem Wesen der Sprache und des Sprechens auch für die Philosophie eine besondere Herausforderung. Aus der Zeit der Antike spielt PLATONS „Kratylos“ eine wesentliche Rolle: In diesem wird, in Dialogform, die Frage diskutiert, ob Wörter und Namen (also die Elemente der Sprache) etwas über die bezeichneten Dinge und ihre Eigenarten verraten. ARISTOTELES, der Schüler PLATONS führt in seinem Werk „Politik“ die These an, dass der Mensch gerade durch seine Fähigkeit zum Sprechen zum Menschen wird. Der Grund dafür besteht nach ARISTOTELES darin dass moralische Kategorien wie „Gut“ oder „Schlecht“, „Gerecht“ oder „Ungerecht“ ausgedrückt werden können. Die Menschen, so die These, seien dafür geschaffen, dass sie in der Gemeinschaft (in der Polis) im Originaltext leben. Voraussetzung für diese typisch menschliche Lebensweise sei die menschliche Sprache und das Sprechen (vgl. WINTERLING 1987: 22 ff.).

Die jüdisch-christlich geprägte Schöpfungsgeschichte – der in unserem europäischen Kulturkreis eine dominante Rolle zukommt – beginnt mit dem Satz: „Am Anfang war das Wort“. Unabhängig davon, wie der Originalbegriff (griechisch *Logos*) ausgelegt wird, eines wird klar ersichtlich: Die Frage nach dem Ursprung der Sprache ist eine der grundlegendsten in der menschlichen Gesellschaft und Kultur. Aus der Bibel kennen wir die theologisch ausgelegte Antwort auf die Frage nach den unterschiedlichen Völkern und ihren unterschiedlichen Sprachen: Den Turmbau zu Babel. Der Name „Babel“, der ver-

mutlich soviel wie „Tor Gottes“ bedeutete, klang im Hebräischen dem Verb „balal“, also „verwirren, vermengen“ ähnlich. So konnten die babylonischen Tempeltürme, die eine kultische Grundlage hatten, als Frevel gegen Gott bezeichnet werden. Die Strafe bestand darin, dass den Menschen unterschiedliche Sprachen auferlegt wurden und sie in der Welt zerstreut wurden. In diesem Zusammenhang ist auch das Johannes-Evangelium bedeutend, in dem der *Logos* mit Gott gleichgesetzt wird (vgl. WINTERLING 1987: 15 ff.).

Die Diskussion über den Ursprung der Sprache wurde im 18. und 19. Jahrhundert, u.a. durch bedeutende deutsche Philosophen fortgesetzt. HERDER vertritt in seinem Werk „Über den Ursprung der Sprache“ (1771) die Ansicht, dass weder „Schälle der Leidenschaft“, noch die Sprechwerkzeuge oder das Prinzip der Nachahmung den Ursprung der Sprache im Wesentlichen beeinflussten, sondern der Zusammenhang zwischen der Vernunft und der Sprache. Diese beiden Fähigkeiten waren dementsprechend von Natur aus Teil des Menschseins. FICHTE bezeichnet die Sprache als ein Vermögen (also eine menschliche Fähigkeit), seine Gedanken willkürlich zu bezeichnen. Gemeint ist damit die Zeichenhaftigkeit der Sprache sowie die prinzipielle Willkürlichkeit der benutzten Zeichen in der Form, wie dies später in der strukturalistischen Sprachtheorie dargestellt wurde. Der wohl wichtigste deutsche Sprachphilosoph dieser Epoche, Wilhelm VON HUMBOLDT (1886: 63), kommt in seinen Überlegungen zu dem Schluss, dass „die Sprache das bildende Organ des Gedanken[s]“ sei. „Die intellektuelle Tätigkeit, durchaus geistig, durchaus innerlich und gewissermaßen spurlos vorübergehend, wird durch den Laut in der Rede äußerlich und wahrnehmbar für die Sinne. Sie und die Sprache sind daher eins und unzertrennlich.“ Gewöhnlich wird seine Auffassung über das Wesen der Sprache mit dem griechischen Ausdruck „Energiea“ (dt. etwa „Tätigkeit“) bezeichnet, womit HUMBOLDT betonte, dass die Sprache das Resultat der menschlichen Geistestätigkeit sei.

Auch die biologischen Grundlagen und die Evolution sind bei dieser Frage von Bedeutung. JONAS & JONAS führten in ihrem Werk „Das erste Wort. Wie die Menschen sprechen lernten“ die Hypothese an, dass die weibliche und die männliche Sprachkompetenz voneinander abweichen. Laut dieser Theorie entstand diese Abweichung sowohl auf Grund der teilweise unterschiedlichen biologischen Grundlagen als auch infolge der unterschiedlichen Entwicklungen der weiblichen und der männlichen Rollen in der (früh)menschlichen Gesellschaft. Die Mädchen (dies ist heute noch bei Schimpansenkindern der Fall) blieben länger in der Obhut der Mutter und mussten sich später auch um die jüngeren Kinder kümmern. Für die Jungen ging es hingegen darum, so früh wie möglich auf die Jagd zu gehen. Aus diesem Grund erreichte die weibliche Sprachkompetenz einen höheren Grad, bei den Männern hingegen zeigt sich eine tendenziell bessere Raumorientierung. Diese Unterschiede sind auch beim heutigen Menschen nachweisbar. Die Autoren leiteten daraus ab, dass der Ursprung der Sprache in einer Phase vor der rapiden Entwicklung der großen Hirnrinde zu suchen sei. Im nächsten Schritt verstärkten sich beide Tendenzen nach und nach, die Entwicklung des zentralen Nervensystems und die der Sprachkompetenz erreichte das für den modernen Menschen charakteristische hohe Ausmaß (vgl. WINTERLING 1987: 60 ff.).

Als nächsten Punkt führen wir die sog. Uk-uk-Theorien an. Mit diesem Ausdruck bezeichnet man Entstehungshypothesen, die davon ausgehen, dass eine prinzipiell tierische Rufform (eben das angenommene Uk-uk) die Grundlage zur informationsvermittelnden gesprochenen Sprache bildete. Ein weiterentwickelter und populärer Erklärungsversuch ist die Theorie von Rudi KELLER (1990). In einem vom ihm verfassten Märchen ist es ein Menschenaffe namens Karlheinz, der einem Anführer seiner Herde die erste kommunikative Handlung entlockt. Ausgang der Handlung ist ein imitierter Angstschrei, dieser dient dem schwachen Karlheinz dazu die ihm körperlich überlegenen Mensch-

affen von der Beute zu verscheuchen. Als der Anführer mit eben dem gleichen imitierten Angstschrei zum Ausdruck brachte, dass die „kleinen Diebe“ verschwinden sollen, war der Schritt „vom naturhaften Angstschrei zum intentionalen kommunikativen Akt“ (KELLER 1990: 38) vollbracht. Sprache ist – nach der Theorie des Autors – ein Phänomen der dritten Art. Dies bedeutet folgendes: Es gibt die so genannten Naturphänomene, die unabhängig vom Willen der Menschen existieren (Phänomene der ersten Art), des Weiteren gibt es Phänomene, die KELLER „Artefakte“ nennt und die das Resultat des menschlichem Handelns und des Willens sind (Phänomene der zweiten Art). Laut KELLER (1990: 86) ist aber zusätzlich eine dritte Art von Phänomenen vorhanden, die in der Regel kollektive Phänomene sind, dazu gehört auch die menschliche Sprache.

Folgender Hinweis soll den Abschluss dieser Überlegungen bezüglich des Ursprungs und der Entstehung der menschlichen Sprache bilden: Die sprachlichen Fähigkeiten eines menschlichen Individuums werden sowohl durch genetische Grundlagen als auch durch die interne Struktur des Gehirns beeinflusst. Diesbezüglich sind viele Erscheinungen nur ansatzweise erforscht, allerdings haben die modernsten computergestützten Untersuchungsmethoden viele neue Erkenntnisse sowohl über die Gehirntätigkeit, als auch über die neuronalen Grundlagen derselben gebracht. Dieses einleitende Kapitel diene dazu, die mit der Sprache und dem Sprechen verbundenen Kernfragen in den Raum zu stellen. Im folgenden Kapitel werden grundsätzliche Aussagen über die deutsche Sprache getroffen, um die Eigenart des Deutschen auch den Studentinnen und Studenten mit Ungarisch als Ausgangssprache näher zu bringen.

2. Die deutsche Sprache in ihrer gesprochenen Form

Die menschliche Sprache tritt für die Sprachbenutzer im Allgemeinen in zwei Formen in Erscheinung, erstens als gesprochene (hörbare), zweitens als geschriebene (sichtbare) Sprache. Folgende Argumente sprechen für das Primat der gesprochenen Sprache:

1) beim Spracherwerb geht Sprechen zeitlich dem Schreiben voraus, entwicklungsgeschichtlich gilt das Schreiben im Vergleich zum Sprechen als sekundär.

2) die mündliche Äußerung ist reicher an Ausdrucksmitteln und -möglichkeiten als die Schrift (Intonation, Akzentuierungsmöglichkeiten, Sprechpausen usw.)

3) bei der mündlichen Kommunikation ist die Möglichkeit der unverzüglichen Rückkopplung gegeben, bei der schriftlichen kann diese dagegen (oft) erst verzögert erfolgen

4) bei der mündlichen Kommunikation braucht das Situative nicht erst versprachlicht zu werden, sondern es ist unmittelbar wahrnehmbar und wirksam.

Das Schriftliche ist dem Mündlichen jedoch in einigen Punkten überlegen: 1) Zumindest vor der mit der elektronischen Datenverarbeitung (EDV) eingetretenen Revolution war es zuverlässiger konservierbar und tradierbar, deshalb galt es – von ebenfalls wichtigen mündlichen Überlieferungen abgesehen – als eine wichtige (wenn nicht die wichtigste) Grundlage der menschlichen Kultur;

2). Oft ist die geschriebene Sprache eindeutiger, weil sie auf einem optischen Informationsträger basiert, dies gilt besonders wenn der Empfänger homophone (gleichklingende) Äußerungen mit Hilfe nicht homographischer (unterschiedlich geschriebenen) Ausdrucksmittel sofort auseinander halten kann, z.B. *Mohr – Moor, viel – fiel, die alten Sagen – die Alten sagen* usw. Allerdings gibt es auch Homographie, die durch die Aussprache unterscheidbar werden, z.B. *umfahren, übersetzt, unmöglich, modern*.

Für die Phonetik ist die geschriebene Sprache zwar nicht unwichtig, da die menschliche Sprache lediglich in ihren beiden Erscheinungsformen zu deuten ist, sie wird jedoch als sekundär, die gesprochene Sprache dagegen als primär angesehen (u.a. aus den oben angeführten Gründen). Seit den 60er Jahren des 20. Jahrhunderts ist der Forschungsgegenstand „gesprochene Sprache“ immer mehr in den Mittelpunkt linguistischer Analysen gerückt. Dies ist umso erfreulicher, da sich synthetisierende Grammatikdarstellungen bis vor kurzem stillschweigend vor allem auf die von Varianten möglichst gereinigte Schriftsprache konzentrierten. SCHWITALLA (1997: 16) prägte z.B. die Begriffe „konzeptionelle Mündlichkeit und Schriftlichkeit“, um die Abgrenzungsschwierigkeiten zwischen denselben bei diversen Textsorten, die zwar als gesprochene bzw. geschriebene Sprache realisiert werden, aber denen eine konzeptionelle Schriftlichkeit bzw. Mündlichkeit zugrunde liegt, näher zu erläutern. HENNIG (2000) weist darauf hin, dass die Heterogenität des Begriffs „gesprochene Sprache“ Aussagen über dieselbe zwar äußerst schwierig macht, die von der Autorin vorgeschlagene „Prototypenmethode“ soll allerdings die Auswahl von Textsorten im Rahmen von Untersuchungen erleichtern (vgl. HENNIG 2000: 122). Andere Autoren wie z.B. FIEHLER (2000) sehen den Begriff als wenig brauchbar an und schlagen eher das Konzept der kommunikativen Praktiken vor (vgl. FIEHLER 2000: 103). Trotz dieser theoretischen Unsicherheiten ist

es nützlich, sich die prinzipiellen Unterschiede der mündlichen und der schriftlichen Kommunikation immer vor Augen zu halten.

Sprachliche Varianz ist eine allgemeine Erscheinung, die es in allen natürlichen Sprachen der Welt gibt (vgl. BARBOUR & STEVENSON 1998). In künstlich zusammengestellten Sprachen wie in der Esperantosprache, fehlt gerade diese Varianz, damit sie leichter erlernbar ist. Da viele Benutzer dieses Lehrbuchs Ungarisch als Ausgangssprache sprechen und die deutsche Standardsprache häufig im institutionellem Rahmen gelernt haben, muss betont werden, dass sie „nur“ die deutsche Standardsprache gelernt haben, die aber nicht mit dem Deutschen identisch ist. Synthetisierende Grammatikdarstellungen zeigen zwar ein homogenes Bild auch bezüglich der deutschen Sprache. Trotzdem müssen vor allem im phonetischen Bereich die Varianzerscheinungen immer berücksichtigt werden. Wenn nun das Deutsche und das Ungarische miteinander verglichen werden, sticht gleich ins Auge, dass es im Ungarischen im Vergleich weniger Varianz existiert. Die Variation im Deutschen erreicht „erstaunliche Ausmaße und vielfältige Abstufungen“ (BARBOUR & STEVENSON 1998: 5). Diese Unterschiede müssen deswegen hervorgehoben werden, weil diese Varianz und dialektale Färbung im Ungarischen in dieser Ausprägung unbekannt ist. Natürlich gibt es auch im Ungarischen Dialekte, allerdings ist ihre sprachliche Varianz nicht so gross wie im Deutschen es der Fall ist. Die deutsche Sprache ist – im Gegensatz zum Ungarischen – eine plurizentrische Sprache. Dies bedeutet, dass im geschlossenen deutschen Sprachraum während der sprachgeschichtlichen Entwicklung verschiedene Zentren (politische, kulturelle, wirtschaftliche, regionale usw.) existierten und auch heute noch existieren. Aus diesem Grunde bezeichnet der Begriff „die deutsche Sprache“ kein einheitliches Phänomen. Das Deutsche ist also auch in Bezug auf die phonetische Komponente eine plurizentrische Sprache. Die

Zentren innerhalb des deutschen Sprachraums sind jedoch nicht identisch mit den vorhandenen (oder früheren) politischen Grenzen.

Diese regionalen Unterschiede sind nicht im Zusammenhang mit den Staaten zu sehen (Bundesrepublik Deutschland, Österreich, Schweiz, Luxemburg, Belgien), in denen Deutsch als offizielle Sprache benutzt wird, sondern es ist eine typische Erscheinung in der deutschen Sprachgemeinschaft. An dieser Stelle muss auch berücksichtigt werden, dass mit der Ausnahme des niederdeutschen Sprachraums (hier dominiert eine nördliche geprägte Standardsprache) und der deutschsprachigen Schweiz (hier wird fast in jeder Situation ein Dialekt gesprochen), in den anderen Regionen des deutschen Sprachraums in der gesprochenen Form in vielen Lebensbereichen und Sprachdomänen eine – von der örtlich oder regional gebundenen – Dialektform bzw. dialektal geprägte Sprachform benutzt wird. Die Sprecher des Deutschen in diesen Regionen können meistens ihre Sprachform variieren, je nach Gesprächssituation oder Gesprächspartner. Des weiteren müssen auch die Sprachbenutzer des Deutschen erwähnt werden, die in anderen Staaten als Sprecher des Deutschen als Minderheitensprache (DaM) leben, wie auch in Ungarn die hiesige deutsche Minderheit (ihre geschätzte Zahl beträgt 200-220 000 Menschen, bei der letzten offiziellen ungarischen Volkszählung 2001 gaben ungefähr 90 000 Personen eine Bindung zur deutschen Kultur an). Für diese kompetenten Sprachbenutzer des Deutschen ist es ebenfalls typisch, dass sie vor allem dialektal beeinflusste Sprachformen verwenden.

Die gesprochene deutsche Sprache zeigt in ihrer Entwicklung z.T. andere Tendenzen als die geschriebene Sprache, eine Entwicklung die auch zeitlich gesehen nicht vollkommen parallel verläuft. Die gesprochene Sprache erreichte nach einem langwierigen, von unterschiedlichen Faktoren beeinflussten Prozess bis zum 17. Jahrhundert einen relativ hohen Einheitlichkeitsgrad. Damit

war eine Ausgangssituation vorhanden, in der es für die geschriebene deutsche Sprache einen relativ einheitlichen Standard gab. In ihrer gesprochenen Form jedoch waren die örtlich gebundenen deutschen Dialekte in den meisten Regionen bis zum Ende des 19. Jahrhunderts das wichtigste Kommunikationsmittel. Erst nach der Gründung des ersten einheitlichen deutschen Staates, also der Reichsgründung 1871 (allerdings mit der sog. kleindeutschen Lösung, ohne Österreich) kommt es dazu, dass auf Grund der einheitlichen deutschen Schriftsprache die ersten Normierungsversuche der Aussprache gestartet werden. Die wissenschaftliche Grundlage und das Leitbild der sog. Bühnenaussprache dafür wurde von Hermann PAUL (1880/1909) in seinem Werk *Prinzipien der Sprachgeschichte* festgehalten. Es gibt also im deutschen Sprachraum keine bestimmte Region, deren Sprache oder Aussprache als Musterbeispiel für die späteren Normierungsversuche verwendet wurde. Die heutige sprachliche Situation im Norden des deutschen Sprachgebietes, dass also die meisten eine nördlich gefärbte standardnahe Aussprache in den meisten Lebensbereichen und Sprachdomänen benutzen, ist die Folge von langen sprachgeschichtlichen Entwicklungen. Die ursprünglichen niederdeutschen (auch Plattdeutsch genannten) Dialekte dieser Gegend wurden nämlich immer mehr zurückgedrängt, da sie sehr weit von der Hochdeutsch geprägten deutschen Schriftsprache (die nämlich auf der Basis von mittel- und oberdeutsch Dialekten entstand) waren. Erst in den letzten Jahrzehnten verstärkten sich auch hier die Anstrengungen, diese typischen Dialekte des deutschen Nordens wieder neu zu beleben und zu schützen.

Für die deutsche Standardaussprache existieren zwei Leitbilder: Die von SIEBS (1898/1969) kodifizierte Bühnenaussprache und die Rundfunkaussprache. Allerdings konnte sich die Bühnenaussprache nicht bei der breiten Masse der kompetenten Sprecher etablieren und wird heutzutage eher als wirklichkeitsfern und übersteigert empfunden. Das Leitbild Rundfunkaussprache wird

im Allgemeinen in allen Ländern, in denen die deutsche Sprache die Erstsprache bildet, als die anzustrebende Norm akzeptiert. Dabei ist anzumerken, dass die regional-dialektale Färbung dieser Rundfunkaussprache vor allem in Gebieten wie Bayern, Österreich, der Schweiz oder aber auch Sachsen ausgeprägter ist als in anderen Gebieten des deutschen Sprachraums. Auf Grund dieses Leitbilds ist der erste Normierungsversuch im Jahre 1964 mit dem „Wörterbuch der deutschen Aussprache“ kodifiziert worden, die Duden-Redaktion übernahm in der 2. Auflage des Aussprachewörterbuchs im Jahre 1974 dieses Leitbild und kodifizierte es nach dem bereits vorhandenen Muster (vgl. VALACZKAI 1984: 16). Diese drei Aussprachewörterbücher beschreiben also eine eher hypothetisch vorhandene deutsche Standardsprache in der gesprochenen Form. Als sprecherbezogene Realität herrschen im gesamten deutschen Sprachraum unterschiedliche, teilweise von der Medienlandschaft, teilweise von den regionalen sprachlichen Standards und dialektal in unterschiedlichem Masse geprägten standardnahen Ausspracheformen. Aus diesem Grund stand in den letzten Jahren die Problematik der Umgangssprache im Mittelpunkt vieler Untersuchungen. Die genaue Unterteilung dieses Kontinuums in eine – m.E. eher als Idealbild denn als eine sprecherbezogene Realität existierenden – Standardsprache und der in der modernen Zeit sich auflösenden deutschen Ortsdialekte, ist allerdings ebenfalls schwierig.

In den Kapiteln 5.4. und 5.8. werden zwar die Lauttypen dieser (eigentlich nur in den Aussprachewörterbüchern existierenden) deutschen Standardvarietät nach artikulatorischen Kriterien beschrieben, allerdings dient diese Darstellung lediglich zu einer ersten Orientierung. Bei den Beschreibungen sind zwar einige Hinweise zur oben angeführten Varianz vorhanden, aber es konnte nicht ausführlich der Versuch unternommen werden, wichtige regionale Standards oder dergleichen detailliert zu beschreiben. Im gesamten Kapitel 5 wurden – eher aus didaktisch-methodischen Gründen – dem Leser praktische Grund-

kenntnisse vermittelt, die zu einer einführenden Phonetik, vor allem bei DaF-Lernenden, unbedingt dazu gehören. Sie sind aber keineswegs als eine unbedingt nachzuahmende Norm zu verstehen, da es ja bis heute eine nicht abgeschlossene Diskussion darüber gibt, ob es eine Aussprachenorm des Standarddeutschen gibt oder geben sollte. Mit diesen Einschränkungen sind also die Ausführungen in den erwähnten Kapiteln zu verstehen und zu deuten.

3. Die Phonetik als interdisziplinäre Wissenschaft

3.1. Forschungsgegenstand der Phonetik

Die Phonetik als Wissenschaft wurde und wird ihrem Wesen und Status nach unterschiedlich beurteilt und im System der Wissenschaften dementsprechend unterschiedlich eingestuft. Zwei Eckpunkte (Extrempunkte) sind erkennbar, zwischen denen sich alle anderen Bewertungen (mit unterschiedlicher Gewichtung) bewegen. VON ESSEN (1966: 1) hält die phonatorischen Erscheinungen für Naturvorgänge, und erklärt sie zum Gegenstand der Phonetik. Aus diesem Grund rechnet er die Phonetik zu den Naturwissenschaften. PÉTRUSSON & NEPPERT (1991: 14) halten die Phonetik hingegen für eine sprachwissenschaftliche Disziplin. Der Grund dafür liegt darin, dass sowohl die Laute als auch die sich aus den Lauten zusammensetzenden Strukturen, sowie die diese überlagernden Erscheinungen (wie der Akzent und die Intonation in die Sprache) als Gesamtphänomen eingegliedert werden. Für TRUBETZKOY (1989: 13) gehört nur die Untersuchung der materiellen Seite (d.h. die Merkmalsseite) zu den Aufgaben der Phonetik, nicht aber die Bedeutungsseite (die Bedeutungsebene), vielmehr schaltet er jede Beziehung zur sprachlichen Bedeutung der untersuchten Lautkomplexe aus. Diese Auffassung von TRUBETZKOY erwies sich als fehlerhaft und wurde durch neuere Erkenntnisse modifiziert. Die phonetischen Methoden erachtet TRUBETZKOY (1989: 13) als rein naturwissenschaftlich und erklärt die naturwissenschaftliche Betrachtungsweise als unentbehrlich für die Phonetik (vgl. VALACZKAI & BRENNER 1999).

PÉTRUSSON & NEPPERT (1996: 14) stellen den Gegenstand und die Aufgabe der Phonetik in ein neues Licht, indem sie erklären: „Das zentrale Anliegen der Phonetik ist die Frage nach der Rolle der lautlichen Substanz im

sprachlichen Kommunikationsprozess, wobei unter lautlicher Substanz materiell-energetische Vorgänge dieses Kommunikationsprozesses verstanden werden.“ Im Gegensatz zur früheren Auffassung, dass nur die Merkmalsseite der Sprache (die „Klanggestalt der Rede“) den Gegenstand der Phonetik darstellt, erachtet man gegenwärtig die Vorgänge des Sprechens, die Strukturen des Sprechschalls und die Vorgänge im Hörorgan bis zum Verstehen der Rede (also alle materiell-energetischen Vorgänge) als den Gegenstand der Phonetik, all dies vollzieht sich im Bereich des so genannten signalphonetischen Bandes. In Anlehnung an eine Arbeit von TILLMANN bzw. TILLMANN & MANSELL (zitiert von PÉTURSSON & NEPPERT 1996: 15) reicht das gesamte signalphonetische Band von der Hirnrinde und dem Sprechapparat des Senders über das Sprechschallübermittlermedium (in der Regel die Luft), über das Hörorgan des Empfängers bis zu seiner Hirnrinde, wo die Dekodierung der im Sprechschall kodierten Mitteilung, das Sprachverstehen, stattfindet.

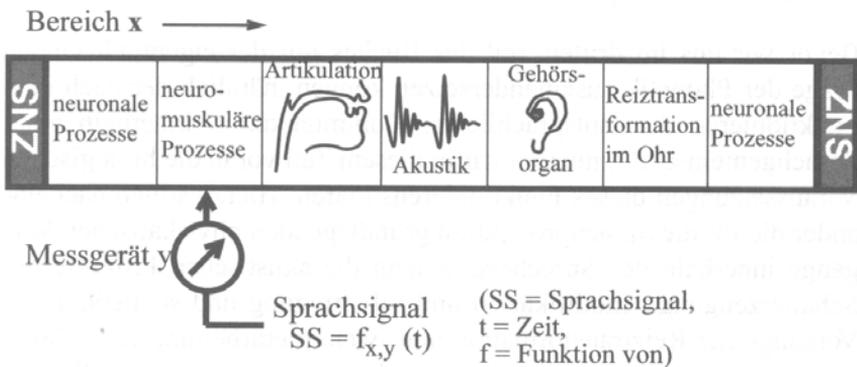


Abbildung 1: Das signalphonetische Band
(POMPINO-MARSCHALL 2003: 14)

Die sehr vereinfachte Formulierung, die Phonetik sei lediglich ein Teilgebiet der Linguistik, dessen Aufgabe darin bestehe, die Sprechlaute unter dem Aspekt der Erzeugung und der Wahrnehmung zu untersuchen, hat, wie die bishe-

rigen Ausführungen zeigen, heute keine Gültigkeit mehr: Die Phonetik ist vielmehr als ein interdisziplinärer Wissenschaftsbereich anzusehen, der auf Grund der Verwendung instrumenteller Messmethoden und der damit verbundenen digitalen Signalverarbeitung unmittelbar auch mit der Physik, der Nachrichtentechnik und schließlich auch mit der Informatik verbunden ist, zudem gibt es organische Bezüge zur Biologie, zur Psychologie und zur Medizin. Damit könnte das signalphonetische Band im Sinne von TILLMANN & MANSELL (1980) folgendermaßen weiter präzisiert werden:

Als zu untersuchende Aspekte bieten sich die Elemente der akustischen Kommunikationskette (engl. *speech chain*) an: **a)** die neurophysiologischen Vorgänge im Gehirn des Sprechers, die die Artikulation in den Sprechorganen vorbereiten; **b)** die die Muskeln der Sprechorgane steuernden elektrischen Vorgänge in den efferenten Nervenbahnen des Sprechers; **c)** die daraus resultierenden Bewegungen und Stellungen der Sprechorgane; **d)** die akustische Erzeugung der Sprechsignale in der durch die Sprechorgane gebildeten Hohlraumgestaltung; **e)** die akustische Übertragung zwischen Sprecher und Hörer; **f)** die mechanischen Vorgänge im Mittelohr, resp. die hydromechanischen Vorgänge im Innenohr des Hörers; **g)** die elektrischen Signale auf den afferenten Hörnervenbahnen; **h)** die neurophysiologischen Vorgänge im Gehirn des Hörers, die der Verarbeitung der in den Signalen kodierten sprachlichen Informationen dienen; **i)** die akustische Rückkopplung, die der Sprecher über sein Gehör erhält und perzipiert (vgl. DENES & PINSON 1972: 5; hier nach HESS 2002: 2).

Die Entwicklung der deutschen Phonetik zur selbständigen Wissenschaft wurde im 19. Jahrhundert entscheidend durch die Formulierung der Gesetze des historischen Lautwandels im Entwicklungsprozess des Indogermanischen zum Germanischen bzw. zum Deutschen angeregt. Zum genauen Verständnis und

zur richtigen Erklärung dieses Lautwandels brauchte man möglichst präzise Kenntnisse der Lautbildung. Auf Grund dieser Notwendigkeit erlebte die physiologische Phonetik einen Aufschwung. Dieser Entwicklung schlossen sich die akustische, die auditive und zuletzt die perzeptive Phonetik an.

Auch die wissenschaftlichen Forschungsmethoden haben sich weiterentwickelt: Die Erwägungen, Erkenntnisse und Ergebnisse der theoretischen (spekulativen) Phonetik wurden durch Erkenntnisse und Ergebnisse aus der Experimental- und der Instrumentalphonetik vielfach ergänzt, präzisiert und zum Teil ersetzt. Fest steht, dass sich diese beiden großen Bereiche der Phonetik gegenwärtig durch ein direktes Abhängigkeitsverhältnis auszeichnen: Die Instrumentalphonetik setzt theoretische Annahmen über das Wesen und die Funktionsweise der phonetischen Komponente der Sprache voraus, und eine generelle phonetische Theorie wird durch die Erkenntnisse und Ergebnisse der Instrumentalphonetik vorangetrieben, bestätigt oder modifiziert. Wie bereits angedeutet, haben sich auch die Auffassungen über den Status der Phonetik im System der Wissenschaften verändert. Da die Phonetik die Wechselbeziehung zwischen der Lautsubstanz der menschlichen Sprache und deren sprachlich-kommunikativen Funktion untersucht, gilt sie nun als eine Disziplin der Sprachwissenschaft. Im System der Wissenschaften ist die Phonetik allerdings auch durch einen interdisziplinären Standort gekennzeichnet. Die Vorgänge der Erzeugung, der Abstrahlung, der Übermittlung und der perzeptiven Verarbeitung des Sprachschalls sind physikalische und psychische Prozesse und werden in ihrer Eigenart als materiell-energetische Vorgänge des signalphonetischen Bandes von der Phonetik untersucht. Zu dieser Untersuchung braucht die Phonetik u.a. Daten, Kenntnisse, Methoden und Apparaturen der naturwissenschaftlichen Disziplinen wie der Anatomie, der Physiologie, der Physik (insbesondere der Akustik) der Strömungstheorie und Strömungsdynamik, der Neurologie, der Psychologie usw. Diese interdisziplinäre Eigenart ist die wich-

tigste Eigenschaft der Phonetik und macht sie in unserer modernen Zeit zu einer der produktivsten Ansätze in den Wissenschaften.

3.2. Die wichtigsten Bereiche der Phonetik

Die Phonetik beschreibt die Bedingungen und die Vorgänge der Bildung, der Übermittlung und der perzeptiven Verarbeitung des Sprachschalls. Ziel ist es zu erklären wie diese materiell-energetischen Vorgänge im signalphonetischen Band ihren sprachlich-kommunikativen Funktionen Rechnung tragen. Demnach ist die Phonetik eine **deskriptive** und zugleich eine **erklärende**, aber keine **präskriptive** Wissenschaft. Für den präskriptiv-normativen Gebrauch gibt es lediglich Anwendungsgebiete wie z.B. den Ausspracheunterricht und die korrektive Phonetik, wo viele Ergebnisse der deskriptiven Phonetik in der Form einer **kodierten Norm** Verwendung finden. Wie bereits im Kapitel 2 darauf hingewiesen wurde, ist aber diese kodierte Norm im Falle der deutschen Standardvarietät umstritten. Die Problematik des Normbegriffs und die spezifischen Problembereiche bezüglich der deutschen Standardvarietät in ihrer Gesamtheit darzustellen würde den Rahmen dieser Arbeit bei weitem übersteigen. Die Phonetik wird also in in verschiedene wissenschaftliche Disziplinen unterteilt, die im Zusammenhang mit dem jeweiligen Anliegen stehen.

Die allgemeine Phonetik geht davon aus, dass alle (bekannten) Sprachen trotz ihrer vielfachen Unterschiede zahlreiche phonetische Wesenszüge, Eigenschaften und Merkmale (sowohl in der Begriffsbasis und im Operationssystem als auch in den Realisierungsformen und in deren perzeptiver Verarbeitung) gemeinsam haben, diese gelten daher als Universalien und sind Forschungsgegenstand der allgemeinen Phonetik.

Die physiologische oder artikulatorische Phonetik beschreibt die artikulatorisch-organgenetische Erzeugung des Sprachschalls, d.h. sie beschäftigt sich mit dem Aufbau und dem Funktionieren des Sprechapparates sowie den koordinierten komplexen Bewegungstypen der Sprechwerkzeuge bei der Erzeugung des Sprachschalls.

Die akustische Phonetik beschreibt die Beschaffenheit, insbesondere die Strukturtypen des zur Übermittlung von Sprachinhalten erzeugten Sprachschalls (der Schallwelle).

Die auditive Phonetik beschreibt die Vorgänge der Hörwahrnehmung des Sprachschalls durch den Empfänger. Sie kann im Zusammenhang mit der perceptiven Phonetik oder der Wahrnehmungsphonetik gesehen werden. Ihr zentrales Anliegen ist 1) die Bestimmung der der Schallwahrnehmung zugrunde liegenden Reize, 2) die Klärung und die Typisierung der Vorgänge, die sich beim Sprachverstehen, d.h. bei der Dekodierung (Entschlüsselung) der im Sprachschall kodierten Äußerungsinhalte (Informationen) vollziehen.

Die Psychophonetik weist gewisse Überschneidungen mit der auditiven Phonetik bzw. mit der Wahrnehmungsphonetik auf, es überwiegen jedoch die Unterschiede. Sie beschreibt die psychischen Vorgänge, die sich bei der Planung und Durchführung der Kodierung und Dekodierung der durch den Sprachschall vermittelten Inhalte beim Sender bzw. beim Empfänger vollziehen. Sie ist ferner bestrebt, das Verhältnis zwischen Sprechen und Sprachverhalten, die Zusammenhänge zwischen den Sprechhandlungen und den Einstellungen der Kommunikationsteilnehmer, deren Motivationen sowie die Bedeutung der Existenz verschiedener Motivationstypen zu ermitteln. In dieser Hin-

sicht hat sie gewisse Charakteristika mit der Pragmatik als linguistische Disziplin gemeinsam.

Die Neurophonetik beschreibt die Vorgänge der Steuerung der Sprechprozesse durch das zentrale Nervensystem (kortikale Steuerung) und durch das periphere Nervensystem (s. 4.1). Sie untersucht ferner, welche Rolle die anatomischen und physiologischen Voraussetzungen, Gegebenheiten und Faktoren wie der Hörnerv und die Nervenbahnen bei der Vermittlung des wahrgenommenen auditiven Signals vom Gehörorgan zu den entsprechenden Bereichen der Hirnrinde und bei der Verarbeitung dieses Signals spielen.

Die Experimentalphonetik gilt als ein Sammelbegriff für experimentelle bzw. instrumentelle Forschungsmethoden und Verfahren, die bei der Untersuchung jenes Bereiches der Wirklichkeit Verwendung finden, die für die Phonetik von Belang sind. So sind technisch-maschinelle Verfahren wie die Sprechsynthese oder die Reihenuntersuchungen mit dem Audiographen zur Feststellung von Gehörschäden und deren Typen ohne Instrumente nicht denkbar, geschweige denn durchführbar. Die eingesetzten Instrumente sind leistungstark, die ermittelten Ergebnisse zuverlässig und informativ. Die Experimentalphonetik zeichnet sich durch die Verwendung empirisch-experimenteller Methoden und Verfahren aus, hierbei werden aber auch geeignete Instrumente verwendet. Dieser Zweig der Phonetik erlebte seit der technischen Revolution im EDV-Bereich in den letzten Jahrzehnte (wieder) einen bedeutsamen Aufschwung.

Die Mensch-Maschine-Kommunikation beinhaltet ein breites Feld von Forschungsansätzen, die sich mit den Fragestellungen der modernen Kommunikationstechnischen Möglichkeiten und ihrer phonetischen Belange auseinandersetzen. Die kommunikative Verbindung und Interaktion zwischen den Men-

schen und dem Computer wird untersucht, die Fragen der Signalverarbeitung, der Nachrichtentechnik und der künstlichen Intelligenz müssen mit den sprachwissenschaftlichen Erkenntnissen kombiniert werden.

Die funktionelle Phonetik oder Phonologie beschreibt im Gegensatz zur Phonetik nicht die Laute schlechthin, sondern ihre Funktion, ihre Stellung, ihren Stellenwert beim Aufbau, in erster Linie jedoch bei der Unterscheidung der Lautformen der sprachlichen Zeichen. Diese Analyse erfolgt auf Grund der so genannten phonologischen Opposition und mit Hilfe derselben werden die phonologischen Systeme der Sprachen bzw. sprachlichen Varietäten zusammengestellt.

Zur angewandten Phonetik gehören alle Anwendungsgebiete und Bereiche, in denen die Ergebnisse der einzelnen Zweige der Phonetik Verwendung finden. Solche Bereiche sind u.a.: die klinische Phonetik oder Sprechpathologie, die Sprechentwicklungsphonetik oder Pädophonetik, die korrektive Phonetik oder Logopädie.

3.3. Kurze Geschichte der Phonetik

Wie in Kapitel 1 schon erwähnt wurde, gehen die ersten Ansätze einer phonetischen Sichtweise bei wissenschaftlich-philosophischen Fragestellungen auf die Zeit der alten Ägypter und Inder zurück. Bereits die Griechen versuchten die Laute nach ihren Eigenarten zu gruppieren. Dank dieser Bemühungen wurden Lautsysteme entwickelt, die in der anerkannten Disziplin der Gesellschaft der Zeit – der Rhetorik – Anwendung fanden. In der Antike schufen auch bekannte Persönlichkeiten wie PYTHAGORAS und HIPPOKRATES

Werke, in denen auch phonetische Belange erörtert wurden. Der griechische Arzt Galen (er arbeitete allerdings vor allem in Rom) beschrieb im biologisch-medizinischen Bereich bedeutende Grundkenntnisse, die auch für die Phonetik relevant waren. Über die Römer gelangten die Idee der Lautsysteme und die Redekunst (s. *Libri XII Institutionis oratoriae* von QUINTILIANUS) in den (west)europäischen Sprachraum, wo die Erbschaft der klassischen Autoren in grammatischen Traktaten jahrhundertlang unverändert erhalten blieb. Der Wissenschaftler des Mittelalters beschränkte sich nämlich auf das langwierige Studium alter Titel zur Bildung und zu den Eigenschaften der Laute, obwohl sich die Rhetorik als Bestandteil des Universitätsstudiums etabliert hatte. Einen erheblichen Aufschwung erlebte die dahin schlummernde Forschung der Lautsysteme im 16. und 17. Jahrhundert: Man führte Beobachtungen in Verbindung mit Fragen der Orthoepie und der Orthographie des Lateins (*Grammatica [latina]* von RAMUS 1559) oder der eigenen Muttersprache (*Teütsche Grammatica* von ICKELSAMER 1527–1534) durch. Das bedeutendste Buch in diesem Zusammenhang ist zweifellos *De literis libri duo* (1586) von MADSEN aus Aarhus, dessen erster Teil (*De vera literarum doctrine*) die Laute in die Betätigung von neun Sprechwerkzeugen² einteilt und nach Bildungsmerkmalen systematisiert (*ordo literarum naturalis*). Der dänische Wissenschaftler befasste sich auch mit der Problematik des Stimmklanges (*vox*), der Mundhöhlen-Kalibration (*certa oris figura*) und der Zungenstellung (*linguae situs*) bei Vokalen. Gefördert wurden diese Entwicklungen durch die Erkenntnis, dass das Studium verschiedener Lautsysteme als ein Schlüssel zur synchronen Untersuchung von Sprachen verstanden werden kann.

Im 17. und 18. Jahrhundert lenkte sich die Aufmerksamkeit auf die Schulung von Taubstummen (*Elements of Speech* von HOLDER 1669) bzw. auf den

2 guttur, os, nasus, maxilla inf., maxilla sup., lingua, labia, palatum, dentes

Gesangunterricht (*Sur les causes de la voix de l'homme et de ces différents tons* von DODART 1703); diese und die damals betriebene „Lautwissenschaft“ wurden durch die Physiologie (*Elementa physiologiae* von HALLER 1757–1766) und die Physik (*Del suono* von BARTOLI 1680) mitgeprägt. Ende des 18. Jahrhunderts wurde das altindische, das Sanskrit „entdeckt“. Dies gab nicht nur einen erneuten Anstoß zur Betreibung der synchronen Sprachwissenschaft, sondern konnte auch zur Anreicherung der phonetischen Kenntnisse beitragen, da das Sanskrit über ein sehr reiches, von dem europäischen in vielen Punkten differentes Lautsystem³ verfügt. Zu dieser Zeit erschienen außerdem zwei weitere bahnbrechende Werke. Das eine ist HELLWAGs Abhandlung über die Systematik der Laute (*Dissertatio inauguralis de formatione loquelaе*, 1871), in der das dreieckige Schema der Anordnung der deutschen Vokale (d.h. das Vokaldreieck) das erste Mal entworfen wird; das andere ist von KEMPELEN (*Mechanismus der menschlichen Sprache*, 1791), der neben der Erörterung physiologischer Faktoren beim Sprechen auch den Weg der technischen Sprachproduktion beschriftet: der Autor baute eine – die erste – „sprechende Maschine“, und legte damit, ohne dies zu wissen, den Grundstein der nachherigen Experimental- oder Instrumentalphonetik.

Im 19. Jahrhundert erfuhr die Phonetik eine rasante Entwicklung, besonders in der Erforschung der allgemeinen Physiologie der Laute (*Grundzüge der Physiologie und Systematik der Sprachlaute* von BRÜCKE 1856) und der Stimm lippenbetätigung (*Der Kehlkopfspiegel und seine Verwendung für Physiologie und Medizin* von CZERMAK 1854). Nun schlossen sich auch Sprachwissenschaftler der phonetischen Forschung an, es wurden Zeitschriften gegründet wie die *Phonetischen Studien* (Marburg, 1888-1893, unter Leitung von VIÈ-

3 Im Sanskrit gibt es auch sog. Retroflexlaute auch kakuminal genannt, die durch das Zurückbiegen der Zunge gegen den harten Gaumen gebildet werden.

TOR) oder die *La Maître Phonétique* der APhI⁴. In die Arbeit wurden auch Akustiker wie CHLADNI (*Traité d' Acoustique*, 1809) oder HELMHOLTZ (*Die Lehre von den Tonempfindungen*, 1862) mit einbezogen. Es war die Geburtsstunde einer neuen Schule: die der Experimental- oder Instrumentalphonetik, deren Begründer Abbé P. J. ROUSSELOT war, der bei seinen Feldaufnahmen die Ungeeignetheit der Sinnesorgane bei der genauen Erfassung der Laute erkannte. Er sammelte daher ein Instrumentarium und gründete das erste instrumentalphonetische Laboratorium am *Institut Catholique de Paris*. Mit dem Beginn der Anwendung naturwissenschaftlicher Mittel und Methoden in der Phonetik endete die Zeit der so genannten Ohrenphonetik (in der das Gesprochene nur mit Sinnesorganen, zumeist mit dem Ohr, beurteilt wird). SIEVERS (1924) als typischer Ohrenphonetiker hat allerdings mit seinem Schaffen in der deutschen Sprachwissenschaft eine – vor allem in der deutschen Dialektologie – häufig befolgte, subjektive Analyseverfahren entwickelt. Trotz der Abneigung von Seiten der Sprachwissenschaft hatte die neue Richtung bald viele Anhänger (*Elements of Experimental Phonetics* von SCRIPTURE 1902; *Einführung in die angewandte Phonetik* von PANCONCELLI-CALZIA 1914). Anfänglich waren den Leistungen notgedrungen Grenzen gesetzt, die jedoch bald durch die Möglichkeit zur Speicherung und Wiedergabe der Stimme (Schallplatte, Tonfilm, Radio etc.) beseitigt werden konnten: Ein neues Kapitel nahm seinen Anfang (vgl. LAZICZIUS 1979: 6ff.; DIETH 1968ff.; LINDNER 1981: 1f.). Dieses neue, moderne Zeitalter der Phonetik wird von der technischen Revolution im EDV-Bereich der letzten Jahrzehnte in Richtungen und zu Ergebnissen geführt, die eine quantitative und qualitative Steigerung repräsentiert. Der „Siegeszug“ der Disziplin ist noch nicht gänzlich vollzogen.

4 APhI = *Association Phonétique Internationale*

4. Neurophysiologische Vorgänge der Sprechschallerzeugung

Die kinetischen Organstereotypien (= in approximative gleicher Form wiederholenden Bewegungskomplexe) bei der Erzeugung des Sprechschalls werden zentralnervös innerviert (= gesteuert) und z.T. periphernervös ausgeführt. Im folgenden Kapitel wird kurz der Aufbau des Nervensystems umrissen und es werden elementare nervöse Ebenen der Artikulationssteuerung angesprochen.

4.1 Eine grobe Skizze des Nervensystems

Das komplizierteste funktionelle System des menschlichen Körpers, das Nervensystem (*Systema nervosa*), besteht aus zwei Teilen, (1) aus dem (für Wirbeltiere kennzeichnenden) zentralen Nervensystem (ZNS) (*Systema nervosum centrale*) und, weil die Notwendigkeit besteht, Verbindungen zwischen diesem und der Kopf- bzw. Körperperipherie aufzubauen, (2) aus dem peripheren Nervensystem (PNS) (*Systema nervosum periphericum*).

Das ZNS umfasst Gehirn (*Cerebrum*) und Rückenmark (*Medulla spinalis*), die durch Knochen des Schädels (*Cranium*) und der Wirbelsäule (*Columna vertebralis*) geschützt, von Hirn- und Rückenmarkshäuten (*Meninges*) umhüllt werden; es schwebt in einer Flüssigkeit, dem sog. Hirnwasser (auch: Nervwasser) (*Liquor cerebrospinalis*) in der harten knöchernen Hülle. Das ZNS selbst besteht aus zwei Substanzen, (1) der grauen (*Substantia grisea*) und (2) der weißen (*Substantia alba*). Die graue Substanz baut sich vorwiegend aus Nervenzellen (sowie wenigen Nervenfasern) auf und enthält viele Blutgefäße.

In der weißen Substanz konfluieren hauptsächlich Nervenfasern, die die Nervenregung in Gehirn und Rückenmark (und natürlich auch zurück) projizieren; ihre Farbe ist die Folge der Myelinisierung. (Das Myelin ist eine aus wasserlöslichen Lipoiden [= Sammelbezeichnung für fettähnliche Substanzen] sowie Eiweißstoffen bestehende und Nervenfasern umhüllende bzw. deren Leitungsgeschwindigkeit steigernde lichtbrechende dünne Scheide.) Das Gehirn hat ein Gewicht von etwa 1000–2000 g, bei Frauen ist es etwas leichter als bei Männern. Es ist gewöhnlich bilateralsymmetrisch angelegt; (homo- und heterotope) Kommissuren (= Faserbündel) verbinden korrespondierende Areale in den beiden Hemisphären (= Hirnhälften) über die mediane Mantelspalte (d.h. die Mittellinie) (*Fissura longitudinalis cerebri*) hinweg. Allerdings können gewisse Funktionen nur jeweils einer Hemisphäre (bes. auf dem Niveau des Endhirns) zugeordnet werden, d.h. die sog. Lateralisation (auch: Asymmetrie); diese mag anatomisch und/oder funktionell ausgeprägt sein. Es ist z.B. nach ZILLES & REHKÄMPER (1998: 9) grob wie folgt zu gliedern:

I. Rautenhirn (*Rhombencephalon*)

1. Nachhirn (auch: Markhirn) (*Myelencephalon*)
2. Hinterhirn (*Metencephalon*) mit Kleinhirn (*Cerebellum*)
3. Mittelhirn (*Mesencephalon*) mit Vierhügelplatte (*Tectum*)

II. Vorderhirn (*Prosencephalon*)

1. Zwischenhirn (*Diencephalon*)
 - a) *Hypothalamus* mit *Hypophyse*
 - b) *Subthalamus*
 - c) *Thalamus (dorsalis)* mit *Metathalamus*
 - d) *Epithalamus* und *Epiphyse*

2. Endhirn (*Teleencephalon*)

- a) Streifenkörper (*Corpus striatum*) und Pallidum (*Globus pallidus*)
- b) Hirnmantel (*Pallium*)

Das Rückenmark, dieser segmentierte Strang sitzt im Wirbelkanal (*Canalis vertebralis*) und ist, wie bereits angesprochen, vom Hirnwasser umgeben. Es reicht vom großen Hinterhauptsloch (*Foramen magnum*) etwa bis zum ersten oder zweiten Lendenwirbelkörper (LWK 1/2), und weist zwei Intumescenzen (= Verdickungen) auf: Im Bereich der Halswirbel (*Intumescencia cervicalis*) und im Bereich der Lendenwirbel (*Intumescencia lumb/osacr/alis*). (Die graue [*Substantia grisea medullae spinalis*] und die weiße Substanz [*Substantia alba medullae spinalis*] des Rückenmarks sind an diesen Stellen besonders stark ausgebildet.) Hier verlassen die zervikalen (= Nacken-, Hals-) und die lumbalen (= Lenden-) Spinalnerven (*Nervi spinales*) das Rückenmark, und von hier aus werden auch die Extremitäten (= Gliedmaße) innerviert. Demgemäß erfüllt das Rückenmark zwei Funktionen: Spinale (muskuläre und autonome Reaktionen auf Reize) und supraspinale (das Gehirn einbeziehende) Reflexe. Gehirnstrukturen, die für die Kontrolle von Körperbewegungen zuständig sind, senden Informationen über das Rückenmark zu den mit den Muskeln in Verbindung stehenden Motoneuronen, und umgekehrt: Über das Rückenmark erhält das Gehirn alle Sinneswahrnehmungen (vgl. THOMPSON 2001: 15; TREPEL 2004: 81).

Unter PNS sind die außerhalb von Gehirn und Rückenmark liegenden Anteile des Nervengewebes zu verstehen (vgl. ZILLES & REHKÄMPER 1998: 9). Das PNS im Kopfbereich bilden zwölf Hirnnerven (*Nervi craniales*), in bestimmten Kopf- und Halsbereichen sowie im Rumpf- und Extremitätenbereich Rückenmarksnerven. Periphere Nerven enthalten sowohl efferente (motori-

sche) Fasern, in denen signalförmige Befehle zu Organen, bes. zur quer gestreiften Muskulatur (aber auch zu Drüsen) geleitet werden, als auch afferente (sensorische) Nervenfasern, in denen Rückmeldungen über den Ausführungszustand der Befehle zum ZNS laufen. PÉTURSSON & NEPPERT (2002: 180) halten es für wichtig, dass „[d]iese Regelvorgänge weitestgehend unterhalb der Bewußtseinsschwelle voll automatisch [ablaufen]“. Für sprechscherzeugende Bewegungen sind zwei periphere (voneinander nicht scharf abgrenzbare) Faserkategorien von besonderem Belang: (1) somatomotorische (Versorgung der Skelettmuskulatur) und (2) speziell-viszeromotorische Fasern (Innervation der Kiemenbogenmuskulatur, d.h. der Gesicht- und der Kehlkopfmuskulatur u.a.).

4.2 Zellstrukturen im Nervensystem

Das Gehirn ist aus Milliarden von einzelnen Nervenzellen, den sog. Neuronen aufgebaut. Nach der fertigen Ausbildung (beim Menschen direkt nach der Geburt) sind diese unfähig, sich zu teilen und zu regenerieren (sensorische und die quergestreifte Muskulatur versorgenden Motoneurone bilden eine Ausnahme); sie existieren ebenso lange wie der Organismus (sofern sie wegen negativer Einflüsse von außen nicht früher absterben). Je nach ihrer Funktion können sie in Größe und Form enorm differieren. Sie sind miteinander vernetzt und bilden folglich einen Schaltplan im Gehirn; ihre Funktion besteht darin, Informationen auf andere Neurone oder aber auf Muskel-/Drüsenzellen zu übertragen. Sie haben jedoch keinen direkten Kontakt miteinander, sondern sind durch schmale Zwischenräume voneinander getrennt, die Kommunikation erfolgt auf dem Weg chemischer Wirkstoffe (Neurotransmitter).

Neurone sind wie folgt aufgebaut: Sie sind von einer semipermeablen (= halbdurchlässigen) Membran umgeben, die aus einer molekularen Doppelschicht besteht. In die Membran eingebettet sind Proteine zu finden; diese fungieren als Kanäle, in denen der Substanztransport von innen nach außen bzw. umgekehrt, und mithin die Veränderbarkeit des elektrischen Ladungsgefälles gesichert ist. Das Neuron hat ein Soma (= Zellkörper) (auch: Perikaryon), das mit Zytoplasma gefüllt ist und den Zytoblast (= Zellkern) (*Nucleus*) und andere Organellen (endoplasmatische Reticula, Golgi-Apparat, Mitochondrium u.a.) einschließt, deren wichtige Funktionen die Herstellung, der Transport oder die Speicherung von Proteinen (und anderer Substanzen wie Glukose und Kalzium) sind. Der Zytoblast enthält die Chromosome, die aus dem genetischen Material, der Desoxyribonukleinsäure (DNS) bestehen, und einen Nucleolus (= Kernkörperchen), der die an der Proteinbildung beteiligten ribosomalen Ribonukleinsäure (rRNS) produziert. Vom Soma gehen zahlreiche stark verzweigte Fortsätze, die Dendriten, aus, auf denen dendritische Dornen (auch: Spines) sitzen.

Die Variationsbreite der Dendriten in Anzahl und Länge ist groß: Einige Neurone besitzen mehr als 20 von ihnen, andere dagegen nur wenige; einige Dendriten erreichen sogar Millimeterlänge, während andere lediglich wenige Mikrometer lang sind. Dendriten sind fähig, während des ganzen Menschenlebens zu wachsen; sie bauen außerdem immer mehr Verknüpfungen mit anderen Neuronen aus (Grundlage für das Lernen). Dendritische Dornen verkörpern synaptische (= Berührungs-, Fügungs-) Informationseingänge von anderen Neuronen. Jedes der Neurone besitzt einen Fortsatz (von wenigen Mikrometern bis zu mehr als einem Meter!), das Axon (auch: Neurit), das einer Soma-Auswölbung, dem Axonhügel, entspringt. Es umgibt eine aus Gliazellen (auch: Neurogliazellen) bestehende Myelinscheide; zwischen den Gliazellen befinden sich ringförmige Räume, die Ranvierschen Schnürringen und extra-

zelluläre Flüssigkeit. Am anderen Ende des Axons sind Kollateralen (= Verzweigungen), die Telodendrien zu finden, die in synaptischen Anschwellungen, den Endknöpfchen, enden. Diese synaptischen Endungen, die so genannten Synapsen, setzen chemische Substanzen (Neurotransmitter) frei, mit deren Hilfe sie mit anderen Nervenzellen kommunizieren; wenn auch indirekt: Zwischen den Neuronen gibt es einen schmalen Spalt; nur die elektrischen Synapsen, die so genannten *gap junctions* bilden hier eine Ausnahme (vgl. KOLB & WHISHAW 1996: 55–60).

4.3 LULLIES' Drei-Ebenen-Modell der Artikulationsteuerung

Die zentrale Innervation der Artikulation fasst laut der LULLIESschen Theorie (s. SURJÁN & FRINT 1982: 50) ein- und auslaufende Reize auf drei hierarchischen Ebenen zusammen:

1. Das untere Zentrum bildet das verlängerte Mark (*Medulla oblongata*), die kraniale (= obere) Fortsetzung des Rückenmarks. Strukturell ist es nach kaudal (= unten) zum Rückenmark nicht scharf abgrenzbar, *per definitionem* reicht es jedoch bis zum Abgang des ersten Zervikalnervenpaares. Nach kranial grenzt es sich durch den *Sulcus bulbopontinus*, reicht also bis zum Beginn der quer verlaufenden und wulstigen Fasern der Brücke (*Pons*) (vgl. TREPEL 2004: 108). Tierexperimente von DU BOIS-REYMOND und KATZENSTEIN (s. SURJÁN & FRINT 1982: 50) bezeugten, dass die Reizung diverser Gebiete des verlängerten Marks die bilaterale (= zweiseitige) Kontraktion (= das Zusammenziehen) von Kehlkopfmuskeln hervorruft.

2. Das mittlere der drei Zentren bilden die motorischen Kerne (*Nukleus reticularis*) des Mittelhirns (*Mesencephalon*), die Basalganglien (auch: Stammganglien) sowie das Kleinhirn. Die Basalganglien werden (neben zahlreichen Faserbahnen, die die motorische Großhirnrinde, die motorischen Hirnstammkerne und das Rückenmark miteinander polysynaptisch verbinden) dem extrapyramidal-motorischen System zugeordnet. Den Basalganglien fällt u.a. eine Rolle in der Ausführung von sequentiellen und (durch Imitation) erlernten Bewegungen zu; ihre Schädigung kann zu Hyperkinese (Frühstadium; häufige und unwillkürliche Muskelzuckungen von kurzer Zeitdauer, verminderter Muskeltonus) und Hypokinese (Spätstadium; Bewegungsarmut, angelegener Muskeltonus) führen. Das Kleinhirn befindet sich in der hinteren Schädelgrube und ist dem verlängerten Mark und der Brücke von hinten her aufgelagert. Es ist auch für das Erlernen, die Koordination und die Feinabstimmung von Bewegungsabläufen zuständig (vgl. TREPEL 2004: 149).

3. Das obere Zentrum bildet die Großhirnrinde (*Cortex cerebri*), insbesondere die beiden Frontallappen (*Lobi frontalis*). Sie umfassen das kortikale Gewebe anterior (= vorn) des *Sulcus centralis* (= vom dorsalen Rand der Hemisphären bis etwa zu ihrem Mittelpunkt führende Fissur), d.h. etwa 20 % des gesamten Neocortex (auch: Isocortex) (= phylogenetisch jüngster Anteil der Großhirnrinde). Sie sind funktional asymmetrisch organisiert: Der linke Frontallappen spielt bei Sprechbewegungen eine Rolle, der rechte dagegen eher bei der Mimik. Sie lassen sich in drei Hauptareale, in (1) ein prämotorisches, (2) ein motorisches und (3) ein präfrontales unterteilen. (1) kann über cortikospini-

nale Projektionen (direkt) oder über Verbindungen zum motorischen Cortex (indirekt) Bewegungen kontrollieren. (2) kontrolliert über Verknüpfungen zu den motorischen Hirnnerven die Gesichtsmuskulatur; es projiziert (= sendet Fasern) aber auch zu anderen motorischen Strukturen (Basalganglien, *Nucleus ruber*) (vgl. KOLB & WHISHAW 1996: 257ff.). (3) ist für die Steuerung kognitiver Prozesse zuständig (vgl. SURJÁN & FRINT 1982: 50f.).

5. Materiell-energetische Vorgänge während der Artikulation

Die Erzeugungsvorgänge des Sprechschalls sind das Resultat komplizierter, aufeinander abgestimmter Tätigkeiten der Sprechwerkzeuge. Dies bedeutet ein ständiges Zusammenwirken der unten beschriebenen Organe und Resonanzräume, das letztlich für die Erscheinung „Sprechschall“ verantwortlich ist.

Der Begriff Artikulation (lat. *articulare* – etw. deutlich aussprechen) bezieht sich gewöhnlich auf die Vorgänge im menschlichen Ansatzrohr. In diesem Sinne würde Artikulation eine „Gliederung des Luftstromes, die vor allem durch bestimmte Größen- und Formveränderungen des Ansatzrohres, durch Verschluss- oder Engebildungen, zu akustisch unterscheidbaren Schallerfolgen führen“ (WÄNGLER 1983: 69) bedeuten.

Bei diesen Prozessen sind dieser Definition zufolge der Unterkiefer, die Zunge, die Lippen und das Gaumensegel am aktivsten. Diese Darstellung stimmt im Wesentlichen, wobei zu bemerken ist, dass hier sie die Einheit des Prozesses weniger betont wird, um die prägnantesten Momente hervorzuheben. In

der folgenden Darstellung der Sprechwerkzeuge wird die Betrachtungsweise daher auf eine breitere Basis aufgebaut, um ein Gesamtbild der Erzeugungsvorgänge des Sprechschalls zu entwerfen.

5.1. Biologische Grundlagen der Sprechproduktion

Die Erzeugung des Sprechschalls ist immer auch eine körperliche Leistung (vgl. WÄNGLER 1983: 48). Obwohl der Mensch kein spezifisches Organ besitzt, das speziell zur lautsprachlichen Kommunikation dient, ist er fähig, den materiellen Träger der sprachlichen Inhalte, den Sprechschall zu produzieren. Benutzt werden also Organe, denen primär eine andere Funktion zukommt, erst durch ein kompliziertes Zusammenwirken dieser Organe wird die Erzeugung des Sprechschalls ermöglicht. Hier ist anzumerken, dass andere Säugetieren ähnliche Organe besitzen, deren Entwicklungs- und Spezialisierungsgrad aber meist weit hinter dem des Menschen liegt. Diese Organe und Räume können in drei Gruppen eingeteilt werden:

- (1) die lebenswichtigen Organe der Atmung,
- (2) der Kehlkopf als das Organ zur Erzeugung des Stimmtons durch die Stimmbänder (Stimmlippen),
- (3) das Ansatzrohr.

Zusammenfassend nennt man die Gesamtheit dieser Organe und Räume auch Sprechwerkzeuge. Dieser Begriff wird im engeren Sinne nur für die Organe und Räume im Ansatzrohr verwendet.

5.1.1. Der Atmungsorganismus

Die lebenswichtigen Organe der Atmung befinden sich im Brustraum, der nach unten durch das Zwerchfell abgegrenzt wird. Die Form des Brustraumes wird im Wesentlichen durch den **Thorax** bestimmt. Der knöcherne Thorax besteht aus den **Rippen** (*Costae*) (zwölf Paare), dem **Brustbein** (*Sternum*) und den **Brustwirbeln** (*Vertebrae thoracicae*). Von den drei Arten der Brustmuskeln (eigentliche Brustmuskeln, Extremitätenmuskeln, Zwerchfell) nimmt das **Zwerchfell** (*Diaphragma*) am intensivsten an der Atmung teil. Diese dünne, aber sehr kräftige Muskelplatte liegt quer zwischen der Brust- und der Bauchhöhle und trennt sie voneinander. Es ist der für das Einatmen wichtigste Muskel. Das Ausatmen ist normalerweise ein passiver Vorgang, der vor allem durch die elastische Lungenspannung gewährleistet wird. Kräftiges Ausatmen wird jedoch durch die Bauch- und Rückenmuskulatur unterstützt.

Die **Lungen** (*Pulmones*) befinden sich in den beiden Pleurahöhlen und bestehen aus passivem Gewebe. Die beiden kegelförmigen **Lungenflügel** werden durch Einschnitte in **Lungenlappen** geteilt, auf der linken Seite sind es zwei, auf der rechten drei. Die **Luftröhre** (*Trachea*) verbindet die Lunge mit dem Kehlkopf. Die Luftröhre gabelt sich in die beiden **Stammbronchien**, so dass der linke und rechte Lungenflügel miteinander verbunden sind. Die Stammbronchien teilen sich im so genannten Bronchialbaum in **Luftröhrchen** mit einem kleineren Durchmesser auf. Die aller kleinsten werden **Bronchiolen** genannt, an deren Ende sog. **Alveolargänge** liegen, die mit **Luftbläschen** (*Alveolen*) besetzt sind. Die Luftbläschen sind der Ort des Gasaustausches.

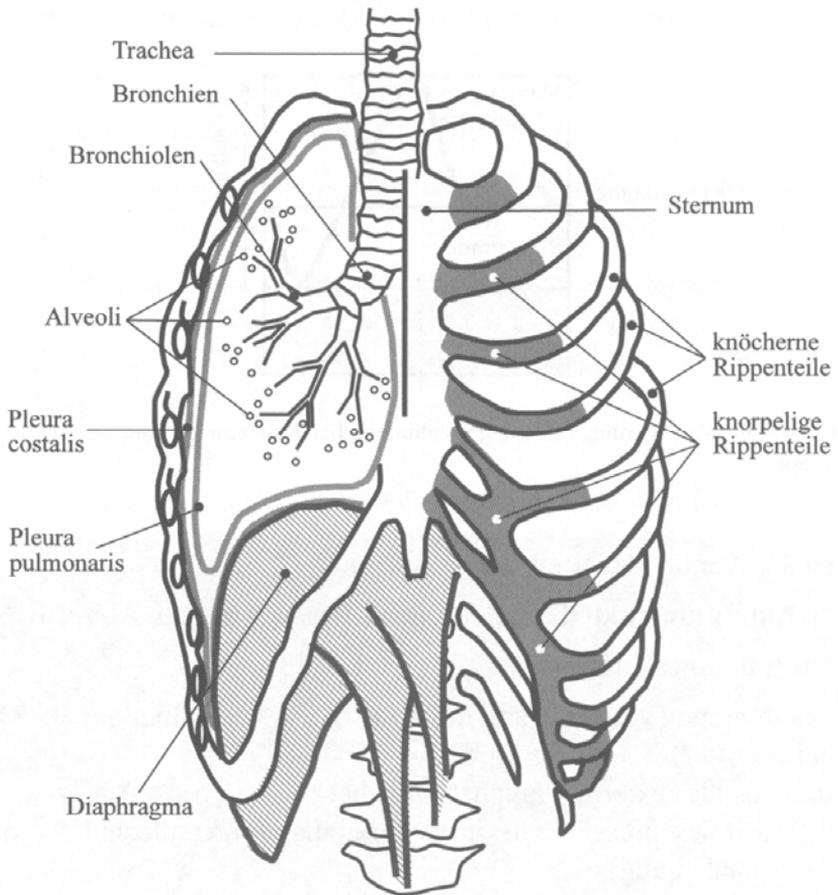


Abbildung 2: Der Atmungsorganismus (POMPINO-MARSCHALL 2003: 21)

An dieser Stelle muss festgehalten werden, dass die Atmung primär dem Gasaustausch (zwischen Kohlendioxid und Sauerstoff) dient. Man unterscheidet dementsprechend zwischen der **Rueatmung** und der **Phonationsatmung**, wo der zeitliche Anteil der Ausatmung im gesamten Atmungsprozess enorm steigt. Trifft die Phonationsluft auf eine geschlossene Stimmritze, erhöht sich der subglottale Luftdruck. Bei der Phonationsatmung ist der gesamte Prozess

wesentlich gespannter und kostet viel mehr Energie als beim ruhigen Atmen, das wiederum sehr effizient den Gasaustausch sichert.

5.1.2. Das Stimmorgan

Die herkömmliche Bezeichnung Stimmorgan bedarf einer Präzisierung. Der **Kehlkopf** (*Larynx*) setzt sich aus verschiedenen Knorpeln zusammen, die durch Bänder miteinander verbunden sind und durch diverse Muskeln bewegt werden können. Die wichtigsten Bestandteile des Kehlkopfes sind: der **Ringknorpel** (*Cartilago cricoidea*), der ihn von unten begrenzt, die **Stellknorpel** (auch: Gießbeckenknorpel) (*Cartilagine arytaenoideae*) sowie der **Schildknorpel** (*Cartilago thyroidea*), der aus zwei zusammengewachsenen Knorpeln besteht und bei Männern als ein Grat gut zu spüren ist (umgangssprachlich auch Adamsapfel genannt). Von oben wird der Kehlkopf von der **Epiglottis** (auch: Kehldeckel) (*Cartilago epiglottica*) abgeschlossen. Sie spielt beim Schlucken eine wichtige Rolle, weil sie dabei den Eingang des Kehlkopfes verschließt und so die Speise in die Speiseröhre leitet. Die **Stimmbänder** (auch: Stimmlippen) (*ligamentum vocale*) sind vorne fest mit den Innenwänden der Schildknorpelplatte, hinten mit den beiden Stellknorpeln verbunden. Sie können eine breitere Öffnungsstellung einnehmen, so dass die Dreiecksform der sog. **Stimmritze** (*Glottis*) entsteht. Die Stimmbänder dienten ursprünglich lediglich dazu, den freien Luftweg sicher nach innen abzuriegeln. Im Zuge der Evolution ist die Fähigkeit, einen Stimmton zu produzieren, erst relativ spät entstanden. Die Absicherung des Luftweges kann als die primäre Funktion der Organe festgehalten werden.

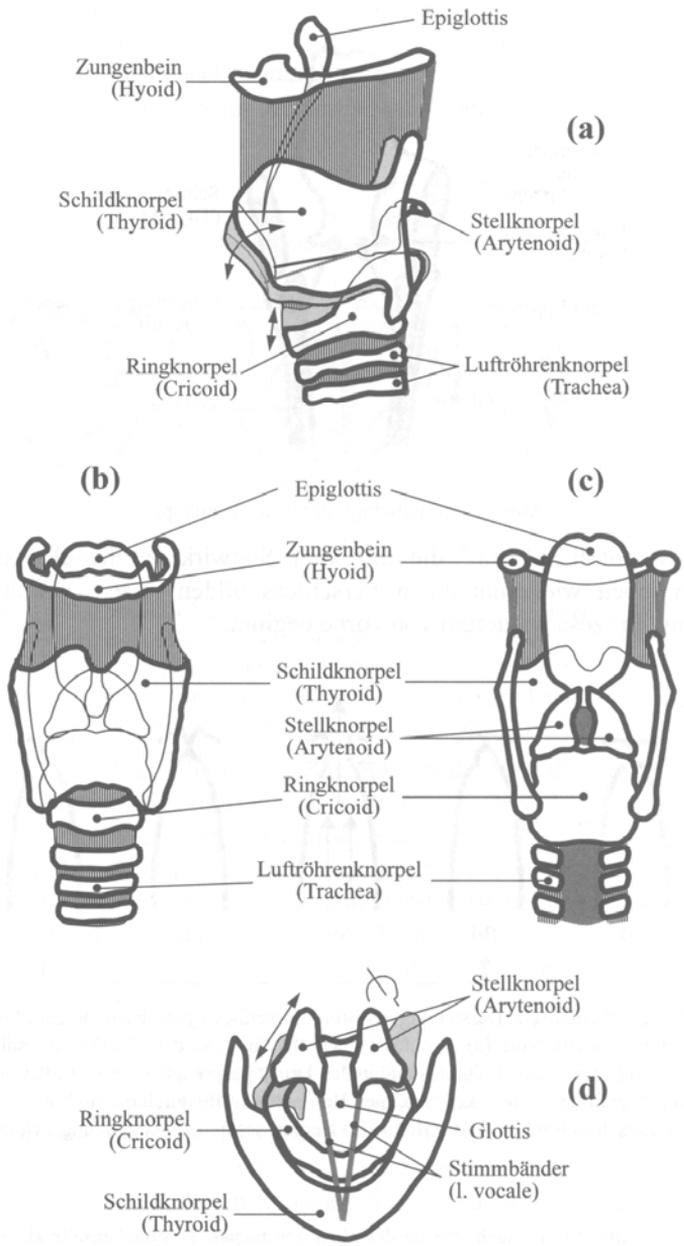


Abbildung 3: Der Kehlkopf und die Glottis (POMPINO-MARSCHALL 2003: 33)

Die Bezeichnung Stimmorgan verweist auf die wichtigste Eigenschaft desselben: hier wird mit Hilfe der Stimmbänder der Stimmklang produziert. Dies kann modellhaft folgendermaßen skizziert werden: Trifft die Phonationsluft auf eine geschlossene Stimmritze, so wird der subglottale Luftdruck größer, bis er stark genug ist, den Spalt von unten aus unter kortikaler Mitwirkung zu öffnen. Nach Entweichen von Phonationsluft schließt sich die Stimmritze auf Grund ihrer Elastizität, da der Luftdruck gleichfalls abgenommen hat. Dieser Vorgang wiederholt sich in kürzester Zeit mehrmals, so dass die Stimmbänder in Schwingung geraten. Das Zusammenspiel der kortikalen Steuerung und des **Bernoulli-Effekts** ermöglicht diesen Prozess. Frequenz und Amplitude der Luftschwingungen werden an dieser Stelle entsprechend der Frequenz und Amplitude, der Schwingungen sowie Länge, Breite und Beschaffenheit der Stimmbänder maßgeblich festgelegt – es entsteht ein Stimmklang.

5.1.3. Das Ansatzrohr

Der Begriff Ansatzrohr (auch: Vokaltrakt) ist eigentlich für Blasinstrumente gebräuchlich, die Anwendung des Begriffs auf menschliche Organe ist in der Fachliteratur üblich. In beiden Fällen stehen Resonanzkörper bzw. Resonanzräume im Mittelpunkt. Daneben sind erhebliche Unterschiede zu vermerken, die gegen eine Verwendung des Begriffs sprechen, wie die Tatsache, dass hier ein Sprechschall erzeugt wird, der sprachliche Inhalte vermitteln kann. Dennoch muss darauf hingewiesen werden, dass bei der Erzeugung des Sprachsignals alle Organe und Räume bei der Resonanz und der Schallformung mitwirken – abhängig von den Artikulationsgegebenheiten in unterschiedlichem Maße.

Zum Ansatzrohr gehören alle Organe und Räume oberhalb der Stimmbänder, bis zu den Lippen und zu den Nasenöffnungen. Die drei größeren Einheiten sind der **Rachenraum** (*Pharynx*), die **Mundhöhle** (*Cavum oris*) und die **Nasenhöhle** (*Cavum nasi*). Das gesamte Ansatzrohr kann durch Muskeln und Eigenelastizität in Form und Volumen verändert werden und ist größtenteils mit Schleimhaut bedeckt. Im Rachenraum findet man über den Stimmbändern zwei sagittale Ausbuchtungen, die sog. **Morgagnischen Taschen**. An der hinteren Rachenwand befindet sich der **Passavantsche Wulst**, der beweglich ist.

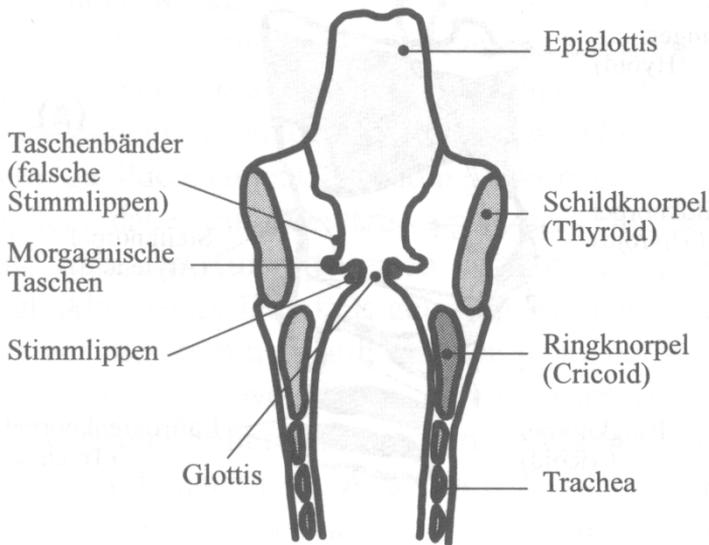


Abbildung 4: Sagittalschnitt des Kehlkopfs
(POMPINO-MARSCHALL 2003: 34)

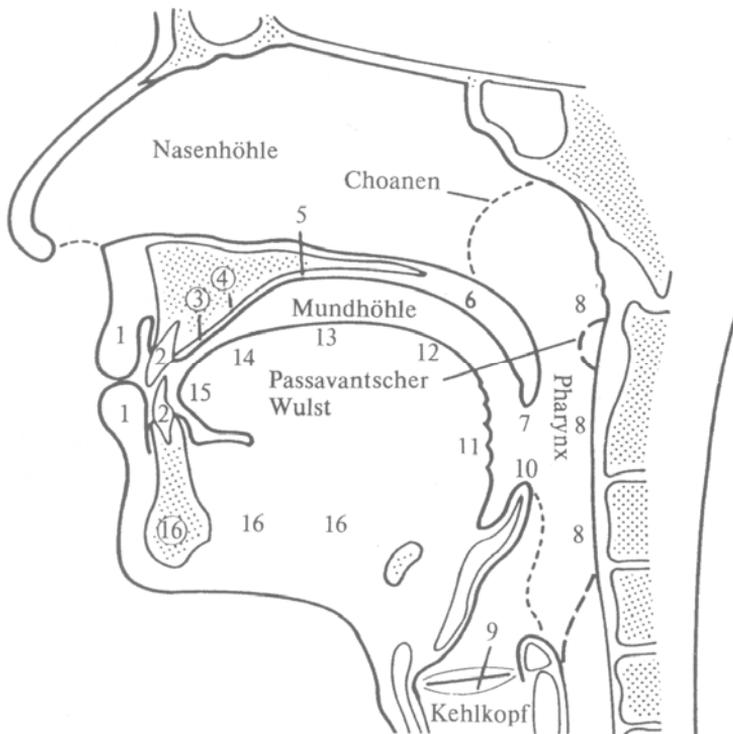


Abbildung 5: Das Ansatzrohr (PÉTURSSON & NEPERT 2002: 84)

Erklärung zur Abbildung 5: in der rechten Spalte befinden sich die lateinischen bzw. altgriechischen, in der mittleren die deutschen Ausdrücke zu den entsprechenden Teilen der Abbildung.

1	Lippe	Labium
2	Zähne	Dentes
3	Zahndamm	Alveolus dentalis
4	Vordergaumen	Praepalatum
5	Gaumen	Palatum
6	Gaumensegel	Velum
7	Zäpfchen	Uvula

8	Rachen	Pharynx
9	Kehlkopf	Larynx
10	Kehldeckel	Cartilago epiglottica
11	Zungenwurzel	Radix
12	hinterer Zungenrücken	Postdorsum
13	mittlerer Zungenrücken	Mediodorsum
14	vorderer Zungenrücken	Prädorsum
15	Zungenspitze	Apex linguae
16	Unterkiefer	Mandibula

In der Mundhöhle wird der Schall weitergeformt. Die Mundhöhle ist beim Weiterformen des Schalls. Ihre Grenzen bilden unten der Mundboden, vorne die **Lippen** (*Labia*), oben der harte und der weiche **Gaumen** (*Palatum durum*, *Palatum molle*), seitlich die Wangen.

Die Lippen sind beweglich, in der Ruhelage liegen sie aufeinander. Im vorderen Teil des Mundraumes befinden sich die **Zähne** (*Dentes*), die in den **Zahn-taschen** (*Alveoli*) verwurzelt sind. Die **Zunge** (*Lingua*) nimmt den größten Teil des Mundraums ein, sie besteht im Wesentlichen aus einem sehr fein strukturierten Muskelgewebe, welches sie sehr beweglich und verlagerbar macht. Im Erzeugungsprozess des Sprechschalls ist die Zunge das aktivste Organ und damit besonders hervorzuheben. Das **Gaumensegel** (*Velum*) versperrt der Phonationsluft in den meisten Fällen den Weg zur Nasenhöhle, ausgenommen sind natürlich die ausgeprägten Nasallaute. Das Velum kann sich dem Passavantschen Wulst nähern, ihn sogar berühren, vor allem bei geschlossenen Artikulationen. An seinem hinteren Ende geht das Velum in das ebenfalls bewegliche **Zäpfchen** (*Uvula*) über. Die ursprüngliche Aufgabe dieses Traktes war und ist die Aufnahme der Nahrung und deren erste Aufarbeitung.

Die Nasenhöhle wird durch die **Nasenscheidewand** (*Septum nasi*) längs geteilt.

Dass das Ansatzrohr so gut zum Modifizieren des Sprechsignals geeignet ist, ist ebenfalls die Folge einer langen Entwicklung. Abschließend ist mit PÉTURSSON und NEPPERT (1991: 86-87) festzuhalten, dass die herkömmliche Darstellung des Ansatzrohres mit der ausschließlich sagittalen Betrachtungsweise die dritte Dimension außer Acht lässt, was in manchen Fällen (z.B. in denen eine seitliche Zungenwölbung vorliegt) die präzise Beschreibung der Artikulation erschwert.

5.2. Vokale

Nach der traditionellen Definition sind Vokale orale, stimmhafte Öffnungslaute. PÉTURSSON & NEPPERT (1986: 89) geben folgende Definition (die ebenfalls noch nicht ganz korrekt ist): Ein Vokal wird ohne Hindernis im Ansatzrohr gebildet und bildet normalerweise den Silbengipfel. Ältere Ausdrücke wie „Selbstlaut“ sind irreführend.

Vokale werden traditionell mit Hilfe der Kriterien Vokalquantität und Vokalqualität beschrieben. Die Quantität ist eine prosodische Einheit und bezeichnet die Länge der einzelnen Vokale, während die Qualität etwas über die Artikulation aussagt. Die Quantität ist bei der normierten deutschen Hochlautung besonders wichtig, nach der Länge unterscheiden wir lange, halblange und kurze Vokale. Die Länge bzw. Quantität darf nicht mit der absoluten Dauer verwechselt werden, da letztere stark variabel ist und von zahlreichen Faktoren wie Sprechtempo, Sprecher, konsonantische Umgebung usw. abhängt. So kann

es auch bei ein und demselben Sprecher vorkommen, dass bei genauer Messung ein Vokal, der an einer bestimmten Stelle des Sprechschalls festgehalten wird, als lang empfunden wird, obwohl er in Wirklichkeit kürzer ist als ein anderer Vokal, der an einer anderen Stelle als kurz empfunden wird oder umgekehrt.

Die Vokalqualität setzt sich aus verschiedenen Faktoren zusammen, die im folgenden Abschnitt detailliert beschrieben werden.

5.2.1. Klassifizierung der Vokale nach primären artikulatorischen Merkmalen

5.2.1.1. Zungenlage

Als horizontale Zungenlage gilt der Abschnitt des Ansatzrohres, in dem die prägnante Verengung oder eine andere, entscheidende artikulatorische Tätigkeit stattfindet. Bei der Bildung von Vokalen ist die Zungenlage in der Regel mit den horizontalen Bewegungen der Zunge verbunden, da die Zunge dadurch den Resonanzraum im Mund verändert.

Je weiter sich die Zunge im Mundraum nach vorne bewegt, desto kleiner wird der Resonanzraum, desto höher ist folglich der zweite Formant (F2) des Vokals. Da die höchste Fläche der Zunge in den meisten Fällen näher dem harten Gaumen positioniert ist, werden diese vorderen Vokale als palatal eingestuft beispielsweise [i] im Wort *Ikone* und [e] im Wort *Egoist*.

Je weiter sich die Zunge im Mundraum nach hinten verlagert, desto größer wird der Resonanzraum, desto tiefer ist entsprechend der zweite Formant (F2). Ist der Zungenrücken dem Velum näher, werden hintere (velare) Vokale gebildet z.B. [u] im Wort *Kusine* und [o] im Wort *Projekt*.

Zwischen diesen beiden Endpunkten der Zungenstellung werden die sog. zentralen Vokale (Mittelzungenvokale) gebildet beispielsweise [ə] im Wort *bitte* und [ɐ] im Wort *bitter*.

5.2.1.2. Zungenhöhe

Die Zungenhöhe bezeichnet üblicherweise den Winkel zwischen der Zunge und dem Gaumen. Je tiefer die Zunge im Mundraum liegt, desto größer wird die Öffnung, d.h. der Öffnungsgrad des Vokals. Je höher die Zunge positioniert ist, desto geschlossener ist der Vokal. Man kann die Zungenhöhe in drei oder vier Stufen einteilen: bei einer Dreiteilung in hoch, mittel und niedrig/tief, bei einer Viererteilung in ganz untere, untere, mittlere und hohe Zungenhöhe, bzw. bezeichnet man die Vokale dementsprechend als offen, halboffen, halbgeschlossen und geschlossen. Die Bewegungen des Unterkiefers sind in der Regel mit denen der Zunge koordiniert. Man darf aber nicht vergessen, dass der Kiefer nicht nur nach vorne bzw. nach hinten, sondern auch seitlich beweglich ist.

Die Zungenhöhe korreliert mit dem ersten Formanten (F1) auf folgende Weise. Je offener ein Vokal ist, desto höher der dazugehörige Frequenzwert. Je geschlossener ein Vokal ist, umso tiefer ist der erste Formant. Zum Formantenbegriff siehe Kapitel 6.2.!

5.2.1.3. Lippentätigkeit

Nach der Beteiligung der Lippen an der Artikulation unterscheiden wir gerundete (also mit Rundung gebildete) und ungerundete (also ohne Rundung gebildete) Vokale. Gerundete Vokale werden auch als labiale, ungerundete als illabiale Vokale bezeichnet. Entsprechend dem Abstand zwischen den Lippen gibt es hier ebenfalls vier Stufen: ganz groß, groß, schmal, ganz schmal. Wenn die Lippen gerundet und vorgestülpt werden, sprechen wir von bilabialen Lauten.

5.2.2. Ein Aspekt der sekundären Modifikation: die Nasalisierung

Unter sekundärer Modifikation der Vokale ist zu verstehen, dass neben den angeführten primären Merkmalen bei der Klassifizierung weitere Aspekte hervortreten können. Diese Aspekte, die das Gesamtbild eines Vokals also maßgebend beeinflussen und abrunden, werden als sekundäre Modifikation bezeichnet. Die Nasalisierung bei den Vokalen ist ein solcher Aspekt.

Die deutschen Vokale werden prinzipiell oral gebildet. Man darf aber nicht vergessen, dass die Resonanz im Ansatzrohr nicht nur auf einen Teil des solchen beschränkt ist. Es gibt nicht immer ganz feste Grenzen zwischen den Orten, wo die Resonanz prägnant ist und wo sie weniger bemerkbar ist. Durch die velopharyngale Öffnung zur Nasenhöhle können die Vokale also „eine Nasalresonanz erhalten“ (PÉTURSSON & NEPPERT 1991: 105). Dieser neue Parameter ist demnach zur ausreichenden Charakterisierung der betroffenen Vokale ebenfalls erforderlich.

5.3. Allgemeine Tendenzen bei der Realisierung der Vokale in der deutschen Standardvarietät

Bei dieser Frage gilt der Grundsatz, dass die Realisierungen in verschiedenen phonetischen Positionen berücksichtigt werden müssen. Von der Fülle der Kriterien sollen an dieser Stelle die aus phonetischer Sicht Wichtigsten angeführt werden:

5.3.1. Stellungsbedingte Variation

Stellung des Vokals innerhalb des Lautkörpers (phonetische Position)

Die Realisierung eines Vokals wird dadurch beeinflusst, ob der Vokal im Anlaut, im Inlaut, oder im Auslaut eines Wortes steht. Des Weiteren können Vokale im Silbenanlaut von zusammengesetzten und abgeleiteten Wörtern neu eingesetzt werden. In solchen Wörtern wird der Vokal nach einer Phonationspause, in der kein Stimmtone produziert wird, wie am Wortanfang produziert.

Die Stellung zum Wortakzent

Je nach Stellung zum Wortakzent können verschiedene Realisationen auftreten. Das Paradebeispiel hierfür ist die breite Skala der Realisierungsmöglichkeiten der e-Laute.

Entfernung von der akzentuierten Silbe

Dieser Aspekt ist z.B. für den Reduktionsgrad des Murrelvokals [ə] enorm wichtig, wobei natürlich auch die Stellung zum Wortakzent und die umliegenden

den Segmente eine Rolle spielen z.B. *die Geretteten* [di gə'ɾɛtətɪŋ], aber: *die Rettenden* [di 'ɾɛtndən].

Die phonetische Struktur des Wortes

Der Murrelvokal [ə] wird nach Nasalen, Liquiden und nach Vokalen im Allgemeinen realisiert, z.B. *nennen* ['nɛnən], *sehen* ['zɛ:ən]. Vor einem Doppelkonsonanten, vor einer Konsonantengruppe, vor einer Affrikate oder <sch> ist der Vokal gewöhnlich kurz zu realisieren, wenn andere Einflüsse wie Flexion oder Ableitung dies nicht überlagern.

5.3.2. Quantitätstendenzen bei den Vokalen

Der Vokal wird im Allgemeinen lang realisiert

- in betonter offener Silbe deutscher oder eingedeutschter Wörter. Die Silbe ist offen, wenn sie auf einen Vokal auslautet, z.B. *Re-gen*, *Mu-mie*,
- in betonter geschlossener Silbe, wenn die Silbe auf einen einfachen Konsonanten ausgeht und dieser im Gebrauch des Wortes zum Anfangskonsonanten der nächsten Silbe werden kann. Solche Fälle treten durch die Erweiterung des ursprünglichen Lautkörpers ein, z.B. *Tag*, *Ta-ges*, *Hof*, *Ho-fes*,
- vor einem Dehnungszeichen oder als Doppelvokal, z.B. *Saal*, *ihm*, *hier*, *Treptow*
- vor <ß>, z.B. *Straße*, *mäßig*.

Der Vokal wird im Allgemeinen kurz realisiert

- in geschlossener Silbe mit Akzent, wenn die Silbe zwar auf einen einfachen Konsonanten ausgeht, aber dieser nicht auf die nächste Silbe übertreten kann z.B. *in*, *vom*,
- in betonter Silbe vor Doppelkonsonanz oder vor einer Konsonanten-
gruppe, z.B. *alle*, *Licht*,
- in betonter Silbe vor <sch>, <pf>, <tz> z.B. *löschen*, *Katze*. Vor <ß>, <st>, <ch> und vor der Verbindung <r> + Dental schwankt die Realisierung des Vokals, auch wenn er betont wird. Der Vokal ist z.B. lang in *husten*, *nach*, aber kurz in *Hast*, *Bach*,
- – der unbetonte Vokal ist vor der Akzentsilbe in eingedeutschten Fremdwörtern kurz, z.B. *Filiale*, *Elefant*. Außerdem sind unbetonte Vokale fast immer kurz, außer in einigen Endungen wie *-sam*, *-sal* usw.

5.4. Die Vokale der deutschen Standardvarietät

5.4.1. Monophthonge (Einzelvokale)

[i:]

Das lange geschlossene [i:] wird durch die Grapheme <i>, <ie>, <ieh>, wie beispielsweise in *Igel*, *Dieb* und *Vieh* und seltener, vor allem in Namen, durch <y> bezeichnet. An-, in- und auslautend ist es Bauelement im Lautkörper von Sprachzeichen. Nach der Phonationspause wird es hart eingesetzt. Der Weg der Phonationsluft zum Nasenraum ist gesperrt, die orale Artikulation domi-

niert. Die Zungenseiten berühren die obere Zahnreihe, der Zungenrücken legt sich in breiten lateralen Streifen dem harten Gaumen an. Die Zungenspitze liegt in „Kontaktstellung“ mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [i:] ein geschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[i]

Das kurze geschlossene [i] wird durch das Graphem <i> in Fremdwörtern und einigen Lehnwörtern bezeichnet, wie beispielsweise in *Titan*. Es ist in unbetonten Positionen im An- und Inlaut Bauelement des Lautkörpers des Sprachzeichens. Nach Sprechpausen ist der harte Einsatz möglich, jedoch ist er im Vergleich zu den akzentuierten Vokalen weniger intensiv. Das Velum versperrt den Weg der Phonationsluft zur Nasenhöhle, die orale Bildung dominiert. Die Zungenseiten berühren die obere Zahnreihe, die zwei lateralen Streifen des Dorsums stehen mit dem harten Gaumen in Kontakt. Die Zungenspitze hat Kontakt mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [i] ein geschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[ɪ]

Das kurze offene [ɪ] wird durch das Graphem <i> bezeichnet, wie beispielsweise in *mit*. An- und inlautend ist es ein Bauelement des Lautkörpers des Sprachzeichens. Nach der Phonationspause wird es in der Regel mit einem Glottisschlag eingesetzt. Der Weg der Phonationsluft zum Nasenraum wird durch das Velum gesperrt, die orale Artikulation dominiert. Die obere Zahn-

reihe wird bilateral berührt, beide lateralen Streifen des Dorsums liegen sich am harten Gaumen an. Die Zungenspitze berührt die unteren Schneidezähne. Auf Grund der Zungenhöhe ist [ɪ] ein geschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[e:]

[e:] wird durch die Grapheme <e>, <ee>, <eh> in betonter Position im An-, In- und Auslaut bezeichnet, wie beispielsweise in *ewig*, *leer* und *mehr*. Nach der Phonationspause wird es gewöhnlich mit hartem Glottisschlag neu eingesetzt. Der velopharyngale Nasenzugang ist gesperrt, die orale Bildung dominiert. Die Zungenränder wölben sich zur oberen Zahnreihe auf, beide lateralen Zonen des Zungenrückens haben Kontakt mit den gegenüberliegenden lateralen Zonen des harten Gaumens. Auf Grund der Zungenhöhe ist [e:] ein halbgeschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[e]

Das kurze [e] wird durch das Graphem <e> in unbetonter Position einge-deutscher Wörter im An- und Inlaut vor dem akzentuierten Vokal bezeichnet, wie beispielsweise in *Egoist*. Nach einer Phonationspause kann es hart eingesetzt werden, der Glottisschlag ist dabei im Vergleich zu den akzentuierten Vokalen weniger intensiv. Der velopharyngale Nasenzugang ist versperrt, die orale Bildung dominiert. Durch ihre Aufwölbung stehen die Zungenränder mit den oberen Zähnen, die lateralen Zonen des Zungenrückens ebenfalls beiderseits mit den entgegengesetzten Streifen des harten Gaumens in Kontakt. Auf

Grund der Zungenhöhe ist [e] ein halbgeschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[ɛ:]

Das lange offene [ɛ:] wird durch die Grapheme <ä>, <äh>, wie beispielsweise in *Bär* und *während* (in aus dem Französischen entlehnten Wörtern auch durch <ai>, <ei>) bezeichnet, das nur in der nördlich geprägten deutschen Standardsprache realisiert wird. In den restlichen Teilen des deutschen Sprachgebiets wird [e:] statt [ɛ:] gesprochen beispielsweise *ähnlich* [e:nliç]. Es ist in betonter Position an- und inlautend bzw. (ziemlich selten) auslautend ein Bauelement des Lautkörpers des Sprachzeichens. Nach der Phonationspause wird es gewöhnlich mit hartem Glottisschlag neu eingesetzt. Der Weg der Phonationsluft zum Nasenraum wird durch das Gaumensegel gesperrt, die orale Bildung dominiert. Die Zungenseiten berühren die obere Zahnreihe, die lateralen Zonen des Zungenrückens legen sich beiderseits des weichen Gaumens an. Die Zungenspitze kommt mit den unteren Schneidezähnen in Berührung. Auf Grund der Zungenhöhe ist [ɛ:] ein halbgeschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[ɛ]

Das kurze [ɛ] wird durch die Grapheme <e>, <ä> bezeichnet, wie beispielsweise in *eng* und *lässt*. Es ist in betonter Position im An- und im Inlaut Bauelement des Lautkörpers des Sprachzeichens. Nach einer Phonationspause wird es gewöhnlich mit einem harten Glottisschlag eingesetzt. Der pharyngo-orale Trakt wird gegen den Nasenraum durch das Gaumensegel abgesperrt, die

orale Bildung dominiert. Durch die Aufwölbung der Zungenränder entsteht ein Kontakt mit den Oberzähnen sowie beiderseits zwischen einem Streifen des Zungenrückens und dem harten Gaumen. Die Zungenspitze ist in Kontaktstellung. Auf Grund der Zungenhöhe ist [ɛ] ein halboffener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein palataler Vokal.

[y:]

Das lange geschlossene [y:] wird durch die Grapheme <ü>, <üh>, <y>, <ui> in betonten Silben bezeichnet, wie beispielsweise in *Tür*, *kühl*, *Mythos* und *Duisburg*. Es ist im An-, bzw. im Inlaut Bauelement des Lautkörpers des Sprachzeichens (in französischen Entlehnungen bei Schreibung von <u> und <ue> auch auslautend möglich). Nach einer Phonationspause wird es gewöhnlich mit einem Glottisschlag eingesetzt. Bei der Prägung des Lautes dominiert die orale Bildung: Das Velum versperrt den Weg zum Nasenraum, die Zungenseiten berühren die obere Zahnreihe, die beiden seitlichen Streifen des Zungenrückens legen sich den lateralen Streifen des harten Gaumens an. Die Zungenspitze hat Kontakt mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [y:] ein geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[y]

Das kurze geschlossene [y] wird durch die Grapheme <y> und <ü> in unbetonten Silben deutscher und eingedeutschter Wörter (in französischen Lehnwörtern auch durch <u>, z.B. *Guyenne* [gy'jen]) bezeichnet, wie beispielsweise in *Physik* und *überzeugen*. Nach Phonationspausen ist der harte Einsatz mit

Glottisschlag möglich, er ist jedoch weniger intensiv als bei den akzentuierten Vokalen. Der Weg zum Nasenraum wird durch den weichen Gaumen abgesperrt, die orale Bildung dominiert. Die Zungenseiten haben Kontakt mit dem harten Gaumen und mit den seitlichen Oberzähnen. Die Zungenspitze liegt in Kontaktstellung. Auf Grund der Zungenhöhe ist [y] ein geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[y]

Das kurze offene [y] wird durch die Grapheme <ü> und <y> bezeichnet, wie beispielsweise in *üppig* und *Ägypten*. Das [y] ist im An- und Inlaut Bauelement des Lautkörpers des Sprachzeichens. Nach einer Artikulationspause wird es gewöhnlich hart eingesetzt. Die Zungenseiten legen sich den Seiten des harten Gaumens, der oberen Zahnreihe und einem Streifen des weichen Gaumens an. Die Zungenspitze ist in Kontaktstellung. Auf Grund der Zungenhöhe ist [y] ein geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[ø:]

Das lange geschlossene [ø:] wird durch die Grapheme <ö>, <oe>, <(i)eu> bezeichnet, wie beispielsweise in *Öl*, *Goethe* und *Friseur*. An- in- und auslautend ist es in betonter Position Bauelement des Lautkörpers des Sprachzeichens. Nach der Phonationspause wird es gewöhnlich mit einem Glottisschlag eingesetzt. Der Weg der Phonationsluft in den Nasenraum wird durch das Gaumensegel versperrt, die orale Bildung dominiert. Die Zungenseiten berüh-

ren bilateral die obere Zahnreihe, die beiden lateralen Streifen des Zungenrückens liegen dem harten Gaumen an. Die Zungenspitze liegt in der Kontaktstellung. Auf Grund der Zungenhöhe ist [ø:] ein halbgeschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[ø]

Das kurze geschlossene [ø] wird durch das Graphem <ö> bezeichnet, wie beispielsweise in *Ökonom*. An- und inlautend in unbetonten Silben eingedeutschter Wörter vor dem akzentuierten Vokal ist es Bauelement des Lautkörpers des Sprachzeichens. Der harte Einsatz nach der Phonationspause ist weniger intensiv als bei akzentuierten Vokalen. Das Gaumensegel versperrt den Weg in die Nasenhöhle, die orale Bildung ist dominant. Die Zungenseiten berühren die obere Zahnreihe und die Ränder des harten Gaumens. Die Zungenspitze berührt die unteren Schneidezähne. Auf Grund der Zungenhöhe ist [ø] ein halbgeschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[œ]

Das kurze offene [œ] wird durch das Graphem <ö> bezeichnet, wie beispielsweise in *zwölf*. Es ist im An- und Inlaut in betonten geschlossenen Silben Bauelement vom Lautkörper des Sprachzeichens. Nach einer Phonationspause wird es gewöhnlich mit neuem Glottisschlag hart eingesetzt. Das Gaumensegel versperrt den Weg zum Nasenraum, die orale Bildung dominiert. Die Zungenseiten berühren den seitlichen harten Gaumen. Die Zungenspitze hat Kontakt mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [œ] ein halb-

geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[a:]

Das [a:] wird durch die Grapheme <a>, <aa>, <ah> im An-, In- und Auslaut bezeichnet, wie beispielsweise in *Rat*, *Paar* und *Zahn*. Nach einer Phonationspause wird es gewöhnlich mit hartem Glottisschlag neu eingesetzt. Der velopharyngale Nasenzugang ist durch das Velum gesperrt, die orale Bildung dominiert. Weder wölben sich die Zungenränder bis zu den hinteren Oberzähnen auf, noch berührt das Postdorsum das Velum. Auf Grund der Zungenhöhe ist [a:] ein offener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[a]

Das [a] wird durch das Graphem <a> im An- und im Inlaut bezeichnet, wie beispielsweise in *Acker*. Nach einer Phonationspause wird es gewöhnlich durch einen harten Glottisschlag neu eingesetzt. Der velopharyngale Nasenzugang ist durch das Velum gesperrt, die orale Bildung ist dominant. Weder haben die Zungenränder einen Kontakt mit den Oberzähnen, noch das Postdorsum mit dem Velum. Auf Grund der Zungenhöhe ist [a] ein halboffener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein zentraler Vokal.

[u:]

Das [u:] wird durch die Grapheme <u>, <uh> (seltener <ue>, <ou>) im An-, In- und Auslaut bezeichnet, wie beispielsweise in *gut* und *fuhr*. Nach einer Phonationspause wird es gewöhnlich mit hartem Glottisschlag neu eingesetzt. Der velopharyngale Nasenzugang ist durch das Velum abgesperrt, die orale Bildung ist dominant. Die beiden Zungenränder haben Kontakt mit den hinteren Oberzähnen, das Postdorsum berührt beiderseits die entgegengesetzten Flächen des Velums. Die Zungenspitze liegt in Kontaktstellung. Auf Grund der Zungenhöhe ist [u:] ein geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein velarer Vokal.

[u]

Das [u] wird durch das Graphem <u> im Anlaut und vor einem akzentuierten Vokal im Inlaut bezeichnet, wie beispielsweise in *Kusine*. Meistens kommt es in eingedeutschten Wörtern (aber: zuvor, zufiel!) vor. Nach einer Phonationspause ist der Neueinsatz mit hartem Glottisschlag möglich, im Vergleich zu den akzentuierten Vokalen ist er allerdings weniger intensiv. Der velopharyngale Nasenzugang ist durch das Velum gesperrt, die orale Bildung ist dominant. Die Zungenränder berühren die hinteren Oberzähne, das Postdorsum hat beiderseits einen schmalen Kontakt mit dem Velum. Auf Grund der Zungenhöhe ist [u] ein geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein velarer Vokal.

[u]

Das [u] wird durch das Graphem <u> im An- und Inlaut in betonten geschlossenen Silben bezeichnet, wie beispielsweise in *Bund*. Nach einer Phonationspause wird es gewöhnlich mit hartem Glottisschlag neu eingesetzt. Die Zungenränder haben Kontakt mit den Oberzähnen, ein kleiner Streifen des hinteren Zungenrückens berührt die entgegengesetzten schmalen Bereiche des Gaumensegels. Der velopharyngale Nasenzugang ist versperrt, die orale Bildung dominiert. Die Zungenspitze hat Kontakt mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [u] ein geschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein velarer Vokal.

[o:]

Das [o:] wird durch die Grapheme <o>, <oo>, <oh> (in Eigennamen auch durch <oe>, <oi>, bzw. <ow>) im An-, In- und im Auslaut bezeichnet, wie beispielsweise in *Hof*, *Boot* und *wohl*. Nach einer Phonationspause wird es gewöhnlich mit hartem Glottisschlag neu eingesetzt. Der velopharyngale Nasenzugang wird durch das Velum abgesperrt, die orale Bildung dominiert. Die Zungenränder haben durch ihre Aufwölbung Kontakt mit den oberen hinteren Backenzähnen. Die Zungenspitze hat Kontakt mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [o:] ein halbgeschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein velarer Vokal.

[o]

Das [o] wird durch das Graphem <o> im An- und Inlaut vorwiegend einge-deutscher Wörter in unbetonter Position bezeichnet, wie beispielsweise in *Lokal*. Nach einer Phonationspause kann es neu eingesetzt werden, allerdings ist der harte Glottisschlag im Vergleich zum akzentuierten Vokal weniger intensiv. Der velopharyngale Nasenzugang ist durch das Velum abgesperrt, die orale Bildung ist dominant. Die Zungenränder wölben sich auf und berühren die oberen Molarzähne. Die Zungenspitze ist in Kontaktstellung. Auf Grund der Zungenhöhe ist [o] ein halbgeschlossener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein velarer Vokal.

[ɔ]

Das [ɔ] wird durch das Graphem <o> im An- und Inlaut in geschlossenen Silben bezeichnet, wie beispielsweise in *voll*. Nach einer Phonationspause wird es mit hartem Glottisschlag neu eingesetzt. Der velopharyngale Nasenzugang ist durch das Velum abgesperrt, die orale Bildung dominiert. Weder die Zungenränder noch der Zungenrücken haben Kontakt mit den Oberzähnen bzw. mit dem Velum. Die Zungenspitze liegt in Kontaktstellung. Auf Grund der Zungenhöhe ist [ɔ] ein halboffener, auf Grund der Lippentätigkeit ein gerundeter, auf Grund der Zungenlage ein zentraler Vokal.

5.4.2. Reduktionsvokale

Die Reduktionsvokale haben im Deutschen diverse gemeinsame Eigenschaften, z.B. dass sie nur in unbetonter Position vorkommen können. Des Weiteren ist die Streuung ihrer konkreten Realisierungen im Vergleich zu anderen Vokaltypen größer, im artikulatorischen und akustischen Sinne. Diese Schwankungen der Reduktionsvokale sind in unterschiedlichen phonetischen Positionen ebenfalls zu bemerken, die folgenden Beschreibungen sind also nur als Orientierungsversuche zu verstehen.

[ə]

Das [ə] wird durch das Graphem <e> in den unbetonten Präfixen <be->, <ge-> sowie in unbetonter Position hinter dem akzenttragenden Vokal bezeichnet. In den unbetonten Silben nach dem akzentuierten Vokal ist der Realisierungs- bzw. Reduktionsgrad des [ə] von seiner Entfernung vom Akzent, von der lautlichen Umgebung, besonders jedoch von der Zahl und der Qualität der darauf folgenden Konsonanten abhängig. Der velopharyngale Nasenzugang ist versperrt, die orale Bildung dominiert. Durch Aufwölbung der Zungenränder entsteht ein Kontakt zwischen ihnen und den Oberzähnen. Die Zungenspitze hat Kontakt mit den unteren Schneidezähnen. Auf Grund der Zungenhöhe ist [ə] ein halbgeschlossener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein zentraler Vokal.

Wegen der mannigfaltigen nuancenreichen Aussprachebesonderheiten des [ə] sollen an dieser Stelle einige Hinweise zur Orientierung des Sprechers stehen,

entsprechend der Ausführungen des *Großen Wörterbuchs der deutschen Aussprache* (1982: 35 – 36):

Das [ə] wird benutzt in den Präfixen <be->, <ge->, im Auslaut und in unbetonten deutschen Suffixen. In der Endung <-en> verhält es sich folgendermaßen:

- der Murmelvokal ist zu benutzen nach den Nasalen [m], [n], [ŋ], nach [l], [r] und [j] oder nach Vokalen
- der Murmelvokal fällt trotz <e>-Graphems nach den Engelaute[n] [f], [v], [s], [z],[ʃ],[ç], [x] und nach den Affrikaten [pf], [ts] aus. Ähnliches gilt auch für die Endung <-em>. In der Diminutivendung <-chen> wird [ə] aber immer gesprochen
- der Murmelvokal fällt in der Regel nach den Verschlusslauten [p], [b], [t], [d], [k], [g] aus. In diversen Sprechsituationen kann er aber realisiert werden, dementsprechend können im Wörterverzeichnis des „Großen Wörterbuchs der deutschen Aussprache“ beide Formen gefunden werden, wobei immer die Form ohne Murmelvokal an erster Stelle steht. Erfolgt der Ausfall nach [p], [b], wird der Verschluss nasal gelöst und es erscheint statt dem darauf folgenden [n], ein silbenbildendes [m]. Bei Ausfall nach [k], [g] wiederum wird das [n] zu einem silbenbildenden [ŋ] assimiliert.

Ausnahmen: In der Endsilbenverbindung <-igen> muss der Murmelvokal realisiert werden, in der Verbindung <-iken> nur dann, wenn der Akzent im Wort auf die drittletzte Silbe fällt z.B. *berichtigen*, *Basiliken*. Bei Endsilbenhäufun-

gen darf der Murrelvokal, wie oben bereits gesagt, nur in der ersten Endsilbe und nach Enge- und Verschlusslauten ausfallen, z.B. *rasenden* ['ra:zndən].

In der Endung <-el> verhält sich das [ə] wie folgt: Es fällt nach den Nasalen [m], [n], [ŋ], nach den Englauten [f], [v], [s], [z], [ʃ], [ʒ], [ç], [x], nach den Affrikaten [pf], [ts] und nach den Verschlusslauten [p], [b], [t], [d] und [k] aus. In diesem Fall wird [l] silbisch, und [t] und das [d] werden lateral gesprengt.

[ə]

Das [ə] wird durch das Graphem <r> nach dem kurzen, unbetonten <e> bezeichnet. In diesem Fall ist [ə] silbisch, wie beispielsweise in *Maler* ['ma:lə]. [ɐ] ist die Realisierung nach Langvokalen wie [i:], [y:], [e:], [ø:], [o:] und [u:]. In letzten Fällen ist [ɐ] unsilbisch. Hier bildet [a:] eine Ausnahme, nach dem [ɐ] nicht erscheint z.B. *Jahr* [ja:].

Die Lippen haben keinen Kontakt. Die vorderen Zungenseiten berühren die unteren Schneidezähne. Der mittlere Zungenrücken wölbt sich geringfügig zum harten Gaumen auf, so dass ein neutraler Vokalklang entsteht. Der velopharyngale Nasenzugang ist versperrt, die orale Bildung dominiert.

Auf Grund der Zungenhöhe ist [ə] ein halboffener, auf Grund der Lippentätigkeit ein ungerundeter, auf Grund der Zungenlage ein zentraler Vokal.

5.4.3. Diphthonge

Ein Diphthong ist nach PÉTURSSON & NEPPERT (1991: 106) „ein vokalisches Element [...], das aus zwei kontinuierlich ineinander übergehenden Klangfarben besteht und einen Silbenkern bildet.“ Bei dieser Definition ist deutlich zu sehen, dass sie nicht der vereinfachenden Beschreibung „Zusammensetzung aus zwei Vokalelementen“ folgt, sondern gemäß der artikulatorischen und akustischen Realität eher den gleitenden Übergang von Element zu Element betont.

Dieser Auffassung kann hinzugefügt werden, dass der Ausgangspunkt der Artikulationsbewegungen nicht automatisch mit dem Ausgangspunkt des ersten Diphthongelements, der Endpunkt nicht unbedingt mit dem Endpunkt des zweiten Diphthongelements übereinstimmen muss. Die komplexe Bildung der Diphthonge ist überwiegend durch den nicht gleichmäßigen Übergang von Element zu Element geprägt. Deswegen muss betont werden, dass in der folgenden Darstellung die Diphthongelemente nur unter Berücksichtigung dieser Einschränkungen in einer möglichen Form der Transkription stehen, was der Übersichtlichkeit und damit dem besseren Verständnis dienen soll. Dieselben Diphthonge werden in der Fachliteratur auch durch andere Transkriptionszeichenkombinationen ausgedrückt wie z.B. [ao], [ae], [ɔø].

Nach der Bewegung der Zunge im Mundraum werden im Allgemeinen zwei Arten von Diphthongen unterschieden: 1. schließende (auch: steigende) 2. öffnende (auch: fallende). In der deutschen Standardvarietät sind nur schließende Diphthonge vorhanden. In deutschen Dialekten hingegen sind auch öffnende Diphthonge zu finden wie beispielsweise in der bairischen Variante des Wortes *gut* [ġuāt].

[aʊ]

Der schließende Diphthong [aʊ] wird durch das Graphem <au>, wie beispielsweise in *auf*, bezeichnet, der im An-, In- und Auslaut meistens in betonter Position vorkommt. Nach einer Phonationspause ist der neu eingesetzte Glottisschlag für die Bildung charakteristisch. Das Gaumensegel schließt den Weg zum Nasenraum während der gesamten Bildung ab, die orale Bildung ist dominant. Auf Grund der vertikalen Bewegung der Zunge ist eine Schließung in der Bildung vorhanden, die jedoch nicht gleichmäßig abläuft. Das [aʊ] setzt sich aus diversen Elementen des [a] und des [u], bzw. aus den verbindenden Gleitbewegungen der Artikulation zusammen. Die erste Diphthongskomponente ist ungerundet, die zweite wird gerundet realisiert.

[ai]

Der schließende Diphthong [ai] wird durch die Grapheme <ei>, <ai>, <ey>, <ay>, wie beispielsweise in *eins*, *Mai*, *Meyer* und *Bayer*, bezeichnet, der im An-, In- und Auslaut, in der Regel in betonter Position vorkommt. Diese Beschränkung bezieht sich aber nicht auf die Realisierungen im Inlaut. Nach einer Phonationspause wird es gewöhnlich mit hartem Glottisschlag eingesetzt. Das Gaumensegel sperrt während der gesamten Bildung den Weg zum Nasenraum ab, die orale Bildung dominiert. In der Artikulation ist eine Schließung zu beobachten, die Zungenlage verändert sich aber nicht gleichmäßig. Der Diphthong setzt sich aus diversen Elementen des [a] und des [ɪ] bzw. aus Elementen des gleitenden Übergangs zusammen. Alle Bildungselemente sind illabial.

Der schließende Diphthong [ɔɪ] wird durch die Grapheme <eu>, <äu>, <oi>, <oy>, wie beispielsweise in *Äußerung* und *Eule*, bezeichnet, der im An-, In- und Auslaut vorkommt. Nach einer Phonationspause wird es gewöhnlich mit hartem Glotisschlag neu eingesetzt. Das Gaumensegel schließt den Weg in den Nasenraum während der gesamten Bildung ab, die orale Bildung ist dominant. In der Artikulation ist die schließende Bewegung der Zunge am auffälligsten, die aber nicht gleichmäßig abläuft. Der Diphthong setzt sich aus diversen Elementen des [ɔ] und des [ɪ], bzw. aus Elementen des fließenden Übergangs zusammen. Die erste Diphthongkomponente wird gerundet, die zweite ungerundet realisiert.

5.5. Das Vokalsystem der deutschen Standardvarietät auf Grund der artikulatorischen Merkmale

Das System der deutschen Vokale wurde schon früh durch eine schematische Darstellung zusammengefasst. Das sog. HELLWAGsche Dreieck aus dem Jahre 1781 wurde auf Grund der Zungenstellung erstellt, die erweiterte und genauere Form stammt aus dem Jahre 1925 und nennt sich Kopenhagener Viereck:

Zungenhebung Vorderzungenvokale Hinterzungenvokale

hoch	[i]	[u]
	[ɪ]	[ʊ]
mittel	[e]	[o]
	[ɛ]	[ɔ]
tief	[a]	[ɑ:]

In jüngster Zeit wird meistens das sog. Kardinalvokalsystem angeführt, das vom englischen Phonetiker Daniel JONES entwickelt wurde:

1	i·	· u	8
2	e·	· o	7
3	ɛ·	· ɔ	6
4	a·	· ɑ	5

In den Sprachen der Welt kommen diese Vokale am häufigsten vor. Interessant ist dabei, dass die vorderen Vokale immer ungerundet, die hinteren fast immer – mit Ausnahme des niedrigen Vokals – gerundet sind.

Abbildung 6 zeigt ein Vokaltrapez, das auf der Basis der artikulatorischen und akustischen Bezüge eines weiteren Vokaltrapezes erstellt wurde und damit beide Aspekte berücksichtigt. Wie man sehen kann, wurden im Laufe der Zeit die schematischen Darstellungen des Vokalismus immer feiner, und immer mehr Aspekte und Charakteristika der artikulatorischen Gegebenheiten konnten berücksichtigt werden. Dennoch warnen zahlreiche Autoren davor, die Bedeutung solcher Zusammenfassungen zu überschätzen. Diese Schemata

stellen keine absoluten artikulatorischen Zusammenhänge dar, erst recht spiegeln sie nicht die exakten Bewegungen der Sprechwerkzeuge wider. Die angeführten Darstellungen lassen lediglich einen ersten Überblick entstehen, der das Verstehen erleichtert. Sogar das zuletzt angeführte Kardinalvokaltrapez legt nur die Extrempunkte artikulatorisch fest ([i] am einen und [ɑ] am anderen Ende). Alle anderen Vokale sind rein auditiv festgelegt und im Trapez in regelmäßigen Abständen positioniert worden, was den physikalisch-artikulatorischen Tatsachen nicht genau entspricht. Es ist wichtig, auf die Relativität und die eingeschränkte Gültigkeit der Darstellungen hinzuweisen.

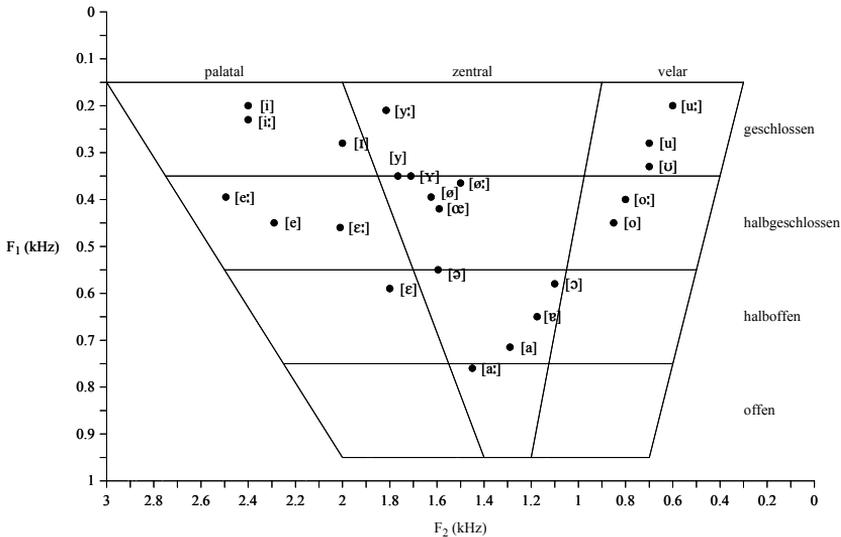


Abbildung 6: Artikulatorisches und akustisches Vokalviereck. Das System der Vokale der deutschen Standardvarietät auf Grund der Zungenlage, der Zungenhöhe sowie der Frequenzwerte F₁ und F₂

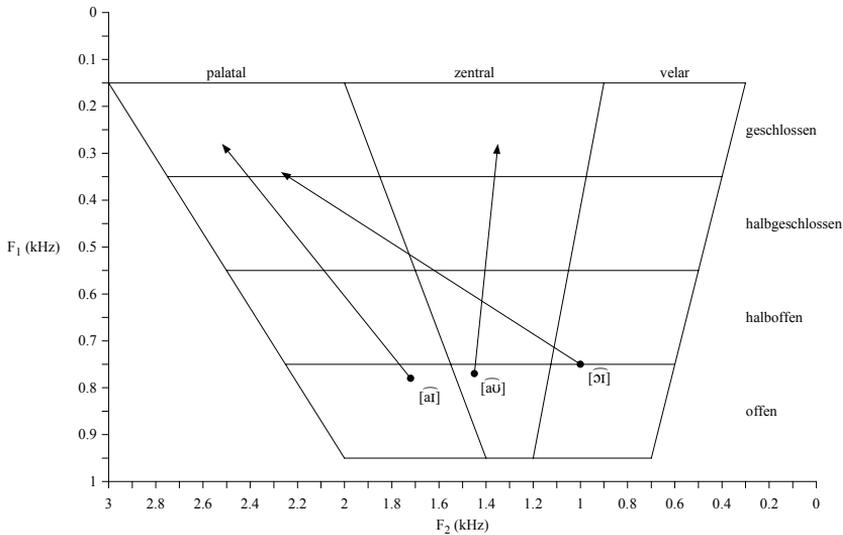


Abbildung 7: Die Diphthonge der deutschen Standardvarietät im artikulatorischen und akustischen Vokalviereck veranschaulicht

5.6. Konsonanten

Die Bezeichnung Konsonant stammt aus dem Lateinischen und beruht auf der früheren Überzeugung, nach der bei dieser Art von Lauten „etwas mittönen“ muss (vgl. WÄNGLER, 1983: 117). Die alte deutsche Form „Mitlaut“ spiegelt ebenfalls eine frühere Auffassung wider nämlich die, dass Vokale gesondert gesprochen werden können, Konsonanten aber nur in Verbindung mit Vokalen. Beide Auffassungen sind nicht zu vertreten, da Nasale und Liquide ohne Beigeräusch produziert werden (d.h. es tönt nichts mit). Interjektionen wie *pst*, *hm* u.Ä. bezeugen zudem, dass Konsonanten auch ohne Vokale „sprechbar“ sind.

Laut WÄNGLER (1983: 119) gibt es in der Gesamtklasse der Konsonanten die Klasse der eigentlichen (stimmhafte und stimmlose) Geräuschlaute oder Konsonanten und die immer stimmhaften Sonanten. Die Sonanten zeichnen sich dadurch aus, dass sie keinen oder einen untergeordneten Geräuschanteil haben bzw. dass sie einen Silbengipfel bilden können. Des Weiteren besteht das Gesamtklassenmerkmal darin, dass die Konsonanten mit Ausnahme des Öffnungskonsonanten [h] dadurch zustande kommen, dass irgendwo im Mundraum an einer oder mehreren Stellen der Strom der Phonationsluft behindert wird.

Diese Definition wird auch von PÉTURSSON & NEPPERT (1991: 89) verwendet, allerdings mit dem Hinweis, dass das Hindernis im Ansatzrohr nur bezogen auf den expiratorischen Luftstrom und relativ zu dessen Stärke zu sehen sei. Mit diesen Einschränkungen ist die Abgrenzung zu akzeptieren, mit der Bemerkung, dass die Zweiteilung der Laute eher eine methodologisch und traditionell begründete Funktion hat.

5.6.1. Klassifizierung der Konsonanten nach primären artikulatorischen Merkmalen

5.6.1.1. Artikulationsstelle

Der Begriff der Artikulationsstelle (auch als Artikulationsort bezeichnet) wurde bereits in 3.2.1.1. erwähnt und umrissen und soll hier auf Grund seiner Bedeutung bei der Klasse der Konsonanten im Prozess der Artikulation noch einmal angesprochen werden.

Als Artikulationsstelle wird herkömmlicherweise der Ort der Lautbildung, d.h. der Ort im Ansatzrohr bezeichnet, der am stärksten im Prozess der Erzeugung des jeweiligen Lautes das Endprodukt bestimmt und so bei der Artikulation eine entscheidende Funktion ausübt. Nach WÄNGLER (1983: 39) werden „an all diesen näher gekennzeichneten Stellen [...] Laute dadurch hervorgebracht, dass sich andere Teile der Sprechwerkzeuge diesen Bereichen mehr oder weniger nähern, Engen oder Verschlüsse bilden“. Nach dieser Auffassung sind die relativ passiven Stellen im Ansatzrohr bei der Bestimmung der Artikulationsstelle zwar wichtig (vgl. auch ESSEN 1966: 71), aber nicht ausschlaggebend bei der Bestimmung der Artikulationsstelle. Diese Auffassung wird von PÉTURSSON & NEPPERT (1991: 98) dadurch ergänzt, dass sie den Ort im Ansatzrohr als Artikulationsstelle bezeichnen, an dem der Abstand zwischen zwei Organen am geringsten ist. Nach dieser Definition sind auch Verschlusslaute, bei denen der Abstand gleich Null wäre und auch Vokale, bei denen dieser Abstand z.T. relativ bedeutend sein kann, ohne weiteres ins System einzuordnen.

5.6.1.2. Artikulationsmodus

Unter Artikulationsmodus versteht LINDNER (1969: 196) die „[...] Art und Weise, wie die Hemmstelle ausgebildet wird.“ Hier wird also festgehalten, dass der Artikulationsmodus (auch: Artikulationsart) besagt, wie die Phonationsluft im Ansatzrohr bzw. in der Glottis je nach Lautklasse weitergeleitet bzw. modifiziert wird. Von PÉTURSSON & NEPPERT (1991: 90) wird diese Definition dadurch erweitert, dass die Autoren neben dem Luftstrom auch noch Aspekte wie Auswirkungen von Druckdifferenzen bzw. resonanzbedingte Qualitätsbeeinflussungen durch Hohlraumveränderungen anführen. Auch

die Bemerkung, dass „es sich um Artikulationseinstellungen oder -vorgänge handelt, die nicht bestimmte einzelne Sprachlaute darstellen, sondern als Merkmale einer Klasse von Sprachlauten fungieren“ muss erwähnt werden, da dies den artikulatorisch-physikalischen Realitäten näher steht. Einige Autoren rechnen des Weiteren auch die Stimmbeteiligung zum Artikulationsmodus, bzw. PÉTURSSON & NEPPERT (1991: 96) sehen sogar die Vokale als einen selbstständigen Artikulationsmodus an. Bei der Beschreibung des deutschen Vokal- und Konsonantensystems scheint es aber aus methodischen Gründen besser, bei der traditionellen Gliederung in Vokale und Konsonanten zu bleiben, auch wenn – wie schon angeführt – beide Begriffe gewisse Fehlerquoten beinhalten, was die Abgrenzung der beiden Kategorien anbelangt.

Mit den angeführten Beschränkungen kommen also in der deutschen Standardvarietät bei Konsonanten folgende Artikulationsmodi vor:

- oral: die Nasenhöhle wird durch das Velum abgesperrt, so dass die pharyngo-orale Bildung dominiert.
- Enge: es erfolgt eine Engebildung zwischen zwei oder mehr Organen im Ansatzrohr,
- die in der Phonationsluft Reibungen und Turbulenzen verursacht.
- Approximant: in der gebildeten Enge entstehen – z.T. wegen der geringeren Turbulenz – keine bedeutenden Reibungen und Turbulenzen.
- Verschluss: die Phonationsluft wird durch einen Verschluss von zwei oder drei Organen zeitweise daran gehindert, durch die Mundöffnung ins Freie hinauszuströmen. Bei Plosiven wird dieser Verschluss in der Lösungsphase ganz abrupt und oral gesprengt, so dass starke Geräuschelemente entstehen.

- Nasal: im Mundraum wird ein Verschluss gebildet, der den Weg der Phonationsluft versperrt, das Velum ist gesenkt, und so kann der Luftstrom durch die Nasenhöhle das Ansatzrohr verlassen. Es entsteht in diesem pharyngo-nasalen Raum eine spezifische Schallqualität, die für diese Klasse von Lauten charakteristisch ist.
- Zentral: die Phonationsluft verlässt das Ansatzrohr durch den mittleren Bereich des Mundraumes, und auch die entscheidenden artikulatorischen Bewegungen finden hier statt.
- Lateral: die Phonationsluft entweicht auf der einen Seite oder beidseitig neben der Zunge und verlässt so den Mundraum.
- Affrikate: die Affrikaten zeichnen sich dadurch aus, dass nach einem Verschlusselement, ein (meistens homorganes) Engeelement folgt und sich so ein klassenspezifischer Artikulationsvorgang vollzieht.
- Vibrant: für diese Klasse von Lauten ist ein durch die Phonationsluft hervorgerufenen Vibrieren eines Organs charakteristisch, das so einen intermittierenden Verschluss mit einem anderen Organ bildet.

5.6.1.3. Stimmbeteiligung

Die Stimmbeteiligung stellt ein Merkmal dar, das bei den deutschen Konsonanten ein prinzipiell binäres Kriterium ist, einerseits auf Grund der allgemeinen Stimmhaftigkeit der deutschen Vokale, andererseits weil in der normierten deutschen Aussprache dies – was die Systematisierung anbelangt – ausreicht, auch wenn durch Koartikulation u.Ä. die Stimmhaftigkeit bestimmter Konsonanten im Redefluss reduziert wird (vgl. PÉTURSSON & NEPPERT 1991: 98 – 99).

Der Laut ist stimmhaft, wenn im Prozess der Erzeugung des Sprechschalls die Stimmbänder durch die Phonationsluft ins Schwingen kommen und so das Endprodukt maßgeblich beeinflussen. Wenn dies nicht der Fall ist, wird der Laut als stimmlos bezeichnet.

Der Überwindungsmodus, der früher als Kategorie zur Charakterisierung von Konsonanten gebraucht wurde, hatte Überschneidungen von diversen artikulationsmodusbezogenen Eigenschaften bzw. von der Kategorie der Stimmbeteiligung in sich vereint. Es erscheint also ratsamer, die Konsonanten nach den verwendeten drei primären Merkmalen zu kategorisieren.

5.6.2. Aspekte der sekundären Modifikationen

Bei den Konsonanten des Deutschen ist ein breites Feld von sekundären Modifikationen anzuführen. Sekundär wird eine Modifikation betrachtet, wenn „eine zusätzliche Verengungsstelle zur eigentlichen Artikulationsstelle“ (PÉTURSSON & NEPERT 1991: 99) existiert, die also die Gesamtheit der Bewegungen der Sprechwerkzeuge beeinflusst. Solche sekundäre Modifikationen sind z.B. die Nasalierung (wie schon erwähnt, ist sie bei den deutschen Vokalen wichtig), Palatalisierung oder Velarisierung. Am häufigsten treten in der deutschen Standardvarietät sekundäre Modifikationen jedoch in Lautverbindungen hervor. Diesbezüglich gibt es aber eine definitionsbedingte Diskussion in der Phonetik, so dass an dieser Stelle auf keinen weiteren Diskurs über diese Problematik eingegangen wird.

5.7. Allgemeine Tendenzen bei der Realisierung der Konsonanten in der deutschen Standardvarietät

Bei der Realisierung der deutschen Konsonanten ist vor allem die phonetische Position derselben im Lautkörper der sprachlichen Zeichen ausschlaggebend. Folgende Aspekte müssen bei der Aussprache deshalb besonders berücksichtigt werden:

- ein einfacher Konsonant wird in den einfachen Wörtern einmal und kurz realisiert. Doppelkonsonanz kommt im Anlaut gar nicht vor.
- die schriftliche Doppelkonsonanz wird in den einfachen Wörtern im In- und Auslaut ebenfalls kurz realisiert, z.B. *essen* [ɛsn̩], *Watte* [vatə], *Lamm* [lam].
- in den abgeleiteten und zusammengesetzten Wörtern sowie im Inneren von Sinneinheiten werden benachbarte gleiche Konsonanten an den Morphemgrenzen einmal und lang realisiert, z.B. *annehmen*, *Lauffeuer*, *die Katze im Sack kaufen*.
- die Auswirkungen der Koartikulation (siehe 5.9.) Bei bestimmten Konsonanten (z.B. [l], [g], [k]) tritt die Erscheinung aber besonders häufig auf
- die stimmlosen Verschlusslaute [p], [t], [k] werden im Allgemeinen mehr oder weniger stark behaucht (aspiriert), wobei die Intensität und die Aspirationstendenz auch regional unterschiedlich sein kann. Diese Behauchungstendenzen werden aber ebenfalls u.a. durch die lautliche Umgebung beeinflusst. Nach dem *Wörterbuch der deutschen Aussprache* (1964: 45) werden die stimmlosen Verschlusslaute in folgenden Fällen meistens nicht behaucht:

- vor unbetontem Vokal, z.B. *atamar, Ecke, hastig*
- vor folgenden Konsonanten, z.B. *absorbieren, Haupt*
- in der Lautverbindung <sp> und <st>, wenn beide Laute der gleichen Silbe angehören, z.B. *Spreu, Stein, Statist*

5.8. Die Konsonanten der deutschen Standardvarietät

5.8.1. Verschlusslaute

5.8.1.1. Plosive

[b]

Das [b] wird durch die Grapheme und <bb> bezeichnet:

- in intervokalischer Position und vor Wortelementen mit [l], [R] im Silbenanlaut ist die Realisierung stimmhaft z.B. *Blei* [blai]
- in anderen anlautenden Positionen ist die Realisierung stimmlos beispielsweise *bald* [balt]. In diesen Fällen wird der Unterschied zum [p] einerseits durch das Verhältnis zum gesamten Lautkörper, andererseits durch gewisse Elemente der akustischen Struktur wie das Fehlen der Aspiration, die Zeitstruktur der Komponenten bzw. Intensität derselben ausgedrückt.

Bei der stimmhaften Bildung schwingen die Stimmbänder mit, im Kehlkopf entsteht ein Stimmklang, während bei der stimmlosen Bildung durch das Vib-

rieren der Stimmbänder kein Stimmklang entsteht. Das Velum schließt den Weg zur Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: die Bildung eines Verschlusses und eine anschließende, explosionsartige Lösung ist charakteristisch.

Artikulationsstelle: die beiden Lippen. Die Zungenseiten können die oberen Zähne berühren. [b] ist ein stimmhafter/stimmloser bilabialer Plosiv.

[p]

Das [p] wird durch das Graphem <p>, <pp>, bezeichnet, wie beispielsweise in *Park* und *lieb*. [p] ist an-, in- und auslautend ein Element des Lautkörpers. Die Stimmbänder lassen die Phonationsluft ohne bedeutendes Vibrieren passieren, im Kehlkopf entsteht also kein Stimmklang. Das Gaumensegel sperrt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: Die Ober- und Unterlippe bilden einen Verschluss, der gelöst wird.

Artikulationsstelle: Die beiden Lippen. Die Zungenseiten spielen keine aktive Rolle in der Artikulation. [p] ist ein stimmloser bilabialer Plosiv.

[d]

Das [d] wird durch die Grapheme <d>, <dd> bezeichnet. In intervokalischer Position und vor einem Wortelement mit [ʀ] im Anlaut ist seine Realisierung stimmhaft z.B. *drei* [dʀai], in anderen anlautenden Positionen stimmlos beispielsweise *Dach* [dax]. In solchen Fällen ist nicht der Stimmklang, sondern das Verhältnis zum gesamten Lautkörper, die fehlende Aspiration und die Zeit- und Intensitätsstruktur der Komponenten ausschlaggebend für den Unterschied zum [t]. Bei der stimmhaften Bildung schwingen die Stimmbänder mit, im Kehlkopf wird ein Stimmklang produziert. Bei der stimmlosen Bildung entsteht durch das Vibrieren der Stimmbänder kein Stimmklang. Das Gaumensegel schließt die Nasenhöhle ab, die orale Bildung dominiert.

Der Artikulationsmodus: Verschlussbildung und anschließendes Sprengen des Verschlusses. Die Zungenseiten berühren die obere Zahnreihe und den Gaumen.

Die Artikulationsstelle ist im alveolaren Bereich mit der Zungenspitze und dem Prädorsum in Konstellation anzusetzen. [d] ist ein stimmhafter/stimmloser alveolarer Plosiv.

[t]

Das [t] wird durch die Grapheme <t>, <tt>, <dt>, <d> bezeichnet, wie beispielsweise in *Tal* und *bald*. [t] ist im An- In- und Auslaut Bauelement des

Lautkörpers. Im Kehlkopf wird kein Stimmklang produziert, das Velum sperrt den Weg der Phonationsluft zur Nasenhöhle ab: Die orale Bildung dominiert.

Artikulationsmodus: die Bildung eines Verschlusses und anschließende Sprengung desselben. Die Zungenseiten berühren die obere Zahnreihe und den Gaumen.

Artikulationsstelle: Konstellation im alveolaren Bereich mit der Zungenspitze und dem Prärdorsum. [t] ist ein stimmloser alveolarer Plosiv.

[g]

Das [g] wird durch die Grapheme <g>, <gg> bezeichnet. Die Bildung ist in intervokalischer Position und vor einem Wortelement mit [l], [r], [n] im Silbenanlaut stimmhaft beispielsweise *Grad* [gra:t], in anderen anlautenden Positionen stimmlos z.B. *Gabe* ['g̥a:bə]. In solchen Fällen trägt die fehlende Aspiration und die akustische Struktur zur Unterscheidung von [k] bei. Nach der normierten Hochlautung ist die Bildung anlautend stimmhaft. Bei der stimmhaften Bildung entsteht durch das Vibrieren der Stimmbänder ein Stimmklang, dieser bleibt bei der stimmlosen Realisierung aus. Das Velum schließt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: die Bildung eines Verschlusses und sein Sprengen. Die Zungenseiten spielen bei der Artikulation keine aktive Rolle.

Die Artikulationsstelle befindet sich im velaren und im postdorsalen Bereich. [g] ist ein stimmhafter/stimmloser velarer Plosiv.

[k]

Das [k] wird durch die Grapheme <k>, <ck>, <kk>, <qu>, <ch>, <g>, wie beispielsweise in *Käfer*, *Bäcker*, *Quelle* und *Tag*, bezeichnet, es ist an-, in- und auslautend ein Bauelement des Lautkörpers des Sprachzeichens. Im Kehlkopf wird kein Stimmklang produziert, das Velum schließt die Nasenhöhle ab, die orale Bildung dominiert.

Der Artikulationsmodus: Verschlussbildung und anschließendes Sprengen desselben. Die Zungenseiten spielen keine aktive Rolle bei der Artikulation.

Die Artikulationsstelle ist im velaren Bereich und im Bereich des hinteren Zungenrückenstreifens anzusetzen. [k] ist ein stimmloser velarer Plosiv. Die Komponenten seiner akustischen Struktur setzen sich aus Effekten der Verschlusslösung und der Aspiration zusammen.

5.8.1.2. Nasale

[m]

Das [m] wird durch die Grapheme <m>, <mm>, wie beispielsweise in *Mut* und *immer*, bezeichnet, es ist an-, in- und auslautend ein Bauelement des Lautkörpers des Sprachzeichens. Die Stimmbänder werden durch die Phonationsluft ins Schwingen gebracht, im Kehlkopf entsteht ein Stimmklang. Das Gau-

mensegel macht den Weg zur Nasenhöhle frei, so dass der nasale Resonanzraum bei der Artikulation dominiert.

Artikulationsmodus: das Bilden, Anhalten und Lösen eines Verschlusses. Während der Haltephase entweicht die Phonationsluft durch die Nasenhöhle.

Die Artikulationsstelle ist bei den Lippen anzusetzen, die Zungenseiten spielen keine aktive Rolle bei der Artikulation. [m] ist ein stimmhafter bilabialer Nasal.

Vor labiodentalen Frikativen wird [m̥] statt [m] realisiert beispielsweise *Amphitheater* [am̥'fi:teat̥ə]

[n]

Das [n] wird durch die Grapheme <n>, <nn>, wie beispielsweise in *Nuss* und *Anne*, bezeichnet, es ist im An-, In- und Auslaut Bauelement des Laukörpers. Die Stimmbänder schwingen mit, im Kehlkopf entsteht ein Stimmklang. Das Gaumensegel macht den Weg der Phonationsluft in die Nasenhöhle frei, so dass diese als dominanter Resonanzraum fungiert.

Artikulationsmodus: das Bilden, Anhalten und Lösen eines Verschlusses. Während der Haltephase verlässt die Phonationsluft das Ansatzrohr durch die Nase. Die Zungenseiten berühren die obere Zahnreihe und den harten Gaumen.

Die Artikulationsstelle ist der alveolare Bereich, in Konstellation mit der Zungenspitze bzw. -saum. [ŋ] ist ein stimmhafter alveolarer Nasal.

[ŋ]

Das [ŋ] wird durch das Graphem <ng> bezeichnet, wie beispielsweise in *Ring*. Bei den graphematischen Zeichen <nk> ist die Realisierung des Lautes ebenfalls [ŋ], wobei hier das [k] ebenfalls realisiert wird, während <g> in <ng> nur orthographisch erscheint. [ŋ] ist in- und auslautend eine Konstituente des Sprachzeichens. Die Stimmbänder schwingen mit, im Kehlkopf entsteht ein Stimmklang. Das Velum macht den Weg zur Nasenhöhle frei, die nasale Bildung dominiert.

Der Artikulationsmodus ist als Bildung, Anhalten und – inlautend – Lösung eines Verschlusses zu beschreiben.

Die Artikulationsstelle ist der velare Bereich in Konstellation mit dem mittleren und hinteren Teil des Zungenrückens. Die Zungenseiten berühren die oberen Zähne. [ŋ] ist ein stimmhafter velarer Nasal.

5.8.1.3. Vibranten

[R]

Das [R] wird durch das Graphem <r>, <rr>, wie beispielsweise in *rot* und *irre*, bezeichnet, im An-, In- und Auslaut des Lautkörpers. In den verschiedenen phonetischen Positionen kann der Reduktionsgrad des Lautes unterschiedlich sein. Die Stimmbänder kommen ins Schwingen, im Kehlkopf wird ein Stimmklang gebildet. Das Gaumensegel schließt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: ein intermittierender Verschluss wird im hinteren Bereich des Mundraumes gebildet.

Die Artikulationsstelle ist im Bereich der Uvula bzw. im hinteren Streifen des Zungerückens. Die Zungenseiten spielen keine aktive Rolle bei der Artikulation. [R] ist ein stimmhafter uvularer Vibrant.

An dieser Stelle muss darauf hingewiesen werden, dass r-Typen auch im alveolaren Bereich gebildet werden können ([r]). Ein weiterer r-Typ ist das uvulare Frikativ [ʀ], das sich von den erwähnten zwei in seinem Artikulationsmodus unterscheidet. Welche Variante vom muttersprachlichen Sprecher benutzt wird, hängt von vielschichtigen Aspekten ab, vor allem ist der Einfluss von oberdeutschen Dialekten in den entsprechenden Gebieten hier zu erwähnen, wo das Zungenspitzen-r typisch ist, auch was die deutsche Standardvarietät angeht. Auslautend wird das /r/ vokalisiert aufgelöst (siehe Mittelzungenvokal unter Reduktionsvokalen).

Im Ungarischen ist das alveolare, apikale [r] der am weitesten verbreiteter r-Typ, so dass an dieser Stelle keine konfrontativen Beispiele angeführt werden. Allerdings sind die Unterschiede zwischen dem Zungenspitzen-r-Typ und der ungarischen Variante kurz darzustellen: Die leichte Breitspannungstendenz der Lippen beim ungarischen Lauttyp fehlt beim Zungenspitzen-r und der entscheidende artikulatorische Moment findet zwar im apikoalveolaren Bereich, aber näher der oberen Zahnreihe statt. Des Weiteren sind die Vibrationen nicht so intensiv und häufig wie bei der ungarischen Variante.

5.8.2. Engelaute

5.8.2.1. Frikative

[v]

Das [v] wird durch die Grapheme <w>, <v> bezeichnet. Das Graphem <w> bezeichnet an- und inlautend in deutschen Wörtern, das Graphem <v> in Lehnwörtern [v], wenn danach ein stimmhafter Laut folgt. Wenn das Graphem <v> inlautend zwischen Vokalen bzw. zwischen [R] und einem Vokal steht, kann es zu Schwankungen zwischen [v] und [f] kommen, z.B. *Eva* [e:va/e:fa]. [v] kommt des Weiteren in der Realisierung der Verbindung <qu> vor. Die Stimmbänder schwingen mit, im Kehlkopf wird ein Stimmklang gebildet. Das Velum sperrt den Weg zur Nasenhöhle ab, die orale Bildung dominiert.

Der Artikulationsmodus: es wird im vorderen Bereich des Mundraumes eine schmale Enge gebildet, durch die die Phonationsluft gepresst wird.

Die Artikulationsstelle ist im Bereich der unteren Kante der oberen Zahnreihe bzw. der Unterlippe anzusetzen. Die Zungenseiten haben keine aktive Rolle bei der Artikulation. [v] ist ein stimmhafter labiodentaler Frikativ.

In intervokalischer Position und in der lautlichen Umgebung von Sonoranten ist die Realisierung von [v] stimmhaft beispielsweise *Uwe* [u:və]. [ʏ] ist entstimmt in anlautender Position z.B. *Wind* [ʏmt]. Zur Klasse der Sonoranten gehören die folgenden Laute: Nasale, Vibranten, Laterale und Approximanten.

[f]

Das [f] wird durch das Graphem <f> an-, in- und auslautend bezeichnet z.B. *vorn*; wenn das Graphem <ff> es im In- und Auslaut bezeichnet, wie beispielsweise in *Affe*; wenn das Graphem <v> es an- in- und auslautend bezeichnet z.B. *positiv*, nach den Regeln, die beim Punkt über das [v] angeführt wurden, bzw. in Fremdwörtern im Wort- und Silbenauslaut, wenn in der abgeleiteten Form das [v] in intervokalischer Position steht; wenn das Graphem <ph> es in Wörtern aus dem Griechischen bezeichnet, wie beispielsweise in *Phonetik*; wenn das Graphem <w> es in Auslautsilben von Wörtern aus slawischen Sprachen bezeichnet z.B. *Tschechow*. Im Kehlkopf wird kein Stimmklang produziert, die orale Bildung dominiert.

Artikulationsmodus: eine schmale Enge wird gebildet, wodurch die Phonationsluft hinausströmt. Die Artikulationsstelle ist im Bereich der unteren Kante der oberen Zähne bzw. der Unterlippe anzusetzen. Die Zungenseiten spielen

keine aktive Rolle bei der Artikulation. [f] ist ein stimmloser labiodentaler Frikativ.

[z]

Das [z] wird durch das Graphem <s> bezeichnet. Im Wort- und Silbenanlaut vor einem Vokal ist es Bauelement des Lautkörpers. Die Stimmbänder schwingen mit, im Kehlkopf wird ein Stimmklang produziert. Das Gaumensegel verschließt den Weg in die Nasenhöhle, die orale Bildung dominiert. [z], das entstimmt ist, kommt anlautend in den Varietäten des Deutschen vor beispielsweise *sagen* [z̥a:ɡ̊n̥]. [z] erscheint stimmhaft nur in intervokalischer Position oder in einer lautlichen Umgebung von Sonoranten in der deutschen Standardvarietät beispielsweise *reisen* [raɪz̥n̥].

Artikulationsmodus: im vorderen Bereich des Mundraumes wird eine schmale Enge gebildet, wodurch die Phonationsluft durchgepresst wird.

Die Artikulationsstelle kann durch die Konstellation des vorderen Abschnittes des Zungenrückens und des Zahndamms charakterisiert werden. Die Zungenseiten stehen im Kontakt mit den oberen Zähnen. [z] ist ein stimmhafter alveolarer Frikativ.

[s]

[s] wird im Wort- und Silbenauslaut, in Wörtern fremden Ursprungs im Anlaut, bzw. im Inlaut vor [t] und [p] als Bauelement des Lautkörpers produziert.

Es wird durch die Grapheme <s>, <ss>, <ß> bezeichnet, wie beispielsweise in *skandinavisch*, *bläss* und *bloß*. Die Stimmbänder vibrieren nicht, es entsteht kein Stimmklang. Das Velum sperrt den Weg in die Nasenhöhle ab, die orale Bildung ist dominant.

Artikulationsmodus: im zentral-vorderen Bereich des Mundraumes wird eine schmale Enge gebildet, durch die die Phonationsluft gepresst wird.

Die Artikulationsstelle ist im Bereich des Zahndamms und dem vorderen Streifen des Zungenrückens anzusetzen. Die Zungenseiten berühren die obere Zahnreihe. [s] ist ein stimmloser alveolarer Frikativ.

[ʒ]

Das [ʒ] wird in Wörtern aus dem Französischen im An- und Inlaut vor palatalen Vokalen [e, i] durch das Graphem <g>, vor velaren Vokalen durch die Grapheme <ge>, <j> bezeichnet. Die Stimmbänder schwingen mit, im Kehlkopf entsteht ein Stimmton. Das Velum schließt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: Bildung einer schmalen Enge im zentralen Bereich des vorderen Mundraumes, durch die die Phonationsluft gepresst wird.

Die Artikulationsstelle ist durch die Konstellation des hinteren Streifens des Zahndamms und des vorderen Bereichs des Zungenrückens zu charakterisieren. Die Zungenseiten stehen in Berührung mit der oberen Zahnreihe. [ʒ] ist ein stimmhafter postalveolarer Frikativ.

Das entstimmte [ʒ̥] kommt anlautend vor z.B. *Journal* [ʒ̥ʊʀ'na:l]. Das stimmhafte [ʒ] ist in intervokalischer Position in der deutschen Standardvarietät zu finden beispielsweise *Regime* [ʀe'ʒi:m].

[ʃ]

Das [ʃ] wird an-, in- und auslautend durch das Graphem <sch>, anlautend durch das Graphem <s> (<sp>, <st>) bezeichnet, wie beispielsweise in *Schal*, *Sparte* und *Strand*. Die Stimmbänder vibrieren nicht, und so entsteht kein Stimmklang im Kehlkopf. Das Velum sperrt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: schmale Engebildung im zentralen Bereich des vorderen Mundraumes, durch die die Phonationsluft gepresst wird.

Die Artikulationstelle ist im Bereich des hinteren Streifens vom Zahndamm und des vorderen Streifens vom Zungenrücken anzusetzen. Die Zungenseiten berühren die obere Zahnreihe. [ʃ] ist ein stimmloser postalveolarer Frikativ.

[ç]

Das [ç] wird durch die Grapheme <ch> bzw. <g> bezeichnet und Bauelement des Lautkörpers im In- und Auslaut nach palatalen Vokalen sowie im Anlaut vor palatalen Vokalen. In diesen Fällen und beim Suffix <-chen> wird es durch das Graphem <ch> bezeichnet. Das Graphem <g> im Suffix <-ig> hat die Realisierung [ç], wenn es im Auslaut steht, oder im Silbenauslaut, wenn

darauf ein Konsonant folgt und in der nächsten Silbe kein „ch“ steht, wie beispielsweise in König. Die Stimmbänder vibrieren nicht, es entsteht im Kehlkopf kein Stimmklang. Das Velum sperrt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: Bildung einer schmalen Enge im zentralen Trakt des mittleren Mundraumes, durch die die Phonationsluft gepresst wird.

Artikulationsstelle: die mittleren Streifen des harten Gaumens bzw. des Zungentrückens bilden eine Konstellation. Die Zungenseiten berühren die obere Zahnreihe. [ç] ist ein stimmloser palataler Frikativ.

[ç] im Anlaut (nur in eingedeutschten Wörtern und vor allem in nördlichen Gebieten des deutschen Sprachraumes. In südlichen Gebieten steht auf Grund der regionalen Varietäten auch in der Standardaussprache an dieser Stelle ein [k], das meistens als fakultative Aussprachevariante anerkannt wird. In der mitteldeutschen Varietät findet man an dieser Stelle ein [ʃ]):

Chirurg, China, Chemie, Chitin, Chinin

Bei dem Suffix „-ig“ ist wieder die regionale Variante [ik] statt des [iç] zugelassen: *wichtig, wuchtig, ewig, süchtig, witzig, flüchtig, tüchtig*

Da der entsprechende Lauttyp im Ungarischen selten ist, wird an dieser Stelle auf eine Gegenüberstellung verzichtet, mit dem Hinweis, dass bei Deutschlernenden mit Ungarisch als Erstsprache der Lauttyp auf Grundlage des vorhandenen [j] eingeführt werden kann, ergänzt durch Erläuterungen, die einerseits die artikulatorischen Unterschiede bei den „j“-Typen in beiden Sprachen darstellen, andererseits die stimmlose Artikulation den Lernern näher bringen. Als

ungarische analoge Beispiele für die Realisierung gelten etwa: *lépj, kapj*, wo „j“ stimmlos gesprochen wird.

[x]

Das [x] wird in deutschen Wörtern durch das Graphem <ch> in- und auslautend nach velaren Vokalen bezeichnet, des Weiteren ist es noch beim graphischen Zeichen <ch> inlautend in Wörtern fremden Ursprungs Bauelement des Lautkörpers. Die Stimmbänder schwingen nicht, im Kehlkopf wird kein Stimmklang produziert. Das Gaumensegel schließt den Weg in die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: es wird im zentralen Trakt des hinteren Mundraumes eine schmale Enge gebildet, durch die die Phonationsluft gepresst wird.

Artikulationstelle: sie ist im Bereich des Gaumensegels und des hinteren Streifens des Zungenrückens anzusetzen. Die Zungenseiten berühren die obere Zahnreihe leicht, spielen aber keine entscheidende Rolle bei der Artikulation. [x] ist ein stimmloser velarer Frikativ.

Laut KOHLER (1999: 88) kommt [x] nur nach den geschlossenen und halbgeschlossenen Vokalen [u:], [ʊ] und [o:] vor, wie beispielsweise in *Buch, Geruch* und *hoch*. Ein [χ] folgt den halboffenen bzw. offenen Vokalen [ɔ] und [a] z.B. *doch* und *Bach*. [χ] ist ein stimmloser, uvularer Frikativ.

Im Ungarischen kommt der Lauttyp sehr selten vor und auch der Phonationsstrom ist weniger intensiv als beim deutschen Lauttyp, z.B. *doh* [dox], *potroh* [ˈpotrox].

[h]

Das [h] wird durch das Graphem <h> im Wort- und Silbenanlaut bezeichnet, wie beispielsweise in *Hut*. Es wird im Kehlkopf kein Stimmklang produziert, das Gaumensegel schließt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: im Kehlkopf passiert die Phonationsluft mit hörbaren Geräuschen die Stimmritze.

Artikulationsstelle: Kehlkopf und Stimmbänder. Die Zungenseiten spielen keine aktive Rolle im artikulatorischen Vorgang. [h] ist ein stimmloser glottaler Frikativ.

5.8.2.2. Approximanten

[j]

Das [j] wird in deutschen Wörtern im An- und Inlaut durch das Graphem <j>, in Wörtern fremden Ursprungs durch das Graphem <y> bezeichnet, wie beispielsweise in *jetzt* und *Yacht*. Die Stimmbänder schwingen mit, es entsteht ein

Stimmklang im Kehlkopf. Das Gaumensegel schließt den Weg zur Nasenhöhle, die orale Bildung dominiert.

Artikulationsmodus: Bildung einer schmalen Enge im zentralen Trakt des mittleren Abschnitts des Mundraumes, durch die die Phonationsluft gepresst wird.

Artikulationstelle: Konstellation zwischen dem mittleren Streifen des harten Gaumens und dem mittleren Streifen des Zungenrückens. Die Zungenseiten stehen in Berührung mit der oberen Zahnreihe, aber sie spielen keine aktive Rolle im artikulatorischen Vorgang. [j] ist ein stimmhafter palataler Approximant.

5.8.2.3. Laterale

[l]

Das [l] wird durch die Grapheme <l>, <ll> bezeichnet, wie beispielsweise in *Land* und *Welle*. Es ist an- in- und auslautend ein Bauelement des Lautkörpers des Sprachzeichens. Die Stimmbänder vibrieren, und es entsteht im Kehlkopf ein Stimmklang. Das Velum sperrt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: im zentralen Trakt des vorderen Mundraumes wird ein Verschluss gebildet, und die Phonationsluft entweicht beidseitig entlang der Zungenseiten.

Die Artikulationsstelle ist der Bereich zwischen der hinteren Zahnfläche der oberen Zahnreihe und dem Zahndamm bzw. beim Verschlusselement die Zungenspitze. Das Engeelement zeichnet sich durch die Beteiligung der Zungen-seiten aus, die in der Artikulation eine besonders aktive Rolle spielen. [l] ist ein stimmhafter alveolarer Lateral.

5.8.3. Affrikaten

Die Affrikaten besitzen die Gemeinsamkeit, dass sie die Verbindung eines Plosivselements mit einem nachfolgenden, sog. homorganen (griechisch *homo* + *organ*, dt. etwa „dasselbe Organ“) Frikativelement repräsentieren. Der Ausdruck „homorgan“ ist so zu verstehen, dass die beiden Elemente in einem nahe liegenden Bereich des Ansatzrohrs ihre pezifische Artikulationsstelle haben. Die Affrikaten sind physiologisch gesehen Einzellaute. Die Quantität der beiden Phasen einer Affrikate also einer Verschluss- und einer Frikativphase machen zusammen nicht mehr aus als die Dauer eines Einzellauts (NEPPERT & PÉTURSSON 1986). Das Quantitätsverhältnis der erwähnten zwei Phasen ist für den Hörer bei der Perzeption relevant, da er auch dann einen Frikativ perzipiert, wenn nach der Verschlusslösung die Quantität der Frikativphase verlängert wird. Auf der Grundlage dieser Annahmen werden hier die Affrikaten als selbständiger Artikulationsmodus und Konsonantentypen aufgefasst.

[pʃ]

[pʃ] ist an-, in- und auslautendes Bauelement des Lautkörpers des Sprachzeichens. Es wird durch das Graphem <pʃ> bezeichnet und besteht aus Elementen

von [p] und [f], wie beispielsweise in *Pfeil*. Die Stimmbänder schwingen nicht mit, im Kehlkopf wird kein Stimmklang gebildet. Das Velum sperrt den Weg in die Nasenhöhle ab, so dass die orale Bildung dominiert.

Artikulationsmodus: Bildung eines Verschlusses im vorderen Bereich des Mundraumes, der in ein schmales Engeelement übergeht.

Artikulationsstelle: sie ist in der Verschlussphase im Bereich der Lippen, in der Engephase im Bereich der unteren Kante der oberen Zähne bzw. der Unterlippe anzusetzen. Die Zungenseiten spielen keine aktive Rolle bei der Artikulation. [pf̥] ist eine stimmlose bilabial-labiodentale Affrikate.

[ts̥]

[ts̥] ist eine aus Elementen von [t] und [s] gebildete Affrikate. Es wird im Anlaut durch die Grapheme <z>, <c>, im Inlaut durch die Grapheme <z>, <zz>, <tz>, <c>, <tv>, im Auslaut durch die Grapheme <z>, <tz> bezeichnet, wie beispielsweise in *Zeit*, *Ritze*, *Skizze*, *Cäsar* und *Information*. Die Stimmbänder vibrieren nicht, im Kehlkopf wird kein Stimmklang produziert. Das Gaumensegel schließt die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: Verschlussbildung im mittleren Trakt des vorderen Mundraumes, Lösung des Verschlusses und ein Übergehen in ein schmales Engeelement.

Artikulationsstelle: im Bereich der hinteren Seite der oberen Zähne und des Zahndamms bzw. des Zungensaums und des vorderen Streifens vom Zungen-

rücken. Die Zungenseiten berühren die obere Zahnreihe. [tʃ] ist eine stimmlose alveolare Affrikate.

[tʃ]

[tʃ] ist eine aus Elementen von [t] und [ʃ] gebildete Affrikate, die an-, in- und auslautend ein Bauelement des Lautkörpers ist. Es wird durch die Grapheme <tsch> und in Fremdwörtern durch <c> und <ch> bezeichnet, wie beispielsweise in *tschüß*, *ciao* und *Sandwich*. Die Stimmbänder schwingen nicht mit, im Kehlkopf entsteht kein Stimmklang. Das Gaumensegel sperrt den Weg der Phonationsluft in die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: es entsteht ein Verschluss im mittleren Trakt des vorderen Mundraumes, er wird gelöst und es erfolgt ein Übergang in ein schmales Engeelement.

Artikulationsstelle: sie ist durch die Konstellation des hinteren Streifens vom Zahndamm und des vorderen Streifens vom Zungenrücken zu charakterisieren. Die Zungenseiten stehen in Berührung mit der oberen Zahnreihe. [tʃ] ist eine stimmlose postalveolare Affrikate.

[dʒ]

[dʒ] ist eine aus Elementen von [d] und [ʒ] gebildete Affrikate, die an- und inlautend ein Bauelement des Lautkörpers ist. Es wird durch die Grapheme <dsch>, <j> und <gi> bezeichnet, wie beispielsweise in *Dschungel*, *Job* und *adagio*. Die Stimmbänder schwingen mit, im Kehlkopf entsteht ein Stimm-

klung. Das Gaumensegel sperrt den Weg der Phonationsluft in die Nasenhöhle ab, die orale Bildung dominiert.

Artikulationsmodus: es entsteht ein Verschluss im mittleren Trakt des vorderen Mundraumes, er wird gelöst und es erfolgt ein Übergang in ein schmales Engeelement.

Artikulationsstelle: sie ist durch die Konstellation des hinteren Streifens vom Zahndamm und des vorderen Streifens vom Zungenrücken zu charakterisieren. Die Zungenseiten stehen in Berührung mit der oberen Zahnreihe. [d̥ʒ] ist eine stimmhafte postalveolare Affrikate.

Artikulationsmodus			Artikulationsstelle									
			Labial	labiodental	alveolar	postalveolar	palatal	velar	uvular	glottal		
nasal		Nasal	stimmhaft	[m]	[ɱ]	[n]				[ŋ]		
oral	Verschluss	Plosiv	stimmhaft	[b]		[d]				[g]		
			stimmlos	[p]		[t]			[k]		[ʔ]	
		Vibrant			[r]					[ʀ]		
	Enge	Lateral	stimmhaft			[l]						
		Approximant						[j]				
		Frikativ			[v]	[z]	[ʒ]			[ʁ]		
	Verschluss+ Enge	Affrikate	stimmlos		[f]	[s]	[ʃ]	[ç]	[x]	[χ]	[h]	
			stimmhaft		[pf]	[ts]	[tʃ]					
						[d̥ʒ]						

Abbildung 8: Die konsonantischen Lauttypen der deutschen Standardvarietät nach artikulatorischen Kriterien dargestellt

5. 9. Die Koartikulation und die Assimilation

Der Sprechschall, als kontinuierliches Trägersignal von Sprachinhalten stellt ein lautliches Geflecht dar. Im Redefluss erscheint dieses lautliche Geflecht durch Pausen unterbrochen und innerlich vielfach fein strukturiert. Bei der Erzeugung dieses Trägersignals wird die jeweilige Sprechereinheit zentral entworfen und gesteuert, „sie entsteht daher aus einer flüssig ablaufenden Dauerbewegung, in der Atmung, Stimmgebung und Artikulation komplex verbunden sind. Dieser kontinuierliche Bewegungsablauf wird als *Koartikulation* oder Synkinese bezeichnet.“ (*Großes Wörterbuch der deutschen Aussprache* 1982: 69)

Der Begriff Koartikulation (lat. con + articulare) soll die eine Form der Ausgleichsprozesse im Redefluss darstellen, die zitierte Definition aus dem *Großen Wörterbuch* bezeichnet sogar die gesamte Erscheinung der Angleichungen und Beeinflussungen im Redefluss als Koartikulation. Diese Auffassung wird von KOHLER (1995: 202) dadurch ergänzt, dass Koartikulation ein „Prinzip wechselnder Koordination artikulatorischer Parameter“ sei. „Beim Zusammenwirken von Organen unterschiedlicher Artikulationspräzision kommt es zu zeitlichen Verschiebungen im gegenseitigen Bewegungseinsatz, d.h. entweder zu einem Beharren oder zu einer Vorwegnahme in einem Artikulationsparameter.“ (KOHLER, 1995: 202) Die Koartikulation soll auf Grund der angeführten Definitionen als der Typ der gegenseitigen Beeinflussung der Laute im Redefluss gelten, der in Sprechereinheiten aus artikulations-dynamischen Gründen abläuft. Einzelne Bildungsmerkmale und Bildungsphasen benachbarter Laute werden mitartikuliert (z.B. ist die Einstellung der Lippen bei der Bildung von [l] unterschiedlich in den Wörtern *leben*, *loben* oder *Lux*). Nach der Auffassung von LINDNER (1981: 249) ist z.B. „die Stelle der Verschlussbildung...

bei [g, k] stark koartikulatorisch von der *Lautumgebung* abhängig.“ Andere koartikulatorisch bedingte Veränderungen wie z.B. der Stimmtonverlust bei Verschlusslauten u.Ä. wurden bei den allgemeinen Ausspracheregeln angeführt. Weitere Beispiele wie koartikulatorisch bedingte Reduktionen oder die sog. „schwachen Formen“ sollen im Folgenden ebenfalls erwähnt werden.

Es sei an dieser Stelle zu bemerken, dass diese Veränderungen nur dann Veränderungen darstellen, wenn ihnen gewisse Ausgangsformen zugrunde liegen, egal ob diese Ausgangsformen abstrakte Wortgestalten oder kodifizierte Einzelwortdarstellungen aus phonetischer Sicht sind. Es muss ferner darauf hingewiesen werden, dass noch immer Überschneidungen mit dem anderen Typ, der Assimilation vorhanden sind. Die Konstatierung dieser beiden Typen allein reicht dabei nicht aus, alle artikulatorische Parameterschwankungen und Reduktionen zu erfassen.

Die Assimilation (lat. ad + similis) ist nach WÄNGLER (1976: 171) „eine notwendige Folge des Sprechvorgangs: Kein Sprechen ohne Assimilation.“ Unter Assimilation verstehen wir den Typ der gegenseitigen Beeinflussung von Lauten, bei dem eine „Angleichung benachbarter Segmente in mindestens einem phonetischen Parameter“ (KOHLENER 1995: 205) geschieht. Wenn wir die in Kapitel 3.2.1. angeführten Beurteilungskriterien der Konsonanten als Ausgangspunkt nehmen, können die Assimilationen in vier wichtige Gruppen eingeteilt werden:

- 1) Assimilation auf Grund der Artikulationsstelle, 2) Assimilation auf Grund des Artikulationsmodus, 3) Assimilation auf Grund der Stimmbeteiligung, 4) Assimilation auf Grund der Nasalität.

Nach der Richtung der Assimilationswirkung unterscheiden wir zwischen progressiver und regressiver Form. Progressiv ist die Wirkung der Assimilati-

on, wenn der nachfolgende Laut vom vorhergehenden beeinflusst wird, z.B. *loben* [lo:bm̩]. Regressiv ist die Wirkung der Assimilation, wenn der vorhergehende Laut vom nachfolgenden beeinflusst wird, z.B. *lobst* [lo:pst].

1) Eine Assimilation nach der Artikulationsstelle geschieht oft z.B. beim Aufeinandertreffen nasaler und oraler Verschlusslaute. Eine progressive Assimilation dieser Lauttypen ist in Fällen zu finden, in denen sie entweder infolge von Tilgungen des Murmelvokals zusammentreffen, oder wenn die ursprünglich oralen Verschlusslaute nicht oral gelöst werden: *haben* [ha:bm̩], *Rappen* [rapm̩], *hängen* [hɛŋŋ]. Eine regressive Assimilation der Lauttypen kann z.B. an Wort- und Morphemgrenzen eintreten wie in *anbeten* [ˈambe:t̩] oder *kann man* [kamman].

2) Assimilation auf Grund des Artikulationsmodus kann z.B. beim Zusammentreffen oraler Verschlusslaute geschehen, wo die Haltephasen der Laute u.U. zusammenschmelzen können: *lebte* [ˈle:pt̩], *bebtte* [ˈbe:pt̩].

3) Assimilation auf Grund der Stimmbeteiligung kommt z.B. beim Aufeinandertreffen stimmhafter Plosive und dem /z/ mit stimmlosen Plosiven oder Frikativen vor. Die Stimmhaftigkeit geht in solchen Fällen verloren: *essbar* [esb̥a:], *ratsam* [rat̥z̥am].

4) Assimilation auf Grund der Nasalität kann regressiv geschehen, wenn nach stimmhaften Plosiven Nasallaute stehen, innerhalb eines Wortes wie in *werden* [ve:n̩], *legen* [le:n̩]. Progressiv kann die Assimilation wirken, wenn nach einem Nasallaut ein Plosiv folgt: *zum Beispiel* [t̩summa̯ʃpi:l̩].

6. Sprechakustik

6.1. Akustische Grundlagen

Die Bewegungen der Sprechwerkzeuge setzen in geringerem Maße die benachbarten Körpergewebe, im Vergleich dazu in wesentlich stärkerem Maße das sich im Sprechapparat befindende Gasgemisch, die Luft, in Schwingungen. Die Schwingungen entstehen grundsätzlich durch die Wechselbewegungen von Körpern und/oder Molekülen um eine potentielle Ruhelage. Diese Wechselbewegungen erzeugen um sich herum Druckschwankungen. Aus der bei der Artikulation durchgeführten schallerzeugenden Tätigkeit der Sprechwerkzeuge resultieren ebenfalls Druckschwankungen in dem diese umgebenden oder an sie angrenzenden Gasgemisch, in der Luft. Diese Druckschwankungen werden als Schallschwingungen in die Atmosphäre abgestrahlt und erreichen als Schallwellen das Gehörorgan. Der Schall ist für Menschen ein hörbares Ereignis aber physikalisch gesehen ist er eine messbare Welle, die zu ihrer Ausbreitung Zeit und ein schallleitendes Medium wie beispielsweise die Luft benötigt. Der Grund dafür liegt darin, dass im Schall Schwingungen existieren können, die von Menschen nicht mehr wahrgenommen werden und deshalb nur mit speziellen Geräten messbar sind. Der Schall ist der Bereich, der in die so genannte Hörfläche fällt. Die Hörschwelle und die Schmerzschwelle bilden die Grenzen der Hörfläche. Der Schall, der in Gasen (beispielsweise Luft) und Flüssigkeiten immer eine Longitudinalwelle ist, breitet sich von der Schallquelle in alle Richtungen hin aus. Bei Longitudinalwellen schwingen die Elemente des Mediums (in unserem Fall die Luftmoleküle) in die Richtung, in die sich die Schallwellen ausbreiten. Schallwellen sind Schwingungen.

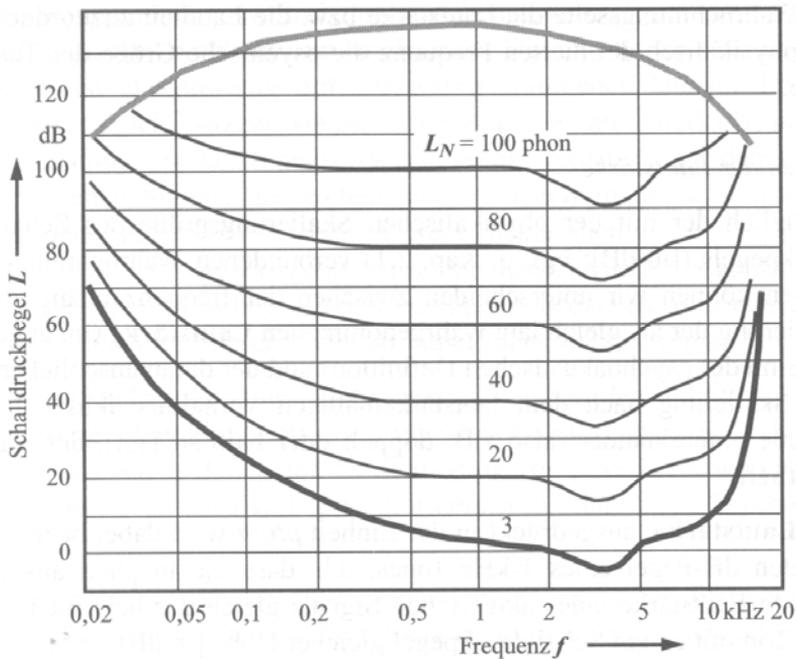


Abbildung 9: Die Hörfäche oben mit der Schmerzschwelle unten mit der Hörschwelle (POMPINO-MARSCHALL 2003: 154)

Der Sprechschall entsteht dadurch, dass die Bewegungen der Sprechwerkzeuge die Luft in Schwingung bringen. Dies erzeugt Druckschwankungen, die in die Luft abgestrahlt werden. Die Luftmoleküle bewegen sich im Ansatzrohr auf parallelen Bahnen, die das Ergebnis des laminaren Luftstroms sind. Dabei entsteht kein Schall. Wenn aber das Ansatzrohr verengt wird wie bei Frikativen, steigt die Luftstromgeschwindigkeit im Bereich des Hindernisses an. Dann schlägt der laminare Luftstrom in einen chaotisch-turbulenten um. Dieser Luftstrom erzeugt an der Stelle des Hindernisses einen Schall. Die Schallgeschwindigkeit ist von der Lufttemperatur abhängig, bei 20°C bewegt sich der Schall mit einer Geschwindigkeit von 340 m/s.

Das Sprechsignal ist ein Träger kodierter Informationen und ein physikalischer Prozess, der sich entsprechend des Inhalts der zu übermittelnden Information zeitlich verändert.

Nach der Klassifizierung der Akustik kann der Schall, was seine Zusammensetzung anbelangt, entweder **einfach** oder **komplex** sein. In der Natur ist nur komplexer Schall zu finden, so wie der Sprechlaut. Unter komplexem Schall ist zu verstehen, dass gleichzeitig nicht nur der Grundton sondern auch Obertöne (auch als: Harmonische) anwesend sind. Heutzutage ist es möglich mit modernen Geräten einfachen Schall, der keine Obertöne hat, herzustellen. Die Obertöne sind immer ganzzahlige Vielfache des Grundtons. Wenn ein Grundton 440 Hz ist, dann hat er die folgenden Obertöne: 880 Hz, 1320 Hz, 1760 Hz usw. Der Grundton 16,35 Hz besitzt die Obertöne: 32,7 Hz, 49,05 Hz, 65,4 Hz usw. Die folgende Feststellung kann an Hand der aufgezählten Beispiele gemacht werden: Je tiefer ein Ton ist, desto größer ist die Dichte der Obertöne. Dieser Sachverhalt wird auch als Reziprozitätsgesetz bezeichnet. Nach dem zeitlichen Verlauf kann der Schall entweder **periodisch** oder **aperiodisch** sein. Bei periodischem Schall sind sich exakt wiederholende Vorgänge in seinem zeitlichen Verlauf zu finden beispielsweise der Kammerton 440 Hz. Bei aperiodischem Schall kann zeitlich gesehen keine Periodizität in den Schwingungen festgestellt werden beispielsweise der Glockenton, das Rauschen von Frikativen und die Verschlusslösung der Plosive. Im aperiodischen Schall sind Töne vorhanden, die nicht die ganzzahligen Vielfachen der Grundfrequenz sind. Die Sprechlaute sind weder periodisch noch aperiodisch sondern **quasiperiodisch** d.h. annähernd periodisch.

Aus dem Sinuston, der ein einfacher, periodischer Ton ist, lassen sich alle anderen Schallformen ableiten, er ist die Basis vieler Theorien in der Akustik. Da die akustische Analyse des Sinustons Sinuswellen zeigt, besitzt er diesen

Namen. Wenn eine Sinuswelle 360° (180° über und 180° unter der horizontalen Achse) beschreibt, hat sie eine **Periode** zurückgelegt. Beim Sinuston ist eine Periode mit einer Schwingung identisch. Die Periodendauer gibt an, wie lange eine vollständige Schwingung dauert. Je öfter sich die Perioden während einer bestimmten Zeit wiederholen, desto höher ist der Ton. Die Anzahl der Perioden innerhalb eines bestimmten Zeitraums wird als **Frequenz** bezeichnet, deren physikalische Messeinheit das **Hertz** (Hz) ist. Es ist nach dem deutschen Physiker Heinrich Rudolf Hertz benannt. Die physikalische Formel für die Frequenz lautet: $f=n/t$, wobei n für die Schwingungszahl und t für die Zeit steht. Für Menschen sind nur die Töne wahrnehmbar, die zwischen 16 und 20000 Hz liegen. Der Schall über 20000 Hz heißt Ultraschall. Unter 16 Hz wird der Schall als Infraschall bezeichnet. Zwischen diesen Grenzwerten befindet sich das für das menschliche Gehörorgan im Normalfall zur Verarbeitung wahrnehmbare hörbare Frequenzband, dessen Breite allerdings auch von persönlichen Gegebenheiten (Lebensalter usw.) beeinflusst wird.

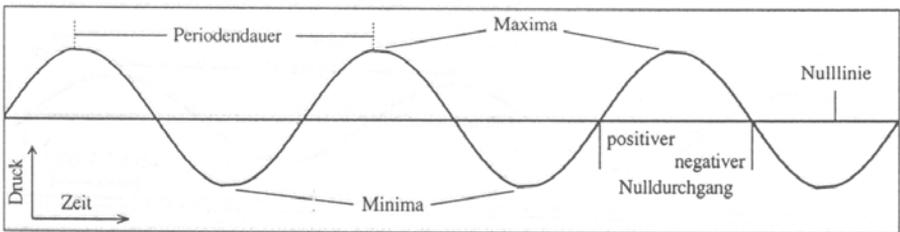


Abbildung 10: Ein einfacher Ton (Sinuston) mit seiner Periodendauer (REETZ 1999: 13)

Die Sinuswellen besitzen auch eine vertikale Bewegung, deren höchste Auslenkung von der horizontalen Achse als **Amplitude** bezeichnet wird. Je größer die Amplitude ist, umso lauter ist der Schall. Mit der Amplitude korreliert die **Lautstärke**, die mit der physikalischen Messeinheit **Dezibel** (dB) angegeben

wird. 130 dB ist die Lautstärke, bei der die Schmerzgrenze für Menschen erreicht wird. Die Lautstärke des Schalls, die von Menschen gerade so wahrgenommen wird, ist die Hörschwelle und liegt bei 4 dB. Sowohl die Schmerzgrenze als auch die Hörschwelle sind frequenzabhängig.

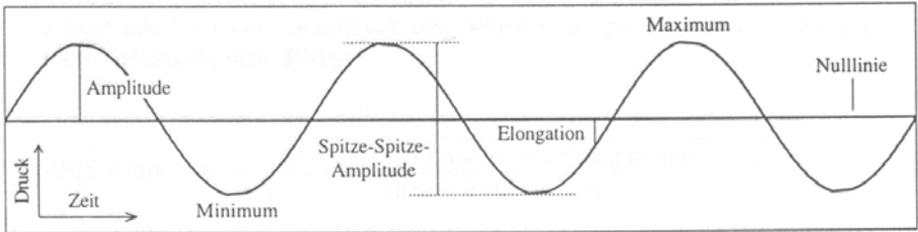


Abbildung 11: Ein einfacher Ton (Sinuston) mit seiner Amplitude und Elongation in einem beliebigen Punkt (REETZ 1999: 19)

Die sprechschallerzeugende artikulatorische Tätigkeit der Sprechwerkzeuge zeichnet sich notgedrungen durch kinetische Stereotypien, d.h. durch sich im Sprechapparat ständig wiederholende, auf Grund ihrer Ähnlichkeit als typisch einstufbare Bewegungskomplexe aus. Diese typischen Organbewegungen erzeugen begrifflicherweise entsprechende typische Schallschwingungen, d.h. Schwingungsstereotypien, die in ihrer Frequenz-, Intensitäts- und Zeitstruktur so gestaltet und gegliedert sind, dass sie sich gut dazu eignen, als Trägersignale der darin kodierte Sprachinhalte (der sprachlichen Informationen) zu dienen.

Als Sprechschall setzt sich dieses Trägersignal aus vier Grundschallformen zusammen.

6.1.1. Die sogenannte „stumme Schallform“

Der Sprechlaut versteht sich als ein hör- und registrierbares Phänomen; ist aber kein Nutzschall (= Der Schall, der vom Perzipienten in einer gegebenen Situation als Element einer sprachlichen Information dekodiert werden soll.) zugehen, wird über stumme Schallform gesprochen. Auf diese Erscheinung, die scheinbar einen Widerspruch in sich birgt, wird nun eingegangen. VÉRTES (1982: 155) behauptet in Anlehnung an eine Erkenntnis von FANT (1960: 218), dass der Sprechschall immer seine artikulatorische Grundlage habe, doch nicht umgekehrt: Die Artikulation solle sich nicht immer in einem hörbaren Schallprodukt manifestieren. Die Relevanz dieser Feststellung lässt sich leicht überprüfen: In der stummen Haltephase des Verschlusses (durch die Hebung des Gaumensegels ist die Öffnung zum Nasenraum versperrt) bei der Bildung des [p] werden die äußeren Lippenkanten zusammengepresst, was offenbar nicht außerhalb der Artikulation erfolgt. Die stumme Schallform ist hier „(...) die akustische Äquivalenz der Haltephase des Verschlusses von Plosivlauten, (...) [die] sich durch einen Nullwert der Frequenz und einen Nullpegel der Intensität, zugleich aber durch eine (...) Zeitdauer (...) [auszeichnet] (VALACZKAI & BRENNER 1999: 104).“

6.1.2. Der Explosionsschall

Infolge eines durch kortikale Steuerung verursachten Druckanstiegs wird der in der Glottis oder aber im Ansatzrohr gebildete luftdichte Verschluss gelöst und die bisher angestaute Luft strömt explosionsartig hervor. Typisch ist dieser Effekt für Plosive, seiner Kürze wegen wird dieser Transiente genannt, und ist als Übergangsstelle zu den stabilen Formantstrukturen (meist des nachfolgen-

den Vokals) zu betrachten. In einigen Fällen kommt es nach der Verschlusslösung zu einer Behauchung von kurzer Zeitdauer, diese wird Aspiration genannt. Die Affrikaten dagegen werden durch die Verschlusslösung in einen wahren Frikativ gebildet, z.B. in *Napf* [napf̥].

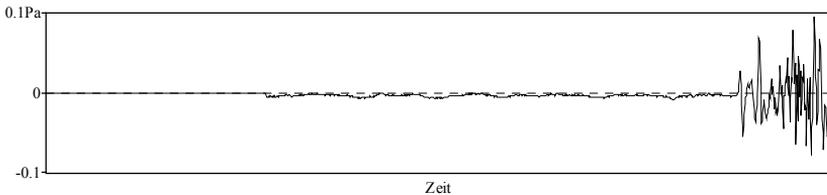


Abbildung 12: Oszillogramm eines Stummschalls (links) und Explosionsschalls (rechts)

6.1.3. Der Frik(a)tionsschall

Im Ansatzrohr herrscht usuell eine laminare Luftströmung (parallele Bewegung von Luftmolekülen), dabei entsteht kein Schall. Verengt sich doch die Passage, steigt die Strömungsgeschwindigkeit im Hindernisbereich. Sollte diese den kritischen Wert überschreiten, entsteht eine chaotisch-turbulente Verwirbelung, die in der Verengung ein Schallsignal erzeugt. Wegen seiner Aperiodizität im Zeitbereich wird dieses Signal auch Rauschen genannt.

Das Ansatzrohr kann sich an mehreren Stellen verengen. Die Lage dieser Verengung bestimmt die geometrische Konfiguration des Resonanzraumes (Filter), Laut TILLMANN (1980), hier nach TILLMANN & SCHIEL 1995: 4) ist die Stärke der Verformung des Rauschsignals der Quelle durch die Hinterverengung (d.h. die sog. artikulatorische Tiefe), im Ansatzrohr beeinflusst. Der

hinter der Engstelle befindliche Resonanzraum übt kaum Einfluss auf das Sprechsignal aus (vgl. TILLMANN & SCHIEL 1995: 4).

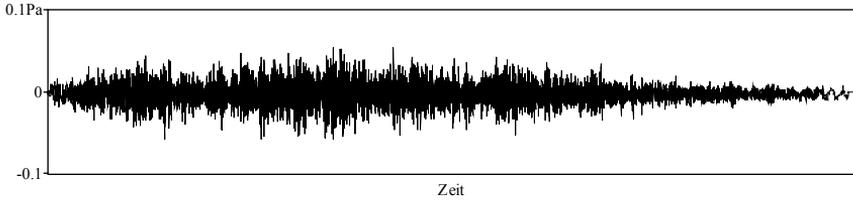


Abbildung 13: Oszillogramm eines Frik(a)tionsschalls

6.1.4. Der (quasi)periodische Schall, der Klang

Durch die nahezu sinusförmigen Schwingungen der Glottis werden periodische Druckänderungen in der sich im Ansatzrohr befindenden Luft erzeugt. Diese (quasi)periodischen Schwingungen der Stimmlippen bilden die Schallgrundlage für Vokale, Nasale und dienen als Zusatz für die stimmhaften oralen Konsonanten. (Auch in dem Explosionsschall, respektive in einigen Teilen des Frikationsschalles ist die Periodizität vorzufinden.)

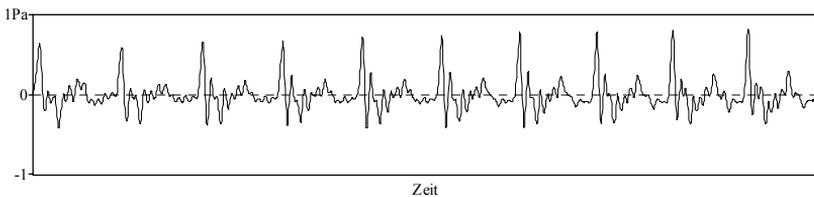


Abbildung 14: Oszillogramm eines (quasi)periodischer Schalls

Die annähernd sinusförmigen periodischen Schwingungen der Stimmlippen erzeugen ebenfalls annähernd sinusförmige periodische Druckänderungen in der die Stimmlippen umgebenden Luft. Diese quasi-periodischen Schwingungen der Stimmlippen, die üblicherweise einfachheitshalber als periodisch angesehen werden, erzeugen die Schallgrundlage der Vokale, der Nasale, und neben dem Frikationsschall und dem Explosionsschall auch die zusätzliche Schallgrundlage der stimmhaften oralen Konsonanten. Die Periodenfrequenz des durch solche Schwingungen der auch als Knallgenerator bezeichneten Stimmbänder erzeugten **Stimmklanges** ist mit der Zyklusfrequenz der Öffnungs- und der Schließungsvorgänge der Stimmlippen identisch. Diese erste, noch unmodulierte harmonische Schwingung heißt wegen ihres Quellencharakters **Grundfrequenz**. Die Schwingungszahl dieser Grundfrequenz ist gleich der Periodenfrequenz des durch die Stimmlippen erzeugten Quellensignals. Der durchschnittliche Wert der Periodenfrequenz in Bezug auf die Grundfrequenz des männlichen Stimmklanges liegt etwa zwischen 100 bis 150 Hz, die des weiblichen Stimmklanges zwischen 190 und 250 Hz, die der Kinderstimme zwischen 350 bis 500 Hz. Die Grundfrequenz hängt von der Länge und der Spannung der Stimmbänder ab. Da bei Männern die Stimmbänder anatomisch gesehen länger sind, produzieren männliche Sprachbenutzer eine tiefere Grundfrequenz.

Der **Stimmklang**, in der früheren phonetischen Literatur „Stimmton“ (ung. „zöngé“) genannt, wird als Primärschall im Ansatzrohr moduliert: gewisse Teilschwingungen des Sprechschalls werden verstärkt, andere dagegen gedämpft. Im Ergebnis solcher Modulationen entstehen die so genannten **Formanten**, d.h. die im zusammengesetzten bzw. kombinierten Schall der entsprechenden Segmente dominierenden Teilschwingungsbereiche (Frequenzbänder) die die betreffenden Segmente in ihrer Eigenart „formieren“, prägen. Die erste, noch unmodulierte harmonische Schwingung (die Grundfrequenz) wird als F_0 ,

die erste, bereits modulierte Teilschwingung F1, die zweite F2, die dritte F3, die vierte F4 ... F_n bezeichnet. Bei der Lokalisierung der Artikulationsstelle spielt F2 eine große Rolle. Problematisch ist aber, dass er nicht immer sichtbar ist. Dies gilt z.B. für stimmlose Frikative und Plosive. Unter **Lokus** wird die angesteuerte Zielfrequenz von F2 zum nachfolgenden Segment verstanden. Der Frequenzwert der Formanten wird in Hz, der Wert der Amplitude, die Maximalauslenkung der Schwingung, d.h. die maximale Abweichung von der potentiellen Ruhelage, die mit der auditiven Lautstärke korrespondiert, mit Hilfe einer dB-Schallpegel-Skala angegeben. Die unterschiedlichen Formantwerte der männlichen und weiblichen Sprachbenutzer sind durch die Länge des Ansatzrohres bedingt. Die weiteren Formanten so wie F4, F5 sind individuell verschieden und für die jeweilige Einzelperson charakteristisch.

Wenn es der Wissenschaft bisher auch nicht gelungen ist, die genaue Äquivalenz zwischen den Bewegungstypen und -phasen der schallerzeugenden Organtätigkeit und den einzelnen Komponenten und Segmenten im Schallprodukt nachzuweisen, so ist es doch möglich, gewisse Zusammenhänge zwischen dem Aufbau des Ansatzrohres und den sich darin vollziehenden schallerzeugenden Vorgängen und den Veränderungen der Formantenfrequenzwerte festzustellen:

Aspekt 1): die senkrechte Bewegung der Zunge.

Der Wert der Mittelfrequenz des Vokalformanten F1 ist zur Höhe der Zungenhebung im Mundraum umgekehrt proportional: je höher der höchste Punkt des Zungenrückens im Mundraum, desto niedriger der Frequenzwert von F1 und umgekehrt: je tiefer die Zunge, desto höher dieser Frequenzwert.

Aspekt 2): die waagrechte Bewegung der Zunge.

Der Wert der Mittenfrequenz des zweiten Formanten (F2) der palatalen Vokale ist höher als der der velaren Vokale.

Aspekt 3): die senkrechte Bewegung des Kiefers.

Der Kieferwinkel ist (im Sinne des Aspektes 1) zum Frequenzwert des ersten Formanten (F1) umgekehrt proportional: je größer der Kieferwinkel, desto niedriger der Frequenzwert und umgekehrt.

Aspekt 4): die Lippentätigkeit.

Die zunehmende Verengung der Öffnung der Lippen und parallel dazu die zunehmende Vorstülpung der Lippen bewirkt niedrigere Frequenzwerte der Formanten. Im Vergleich zum ersten Formanten werden der 2. und der 3. Formant stärker durch die zunehmende Lippenrundung und -vorstülpung beeinflusst.

Aspekt 5): die physiologische Beschaffenheit des Ansatzrohres, vor allem seine Länge (Längeregel). Der Wert der Mittelfrequenz der Vokalformanten ist zur Länge des Ansatzrohres umgekehrt proportional: zum längeren Ansatzrohr gehören niedrigere Frequenzwerte, zum kürzeren höhere.

Auf Grund der genannten Zusammenhänge ist es nicht hilfreich, die Diphthonge nur physiologisch einzustufen, denn dabei wird gerade die umgekehrte Proportion zwischen den genannten Faktoren ausgeschaltet.

6.2. Korrelation zwischen artikulatorischen und akustischen Vorgängen

Damit der Schall visuell analysiert werden kann, muss er aufgezeichnet werden. Die zur Sprechschallanalyse benutzten Geräte sind der **Oszillograph** und der **Sonagraph**, der auch als Spektrograph bezeichnet wird. Wie viele technische Errungenschaften wurde der erste Sonagraph für die amerikanische Armee von der Firma Kay während des Zweiten Weltkriegs entwickelt. Ein Sonagraph ermöglicht die visuelle Darstellung oder Sichtbarmachung der Sprechsignale. Er arbeitet mit der so genannten Fourieranalyse, die das Zerteilen eines beliebigen Signals in Sinus- und Kosinusfunktionen vornimmt. Der Sonagraph zerlegt das Sprechsignal in seine spektralen Komponenten. Er stellt Sonagramme/Spektrogramme her, die dreidimensional sind. Die Zeit, die in Millisekunden (ms) angegeben wird, wird auf der horizontalen Achse dargestellt und muss von links nach rechts gelesen werden. Die von unten nach oben laufende Frequenz (gewöhnlich von 0-8000 Hz angegeben) wird auf der horizontalen Achse abgebildet. Die dritte Dimension ist der Schwärzungsgrad, der von der Amplitude (Intensität) der Schwingungen abhängt. Je lauter ein Ton ist, desto tiefer ist der Schwärzungsgrad. Auf Sonagrammen sind dementsprechend weiße und schwarze Flächen zu sehen. Die schwarzen waagerechten Balken bilden die **Formanten** ab, die durch eine bestimmte Bandbreite gekennzeichnet sind. Sie tragen zur Identifizierung der Segmente im Sprechsignal bei.

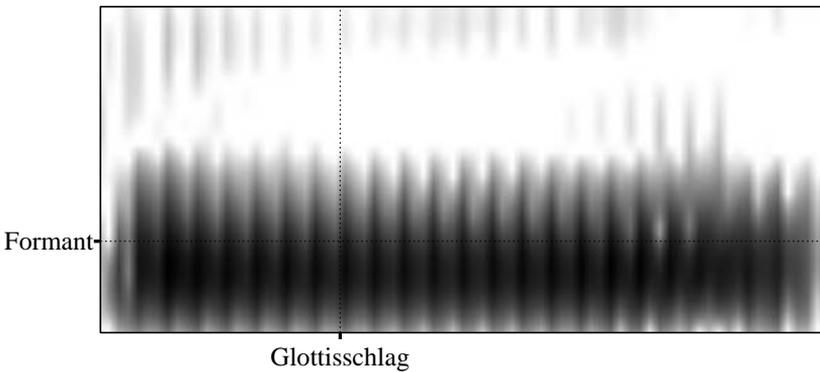
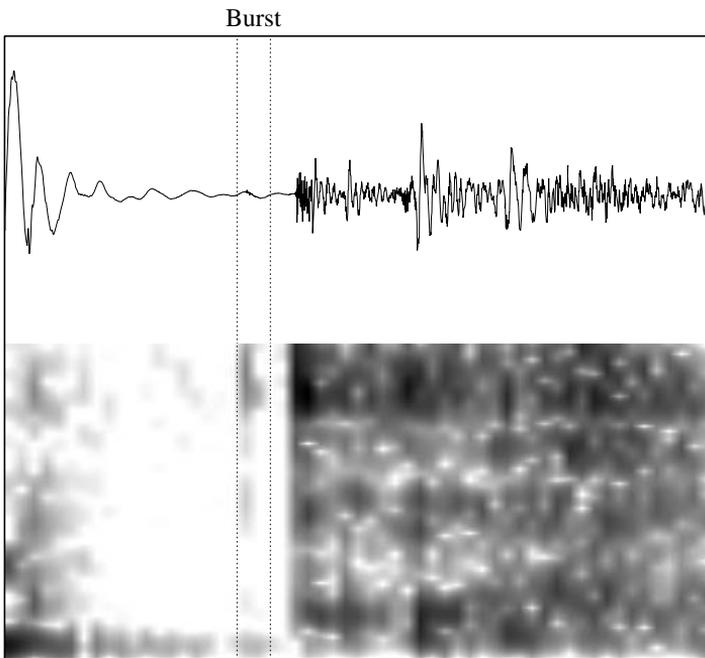


Abbildung 15: Der Formant eines Vokals mit seinen Glottisschlägen auf einem Breitbandsonagramm

Die Formanten der Sprechlaute sind von der Form und der Größe des Ansatzrohrs beeinflusst. Aus der Stellung des Ansatzrohrs kann auf den Sprechschall geschlossen werden, er kann daraus aber nicht eindeutig bestimmt werden. Die Formanten sind das Produkt der Schwingungen im Ansatzrohr. Es existieren aber sog. **Anti-Formanten**, die dann entstehen, wenn das Velum gesenkt ist und damit der Nasenraum als Resonanzraum an das Ansatzrohr gekoppelt ist beispielsweise im Falle der nasalierten Vokale. Durch Schwingungen im Nasenraum wird dem Schall im restlichen Ansatzrohr bei bestimmten Frequenzen Energie entzogen. Auf dem Spektrogramm zeigt sich dies in Form von Einbrüchen bei diesen Frequenzen.

Glottisschläge sind auf dem Sonagramm als vertikale Linien zu sehen, die über die Stimmhaftigkeit des Abschnittes Auskunft geben. Wenn sie fehlen, ist der Sprechlaut als stimmlos einzustufen, wobei bei Konsonanten die Einstufung bezüglich der Stimmbeteiligung komplizierter ist, da im Falle der nicht während der gesamten Artikulationsphase ein Stimmtone produziert wird. Aber bei allen Vokalen des Deutschen sind Glottisschläge zu finden, da sie generell

stimmhaft sind. Die Verschlusslösung eines Plosivs sieht ähnlich aus wie ein Glottisschlag, wobei bei stimmlosen oder Plosiven mit einem Stimmtonverlust vor der Verschlusslösung immer eine akustische Stille während der Verschlussphase vorhanden ist und dadurch unterscheidbar wird. Bei stimmhaften Plosiven aber herrscht oberhalb von zirka 500 Hz eine absolute Stille während der Verschlussphase. Der Burst [bɜːst] ist die englische Bezeichnung der Verschlusslösung, die auf einem Sonagramm in der Form von Geräuscheffekten in einem breitem Frequenzspektrum zu sehen ist.



*Abbildung 16: Der Burst von [kʰ] im Wort *Kuchen* [kʰuːxŋ] auf einem Breitbandsonagramm*

Vokale und Konsonanten passen sich aneinander an, was auf dem Sonagramm gut sichtbar wird. Die Formantübergänge von einem Laut zum nächsten heißen

Transitionen. Sie sind daran zu erkennen, dass die schwarzen Balken in horizontaler Richtung (die Formanten) fallen, steigen oder neutral bleiben. Es werden drei Arten von Transitionen unterschieden. 1. Die negative, bei der die Formantübergänge Richtung nachfolgender Sprechlaut fallen. 2. Die positive, bei der die Formanten Richtung folgender Sprechlaut steigen 3. Die neutrale, bei der sich die waagerechten schwarzen Balken auf dem Sonagramm nicht bewegen.

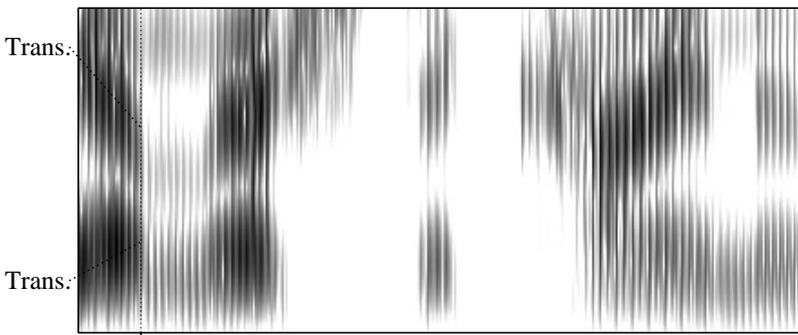


Abbildung 17: Transitionen des Vokals [e:] im Wort *wer* [ve:ɐ̯] auf einem Breitbandsonagramm

Zwei Arten der Sonagramme werden unterschieden. Das eine ist das **Breitbandsonagramm**, das mit einem 300 Hz-Analysefilter hergestellt wird. Es ermöglicht eine hohe Auflösung im Zeitbereich, auf dem die Formantenstruktur und die einzelnen Glottisschläge gut erkennbar sind. Das zweite ist das **Schmalbandsonagramm**, das mit einem 125 Hz-Analysefilter hergestellt wird.

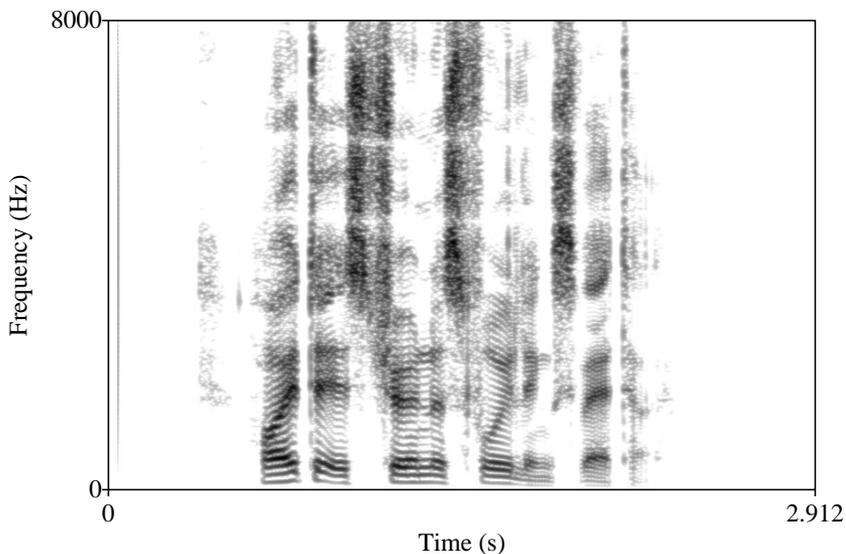


Abbildung 18: Das Schmalbandsonagramm des Satzes: *Heute ist schönes Frühlingswetter*. In der Abbildung sind die Obertöne deutlich sichtbar.

Das Schmalbandsonagramm macht eine hohe Auflösung im Frequenzbereich möglich, auf dem der Tonhöhenverlauf bzw. der Verlauf der Obertöne gut zu beobachten sind. Sie erscheinen auf einem Schmalbandsonagramm als schmale, parallele Linien., deren Abstände immer gleich groß sind.

Der Oszillograph produziert Oszillogramme, auf denen die zeitliche Veränderung des Sprechschalls zu beobachten ist. Oszillogramme sind zweidimensional, auf der horizontalen Achse sind die Zeitwerte, auf der vertikalen Achse die Amplitudenwerte abzulesen. Letztere geben über die Lautstärke des Tons Auskunft. Auf einem Sonagramm heißt die Auslenkung der Schallwelle von der horizontalen Linie in einem beliebigen Zeitpunkt Elongation.

Die Verwendung sowohl des Oszillographen als auch des Sonagraphen hilft dem Phonetiker bei der Bestimmung der Werte von den Komponenten der

segmentalen bzw. der suprasegmentalen Ebene. Auf der segmentalen Ebene geht die Suche nach Formanten und der Vokal- und Konsonantendauer, die mit der sogenannten Segmentierung feststellbar ist. Bei ihr sucht der Fachmann nach den Segmenten des Lautkontinuums. Bei der Segmentierung des Oszillogramms ist so vorzugehen, dass das nächste Segment seinen Anfang immer nach der negativen Phase im Nulldurchgang hat. Was die Vokal- und Konsonantendauer anbelangt, liegt sie in der Regel zwischen 30 ms und 150 ms. Am Ende einer Äußerung (auch im Falle von isolierten Einheiten) lässt sich tendenziell eine zeitliche Dehnung der Segmente erkennen. Heutzutage funktionieren sowohl Sonographen als auch Oszillographen binär, mit anderen Worten digital. Sechs Programme, die entweder nur Sonographen oder auch Oszillographen beinhalten, können aus dem Internet, frei (und legal) heruntergeladen werden: 1. Speech Filing System (SFS) vom UCL 2. STX von der Österreichischen Akademie der Wissenschaften, (bei dieser Software handelt es sich um eine Shareware). Diese ist aber von allen das anspruchsvollste Programm. 3. Speech Tools vom SIL, die das Programm Speech Analyzer enthält. 4 Das heutzutage bekannteste und am meisten verwendete Sprechschallanalyseprogramm heißt Praat. 5. WinCECIL ist das Programm, dessen Leistungsfähigkeit unter den Konkurrenten am Begrenztesten ist. 6. Formant heißt das Programm, mit dem Formantenmanipulation für auditivphonetische und Unterrichtszwecke veranschaulicht werden kann, es wurde von Tamás Bóhm entwickelt.

Bei Vokalen kann ein Zusammenhang von F1 und F2 mit der Zungenhöhe und der Zungenlage festgestellt werden. Der erste Formant korrespondiert mit der Zungenhöhe so, dass sein Frequenzwert mit der Offenheit des Vokals wächst. Je geschlossener ein Vokal ist, desto tiefer ist der erste Formant. Der zweite Formant ist mit der Zungenlage verbunden. Desto weiter vorne der auffälligste Moment des Vokals im Mundraum geschieht, umso höher ist der zweite For-

mant und umgekehrt. Die gerundeten Vokale zeigen einen tieferen F2 als die mit der gleichen Zungenhöhe gebildeten ungerundeten Vokale, beispielsweise [y] und [i]. Der Abstand der ersten zwei Formanten F1 und F2 der palatalen Vokale, der zirka 2000 Hz ausmacht, ist sehr groß. Die Entfernung von F1 und F2 der velaren Vokale, die ungefähr 1000 Hz beträgt, ist dagegen relativ gering. Sie liegen häufig so eng zusammen, dass sie wegen ihrer Bandbreiten kaum auseinandergehalten werden können. F2 und F3 der palatalen Vokale weisen meistens keine große Distanz auf. Im Gegensatz dazu sind F2 und F3 der velaren Vokale voneinander relativ weit entfernt. Alle ersten drei Formanten des Murrelvokals [ə], unter denen sich der Abstand auf 1000 Hz beläuft, liegen voneinander gleich weit entfernt. Der erste Formant der Konsonanten liegt wegen der größeren artikulatorischen Enge immer tiefer als der der Vokale. Wenn ein Vokal einem Konsonanten vorausgeht, steigt der erste Formant grundsätzlich nicht an.

Artikulatorisch werden bei der Bildung eines Plosivs drei Phasen unterschieden: 1. Verschlussbildung 2. Verschlussphase 3. Verschlusslösung. In der Verschlussbildung ist ein plötzlicher Abfall der Energie im gesamten Frequenzbereich zu beobachten. Während der gesamten Verschlussphase ist eine völlige akustische Stille bei stimmlosen Lauten oder Plosiven vorhanden. Bei stimmhaften Plosiven ist die akustische Stille oberhalb von ungefähr 500 Hz zu finden, da unterhalb dieses Wertes die „voice bar“ [vois ba:ɾ] auf dem Sonagramm zu sehen ist.

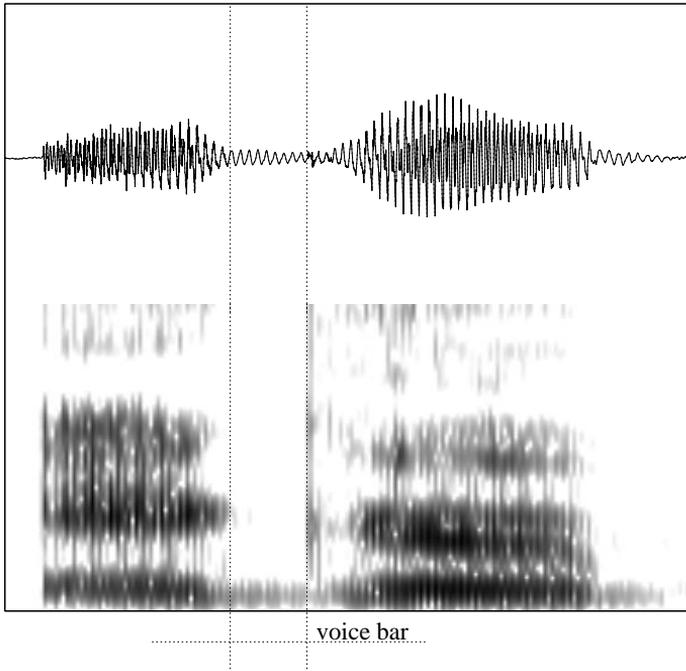


Abbildung 19: Die voice bar von [g] im Wort *Egge* [ɛgə] auf einem Breitbandsonogramm

Die Verschlussphase der stimmhaften Plosive ist meistens kürzer als die der stimmlosen. Bei der Verschlusslösung ist ein plötzlicher Anstieg der spektralen Energie zu sehen. Der Burst hat bei stimmhaften Plosiven einen Wert von 10-20 ms. Ihr Burst ist aber schwächer als der der stimmlosen Verschlusslaute. Im Deutschen sind wirklich stimmhaft realisierte Plosive selten zu finden, da [b], [d], [g] oft stimmlos realisiert werden. Wenn ein stimmhafter Plosiv einem Nasal folgt, reduziert sich die akustische Stille des Plosivs auf wenige ms. Dies gilt insbesondere dann, wenn Nasal und Plosiv homorgan d.h. an der annähernd gleichen Artikulationsstelle gebildet werden.

Die Artikulationsstelle eines Plosivs lässt sich an Hand des angrenzenden Vokals bestimmen. Wenn F1, F2 und F3 eines Vokals zum folgenden Plosiv fallen, ist das ein Hinweis für einen labialen Plosiv. Verlaufen F2 und F3 bei einem palatalen Vokal annähernd parallel, weist dies auf einen alveolaren Plosiv hin. Auch ein scheinbares Zusammentreffen von F2, F3 und F4 bei zirka 2500 Hz am Anfang eines velaren Vokals verweist auf einen alveolaren Plosiv. Falls F2 und F3 bei einem palatalen Vokal nicht zusammenlaufen, ist das ein Grund einen velaren Plosiv zu vermuten. Wenn ein velarer Plosiv einem velaren Vokal vorausgeht, gibt es keine Transition. Die Aspiration, deren Dauer zirka zwischen 40 – 80 ms liegt, folgt den stimmlosen Plosiven im Deutschen. Bei ihnen fehlt die Aspiration in Verbindung mit Frikativen und vor Nasalen beispielsweise in Wörtern wie *Stadt* und *Abmagerung*.

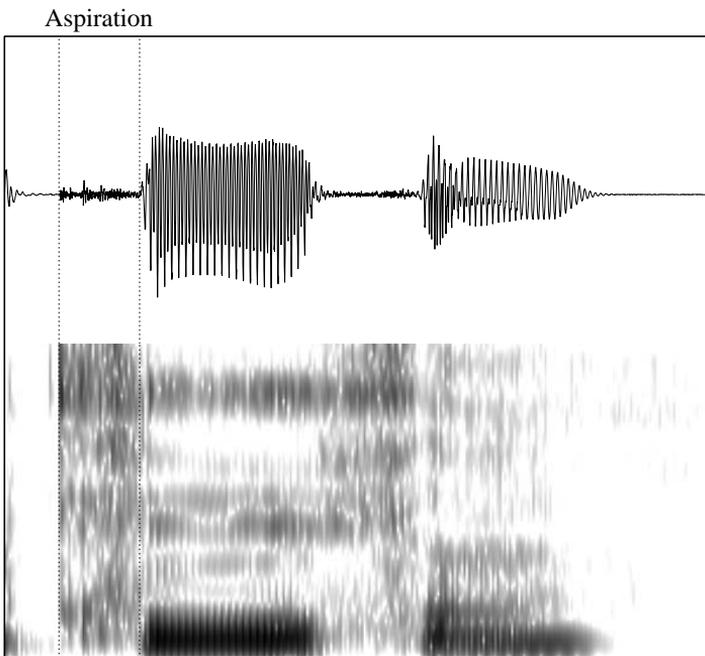


Abbildung 20: Die Aspiration von [k^h] im Wort *Kuchen* [k^hu:xŋ] auf einem Breitbandsonagramm

Die Zeit zwischen der Verschlusslösung und dem Einsetzen der Stimmlippenschwingungen, die so genannte Stimmeinsatzzeit, wird als **VOT** (voice onset time) bezeichnet. Nach der VOT erscheinen aber nicht alle Formanten des nachfolgenden Sprechlauts gleichzeitig, da die Aspiration sich in seine Formantenstruktur noch hineinzieht. Aus diesem Grund wird die sog. Klatt-VOT unterschieden, die die Zeit zwischen der Verschlusslösung und dem Ausklingen der Aspiration angibt. Wenn der Plosiv aspiriert ist, wird über positive VOT gesprochen. Ist der Plosiv unaspiriert oder verliert er seinen Stimmtton, wird die Stimmeinsatzzeit als zero-VOT bezeichnet. Ist der Plosiv stimmhaft, wird der Ausdruck negative VOT benutzt.

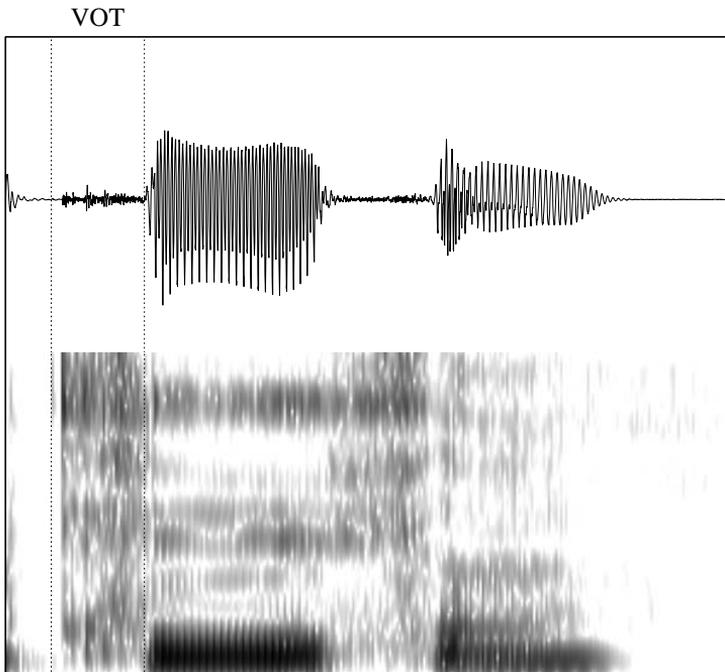
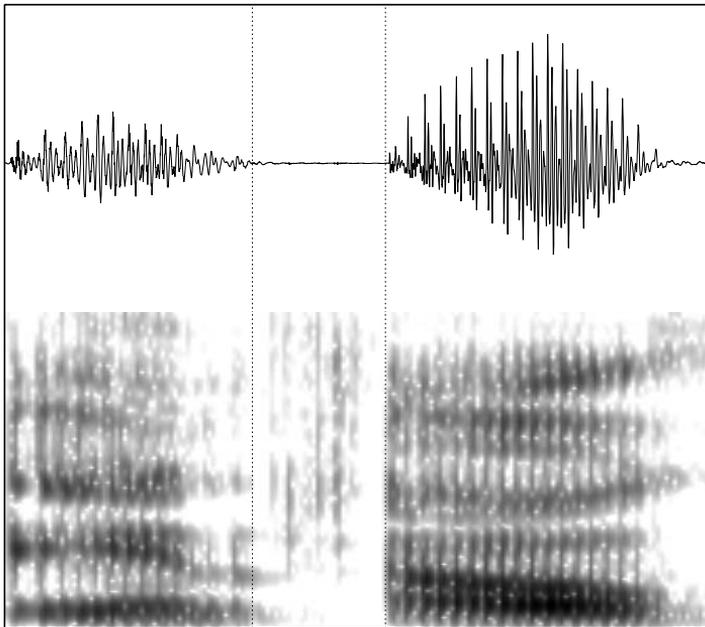


Abbildung 21: Die VOT von [kʰ] im Wort *Kuchen* [kʰu:xŋ] auf einem Breitbandsonagramm

Orthographisch kann ein Wort im Deutschen mit einem Vokal beginnen (...) phonetisch geht jedoch im Allgemeinen ein glottaler Plosiv [ʔ] dem Vokal voraus beispielsweise Ast [ʔast]. Dies kann nicht nur Wortinitial der Fall sein sondern auch silbeninitial z.B. Hiatus [hiʔ'artʊs]. Die akustische Stille beim glottalen Plosiv beträgt meistens weniger als 100 ms. Seine Verschlusslösung zeigt unregelmäßige Stimmlippenschwingungen am Anfang des Vokals, dessen Formantenstruktur der des glottalen Plosivs entspricht. Statt des glottalen Plosivs ist oft ein Glottalisierungsphänomen am Vokalanfang zu beobachten. Unter **Glottalisierung**, deren neudeutsche Bezeichnung creaky voice ist, sind unregelmäßige Glottisschwingungen mit niedriger Frequenz zu verstehen. Bei der Glottalisierung werden die Abstände der Glottisschläge größer und die Grundfrequenz (F0) nimmt plötzlich ab. Dieses Phänomen ist oft in Konsonant-Vokaltransitionen zu finden beispielsweise wenn ein Lateral oder ein Nasal an einer Silbengrenze steht, so in den Wörtern *Schalleffekt* und *Schein-effekt*.



creaky voice

Abbildung 22: creaky voice an der Wortgrenze in Augsburg auf einem Breitbandsonagramm

Bei Nasalen ist ebenfalls (wie bei den Plosiven) ein Verschluss vorhanden, der gelöst wird. Der Unterschied der Verschlusslösung der Nasale zu der der Plosiven besteht darin, dass der Verschluss bei den Nasalen langsamer gelöst wird. Die Nasale im Deutschen sind stimmhaft aber sie beginnen nach stimmlosen Frikativen meistens stimmlos (Beispiel: *öffnen*). Während der Verschlussphase des Nasals entstehen auf Grund des zugeschalteten Nasenraums Anti-Formanten, die Teile der Formantenstruktur abschwächen. Das Hauptmerkmal der Anwesenheit eines Nasals im Sonagramm ist, dass das Formantenspektrum oberhalb zirka 500 Hz stark gedämpft ist. Der Nasenschall hat im Vergleich zu Vokalen oberhalb von 500 Hz eine geringe Intensität und ist deshalb zu vernachlässigen. Der erste Formant dominiert das Spektrum, wobei

F2 schwach ausgeprägt oder überhaupt nicht vorhanden ist. Höhere Formanten sind beobachtbar, deren Intensität ist jedoch gering. Die Auswirkung der Antiformanten der Nasale ist besonders gut in der intervokalischen Position zusehen, bei der am Anfang des Nasals ein großer Energieabfall und am Vokalende ein großer Amplitudenabfall zu erkennen ist, wie beispielsweise in *Emma*. Vokale werden im Deutschen häufig an Nasal-Vokal- bzw. Vokal-Nasaltransitionen nasalisiert, wobei die Nasalisierung vollständig (wie im Französischen) oder partiell sein kann. Das Velum wird oft schon am Vokalende ungefähr um 100 ms vor der Bildung des oralen Nasalverschlusses abgesenkt und erst nach der Verschlusslösung des Nasals erfolgt seine Anhebung, deren Folge eine partielle Nasalisierung der mit dem Nasal benachbarten Vokale ist. Im Spektrum des Vokals ist eine weitere niedrigfrequente Resonanz im Bereich der Grundfrequenz F0 vorhanden, die F0 verstärkt. Die sich unmittelbar darüber befindende Antiresonanz dämpft F1, da bei der Nasalisierung auf Grund des zugeschalteten Nasenraums im Ansatzrohr weitere Resonanzen und Antiresonanzen entstehen. Die Frequenzen der Antiresonanzen hängen stark vom Grad der Gaumensegelöffnung ab.

Der laterale [l], dessen erster Formant F1 tief liegt, weist sowohl vokalische als auch konsonantische Merkmale auf. Bei [l] entweicht die Luft an beiden Seiten der Zunge, weshalb seine Formantenstruktur (F1 zirka 500, F2 1800) der eines Vokals sehr ähnelt. Die Entfernung zwischen den ersten zwei Formanten F1 und F2 von [l] macht zirka 1000 Hz aus. [l] ist aber stark vom Kontext beeinflusst und an einer Koartikulation beteiligt, die sich in seiner Formantenstruktur zeigt. Sie hängt davon ab, in was für einer Vokal- oder Konsonantenumgebung sich [l] befindet. Antiresonanzen schwächen die Intensität der Formanten von alveolarem Lateral ab, wenn auch nicht so stark wie bei den Nasalen. An der Segmentgrenze oder innerhalb des Laterals ist manchmal

eine Verschlusslösung zu finden, die ebenfalls auf ein [l] und nicht auf einen Vokal verweist.

In der deutschen Standardvarietät sind zwei Allophone des Vibranten /r/ zu finden, die beide stimmhaft sind. Sie können aber auf Grund umgebender stimmloser Segmente als Varianten, die ihren Stimmton verloren haben, vorkommen. Das eine /r/-Allophon ist ein alveolarer, das andere ein uvularer Vibrant. Bei ihrer Bildung werden zwei Phasen unterschieden, die einen Schlag ausgeben: 1. Offene Phase 2. Verschlussphase. Bei der offenen Phase erreicht der Sprechschall ein Maximum, bei der Verschlussphase ein Minimum. Die Verschlussphase geht schneller in die offene Phase über als umgekehrt und die letztere ist zeitlich länger als die erstgenannte.

Für die Frikative ist ein Friktionsgeräusch charakteristisch, das im Mundraum bei der Verengung entsteht. Diese teilt ihn in zwei Teile, in einen vorderen und einen hinteren. Das Spektrum des Friktionsgeräusches unterliegt hauptsächlich der Größe und der Form des vorderen Resonanzraums. Je größer er ist, desto markanter ist der Frikativschall. Zu den Spektren der Frikative kann gesagt werden, dass die hinteren Artikulationsstellen einen stärker modulierten Schall aufweisen. Deshalb besitzt das uvulare [χ] beispielsweise im Wort *Bach* ein ausgeprägteres Schallspektrum als das alveolare [s] z.B. im Wort *Quiz*. Das Spektrum der Frikative oberhalb von 2500 Hz zeigt eine größere Intensität als unterhalb dieses Frequenzwerts, wobei stimmlose Frikative im Vergleich zu stimmhaften durch eine größere Intensität gekennzeichnet sind. Stimmhafte Frikative können zum Teil oder ganz entstimmt vorkommen, wenn sie satzinitial oder von stimmlosen Lauten umgeben stehen. Das Spektrum des frikativischen Segments der Affrikaten zeigt eine schief verlaufende Untergrenze, die meistens fallend ist.

Ein suprasegmentales Element ist die **Intonation**, bei der es sich eigentlich um die Veränderung der Grundfrequenz (F0) in der Zeit handelt. Daraus ergibt sich, welche zwei Variablen bei der Veranschaulichung des Intonationsverlaufes auf den Achsen abgebildet werden. Auf der vertikalen Achse wird die Tonhöhe der Grundfrequenz in Hertz (Hz), auf der horizontalen Achse die Zeit in Sekunden (s) angegeben. Der Verlauf der Grundfrequenz lässt sich in den stimmhaften Abschnitten des Sprechschalls beobachten, da nicht nur stimmhafte sondern auch stimmlose Abschnitte und Pausen einander im Lautkontinuum folgen. Der Hörer nimmt den Intonationsverlauf trotz seiner Unterbrechungen als kontinuierlich wahr. Laut VON ESSEN (1979: 209) werden drei Intonationsverläufe unterschieden: 1. Terminaler (abschließender) 2. Progredienter (weiterweisender) 3. Interrogativer (fragender).

Der terminale Intonationsverlauf ist in Ausrufen, Aufforderungen, Aussagen und Ergänzungsfragen zu finden. Für ihn ist charakteristisch, dass er nach der hervorgehobenen Silbe bis zum Satzende sinkt.

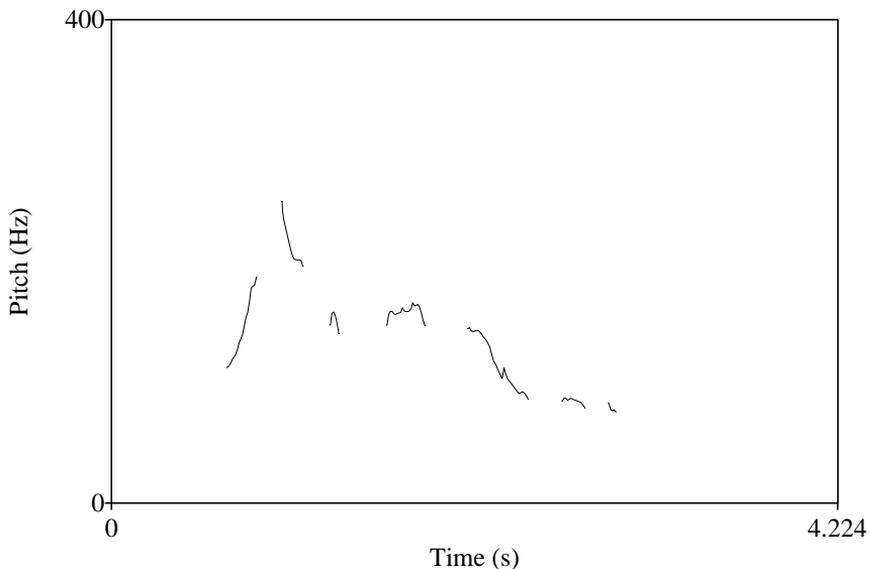


Abbildung 23: Der terminale Intonationsverlauf der Äußerung: Heute ist schönes Frühlingswetter

Der progrediente Intonationsverlauf, bei dem der Hörer mit der Fortsetzung der Äußerung rechnet, kommt in Äußerungen zusammengesetzter Sätze vor. Der Intonationsverlauf geht am Äußerungsende ein bisschen aufwärts oder sinkt ein bisschen, dies geschieht jedoch im geringeren Maße als terminalen Intonationsverlauf.

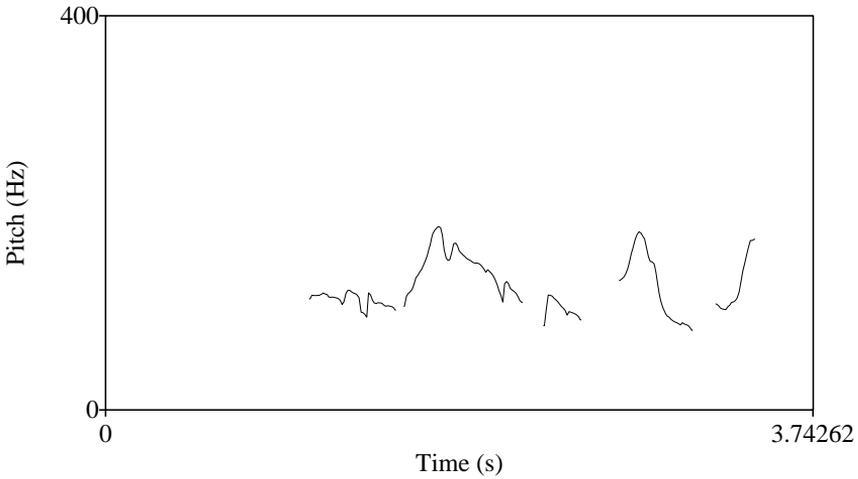


Abbildung 24: Der progrediente Intonationsverlauf der Äußerung: Wenn du den Glühwein durch einen Strohhalm trinkst,...

Der interrogative Intonationsverlauf, der in Fragesätzen erscheint, wird dadurch gekennzeichnet, dass bis zum Satzende ein Grundfrequenzanstieg der hervorgehobenen Silbe folgt.

zugrunde liegen. Jede Hypothese basiert aber quasi auf einem Black-Box-Modell, weil die zu untersuchenden Vorgänge den unmittelbaren Beobachtungen nicht zugänglich sind.

In der Erforschung der Perzeption nimmt die Sprechperzeption eine besondere Stellung ein. Der Grund hierfür liegt darin, dass sich nur die Sprechkommunikation dadurch auszeichnet, dass der Empfänger imstande ist das gleiche (nicht aber dasselbe) informationstragende Signal auszusenden, welches er vom Sender wahrgenommen hat. Diese Tatsache spielt in den Hypothesen über das Sprachverstehen eine zentrale Rolle. Das Sprachverstehen stellt einen Bestandteil der (laut)sprachlichen Kommunikation dar, der – die lautsprachliche Informationsmitteilung vorausgesetzt – mit der sinnlichen Wahrnehmung des Sprechschalls durch das Gehörorgan beginnt und mit den Ergebnissen der Dekodierung der im Sprechschall kodierten sprachlichen Informationen im Gehirn endet. Dementsprechend zerfällt das Sprachverstehen in physiologische, physikalische, psycho-akustische, neurophysiologische, linguistische, assoziative und eventuell in weitere – zum Teil noch unbekannte oder wenig bekannte – Vorgänge.

Im Vergleich zu vielen früheren Erkenntnissen, Auffassungen und Formulierungen über das Sprachverstehen stellt LINDNER (1969: 81) sowohl das Wesen des Sprechschalls als auch die Rolle des Perzipienten in ein neues Licht, welches der Wirklichkeit besser entspricht: „Wörter, Silben und Laute sind Einheiten, die wohl vom Produzenten willentlich erzeugt und vom Perzipienten aus dem lautsprachlichen Zeichen entnommen werden, die aber als Ding an sich keinen Sinn haben. Nur vom Perzipienten aus gesehen besteht das akustische Zeichen aus Silben und Lauten, und nur von seiner Warte her, d.h.

von einer sprachlich-gesellschaftlichen Grundlage, kann gesagt werden, welche Teile des lautsprachlichen Signals zu einer Einheit gehören...“

7.1. Funktionelle (Neuro-)Anatomie des auditorischen Systems

Außenohr (*Auris externa*)

Das Organ der Schallbündelung (nach Frequenzgruppen) und des Richtungshörens, die Ohrmuschel (*Auricula*) wird durch ein elastisches Knorpelgerüst geschützt; ohne Schutz ist lediglich ihr schlaffer, fleischiger und fettreicher Zipfel, das Ohrläppchen (*Lobulus auriculae*). Wenngleich ihre Morphologie (= Aufbau) bei jedem Menschen verschieden ausfällt, finden sich ähnliche Strukturen: Den nach außen gewölbten Rand stellt die Ohrleiste (auch: Ohrkrempe) (*Helix*) dar, weiter innen ist die formähnliche *Anthelix* zu finden. Die Ohrecke (*Tragus*) ist ein von der Gesichtsseite in die Öffnung des Gehörgangs (*Porus acusticus externus*) hereinragender knorpeliger Höcker; ihm gegenüber, unten sitzt der *Antitragus*. Sein äußeres Drittel ist epidermal (= „oberhäutig“) (das trifft auch auf den ganzen Gehörgang zu) und knorpelig ausgeprägt bzw. mit Haaren und Talgdrüsen überzogen. Medial und schräg in ihr ist eine (annähernd) runde, leicht trichterförmige Membran von ca. 1 cm Durchmesser, das Trommelfell (*Membrana tympani*) eingelassen. Dieses hat die *anatomische* Funktion das äußere Ohr vom Mittelohr zu trennen, seine *sprachlichen* Funktion besteht darin den Schall auf die Gehörknöchelchen (*Ossicula auditoria*), d.h. auf Hammer (*Malleus*), Amboss (*Incus*) und Steigbügel (*Stapes*) (s.u.) zu übertragen. Die fehlende Intaktheit des Trommelfells (Riss oder Loch) kann

eine Verminderung des Hörvermögens mittleren Grades verursachen (vgl. TREPEL 2004: 314f.).

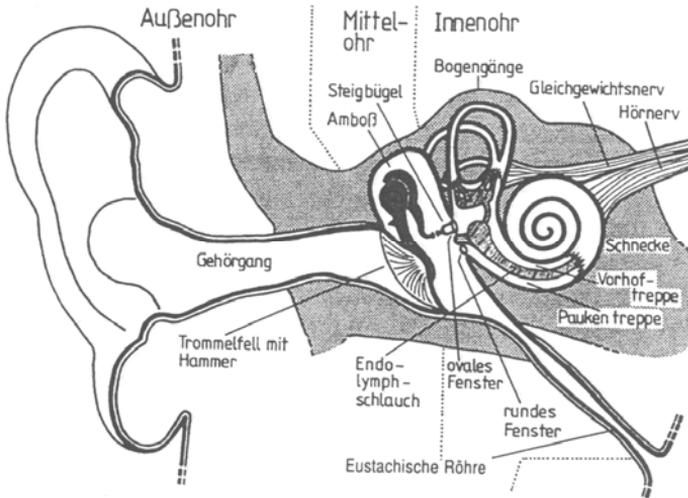


Abbildung 26: Sagittalschnitt des Ohrs (NEPPERT 1999: 273)

Mittelohr (*Auris media*)

Das Mittelohr besteht vornehmlich aus einem schmalen lufthaltigen Hohlraum, der Paukenhöhle (*Cavum tympani*); ihre Verbindungen zu den Warzenzellen (*Cellulae mastoideae*) (über die Warzenfortsatzhöhle) (*Antrum mastoideum*) (h.) und dem Nasopharynx (über die Ohrtrompete) (*Tuba auditiva*)⁵ (v.) werden hierzu gezählt. Die knöcherne Wand der Paukenhöhle hat verschiedene Öffnungen: (1) Lateral (= seitwärts) liegt der (durch das Trommelfell verschlossene) äußere Gehörgang, (2) medial, zum Innenohr hin liegen das (durch eine Membran und den Steigbügel Fuß verschlossene) ovale (*Fenestra vestibuli*) und (3) das (durch eine Membran verschlossene) runde Fenster (*Fenestra*

cochleae). Die oben bereits erwähnten und schleimhaut(epithel)überzogenen Gehörknöchelchen sind die wichtigsten Elemente der Paukenhöhle: Sie übertragen den Schall (1) von der Luft (im äußeren Gehörgang) (2) auf die Flüssigkeit *Perilymphe* (s.u.) (im Innenohr). (1) hat einen niedrigen, (2) aber einen hohen Schallleitungswiderstand; die Reflexion (= Zurückgeworfenwerden) von Schallwellen wird dadurch vermieden, dass sie von einer großen (Trommelfell) auf eine kleine Schwingungsfläche (ovales Fenster) übertragen wird, dabei wird die Schallamplitude enorm (22fach) vergrößert (Dazu trägt auch die Hebelwirkung der Gehörknöchelchenkette bei.). Bei der Schallübertragungsmodulation arbeiten auch die beiden Mittelohrmuskeln, der „Trommelfellspanner“ (*Musculus tensor tympani*) und die „Steigbügelmuskel“ (*Musculus stapedius*), trotz der entgegengesetzten Zugrichtungen, synergetisch (= zusammen) (TREPEL 2004: 315 – 318).

Innenohr (*Auris interna*)

Den Hörbahnanfang, „de[n] weitaus wichtigste[n] Eingangskanal in das kortikale auditorische System“ (TREPEL 2004: 228) stellt das Innenohr dar.

Der häutige (auch: membranöse) Anteil des Innenohrs entwickelt sich aus einer auf Höhe des Rautenhirns (*Rhombencephalon*) bereits in der dritten embryonalen Woche erkennbaren Verdickung der obersten Hautzellschicht (*Epithelium*), der Ohrplakode. Diese sinkt einige Tage später als Labyrinthgrube herunter, verliert den Kontakt zum Oberflächengewebe und wird zu einem mit *Endolymphe* gefüllten Epithelbläschen, dem Labyrinthbläschen. Von mesodermalen Spalträumen eng umgeben baut sich dabei ein mit *Perilymphe*⁶ gefülltes Hohlraumsystem, der Perilymphraum auf.

5 Die ~ dient dem Druckausgleich zwischen Paukenhöhle und Außenluft.

6 *Endolymphe* und *Perilymphe* sind wasserklare Flüssigkeiten, sie unterscheiden sich allein in ihrer chemischen Zusammensetzung, bes. in der Elektrolytkonzentration:

Aus dem Labyrinthbläschen tritt das Labyrinthorgan mit zwei Bogengängen, dem Säckchen (*Sacculus*) und dem Schläuchlein (*Utriculus*) hervor, diese enthalten spezifische Sinnesepithelien für die Wahrnehmung von Linear- und Drehbeschleunigungen (kurzum: Gleichgewichtsempfindung), sowie dem Schnecken gang (*Ductus cochlearis*). Letzterer nimmt bei Erwachsenen eine Länge von ca. 4 cm ein; wegen der großen Raumforderung wickelt es sich jedoch zu einer Raumspirale, einem „Schneckengehäuse“ (*Cochlea*) von ca. 2,5 cm Windungen auf. Säckchen, Schläuchlein und Perilymphraum bilden zusammen das Vestibularapparat (= Gleichgewichtsorgan) (*Vestibulum auris*); den Schnecken gang mit eingerechnet bilden diese Teile zusammen das Innenohr.

Im cochlearen Querschnitt sind drei voneinander nicht ganz separierte Räume zu erkennen: (1) Vorhofstreppe (*Scala vestibuli*), (2) Schnecken gang (*Scala media*)⁷ und (3) Paukentreppe (*Scala tympani*); (1) und (3) sind im „Schneckenloch“ (distale Spitze der Schnecke) (*Helicotrema*) ineinander übergehende perilymphatische Räume, (2) ist mit Endolymphe gefüllt. Zwischen (1) und (2) ist ein sehr dünnes Häutchen, die Reissner-Membran (*Paries vestibularis ductus cochlearis*); während zwischen (2) und (3) ein Stück bindegewebiges (z.T. straffes) Häutchen, die Basilarmembran (*Membrana basilaris*) liegt. Ihre Struktur ist sowohl bezüglich ihrer Dicke als auch bezüglich ihrer Breite ungleichmäßig: In der unteren Windung (= Beginn des Schnecken ganges) ist sie dick, schmal und folglich versteift, in der Spitzenwindung hingegen dünn, breit und verhältnismäßig flexibler. In einem Teil der Basilarmembran, dem

Die Perilymphe ähnelt sich dem Hirnwasser (*Liquor cerebrospinalis*), sie hat nämlich eine hohe Na⁺-Konzentration (140 mval (=Millival)/l) bei geringer K⁺-Konzentration (10 mval/l), während die Endolymphe eine umgekehrte Proportion, eine hohe K⁺-Konzentration (144 mval/l) und eine geringe Na⁺-Konzentration (5 mval/l) zeigt (vgl. TREPPEL 2004: 318; ZILLES & REHKÄMPER 1998: 180).

Corti-Organ (auch: Cortisches Organ) (*Organum spirale*) befinden sich zwei Gruppen von Sinneszellen: innere und äußere Haarzellen (die kein eigenes Axon haben); diese tragen jeweils innere und äußere Phalangenzellen (= Stützzellen).

Die primäre Leistung des auditorischen Systems, besonders der Kerngebieten in der aufsteigenden Hörbahn ist die Mustererkennung. Die bilaterale Ausbildung des Innenohrs und die zentralnervöse Verschaltung ermöglichen Richtungshören, das Corti-Organ sorgt für Frequenzanalyse. Die eintreffende Schalldruckwelle wird über die am ovalen Innenohrfenster angesetzten Gehörknöchelchen (Hammer, Amboss, Steigbügel) und das seitlich im hinteren Schädelbereich liegende Schläfenbein (*Os temporale*), kurz: über „Knochenleitung“ auf den vestibularen Perilymphraum übertragen. Weitergeleitet wird sie über die Paukentreppe, wo das membranverschlossene „Schneckenfenster“ zur Paukenhöhle (*Cavum tympani*) den Ausgleich des Druckes ermöglicht. Aus dem Perilymphraum gelangt die Schalldruckwelle über die Reissner-Membran in den endolymphatischen Innenraum des Schneckenanges.

Da Flüssigkeiten nicht komprimierbar sind, verursachen Schwankungen der Schalldruckwelle eine kraniale (= oben) Auslenkung der Basilarmembran in Richtung der Tektorialmembran (*Membrana tectoria*). Diese erfolgt lediglich im Bereich der äußeren Haarzellen; da die Basilarmembran im Bereich der inneren Haarzellen unbeweglich an einer Knochenleiste, der *Lamina spiralis ossea*, fixiert ist. Die Kodierung der eintreffenden Schalldruckwelle durch Haarzellen ist binär: Laut dem Ortsprinzip erreicht die Basilarmembran die Maximalauslenkung an einem Ort, der, auf Grund seiner ortsspezifischen Schwingungseigenschaften, mit der Frequenz in Beziehung steht, d.h., für tiefe

7 *Ductus cochlearis* und *Scala media* sind Synonyme.

Frequenzen geschieht das apikal (= oben), beim „Schneckenloch“, für hohe Frequenzen basal (= unten), beim ovalen Fenster (*Fenestra vestibuli*). Laut dem Periodizitätsprinzip gibt es einen Zusammenhang zwischen dem Feuermuster der Haarzelle und der Frequenz: Bei Aufwärtsbewegung der Basalmembran depolarisieren die Nervenzellen, wenn auch auf Grund von Refraktärzeiten (= Zeitintervall der Unreizbarkeit) nicht immer synchron. Beide Prinzipien wirken zusammen: Im weitergeleiteten Signal in der Hörrinde ist sowohl die topographische als auch die zeitliche Information kodiert (ZILLES & REHKÄMPER 1998: 179 – 185).

7. 2. Psychoakustik des Gehörs. Wahrnehmungsgrößen⁸

Das menschliche auditorische System ist in Hinsicht auf die Wahrnehmungsverarbeitung akustischer Signale aus der Umwelt optimiert. Die akustische Struktur der sprechsprachlichen Signale ist in der Zeitfunktion schnell veränderlich; das Ohr hat trotzdem die Fähigkeit, die für das Hören wesentlichen Komponenten (wie z.B. Intensität und Frequenz) als subjektive Größen (wie Lautheit und Klanghöhe) wahrzunehmen.

Die Erhöhung der Frequenz führt zur Wahrnehmung einer ansteigenden Klanghöhe. Laut dem WEBER-FECHNERSchen Gesetz ist die wahrgenommene Klanghöhe proportional zum Logarithmus der Frequenz (In der Tat funktioniert dies Gesetz lediglich in einem stark begrenzten Mittelfrequenzbereich.⁹). Die Funktion der subjektiv empfundenen Klanghöhe und der objekti-

8 Die ~ sind immer an jugendliche Normalhörer rektifiziert.

9 Eine Ausnahme bilden die musikalisch ausgebildeten Personen.

ven Frequenz liefert die (auf dem Konzept der Frequenzgruppen¹⁰ basierende) Bark-Skala.

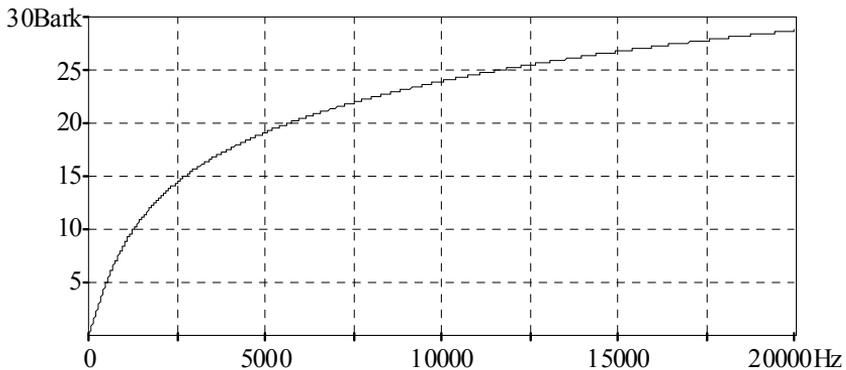


Abbildung 27: Die Funktion von Bark und Frequenz

Exakt messbar sind zudem physikalische Größen wie Schallintensität¹¹ (od. Schallstärke, d.h. Lautheitsempfindung) und Schalldruck (bzw. dessen Effektivwert, der Schallpegel). Sie geben an, was der Perzipient beim Hören als leiser oder lauter empfindet. Ein erhebliches Problem ergibt sich daraus, dass man diese Größen zum wahrnehmbaren Anstieg der Lautheit auf Grund des großen Maßstabs nicht direkt zuordnen kann. Aus diesem Grund bedient man sich der Zehnerpotenzstufen. So ist z.B. für die auditive Zuordnung das Bel (als Grundeinheit der zehnerlogarithmischen Skalenteilung) zu groß, aus die-

10 Frequenzgruppe nennt man diejenige Bandbreite in dem auditorischen System, deren jegliches Signal zu **einem** Erregungspegel gebündelt verarbeitet wird. Laut diesem Konzept ist die Lautheit bandbreitenabhängig: Wird die wahrgenommene Lautheit bei gleicher Leistung pro Hertz, doch bei verschiedener Bandbreite bestimmt, so ist es festzustellen, dass die Verteilung der Leistung auf einen größeren Frequenzbereich den Anstieg der wahrgenommenen Lautheit auf etwa das Doppelte mit sich bringt. Sollte aber der Frequenzbereich kleiner bestimmt werden als die Frequenzgruppe, setzt sich das Gesetz außer Kraft. Dementsprechend ist aus der Bandbreitenabhängigkeit auf die Größe der Frequenzgruppe zu folgern (vgl. KOLLMEIER 2004: 126 f.).

sem Grund verwenden wir in der Messpraxis das Dezibel (dB) (als nächstun-tergeordnete Einheit). Wichtig ist es zu erwähnen, dass das Dezibel eine relative Einheit ist: Gegen einen (beliebigen) Bezugspunkt bewegen wir uns auf einer Skala, deren Aufteilung dem Mehrfachen des Bezugspunktes entspricht. Deshalb können auch negative Dezibelwerte errechnet werden (vgl. NEPPERT & PÉTURSSON 1986: 58 f.).

Der (optimale) Bezugspunkt, genannt Ruhehörschwelle, ist der Pegel eines schon wahrnehmbaren Sinustons von dem vorher bestimmten Wert von 1000Hz (ohne das Vorliegen von Störsignalen). Steigt der Pegel des Sinustons kontinuierlich an, hören wir – nach Überschreiten der Ruhehörschwelle – frequenzabhängig einen immer lauterem Ton. Bei jeder Frequenz können wir den zum gleichen Lautheitseindruck führenden Pegel bestimmen. Daraus bildet sich optisch eine Kurve, die Isophone. Die 10phon-Isophone übersteigt bei 1000Hz die Ruhehörschwelle um 10dB, während die 120phon-Isophone bei 1000Hz mit ihrem 120dB *en bloc* der Schmerzschwelle eines Normalhörenden entspricht¹². Demnach ist Phon eine physikalische Reizgröße, die als Pegel-lautstärke od. Lautstärkepegel bezeichnet wird und die zu der für alle Lautstärken auf einer Isophone gleichen Lautheit führt (vgl. KOLLMEIER 2004: 121 ff., 125 f.).

Der Zusammenhang der Lautheiten wurde von STEVENS (nach PAUKA 1980: 36) mit Hilfe eines mo^{-13} und binauralen¹⁴ Tonvergleichs festgelegt. Zur Maßeinheit der lautheitlichen Verdoppelung wurde sone (eine Empfindungs-

11 „Physikalisch ist die Schallintensität als die Energiemenge definiert, die pro Sekunde auf einer senkrecht zur Ausbreitungsrichtung befindlichen Fläche von 1m² ankommt (NEPPERT & PÉTURSSON 1986: 57).“

12 Der Abstand zwischen der minimalen und der maximalen Lautheit einer Frequenz ist die sog. Dynamikbreite.

13 ~: mit einem Ohr

14 ~: mit zwei Ohren

größe) gewählt, wobei die Parameter 1kHz / 40dB / 40phon eines Referenztons genau 1sone entspricht. Vergleicht man die sone-lautheitliche Skala mit der phon-lautheitlichen Skala, so lässt sich feststellen, dass eine Steigerung von 10phon eine doppelte Steigerung der Lautheit in sone bewirkt, *ergo*: wird ein Ton um 10phon stärker, dann hören wir ihn doppelt so laut (vgl. PAUKA op. cit. 36 f.).

Diese Größen verhelfen uns dazu, die Ergebnisse der Untersuchungen in offiziellen Skalensystemen anzuordnen, was nicht nur den Vergleich der akustischen Konstituenten der Sprechlaute miteinander, sondern auch die Gegenüberstellung mit früheren (evtl. zukünftigen) Ergebnissen des Faches ermöglicht.

7.3. Die Stufen oder Ebenen des Sprachverstehens

Die meisten Hypothesen über das Sprachverstehen stellen die Variationen eines Vorganges dar, der sich aus drei Ebenen zusammensetzt. Diese sind bekannt als

- 1) die auditive,
- 2) die phonetisch-lautliche,
- 3) die phonologische Ebene

7.3.1. Die auditive Ebene

Auf der ersten, der **auditiven Wahrnehmungsebene** erfolgt eine primäre akustische Zerlegung des Sprechschalls innerhalb des Ohrs: der Sprechschall wird in seine unmittelbaren physikalisch-akustischen Bestandteile wie temporale Verhältnisse, Komponenten der Frequenz und des Schallpegels (der Intensität) zerlegt und als bioelektrisches Erregungsmuster im Hörnerv zum Gehirn weitergeleitet. Die Ergebnisse dieser akustischen Zerlegung, die akustischen Sinnesdaten werden im so genannten akustischen Gedächtnis gespeichert und mit den einschlägigen Vorkenntnissen und Erfahrungen, d.h. mit den entsprechenden Bereichen des eigenen sprachlichen Codes, abgeglichen.

7.3.2. Die phonetische Ebene

Auf der zweiten, der **phonetischen Ebene** erfolgt die Analyse und die Verarbeitung der auf der Ebene 1 gewonnenen und im akustischen Gedächtnis gespeicherten akustischen Sinnesdaten. Auf der Ebene 2 werden diese – nach dem vorgenommenen Vergleich mit dem eigenen Code – einer bzw. mehreren Klassen zugeordnet. Eine erfolgreiche Zuordnung aktuell eintreffender akustischer Sinnesdaten zu bereits bekannten, im akustischen Gedächtnis gespeicherten Sprechlauten des eigenen Codes wird in den klassischen Kommunikationsmodellen als „Dekodierung“ bezeichnet. Die Sprechlaute bekommen auf dieser Ebene einen „Namen“ ([v], [f], [a:] oder [u]). Die so erkannten Laute werden dann im phonetischen Gedächtnis gespeichert. Die Existenz dieser Ebene 2 wird nicht nur angenommen, sondern unter anderem auch durch EEG-Versuche bestätigt. In diesen zeigte sich, dass der EEG offensichtlich erst bei

der Unterscheidung von Sprechreizen (bei der Zuordnung von akustischen Merkmalen zu Lauten) eine größere Amplitude in der Reaktion der linken Gehirnhälfte der Versuchsperson zeigte.

7.3.3. Die phonologische Ebene

Auf der dritten, der **phonologischen Ebene** erfolgt die Zuordnung phonetisch ermittelter distinktiver Merkmale und Einheiten zu Phonemklassen. Die Erkennung der Phoneme ist Ergebnis dieser Operation. Die Bestimmung der Phoneme bezieht auch die Reihenfolge der Phoneme und andere wichtige Kriterien mit ein, abhängig von der Vorprogrammierung durch die phonologischen Regeln des Operationssystems: Auf der phonologischen Ebene 3 wird entschieden, welche Laute in welcher Reihenfolge (usw.) einen bedeutungstragenden Zeichenkörper bilden z.B.: /g+r+a:+f/ etwa im Vergleich zu /f+r+a:+g/, nicht aber /r+g+...+/, weil es im Deutschen kein Wort gibt, das mit den Lauten „rg-“ beginnt. (Hier soll die Bemerkung stehen, dass die Perzeptionseinheit bis heute nicht eindeutig bestimmt ist.)

7.4. Perzeptionsmodelle und -theorien

Das Perzeptionsmodell dient der vereinfachten Darstellung des Sprechverstehensvorgangs. Wie in jedem Modell so geht es auch im Perzeptionsmodell darum, das Funktionieren und Zusammenwirken der Teilvorgänge, sowie ihre gegenseitige Beeinflussung zu erschließen und zu beschreiben. Die konkreten Modelle zeichnen sich u.a. durch sehr unterschiedlichen Grundauffassungen

und durch das Herangehen mit sehr unterschiedlichen Methoden aus. Ein gemeinsamer Wesenszug aller Perzeptionsmodelle besteht darin, dass der Vorgang des Sprechverstehens als ein sukzessiver Vorgang angesehen und sein Ablauf von gewissen Voraussetzungen abhängig gemacht wird. Das beschriebene 3-Ebenen-Modell bietet beispielsweise eine Lösung für die Problematik der Unterscheidung und der Identifizierung von Lauten, Phonemen und deren Merkmalen. Der Mensch ist imstande, eine Menge von Lauten zu identifizieren, die Menge derer aber, die er zu unterscheiden vermag, kann das Zweihundertfache betragen. Daraus ergibt sich eine weitere Frage: Was liegt der Identifizierung der Laute zugrunde? Zwei der möglichen Antworten lauten:

- 1) die angeeigneten gleichen Wesenszüge,
- 2) die angeeigneten Unterschiede.

Für den Fall 1) muss der Mensch imstande sein, die in Bezug auf die Kommunikation unwichtigen Unterschiede, zu ignorieren. Für den Fall 2) muss er die Fähigkeit besitzen, das für das Sprachverstehen Bedeutende von allen anderen unwesentlichen Merkmalen, scharf zu trennen. Die Unvollkommenheit dieser Theorie besteht darin, dass sich der Perzeptionsmechanismus der kontinuierlichen Rede nicht eindeutig davon ableiten lässt.

Von den bekannten Perzeptionsmodellen werden hier fünf behandelt.

7.4.1. Analyse mit Synthese

Nach dieser Theorie werden die Sprechlaute mit gespeicherten „synthetisierten“ Lauttypen (die auf früheren auditiven Erlebnissen basieren) konfrontiert: Die Dekodierung erfolgt erst, wenn Sprechlaute des Expedienten (= Sprechers)

mit den intern generierten „synthetischen“ Lauttypen des Rezipienten (= Empfängers) koinzidieren. Die Unvollständigkeit dieser Konzeption ergibt sich v.a. daraus, dass Varianz und Assoziation umgangen werden (vgl. VALACZKAI & BRENNER 1999: 112 f.).

7.4.2. Die Motor-Theorie

Diese Theorie ging aus der Erforschung des Locus, der angesteuerten „Zielfrequenz“ von Transitionen (= Formantenabbiegungen) in den konsonantisch-vokalischen Lautübergängen hervor. Dieser Locus verkörpert nach Meinung der Forscher „das akustische Korrelat der Artikulationsstelle“ (NEPPERT 1999: 203). Die Vertreter der MT behaupten, dass der Rezipient bei der Dekodierung des Sprechschalls diesen (natürlich unterhalb der Bewusstseinschwelle) auf seine eigene artikulationsmotorische Tätigkeit, auf Referenzmuster zurückführt, quasi stumm „mitartikuliert“. (Man sei darauf verwiesen, dass Patienten mit Stimmlippenerkrankung auch das Anhören längerer Reden u.dgl. ärztlich verboten ist.)

7.4.3. Die kategoriale Theorie

Hypothetisch liegt der Lautperzeption der Fakt zu Grunde, dass Konsonanten und Vokale divers identifiziert werden: Bei Konsonanten ist an der Lautgrenze (und deren mentalen Abbildung, der Phonemgrenze) eine sprunghafte Änderung in der Lautidentifikation zu detektieren (= kategoriale Perzeption), während bei Vokal(phonem)en kein jähes Umschlagen in den nachfolgenden Vokal zu konstatieren ist (= nicht-kategoriale Perzeption). Das kann mit der KaT

so erklärt werden, dass Konsonanten mit diskontinuierlichen, Vokale hingegen mit kontinuierlichen Bewegungen der Sprechorgane erzeugt werden. Das würde aber heißen, dass der Perzeption artikulatorische Bewegungen zu Grunde liegen, nicht aber deren akustische Projektion, dies ist in der Tat der Fall (vgl. VALACZKAI & BRENNER 1999: 113 f.).

Die kategoriale Theorie bezieht sich vor allem auf die Lautperzeption. Bezüglich der solchen herrscht die Hypothese vor, dass die Vokale bzw. die Konsonanten unterschiedlich erkannt werden. Dieser Unterscheidung liegt die Annahme zugrunde, dass sich die Identifizierung der Konsonanten nahe der Phonemgrenze sprunghaft verändert, d.h. ein Phonem plötzlich in ein anderes umschlägt. Im Gegensatz dazu zeichnet sich die Identifizierung der Vokale durch einen kontinuierlichen Übergang aus, und in der Nähe der Phonemgrenze tritt kein plötzliches Umschlagen der Vokalphoneme in andere auf.

In diesem Sinne wird von einer kategorialen, d.h. diskontinuierlichen und von einer nichtkategorialen, d.h. kontinuierlichen Lautperzeption gesprochen. Die kategoriale Perzeption ist für Konsonanten, die nichtkategoriale für Vokale charakteristisch.

Diesen Auffassungen liegen Unterschiede in der Artikulation zugrunde: die Vokale unterscheiden sich etwa von den Explosivlauten u.a. gerade dadurch, dass die Artikulationsbewegungen bei der Erzeugung der Vokale kontinuierlich in die anderen Vokale übergehen können, z.B. [a: – o: – u:]. Bei der Mehrheit der Konsonanten ist das nicht möglich, z.B. [b – d – g].

Nach Auffassung der Vertreter der Motor-Theorie besteht ein anderes Verhältnis zwischen der Perzeption und der Artikulation als zwischen der Perzeption und dem akustischen Produkt der Artikulation. Daraus resultiert eine bemerk-

kenswerte Konsequenz: Für die Perzeption existiert nur das, was auch artikulatorisch existiert. Was akustisch existent ist, artikulatorisch jedoch nicht, das ist auch für die Perzeption nicht existent. Die Problematik der kategorialen Perzeption ist nach wie vor umstritten. Erst bei der technischen Synthese von Lauten und der Rede hat man die Bedeutung z.B. der Lautübergänge erkannt und bewiesen: Die Identifizierung gewisser Laute ist stark von den Merkmalen des Lautüberganges abhängig. Laute von ähnlicher akustischer Struktur können erst auf Grund der Unterschiede in ihren Übergängen zu der so genannten Klarphase des nachstehenden Lautes eindeutig identifiziert werden, z.B. die Realisationen der Phoneme /m/ und /n/.

7.4.4. Die Quantaltheorie

Die QT geht – grob formuliert – von der Prämisse aus, dass die Sprachen der Welt *in genere* diejenigen Areale des Ansatzrohres zur Erzeugung des Sprechschalls gebrauchen, in denen *geringe* artikulatorische Bewegungen (Kontinua) eine *große* akustisch-auditive Fluktuation (Diskontinua) hervorrufen. Man nehme als Beispiel den Übergang von [ʃ] und [s], wo an die *kontinuierlichen* Bewegungen des Apex (Präpalatum – Alveolen – Inzisiven) ein quantaler Sprung (Quant = nicht weiter teilbares Segment), ein abrupter akustisch-auditiver „Übersprung“ von [ʃ] zu [s] anknüpft. Leider können in diesem Theorierahmen lediglich solche akustisch-auditiven Fluktuationen erklärt werden, die auf sagittale Veränderungen des Ansatzrohres zurückzuführen sind; räumlichen Bewegungen wie der der Oberfläche des Dorsums wird nicht Rechnung getragen (vgl.: NEPPERT 1999: 300 f.).

7.4.5. Die Kohorten-Theorie

Bedeutungstragende Einheiten (Wörter) im Sprechschallkontinuum sind in Kohorten (= phonotaktischen und semantischen Feldern im Sprachbewusstsein) gebündelt. Das Wahrgenommene wird auf Grund sich in der Zeitfolge sukzessive verengender Kohorten dekodiert. Das Problem dieser Theorie besteht darin, dass die Artikulation vollständig übergangen wird.

Die Kohorten-Theorie oder das Kohorten-Modell gehört zu den Theorien bzw. Modellen der Worterkennung. Zur Lösung von Problemen der Sprachwahrnehmung und des Sprachverstehens entwickelte man Methoden und Modelle zuerst für die visuelle, später (etwa seit den 60er Jahren des 20. Jahrhunderts) auch für die akustische Wahrnehmung und die Verarbeitung von Sprachreizen durch den Perzipienten.

Der Kohorten-Theorie liegt die Vorstellung zugrunde, dass die bedeutungstragenden Einheiten im Kontinuum des Sprechschalls, auf Grund des mentalen Lexikons, in semantische Felder und phonetische Gruppen gebündelt sind. Die phonetischen Gruppen konstituieren sich nach den phonotaktischen Regeln der jeweiligen Sprache. Die in semantischen Feldern gruppierten und auf Grund der phonotaktischen Gesetzmäßigkeiten in phonetischen Klassen zusammengefassten Wörter bilden Kohorten.

Wahrgenommenes wird mit dem Referenzmuster zu den nach diesem Modell im sprachlichen Bewusstsein in Form von Kohorten gespeicherten Wörtern verglichen. Der Perzipient entscheidet, ob die wahrgenommenen Reize ein sinnvolles Wort ergeben oder nicht.

Im Kohorten-Modell wird durch den Input mit Hilfe des Initialreizes eine ganze „Kohorte“ von Wörtern aktiviert. Als Initialreiz gilt z.B. ein gemeinsames Element aller Wörter der „Kohorte“, etwa die erste Silbe, die also phonotaktisch aus der gleichen Lautstruktur zusammengesetzt ist. Bei der Worterkennung sind allerdings auch weitere – semantische, strukturelle usw. – Aspekte von Belang. Auf Grund z.B. der phonotaktischen Struktur funktioniert ein solches Modell wie folgt:

Konto

Kontur

konstituieren konstituieren konstituieren

Kontakt

Konstitution Konstitution Konstitution Konstitution

Konstrukt Konstrukt

Konstitut Konstitut Konstitut

 /kon/ /konst/ /konstitu/ /konstitution/

8. Literatur

- AUTORENKOLLEKTIV (1982): Großes Wörterbuch der deutschen Aussprache. (erste Auflage: Wörterbuch der deutschen Aussprache. 1964) Leipzig.
- AMMON, U. (1992): Varietäten des Deutschen. In: Offene Fragen – offene Antworten in der Sprachgermanistik. (Hg. ÁGEL, V. & HESSKY, R.) (=Reihe Germanistische Linguistik 128) Budapest (Lizenzausgabe für Ungarn), 203-224.
- BARBOUR, S. & STEVENSON, P. (1990): Variation in German: A critical approach to German sociolinguistics. Cambridge.
- BASSOLA, P. (1995): Deutsch in Ungarn – in Geschichte und Gegenwart. Heidelberg.
- BREITUNG, H. (1994): Phonetik – Intonation – Kommunikation. München.
- BRENNER, K. (1999): Magánhangzók kontrasztív akusztikai elemzése [Kontrastive Analyse von Vokalen]. In: Beszédkutatás '99 (Hg. GÓSY, M.), MTA Nyelvtudományi Intézete. Budapest, 44-52.
- BRENNER, K. (2002): Affrikaten konfrontatív – ein Vergleich ungarndeutsch vs. ungarisch. In: „das gueth von alten Lern“ Jugend-Festschrift für Karl Manherz zum 60. Geburtstag. (Hg. ULRICH LANGANKE) Budapest, 31-41.
- BRENNER, K. (2004): Plosive der deutschen Dialekte in West-Ungarn. Budapest (= Budapesti Beiträge zur Germanistik, Band 44).
- BOLLA, K. (1980): Magyar hangalbum. (=Magyar Fonetikai Füzetek 6) Budapest.
- DIETH, E. (1950): Vademekum der Phonetik. Berlin.
- ESSEN, O. von (1966): Allgemeine und angewandte Phonetik. 4. veränderte Auflage, Berlin.
- ESSEN, O. von (1979): Allgemeine und angewandte Phonetik. 5. neubearbeitete und erweiterte Auflage, Berlin.
- FANT, G. (1968). Analysis and synthesis of speech processes. In: Manual of Phonetics (Hg. MALMBERG, B.), Amsterdam, 173-276.
- FIEHLER, R. (2000): Gesprochene Sprache – gibt's die? In: Jahrbuch der ungarischen Germanistik 2000, Budapest/Bonn, 93-104.
- FIUKOWSKI, H. (1967): Sprecherzieherisches Elementarbuch. Leipzig.
- FREDRICH, R.-B. (1985): Übungen zur Phonetik der deutschen Sprache unter kommunikativem Aspekt. Halle (Saale).
- GÓSY, M. (1981): A szegmentális hangszerkezet percepciójáról. In: Magyar Fonetikai Füzetek (Hungarian Papers in Phonetics.) 8, Budapest, 87 – 103.
- GÓSY, M. (2004): Fonetika, a beszéd tudománya. Budapest.

- HARTH, K.-L. (1964): Deutsche Sprechübungen mit Ausspracheregeln. Fünfte, durchgesehene Auflage, Weimar.
- HARTH, K.-L. (1969): Gesprochenes Deutsch. Ein Arbeitsbuch über Wesen, Erscheinung und Technik gesprochenener Information. Weimar.
- HARTH, K.-L. (1973): Deutsch. Sprechen, Lesen, Vortragen, Reden. Ein Arbeitsbuch für Lehrer und Studierende. Weimar.
- HEIKE, G. (1982): Phonologie. Zweite Auflage., Stuttgart.
- HELL, G. (1994): Phonetik und Phonologie der deutschen Standardlautung. Stuhlweissenburg/Székesfehérvár.
- HENNIG, M. (2000): Gesprochene und geschriebene Sprache. In: Jahrbuch der ungarischen Germanistik 2000, Budapest/Bonn, 105-125.
- HESS, W. (2005): Grundlagen der Phonetik. Einführung: Grundlagen der akustischen Phonetik. Im Internet unter: http://www.ikp.uni-bonn.de/dt/lehre/materialien/grundl_phon/gph_1b.pdf (03. 05. 2006).
- HUMBOLDT, W. von (1883): Ueber die Verschiedenheit des menschlichen Sprachbaus und ihren Einfluß auf die geistliche Entwicklung des Menschengeschlechtes. 3. Aufl. Berlin.
- JAKOBSON, R./HALLE, M. (1960): Grundlagen der Sprache. Berlin.
- KELLER, R. (1990): Sprachwandel. Tübingen.
- KELZ, H.P./KUMMER, M./SEEL, H. (1989): Audiograph. Bonn.
- KOHLER, K. (1995): Einführung in die Phonetik des Deutschen. 1. Aufl. Berlin 1977; 2. neubearbeitete Auflage, Berlin.
- KOHLER, K. (1999): German. In: Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet. Cambridge, 86 – 89.
- KOLB, B. & WHISHAW, I. Q. (1996): Neuropsychologie. 2. Auflage. Übersetzung herausgegeben von Monika Pritzel. Aus dem Englischen übersetzt von Marianne Mauch, Monika Niehaus-Osterloch und Markus Numberger. Heidelberg, Berlin, Oxford.
- KOLLMEIER, B. (2006): Audiologie. Im Internet unter: http://medi.uni-oldenburg.de/html/teaching_download.html (gesehen am 19.06.2006).
- LAZICZIUS, G. (1963): Fonetika. Budapest.
- LAZICZIUS, G. (1979): Fonetika. Budapest.
- LINDNER, G. (1969): Einführung in die experimentelle Phonetik. Berlin.
- LINDNER, G. (1975): Der Sprechbewegungsablauf. Eine phonetische Studie des Deutschen. Berlin.
- LINDNER, G. (1976): Leit- und Funktionsschemata des Sprech- und Hörprozesses. In: Einführung in die Sprechwissenschaft, Leipzig, 18 – 67.

- LINDNER, G. (1981): Grundlagen und Anwendung der Phonetik. Berlin.
- MARTENS, C. & MARTENS, P. (1965): Phonetik der deutschen Sprache. München.
- MARTINET, A. (1963): Grundzüge der allgemeinen Sprachwissenschaft. Stuttgart.
- NEPPERT, J. & PÉTURSSON, M. (1986): Elemente einer akustischen Phonetik. Hamburg.
- NEPPERT, J. (1999): Elemente einer Akustischen Phonetik. 4., vollständig neu bearbeitete Auflage. Hamburg: Buske.
- OLASZY, G. (1989): Elektronikus beszédelőállítás. A magyar beszéd akusztikája és formánsszintézise. Budapest.
- PAUKA, K. (1980): Halláslélektan. A beszédmegértés alaptényezői. Budapest: Tankönyvkiadó.
- PAUL, H. (1909): Prinzipien der Sprachgeschichte. 4. Aufl. Halle.
- PÉTURSSON, M. & NEPPERT, J. (1996): Elementarbuch der Phonetik. Hamburg (2., durchgesehene und erweiterte Auflage.
- PÉTURSSON, M. & NEPPERT, J. (2002): Elementarbuch der Phonetik. Hamburg (3., durchgesehene und erweiterte Auflage.
- PIKE, K. (1947): Phonemics. Ann Arbor/Mich..
- POMPINO-MARSCHALL, B. (2003): Einführung in die Phonetik. Berlin, New York.
- RAUSCH, R. & RAUSCH, I. (1988): Deutsche Phonetik für Ausländer. Ein Lehr- und Übungsbuch. Leipzig.
- SCHRAMM, E. & SCHMIDT, L. (1977): Übungen zur deutschen Aussprache. Leipzig.
- SCHUBIGER, M. (1977): Einführung in die Phonetik. 2., überarb. Aufl. Berlin.
- SCHWITALLA, J. (1997): Gesprochenes Deutsch. Eine Einführung. Berlin.
- SIEBS, Th. (1969): Deutsche Aussprache. Reine und gemäßigte Hochlautung mit Aussprachewörterbuch, 19. umgearbeitete Auflage. Berlin.
- SIEVERS, E. (1924): Ziele und Wege der Schallanalyse. Heidelberg.
- STOCK, E. (1976): Laut und Intonation. In: Einführung in die Sprechwissenschaft, Leipzig, 68 – 143.
- STOCK, E. & ZACHARIAS, C. (1972): Deutsche Satzintonation. Leipzig.
- SURJÁN, L. & FRINT, T. (Hg.) (1982): A hangképzés és zavarai, beszédzavarok. [Foniátria.] Második, bővített kiadás. Budapest.
- SZENDE, T. (1982): A mai magyar nyelv fonémái. In: Fejezetek a magyar leíró hangtan köréből. (Hg. BOLLA, K.) Budapest, 233-65.
- TILLMANN, H. & MANSELL, P. (1980): Phonetik. Lautsprachliche Zeichen, Sprachsignale und lautsprachlicher Kommunikationsprozeß. Stuttgart.

- TILLMANN, H. & SCHIEL, F. (1995): Akustische Phonetik. Begleitendes Hypertext-Dokument zur Vorlesung. Im Internet unter: <http://www.phonetik.uni-muenchen.de/AP/APHome.html> (gesehen am 19.06.2006).
- THOMPSON, R. F. (2001): Das Gehirn. Von der Nervenzelle zur Verhaltenssteuerung. 3. Auflage. Berlin.
- TREPEL, M. (2004): Neuroanatomie. Struktur und Funktion. 3. Auflage. München, Jena.
- TRUBETZKOY, N. S. (1989): Grundzüge der Phonologie. 7. Auflage Göttingen.
- VALACZKAI, L. (1976): Abriß einer ungarisch-deutschen kontrastiven Phonetik. Budapest.
- VALACZKAI, L. (1990): Übungen zur deutschen Phonetik. Szeged.
- VALACZKAI, L. (1992): Theoretische Probleme und praktische Darstellung der akustischen Projektion der Artikulation. In: Jahrbuch der ungarischen Germanistik. Budapest/Bonn, 227 – 238.
- VALACZKAI, L. (1998): Atlas deutscher Sprachlaute. Instrumentalphonetische Untersuchung der Realisierung deutscher Phoneme als Sprechlaute. Wien.
- VALACZKAI, L. & BRENNER, K. (1999): Abriss einer deutschen Phonetik. Handbuch für das Germanistikstudium mit deutsch-ungarischen Materialien. Veszprém: Universitätsverlag.
- VENNEMANN, T. (1986): Neuere Entwicklungen in der Phonologie. Berlin/New York/Amsterdam.
- VÉRTES O., A. (1982): Az artikuláció akusztikus vetülete. In: Fejezetek a magyar leíró hangtanból (Hg. BOLLA, K.). Budapest, 155–163.
- WÄNGLER, H.-H. (1976): Atlas deutscher Sprachlaute. 6. Aufl. Berlin.
- WÄNGLER, H.-H. (1983): Grundriß einer Phonetik des Deutschen. Marburg 1960, 4. überarbeitete Auflage.
- WINTERLING, F. (Hg.) (1987): Ursprünge der Sprache. Frankfurt am Main.
- ZILLES, K. & REHKÄMPER, G. (1998): Funktionelle Neuroanatomie. Lehrbuch und Atlas. 3., korrigierte Auflage. Berlin, Heidelberg, New York.

Abbildungsverzeichnis

- NEPPERT, J. M. H. (1999): Elemente einer Akustischen Phonetik. Hamburg, Abb. 26
 PÉTURSSON, M. & NEPPERT, J. (2002): Elementarbuch der Phonetik. Hamburg 1991 (3., durchgesehene und erweiterte Auflage), Abb. 5
 POMPINO-MARSCHALL, B. (2003): Einführung in die Phonetik. Berlin, New York, Abb. 1, 2, 3, 4, 9
 REETZ, H. (1999): Artikulatorische und akustische Phonetik. Trier, Abb. 10, 11

KONSONANTEN (PULMONAL)

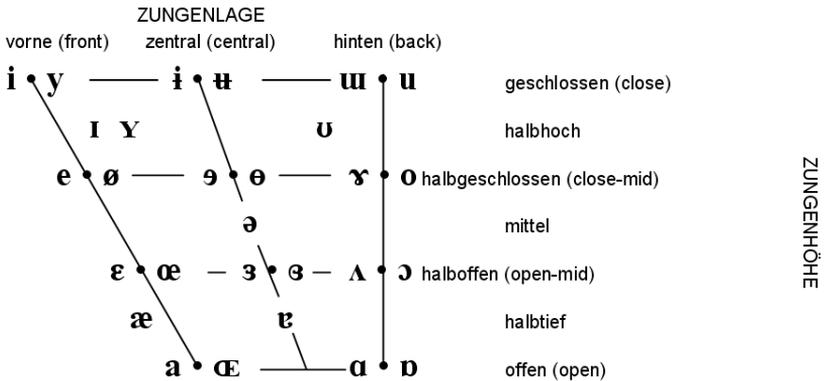
	bilabial		labiodental		dental		alveolar		post- alveolar	retroflex		palatal		velar		uvular		pharyngal		glottal		
plosiv	p	b					t	d			ʈ	ɖ	c	ɟ	k	g	q	ɢ			ʔ	
nasal		m		ɱ				n				ɳ	ɲ		ŋ	ɴ						
vibrant		ʙ						r									ʀ					
gebippt/ geschlagen							ɾ				ɽ											
frikativ	ɸ	β	f	v	θ	ð	s	z	ʃ	ʒ	ʂ	ʐ	ç	ʝ	x	χ	ʁ	ħ	ʕ	h	ɦ	
lateral- frikativ							ɬ	ɮ														
approximant				ʋ			ɹ				ɻ	j		ɰ								
lateral- approximant							l				ɭ	ʎ		ʟ								

Bei paarweisen Symbolen kennzeichnet das rechte den stimmhaften Konsonanten.
 Schraffierte Flächen kennzeichnen unmögliche Artikulationen.

KONSONANTEN (NICHT PULMONAL)

Clicks		stimmhafte Implosive		Ejektive	
⊙	bilabial	ɓ	bilabial	ʼ	Diakritikum, wie in:
	dental	ɗ	dental/ alveolar	pʼ	bilabial
!	(post-) alveolar	ɟ	palatal	tʼ	dental/ alveolar
‡	palato- alveolar	ɡ	velar	kʼ	velar
	alveolar lateral	ɠ	uvular	sʼ	alveolar frikativ

VOKALE



SUPRASEGMENTALIA

'	Hauptbetonung	
ˊ	Nebenbetonung	
	ˌfounə'tɪʃən	
ː	lang	eː
ˑ	halblang	eˑ
◌̣	extrakurz	ě
	kürzere (Takt-/Fuß-)Gruppe	
	größere (Intonations-)Gruppe	
.	Silbengrenze	ˌi.ækt
◌̣	verschliffen (fehlende Grenze)	

TÖNE UND WORTAKZENTE

STUFEN			KONTUREN		
é or	⌈	extra hoch	ě or	⌈	Steigend
é	↑	hoch	ê	∨	Fallend
ē	⊣	mittel	ě	↑	hoch steigend
è	↓	tief	è	⌋	tief steigend
è	⌋	extra tief	ě	∧	steigend fallend
↓		Downstep	↗		globaler Anstieg
↑		Upstep	↘		globaler Abfall



A projekt az Európai Unió támogatásával,
a Nemzeti Fejlesztési Terv keretében valósult meg.
A felsőoktatás szerkezeti és tartalmi fejlesztése
HEFOP-3.3.1-P.-2004-09-0134/1.0