

3 Phonetik und Phonologie: Die Lehre von den Lauten der Sprache

- 3.1 Laute, Schriftzeichen und phonetisches Alphabet
- 3.2 Der Laut – phonetische und phonologische Betrachtungsweise
- 3.3 Phonetische Grundlagen – die Erzeugung der Laute
 - 3.3.1 Erzeugung des Luftstroms
 - 3.3.2 Lautbildung im Kehlkopf
 - 3.3.3 Artikulation der Konsonanten
 - 3.3.3.1 Artikulationsorgane
 - 3.3.3.2 Artikulationsstelle und Artikulationsart
 - 3.3.4 Artikulation der Vokale
- 3.4 Laute im Kontext des Sprechens – Koartikulation
- 3.5 Die phonologische Abstraktion
 - 3.5.1 Segmentierung
 - 3.5.2 Minimalpaarbildung
 - 3.5.3 Kombinatorische und freie Varianten
 - 3.5.4 Phoneme als Bündel distinktiver Merkmale
 - 3.5.5 Phonotaktik
 - 3.5.6 Zum Status phonologischer Analysen

3.1 Laute, Schriftzeichen und phonetisches Alphabet

Natürliche Sprachen sind Kommunikationsinstrumente; sie dienen – grob gesprochen – der Übermittlung sprachlich gefasster Inhalte von einem Kommunikationspartner zum anderen. Um diese Inhalte transportieren zu können, bedarf es eines materiellen Trägers, denn das, was wir mitteilen wollen, muss ja irgendwie aus unseren Köpfen herauskommen. Wenn wir uns schriftlich äußern, verwenden wir dazu die Schriftzeichen unserer Buchstabenschrift; bei der mündlichen Kommunikation produzieren wir dagegen Laute.

In diesem Kapitel geht es um die Laute, die man als das gebräuchlichste und natürlichste Medium der Sprache ansehen kann, und zwar aus den folgenden Gründen: Zum einen ist längst nicht jede Sprache schriftlich fixiert; ja bezogen auf die Gesamtheit der Sprachen stellt die Verschriftung sogar eher den Ausnahmefall dar und ist keineswegs so selbstverständlich, wie es uns als Angehörige einer traditionsreichen Schriftgemeinschaft erscheinen mag. Zum zweiten erwirbt jedes gesunde Kind zunächst die Lautsprache, während der schriftliche Modus erst wesentlich später durch gezielte Unterweisung erlernt wird. In diesem

Sinne ist die Schrift gegenüber der Lautung sekundär. Selbstverständlich heißt das nicht, dass die Schrift weniger wichtig ist, denn sie eröffnet sogar zusätzliche Kommunikationsmöglichkeiten: Während nämlich eine lautsprachliche Äußerung aufgrund ihrer materiellen Eigenschaften bereits im Zuge ihrer Verfertigung kontinuierlich verschwindet, kann eine „schwarz auf weiß“ fixierte Mitteilung räumliche und zeitliche Entfernungen überwinden. Man kann sie an einem beliebigen Ort und zu einer beliebigen Zeit lesen, völlig unabhängig davon, wo und wann sie produziert wurde. Salopp formuliert: „Der Witz der Schrift ist die Ausdehnung des Kommunikationsradius“ (COULMAS 1985, 107).

Streng von der Schrift zu unterscheiden sind **phonetische Umschriften** (auch **phonetische Transkriptionen** genannt), wie man sie aus Sprachlehrwerken und zweisprachigen Wörterbüchern kennt. Der Unterschied liegt in der Funktion: Man verwendet eine phonetische Umschrift nicht für die schriftliche Kommunikation über außersprachliche Sachverhalte, sondern als reines Beschreibungsinstrument, mit dessen Hilfe lautliche Phänomene erfasst und verschriftet werden – z. B. um sich über sie verständigen zu können, wie das in diesem Kapitel geschieht.

Warum man zu diesem Zweck nicht die wesentlich geläufigeren Notierungskonventionen der Buchstabenschrift benutzt, hat einen einfachen Grund: Zwar handelt es sich bei Buchstabenschriften um Schriftsysteme, die systematisch auf Laute bezogen sind (vgl. Kap. 8.1), aber sie sind keineswegs so konzipiert, dass sie die Lautfolge eines gesprochenen Wortes exakt abbilden. So werden in bestimmten Fällen einfache Laute durch Buchstabenfolgen wiedergegeben (etwa der erste Laut in *schwarz* durch die Buchstabenfolge <sch>), oder ein einzelner Buchstabe steht für eine Folge aus mehreren Lauten (etwa der Buchstabe <x> für die Lautfolge [ks] in *Hexe*). Und noch viel hinderlicher für die Zwecke der reinen Lautabbildung ist der Umstand, dass ein und derselbe Laut nicht immer mit ein und demselben Schriftzeichen korrespondiert. So schreiben wir den Buchstaben <g> sowohl in *grau* als auch in *Tag*, obwohl damit – wie jeder leicht nachvollziehen kann – unterschiedliche Laute bezeichnet werden.

Umgekehrt müssen wir ein und denselben Laut bisweilen unterschiedlich ausbuchstabieren, etwa den jeweiligen Anfangslaut in den Wörtern *Vater* und *Fisch*. Kurz: Die Beziehungen zwischen Lauten und Buchstaben sind keineswegs eindeutig. Die Gründe dafür sollen uns an dieser Stelle nicht weiter interessieren (vgl. aber Kap. 8.4.1 – 8.4.4). Wichtig ist nur, dass man für die exakte Wiedergabe des Lautlichen ein Zeichen-

inventar braucht, das für je einen Laut auch jeweils ein graphisches Symbol bereithält. Ein solches Inventar wird als **phonetisches Alphabet** bezeichnet. Am verbreitetsten und gebräuchlichsten ist das der **International Phonetic Association** (kurz **IPA**), dessen erste Fassung bereits Ende des 19. Jahrhunderts erarbeitet worden ist; seitdem hat man es beständig weiterentwickelt und in revidierten Versionen neu publiziert.

Bei der Durchsicht des IPA-Alphabets darf man sich nicht dadurch verwirren lassen, dass die graphischen Symbole z. T. mit den Schriftzeichen des lateinischen Alphabets übereinstimmen. Aus Gründen der leichteren Handhabbarkeit hat man nämlich darauf verzichtet, ein völlig neues Zeicheninventar zu entwickeln, und stattdessen zunächst die Buchstaben des lateinischen Alphabets verwendet. Aber Vorsicht: Zumindest in einigen Fällen ist der Lautwert der IPA-Symbole völlig anders, als wir das aus unserer Buchstabenschrift gewohnt sind. So steht etwa das IPA-Symbol [z] für den stimmhaften Anfangslaut in Wörtern wie *sehen, Sahne, Sonne* etc., der in der Buchstabenschrift mit <s> bzw. <S> wiedergegeben wird. Um die beiden Ebenen nicht zu verwechseln, verwendet man unterschiedliche Notierungskonventionen: Zeichen der Lautschrift stehen in eckigen Klammern, Buchstaben in spitzen.

Da das lateinische Alphabet nur eine begrenzte Anzahl von Schriftzeichen zur Verfügung stellt, die für die exakte Kennzeichnung nicht ausreicht, ist die IPA-Schrift gegenüber dem Ausgangsalphabet von Beginn an systematisch erweitert worden. Zum einen durch Zeichen, die aus den Buchstaben des lateinischen und griechischen Alphabets abgeleitet sind, z. B. [ŋ] (wie in *Wange*) oder [ʃ] (wie in *schwarz*); zum anderen durch so genannte **diakritische Zeichen**, „die an ein alphabetisches Schriftzeichen angehängt oder mit ihm kombiniert werden können, um bestimmte lautliche Unterschiede auszudrücken“ (HENGARTNER/NIEDERHAUSER 1993, 64). So dient im IPA beispielsweise ein Doppelpunkt als Zeichen für einen lang ausgesprochenen Vokal (z. B. [i:]) oder ein hochgestelltes kleines ^h als Zeichen für eine behauchte Aussprache von Konsonanten (vgl. 3.3.2).

Das komplette IPA-Inventar mitsamt diakritischen Zeichen findet man u. a. im DUDEN-Aussprachewörterbuch (vgl. Literaturliste). Noch ein Hinweis: Wenn im Folgenden zu Illustrationszwecken IPA-Zeichen verwendet werden, die von den bekannten Buchstaben des lateinischen Alphabets abweichen oder deren Lautwert nicht mit dem der Buchstaben übereinstimmt, wird durch ein in Klammern gesetztes Beispielwort gezeigt, welcher Laut gemeint ist. Entspricht der Lautwert des IPA-Zeichens dagegen im Großen und Ganzen dem des Buchstabens, verzichten wir auf die Erläuterung.

3.2 Der Laut – phonetische und phonologische Betrachtungsweise

Laute können unter verschiedenen Gesichtspunkten betrachtet werden, wobei zwei Betrachtungsweisen grundlegend sind: die **phonetische** und die **phonologische**. Die Kluft zwischen den beiden Perspektiven ist allerdings nicht so tief, wie man das in der sprachwissenschaftlichen Forschung lange Zeit gesehen hat. Ganz sicher handelt es sich nicht um zwei Welten, sondern um zwei auf vielfältige Weise miteinander verbundene und aufeinander angewiesene Forschungszweige, die in ständiger Wechselbeziehung stehen. Die beiden Disziplinen sollen im Folgenden kurz vorgestellt werden.

Die Phonetik

Die Phonetik nähert sich den Lauten aus einer eher naturwissenschaftlichen Perspektive. Ihr „zentrale[s] Anliegen [...] ist die Frage nach der Rolle der *lautlichen Substanz* im sprachlichen Kommunikationsprozeß“ (PÉTURSSON/NEPERT 1991, 15), wobei unter lautlicher Substanz alle physiologischen und physikalisch messbaren Aspekte der Lautereignisse verstanden werden. Entsprechend den drei „Eckpfeilern“ einer Kommunikationskette – Sprecher, Lautprodukt und Hörer – geht es dabei konkret um die folgenden Fragestellungen:

- (a) Wie werden die Laute vom Sprecher produziert, d. h. welche Organe sind in welcher Weise an der Bildung von Lauten beteiligt?

Wenn Sprecher Laute produzieren, ist das im Grunde nichts weiter als ein biomechanischer Vorgang, den man auch als solchen beschreiben kann. So wird beispielsweise ein [p] dadurch gebildet, dass ein von der Lunge ausgestoßener Luftstrom sich nach oben durch die Luftröhre fortpflanzt, dabei ungehindert den Kehlkopf passiert, im Mundraum vor den Lippen gestaut wird, bis er durch das plötzliche Lösen des Lippenverschlusses aus dem Mund entweichen kann. Eine solche Beschreibung der Lautproduktion, die man in vergleichbarer Weise auch für alle übrigen Laute liefern kann, wird im Rahmen der **artikulatorischen Phonetik** geleistet.

- (b) Welche akustischen Eigenschaften hat das vom Sprecher produzierte Schallereignis?

Bei den vom Sprecher produzierten Lauten handelt es sich physikalisch betrachtet um Schallwellen. Schallwellen wiederum sind mechanische Schwingungen, an denen eine Reihe charakteristischer Größen gemes-

sen werden kann: z. B. die Anzahl der Schwingungen pro Sekunde (**Frequenz**), die in Hertz (Hz) angegeben wird und die u. a. als Tonhöhe wahrgenommen wird; oder die Schwingungsweite, d. h. die vertikale Ausdehnung einer Welle von der Ruhestellung aus gesehen (**Amplitude**), die man in Dezibel (dB) misst und die für das Lautstärkeempfinden verantwortlich ist. Mit dem Sprachschall als akustischem Phänomen ist die **akustische Phonetik** befasst.

(c) Wie verarbeitet der Hörer die vom Sprecher produzierte akustische Reizquelle?

Das letzte Glied in der Kommunikationskette ist der Hörer, der die vom Sprecher produzierte sprachliche Äußerung mit seinem Ohr aufnimmt. Dabei müssen die am Trommelfell ankommenden Schallwellen zunächst über die Gehörknöchelchen ins Innenohr übertragen werden, bevor die akustische Information dort in Nervenimpulse umgewandelt und über den Hörnerv ins Gehirn weitergeleitet wird. Erst dort, im so genannten **Wernicke-Areal**, einer ziemlich ausgedehnten Gehirnregion in der Nähe des linken Ohrs, findet die eigentliche **Lauterkennung** statt. Mit den eher physiologischen Aspekten dieses Prozesses beschäftigt sich die **auditive Phonetik**; die mehr psychologischen Momente werden von der Sprachperzeptionsforschung bzw. der Wahrnehmungspsychologie bearbeitet, wobei die Grenzen zur auditiven Phonetik natürlich fließend sind.

Die Phonologie

Mit dem Stichwort Lauterkennung haben wir im Grunde schon den Übergang zur phonologischen Betrachtungsweise geschaffen, denn die menschliche Sprachwahrnehmung funktioniert nicht wie ein Automat, der passiv alle physikalisch messbaren Lauteigenschaften registriert; die Laute gehen bei der Sprachwahrnehmung vielmehr durch einen phonematischen Filter.

Damit ist Folgendes gemeint: Laute werden in einer unendlichen Bandbreite von Variationen produziert. Das fängt bei scheinbaren Trivialitäten an: Beispielsweise unterscheidet sich eine „hohe“ Frauenstimme rein physikalisch-akustisch betrachten erheblich von einer „tiefen“ Männerstimme; das gleiche gilt für die Stimme eines alten und eines jungen Menschen, und selbst eine banale Erkältung wirkt sich in erkennbarer Weise auf das „akustische Gesicht“ der produzierten Laute aus. Streng genommen sind sogar die Lautäußerungen ein und derselben Person nie identisch, denn wir sprechen immer unterschiedlich, je nachdem, ob wir freudig, traurig, müde, nervös oder sonst wie gestimmt sind.

Deswegen kann selbst der Phonetiker, dem es ja gerade darum geht, die Laute in ihren substantiellen Eigenschaften zu beschreiben, unmöglich alle potentiell registrierbaren lautlichen Nuancen erfassen. Wenn er beispielsweise ein Lautereignis als ein mit der Zungenspitze gebildetes [r] klassifiziert, dann sind für ihn Geschlecht, Alter oder Befindlichkeit des Sprechers ohne Bedeutung. Ebenso unwichtig ist es für ihn, ob der Laut langsam oder schnell, mit Emphase oder eher gedämpft artikuliert wird, ob er laut oder leise gesprochen wird etc. Mit anderen Worten: Der Phonetiker erbringt bei seiner Arbeit eine gehörige Abstraktionsleistung, denn er lässt bestimmte, physikalisch durchaus messbare Eigenschaften der Laute unberücksichtigt und fasst unterschiedliche Lautereignisse zu einer Kategorie zusammen.

Eine prinzipiell vergleichbare Abstraktionsleistung erbringt auch der Hörer im normalen Kommunikationsprozess. Während aber der Phonetiker von der lautlichen Variation nur insoweit absieht, als sie auf individuelle Dispositionen im eben skizzierten Sinne zurückzuführen ist, reicht die Abstraktionsleistung des Hörers wesentlich weiter. So wird der Phonetiker darauf achten, in welcher der drei folgenden Varianten der Anfangslaut des Wortes *Ratte* gebildet wird:

- a) durch die bereits erwähnte schlagende Bewegung der Zungenspitze an den Zahndamm (Zungenspitzen-r),
- b) durch eine ähnliche Bewegung des Gaumenzäpfchens gegen den hinteren Zungenrücken (Zäpfchen-r) oder
- c) durch eine von Zäpfchen und hinterem Zungenrücken gebildete Engstelle, die der ausströmenden Atemluft ein charakteristisches Reibegeräusch verleiht.

In jedem einzelnen Fall wird der Phonetiker das Ergebnis der Lautproduktion anders werten und entsprechend mit einem eigenen Symbol des phonetischen Alphabets festhalten: das Zungenspitzen-r mit dem Symbol [r], das Zäpfchen-r mit [ʀ] und das Reibe-r mit [ʁ]. Aus seiner Perspektive ist das auch notwendig, denn es handelt sich ja ganz offenkundig um unterschiedliche Lautereignisse, wobei die lautliche Varianz auf keinen der oben angeführten Faktoren (Geschlecht, Alter, Befindlichkeit etc.) zurückzuführen ist.

Für den Hörer ist der Unterschied, soweit er am Inhalt der Äußerung interessiert ist, ebenso belanglos wie die Frage, ob er mit einem alten oder jungen, einem weiblichen oder männlichen, einem erkälteten oder gesunden Menschen spricht. Er nimmt gewissermaßen ein r „an-sich“

wahr, ganz gleich, welcher von den drei r-Lauten produziert wird. Der Grund liegt darin, dass für den Hörer lautliche Unterschiede unwichtig sind, solange damit keine Unterschiede in der übermittelten Bedeutung verbunden sind. Bei den verschiedenen Realisierungsweisen des r-Lautes ist das ganz offenkundig nicht der Fall, denn ganz gleich, mit welchem r-Laut *Ratte* ausgesprochen wird – für den Hörer bleibt es ein und dasselbe Wort mit ein und derselben Bedeutung.

Das klingt für sich genommen ziemlich trivial. Wie wenig selbstverständlich dieser Befund ist, erkennt man erst, wenn man einmal über den engen Tellerrand der eigenen Sprache blickt. Im Portugiesischen beispielsweise sind die Verhältnisse grundsätzlich anders. Hier kann man zwei der genannten r-Varianten, nämlich Zungenspitzen- und Zäpfchen-r, nicht bedenkenlos gegeneinander austauschen, ohne dass damit eine Bedeutungsveränderung einhergeht. So entnimmt der portugiesische Hörer der mit Zungenspitzen-r gesprochenen Lautfolge [karu] die Bedeutung 'teuer', 'lieb'. Die mit Zäpfchen-r gesprochene, ansonsten aber identische Lautfolge [karu] verbindet er dagegen mit der Bedeutung 'Karren', d. h. er erkennt darin die lautliche Seite eines völlig anderen Wortes. Anders gesagt: Für ihn ist der Unterschied zwischen Zungenspitzen-r und Zäpfchen-r bedeutungsunterscheidend.

Ein deutscher Hörer ohne entsprechende Portugiesischkenntnisse würde den lautlichen Unterschied dagegen entweder gar nicht bemerken oder ihn in Analogie zum Deutschen unter der Rubrik „lautliche Variante“ buchen.

Daraus wird ersichtlich, dass die Sprecher/Hörer unterschiedlicher Sprachen Laute in ganz unterschiedlicher Weise Kategorien zuordnen. Laute, die für den deutschen Sprecher lediglich Varianten ein und derselben Kategorie sind, behandelt der portugiesische Sprecher/Hörer als Vertreter zweier unterschiedlicher Kategorien – und das unbeschadet der Tatsache, dass sich die vergleichbaren Laute der beiden Sprachen in ihrer Bildungsweise nicht unterscheiden.

Ein ganz ähnlicher Abstraktionsprozess liegt der Arbeit eines Sprachwissenschaftlers zugrunde, der die Laute einer Sprache aus der phonologischen Perspektive betrachtet. Denn in der **Phonologie** geht es letztlich darum, die prinzipiell unendliche Zahl lautlicher Varianten auf eine endliche Zahl lautlicher Kategorien zu reduzieren. Und das sind genau diejenigen, die vom Standpunkt der jeweiligen Sprache bedeutungsunterscheidend sind. Diese Lautkategorien nennt man **Phoneme**. Wie man sieht, ist Phonologie also „nicht einfach ein linguistisches Schreibtischkonstrukt“, sondern „im Prinzip nichts weiter als der Versuch, [...] Regularitäten der Sprachwahrnehmung [...] bewusst zu machen, zu systematisieren und mittels sprachwissenschaftlicher Methoden zu beschreiben“ (HENGARTNER/NIEDERHAUSER 1993, 44). Welche Methoden das sind, werden wir in Abschnitt 3.5.2 noch sehen. Zunächst wollen wir uns aber der phonetischen Ausgangsbasis zuwenden, auf deren Grundlage der phonologische Abstraktionsprozess vollzogen wird.

3.3 Phonetische Grundlagen – die Erzeugung der Laute

Die phonetische Beschreibung der Laute kann – wie oben angesprochen – auf drei unterschiedliche Weisen geschehen, artikulatorisch, akustisch und auditiv. Für unsere Zwecke reicht die artikulatorische Beschreibung völlig aus, denn auch die Phoneme, auf die wir letztlich hinauswollen, werden in der Regel durch artikulatorische Merkmale definiert. Wir müssen also zunächst die Frage klären, wie Laute überhaupt erzeugt werden. Dabei sind drei zeitlich aufeinander folgende Prozesse zu unterscheiden:

- a) die Erzeugung des Luftstroms in der Lunge,
- b) die Lautbildung bzw. Bildung von Lautmerkmalen im Kehlkopf und
- c) die Artikulation der Laute im Rachen-, Mund- oder Nasenraum.

3.3.1 Erzeugung des Luftstroms

Laute sind – wie oben bereits erwähnt – Schallwellen, die sich in elastischen Medien wie Gasen oder Flüssigkeiten ausbreiten. Das für die Lautproduktion relevante Trägermedium ist die Atemluft, die entweder ausgestoßen oder eingesogen werden kann. Im Deutschen und in den allermeisten anderen Sprachen werden die Laute fast ausschließlich beim Ausatmen erzeugt; man spricht von **expiratorischer Lautbildung**. Das Pendant, die **inspiratorische Lautbildung**, bei der die Laute beim Einatmen produziert werden, kommt in diesen Sprachen dagegen nur in Ausnahmefällen vor. Meistens hat diese Art der Lautbildung dann eine über die reine Sachverhaltsvermittlung hinausgehende kommunikative Funktion. So soll ein mit eingesogener Luft ausgesprochenes *ja* oder *nein* dem Gesprächspartner gewöhnlich signalisieren, dass seine Frage wenig Sinn macht oder die Antwort ohnehin selbstverständlich ist.

Verschiedene Sprachen machen im Übrigen von der Möglichkeit Gebrauch, „ohne Nutzung des Luftstroms Laute zu erzeugen“ (CRYSTAL 1993, 126); so z. B. eine Reihe südafrikanischer Sprachen, in denen so genannte **Schnalzlaut**e vorkommen, die als scharfe, unter Beteiligung der Zunge und des Gaumens gebildete Saugeräusche wahrgenommen werden.

3.3.2 Lautbildung im Kehlkopf

Bei der expiratorischen Lautbildung pflanzt sich der von der Lunge ausgestoßene Luftstrom durch die Luftröhre (**Trachea**) nach oben fort und trifft dabei auf den Kehlkopf, der sich am oberen Ende der Luftröhre befindet. Der Kehlkopf ist ein Knorpelgerüst, in dem sich zwei nach hinten

gezogene Schleimhautausstülpungen befinden, die **Stimmlippen**, deren Innenränder man **Stimmbänder** nennt. Die Stimmlippen sind variabel; sie können weit auseinander liegen, zu einem Spalt verengt sein oder sich ganz schließen. Diesen variablen Raum zwischen den Stimmlippen bezeichnet man als **Stimmritze** bzw. **Glottis**. Je nach Stellung der Glottis werden im Kehlkopf die folgenden Laute bzw. Lautmerkmale erzeugt:

(1) Bei leicht geöffneter Glottis reibt die ausströmende Luft an den Stimmbändern, und es bildet sich ein Hauchgeräusch, das dem Laut [h] entspricht, wie er am Anfang der Wörter *Hund, Hemd, Hirte* etc. gesprochen wird.

(2) Ist die Glottis vollkommen geschlossen und wird dann durch den unterhalb der Glottis entstehenden Luftdruck plötzlich gesprengt, entweicht die gestaute Luft explosionsartig. Dadurch entsteht ein **Knacklaut** (auch: **glottaler Verschlusslaut** oder **Glottisschlag**), der im Deutschen einen **gepressten Vokaleinsatz** bewirkt. Man nimmt diesen Laut gewöhnlich nicht wahr; man kann sich seine Existenz aber bewusst machen, wenn man sich das Wort *Spiegelei* und den zweiten Bestandteil von *(Eulen)spiegelei* langsam vorspricht: Der Unterschied besteht bei ansonsten gleicher Lautkette nur darin, dass das *-ei* in *Spiegelei* „hart“ einsetzt, die nämliche Lautfolge in *(Eulen)spiegelei* aber „weich“ an das vorangehende[l] gebunden ist (das Symbol für den Knacklaut im IPA-Kode ist [ʔ]).

Eine weitere „gute Selbstkontrolle zur Ermittlung des Knacklautes liefert das Befühlen des Kehlkopfs mit Daumen und Zeigefinger: Der Knacklaut ist bei hartem Stimmeinsatz vor Vokalen am Silbenanfang deutlich als kleine Schlagbewegung spürbar“ (HENGARTNER/NIEDERHAUSER 1993, 30).

(3) Ist die Glottis zunächst ganz geöffnet, schließt sich dann aber langsam, entsteht ein leichter Hauchlaut. Dieser Hauchlaut „steht aber nicht für sich“, sondern begleitet im Deutschen in bestimmten Stellungen die Artikulation der Konsonanten [p], [t] und [k], z. B. am Wortanfang vor Vokal (*Piste, Taste, Küste* etc.). Man nennt sie folgerichtig **behauchte** bzw. **aspirierte Konsonanten** und markiert sie bei der Transkription mit dem entsprechenden diakritischen Zeichen als [p^h], [t^h] und [k^h]. In anderen Stellungen, z. B. vor [s], sind die nämlichen Konsonanten dagegen in der Regel unbehaucht (*Raps, mopsen, Erbse* etc.).

Die **Aspirierung** wird von deutschen Sprechern/Hörern ähnlich wie der Knacklaut gewöhnlich nicht bewusst wahrgenommen. Das deutet darauf hin, dass sie zu den phonetischen Eigenschaften deutscher Laute zählt, die man bei der Sprachwahrnehmung getrost „überhören“ kann, weil sie bei der Lauterkennung durch das phonologische Sieb fallen. Wir werden in 3.5.2 darauf zurückkommen.

(4) Der wichtigste Effekt, der mit den Stimmlippen und der Glottis erzielt werden kann, ist aber zweifellos die **Stimmbildung** oder **Phonation**. Sie entsteht dadurch, dass die Stimmlippen an der zu einem winzigen Spalt geöffnete Glottis durch ausströmende Luft in Schwingungen versetzt werden. Genau genommen sind die Schwingungen nichts weiter als periodisch auftretende

Verschlusslösungen, die „in einer eine gewisse Mindestzeit andauernden Folge als Stimmklang wahrgenommen werden“ (PÉTURSSON/NEPPERT 1991, 72).

Man kann die Stimmlippenschwingungen leicht spürbar bzw. hörbar machen, und zwar auf zweierlei Weise. Zum einen, indem man Zeigefinger und Daumen an den Kehlkopf legt und nacheinander Wortpaare wie *leasen* und *ließen* ausspricht. In *leasen* fühlt man beim s-Laut in der Wortmitte ein deutliches Vibrieren der Stimmlippen, bei *ließen* dagegen nicht. Zum anderen, indem man die Ohren mit den Fingern schließt und dabei die genannten Wörter artikuliert. Auch hier sind die Schwingungen beim entsprechenden Laut in *leasen* deutlich spürbar (Symbol im IPA-Kode: [z]).

Bei der Erzeugung des stimmlosen Gegenstücks (Symbol im IPA-Kode: [s]) bildet die Glottis dagegen wie beim normalen Ausatmen eine dreieckige Öffnung, sodass die Luft ohne Widerstand entweichen kann. Der Unterschied stimmhaft/stimmlos zieht sich durch eine ganze Reihe von Konsonantenpaaren, z. B. [d] und [t], [b] und [p], [g] und [k]. Jeweils ein Glied des Paares wird stimmhaft gesprochen, das andere stimmlos bei ansonsten gleicher Bildungsweise.

Bei anderen Lauten gibt es diese systematische Unterscheidung dagegen nicht. So haben die stimmlos ausgesprochenen Laute [ç] (wie in *Dichte*), [x] (wie in *Nacht*), [h] (wie in *Haut*), [ts] (wie in *Zahn*) kein stimmhaftes Gegenstück, und umgekehrt gibt es zu den stimmhaft artikulierten Lauten [m], [n], [ŋ] (wie in *Wange*), [l], [r], [ʀ] und [ʁ] kein stimmloses Pendant. Das gleiche gilt für die Vokale; sie werden in aller Regel stimmhaft ausgesprochen.

3.3.3 Artikulation der Konsonanten

Wenn die Luft die Glottis passiert hat, ist die Lautbildung (abgesehen vom [h]-Laut und vom glottalen Knacklaut) noch längst nicht abgeschlossen. Der wichtigste Teil vollzieht sich erst jetzt im so genannten **Vokaltrakt**, der in Anlehnung an einen Terminus aus der Instrumentenkunde auch **Ansatzrohr** genannt wird. Der Vokaltrakt umfasst Rachen, Mund und Nase, die zusammen ein System von Hohlräumen bilden, in denen der Luftstrom durch das Zusammenspiel der **Artikulationsorgane** zu Lauten geformt wird. Diesen Teilprozess der Lautproduktion nennt man **Artikulation**.

Bei der Beschreibung der Artikulation wollen wir zunächst eine grundsätzliche Unterscheidung zwischen zwei Lautklassen vornehmen, die wir oben schon en passant eingeführt, aber noch nicht erklärt haben: zwischen den **Vokalen** und den **Konsonanten**.

Vokale werden auch als **Öffnungslaute** bezeichnet, womit im Grunde schon ihre charakteristischen artikulatorischen Eigenschaften genannt sind. Vokale werden nämlich erzeugt, indem der durch die Stimmlippen

in periodische Schwingungen versetzte Luftstrom weitgehend ungehindert durch den Mund oder die Nase ausströmt. Weitgehend ungehindert meint, dass der Luftstrom ohne wahrnehmbares Reibegeräusch den Vokaltrakt verlässt.

Konsonanten sind dagegen als Laute definiert, bei denen der Luftstrom Hindernisse überwinden muss, bevor er aus Mund oder Nase austritt. Ein solches Hindernis kann ein kompletter Verschluss sein, der die Luft staut, bevor sie entweichen kann, oder eine starke Verengung, die beim Austritt der Luft ein hörbares Reibegeräusch verursacht. Bei dieser Unterscheidung gibt es natürlich eine Reihe von Grenzfällen, auf die wir hier nicht näher eingehen können (vgl. PÉTURSSON/NEPPERT 1991, 89). Beginnen wir mit den Konsonanten.

3.3.3.1 Artikulationsorgane

Die wichtigsten **Artikulationsorgane**, die natürlich primär anderen körperlichen Funktionen dienen als der Lauterzeugung, befinden sich im Mundraum. Dabei muss man zwischen den **unbeweglichen** und **beweglichen** (passiven und aktiven) **Artikulationsorganen** unterscheiden.

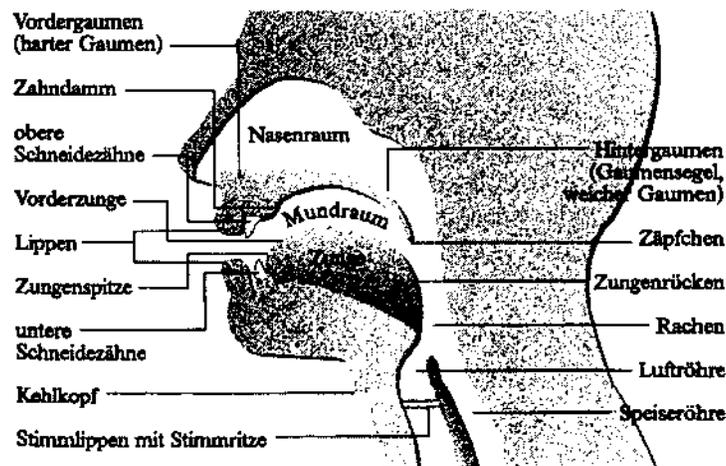
[In der folgenden Auflistung stehen in den Klammern zunächst die lateinischen Bezeichnungen für die Artikulationsorgane; danach werden die daraus abgeleiteten Adjektive genannt, die man – wie wir unten noch sehen werden – für die Angabe der Artikulationsstelle eines Lautes verwendet.]

- Die Lippen (**Labies; labial**), die mit Hilfe des Mundmuskelringes bewegt werden können. Sie lassen sich öffnen oder schließen, runden oder spreizen. Die beiden letztgenannten Bewegungsmöglichkeiten sind allerdings für die Bildung der Konsonanten irrelevant; sie spielen aber – wie noch zu zeigen sein wird – eine große Rolle bei der Vokalerzeugung.
- Die Zähne (**Dentes; dental**), bei denen in erster Linie die obere, unbewegliche Zahnreihe für die Artikulation der Konsonanten wichtig ist. Dagegen sind die Zähne des Unterkiefers, die durch entsprechende Bewegungen des Unterkiefers gehoben oder gesenkt werden können, für die Lautbildung ohne große Bedeutung.
- Der unmittelbar hinter den oberen Zahnreihe gelegene wulstartige Zahndamm (**Alveolen; alveolar**), den man mit der Zunge ertasten kann. Der Zahndamm ist unbeweglich.
- Der Gaumen, der sich an den Zahndamm anschließt und den Mundraum gegen den Nasenraum abgrenzt. Er besteht aus dem unbeweglichen harten Gaumen (**Palatum; palatal**) und einer sich daran

anschließenden segelartigen Membran, dem weichen Gaumen (**Gaumensegel** oder **Velum; velar**). Das Velum ist beweglich und arbeitet wie ein Klappventil: Wenn es gehoben wird, schließt es den Zugang zum Nasenraum ab, sodass Luft nur durch den Mund entweichen kann. Wird es dagegen abgesenkt, kann die Luft nur aus der Nase ausströmen, und es entstehen die so genannten **Nasallaute** wie z. B. [m] oder [n]. An seinem hinteren Ende geht das Velum in das ebenfalls bewegliche Gaumenzäpfchen (**Uvulum; uvular**) über.

- Die Zunge (**Lingua**) ist das Artikulationsorgan schlechthin. Sie ist an der Bildung der meisten Laute beteiligt und zeichnet sich durch höchste Beweglichkeit aus: Sie kann nach vorne oder nach hinten bzw. nach unten oder nach oben gezogen werden. Anatomisch ist die Zunge nicht in deutlich voneinander abgrenzbare Abschnitte unterteilt; im Hinblick auf die Bildung der Konsonanten unterscheidet man aber zwischen den folgenden Teilen: a) Zungenspitze (**Apex; apikal**), b) Zungenkranz (**Korona; koronal**) und c) Zungenrücken (**Dorsum; dorsal**). Die prominente Rolle der Zunge bei der Lautproduktion spiegelt sich im Übrigen schon darin wieder, dass in vielen Sprachen ein und dasselbe Wort für 'Zunge' und 'Sprache' verwendet wird (z. B. ital. *lingua*, franz. *langue* oder engl. in der Wendung *mother tongue*).
- Neben den im Mundraum befindlichen Artikulationsorganen muss schließlich der Vollständigkeit halber noch die Rachenwand (**Pharynx; pharyngal**) genannt werden, die allerdings bei der Lautproduktion im Deutschen nur eine untergeordnete Rolle spielt. Dazu kommt die **Glottis (glottal)**, auf deren Funktion wir oben bereits eingegangen sind.

Abb. 1



aus DUDEN-Grammatik 1998, 19: Artikulationsorgane

3.3.3.2 Artikulationsstelle und Artikulationsart

Durch die Angabe der an der Lautproduktion beteiligten Artikulationsorgane kann man denjenigen Ort exakt bestimmen, an dem aus dem lautlichen Rohmaterial ein bestimmter Konsonant gebildet wird. Man spricht in diesem Zusammenhang auch von der **Artikulationsstelle** oder dem **Artikulationsort**. Die Artikulationsstelle ist als derjenige Punkt definiert, an dem der Abstand zwischen zwei Artikulationsorganen am geringsten ist. Dabei reicht die Palette von einer mehr oder minder starken Annäherung bis hin zum vollständigen Kontakt.

So wird z. B. das [t] dort gebildet, wo der Zungenkranz (Korona) an den Zahndamm (Alveolen) stößt. Bei einer exakten Angabe der Artikulationsstelle müsste man natürlich beide Artikulationsorgane berücksichtigen und das [t] als **koronal-alveolaren Laut** kennzeichnen. In den meisten Fällen verzichtet man aber auf die exakte Bezeichnung und begnügt sich mit der Nennung des unbeweglichen Artikulators (für [t] z. B. **alveolar**), weil die Artikulationsstelle dadurch im Vergleich zur Artikulationsstelle anderer Laute hinreichend charakterisiert ist.

Neben der Angabe des Artikulationsortes muss für die phonetische Beschreibung eines Konsonanten noch ein zweites Merkmal bestimmt werden: die **Artikulationsart** bzw. der **Artikulationsmodus**. Darunter versteht man „die Art und Weise, wie der Luftstrom im Ansatzrohr oder in der Glottis gehemmt bzw. modifiziert wird“ (PÉTURSSON/NEPPERT 1991, 90). Zwei Artikulationsarten sind bereits in der oben angeführten Definition der Konsonanten genannt worden: Verschluss- und Engebildung. Dazu kommen: nasale Öffnung, seitliche Engebildung, Vibration und Verschlussbildung mit anschließender leichter Öffnung.

Was sich dahinter im Einzelnen verbirgt, werden wir in der folgenden Aufstellung sehen, bei der die Artikulationsarten das oberste Klassifikationskriterium für die artikulatorische Beschreibung der Konsonanten bilden. Im Verbund mit der Artikulationsstelle und dem Merkmal Stimmbildung haben wir dann das Instrumentarium zusammen, das für die artikulatorische Charakterisierung der Konsonanten ausreicht. (Vom Merkmal der Aspirierung, das ohnehin nur für die stimmlosen Verschlusslaute von Belang ist, werden wir hier absehen.)

Verschlusslaute

Ein Verschluss entsteht, wenn zwei Artikulationsorgane so eng zusammengebracht werden, dass der Luftstrom hinter der Kontaktstelle gestaut wird. Dort wird er für einen kurzen Moment angehalten, bis er durch die plötzliche Öffnung des Verschlusses entweichen kann, was ein charakte-

ristisches Sprengeräusch hervorruft. Die in dieser Artikulationsart gebildeten Laute werden deswegen auch **Sprenglaute** oder **Plosive** genannt.

Artikulationsstelle

Je nach Ort, an dem der Verschluss gebildet wird, unterscheidet man zwischen den folgenden Lauten:

- a) dem **glottalen Plosiv** [ʔ], der – wie oben bereits beschrieben – durch den Verschluss der Stimmlippen gebildet wird;
- b) den **bilabialen Plosiven** [p] und [b], die durch einen von Ober- und Unterlippe gebildeten Verschluss zustande kommen;
- c) den **alveolaren Plosiven** [t] und [d], gebildet durch einen Verschluss zwischen Zungenkranz und Zahndamm;
- d) den **velaren Plosiven** [k] und [g], bei denen der Verschluss zwischen Zungenrücken und weichem Gaumen liegt.

Stimmbeteiligung

Je nach Beteiligung der Stimmlippen unterscheidet man zwischen:

- a) den **stimmlosen Plosiven** [p], [t] und [k] und
- b) den **stimmhaften Plosiven** [b], [d], [g].

Der glottale Plosiv ist stimmlos.

Engelaute

Im Unterschied zur Verschlussbildung bleibt bei dieser Art der Artikulation ein enger Spalt zwischen den beteiligten Artikulatoren bestehen. Durch diese Engstelle kann die Luft entweichen; dabei entsteht ein Reibegeräusch. Die dadurch produzierten Laute werden folgerichtig auch als **Reibelaute** bzw. **Frikative** oder **Spiranten** bezeichnet.

Artikulationsstelle

Man unterscheidet zwischen den folgenden Lauten:

- a) den **labio-dentalen Frikativen** [f] und [v] (wie in *wohnen*), bei deren Artikulation die Unterlippe leicht die Schneidezähne der oberen Zahnreihe berührt;
- b) den **alveolaren Frikativen** [s] und [z], die gebildet werden, indem der Zungenkranz und der Zahndamm eine Engstelle bilden;
- c) den **palato-alveolaren Frikativen** [ʃ] (wie in *Schal*) und [ʒ] (wie in *Garage*), die durch die Bildung einer Engstelle zwischen Zungenkranz und der Übergangsstelle von Zahndamm und hartem Gaumen produziert werden. Die Artikulationsstelle ist also gegenüber den alveolaren Frikativen leicht nach hinten in Richtung Palatum verschoben;
- d) den **palatalen Frikativen** [ç] (wie in *nicht*) und [j], die durch eine Engbildung zwischen Zungenrücken und hartem Gaumen zustande kommen, wobei die Engbildung bei [j] weniger stark sein kann;

- e) dem **velaren Frikativ** [x] (wie in *noch*), bei dessen Bildung die Engstelle zwischen Zungenrücken und weichem Gaumen verläuft;
- f) dem **uvularen Frikativ** [ʁ], dem so genannten Reibe-r (s. o.), das durch eine Engstelle zwischen Zungenrücken und Zäpfchen gebildet wird;
- g) dem **glottalen Frikativ** [h], der – wie oben bereits erwähnt – durch die Engbildung der Stimmlippen verursacht wird.

Stimmbeteiligung

Nach den Merkmalen Stimmlosigkeit/Stimmhaftigkeit unterscheidet man die Paare {f} und {v}, {s} und {z} sowie {ʃ} und {ʒ} – jeweils der erste Laut wird stimmlos, der zweite stimmhaft gesprochen. Von den übrigen Frikativen sind [ç] und [x] stimmlos, [j] und [ʁ] stimmhaft.

Nasallaute

Nasallaute entstehen durch die Bildung eines Totalverschlusses an irgendeiner Stelle des Mundraumes bei gleichzeitiger Senkung des Gaumensegels, sodass die Luft nur durch die Nase entweichen kann. Wegen des Totalverschlusses im Mund werden die Nasallaute bisweilen auch den Verschlusslauten zugeschlagen und als **nasale Plosive** bezeichnet (vgl. z. B. HENGARTNER/NIEDERHAUSER 1993, 31). Wir bleiben hier aber bei der traditionellen Einteilung, bei der die Nasallaute eine eigene Kategorie darstellen.

Artikulationsstelle

Im Einzelnen unterscheidet man zwischen:

- a) dem **bilabialen Nasal** [m], bei dem die Lippen geschlossen sind;
- b) dem **alveolaren Nasal** [n], bei dem die Zungenspitze mit dem Zahndamm einen Verschluss bildet;
- c) dem **velaren Nasal** [ŋ] (wie in *Bank*), bei dem der Zungenrücken am weichen Gaumen anliegt.

Stimmbeteiligung

Die Nasale werden alle stimmhaft artikuliert.

Laterallaute

Bei Laterallauten kann der Luftstrom nur durch die Seiten der Mundhöhle entweichen, da der Zungenkranz den Zahndamm berührt und die Zungenseiten mit den Backenzähnen eine Engstelle bilden. Dadurch wird der „Luftstrom zwar behindert, aber nicht in einem solchen Mass, dass dabei (wie bei den Frikativen) ein Geräusch der Reibung entsteht“ (HENGARTNER/NIEDERHAUSER 1993, 32).

Artikulationsstelle:

Im Deutschen gibt es nur einen Laterallaut, nämlich [l]. Er wird immer stimmhaft gesprochen. Als Artikulationsstelle gilt die Kontaktstelle zwischen Zungenkranz und Zahndamm.

Vibrationslaute (Vibranten)

Diese Konsonanten werden durch schlagende Bewegung eines beweglichen Artikulators gegen einen unbeweglichen gebildet. Im Grunde handelt es sich um einen „Wechsel zwischen Verschluss oder Beinahe-Verschluss einerseits und Öffnung andererseits“ (PÉTURSSON/NEPPERT 1991, 94). Deswegen spricht man auch von einem **intermittierenden Verschluss**.

Artikulationsstelle

Man unterscheidet zwischen:

- a) dem **apikalen Vibranten**, dem so genannten Zungenspitzen-r (s. o.), bei dem die Zungenspitze gegen den Zahndamm schlägt, und
- b) dem **uvularen Vibranten**, dem so genannten Zäpfchen-r (s. o.), bei dem das Zäpfchen gegen den weichen Gaumen schlägt.

Stimmbeteiligung

Die Vibranten werden stimmhaft gesprochen.

Affrikaten

Affrikaten werden durch eine Kombination aus Verschluss und Engebildung produziert: Zunächst wird die Atemluft wie bei den Plosiven kurz gehemmt, bevor sich der Verschluss löst und gewissermaßen „in einem Zug“ an der gleichen Artikulationsstelle in eine Enge übergeht.

Artikulationsstelle

Unterschieden wird zwischen:

- a) der **alveolaren Affrikate** [ts] (wie in *Zahn*) und
- b) der **labiodentalen Affrikate** [pf].

Stimmbeteiligung

Die Affrikaten sind stimmlos.

3.3.4 Artikulation der Vokale

An der Artikulation der Vokale sind in der Hauptsache die Zunge, der Gaumen und die Lippen beteiligt. Durch deren Zusammenwirken wird der von den Stimmlippen in Schwingungen versetzte Luftstrom zu unterschiedlichen Vokalqualitäten geformt. Dabei sind die folgenden vier Faktoren relevant:

Faktor 1

Die Lage des Zungenrückens in der vertikalen Dimension (oben-unten) in Bezug auf den Gaumen.

Man unterscheidet drei Stellungen:

a) hoch (gemeint ist die höchste Position, die die Zunge in Bezug auf den Gaumen einnehmen kann, ohne dass durch die ausströmende Luft ein Reibegeräusch entsteht),

b) tief,

c) mittel,

wobei die mittlere Lage manchmal noch in halbhoch und halbtief unterteilt wird.

Bei diesen Stellungen handelt es sich – wie wir in Abb. 2 (nächste Seite) sehen werden – um grobe Fixpunkte, anhand derer die genauen Artikulationspunkte der einzelnen Vokale verortet werden können. Die Lage des Zungenrückens in der vertikalen Dimension korrespondiert in der Regel mit dem Öffnungsgrad des Mundraumes: Je höher die Zungenposition, desto geschlossener der Mundraum. Deswegen spricht man auch von geschlossenen, offenen, halbgeschlossenen und halboffenen Vokalen.

Faktor 2

Die Position des Zungenrückens in der horizontalen Dimension (vorn-hinten).

Hier geht es um die Frage, welche Stelle des Zungenrückens dem Gaumen am nächsten kommt. Faktor 2 ist also vergleichbar mit der Artikulationsstelle bei den Konsonanten, die ja oben als derjenige Punkt definiert wurde, an dem der Abstand zwischen den Artikulationsorganen (hier Zungenrücken und Gaumen) am geringsten ist.

Im Hinblick auf die Artikulation der Vokale werden drei Zungenpositionen als Fixpunkte unterschieden, und zwar

a) vorn,

b) hinten und

c) zentral.

So ist beispielsweise bei der Bildung des langen i-Lautes der vordere Zungenrücken dem Gaumen am nächsten, bei der Erzeugung des lang gesprochenen u-Lautes der hintere und bei der Artikulation des so genannten **Schwa-Lautes** [ə] (wie in *manche*) der mittlere.

Faktor 3

Die Lippenformung, bei der zwischen gerundet und ungerundet differenziert wird.

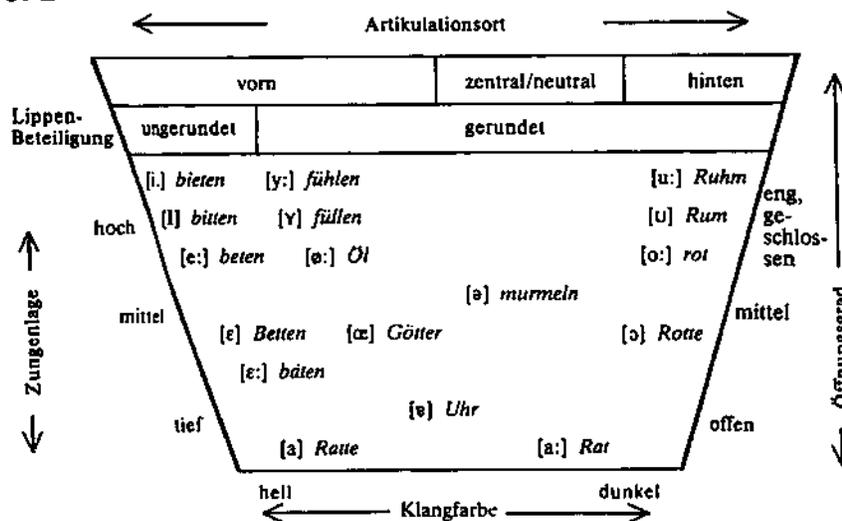
Faktor 4

Die Stellung des Gaumensegels, das entweder gehoben oder gesenkt sein kann.

Durch diesen Faktor werden die **Nasalvokale** (gesenkt; die Luft entweicht durch die Nase wie in frz. *pain* = Brot) von den **Oralvokalen** (gehoben; die Luft entweicht durch den Mund) abgegrenzt. Für das Deutsche ist diese Unterscheidung allerdings irrelevant, weil nur Oralvokale vorkommen.

Zur Veranschaulichung der Bildungsweise von Vokalen verwendet man gewöhnlich eine schematisierte Darstellung des Mundraumes, ein so genanntes **Vokaltrapez**, auch **Vokalviereck** genannt. Zwar können in einem solchen zweidimensionalen Trapez nur die Faktoren vertikale und horizontale Zungenstellung grob schematisch dargestellt werden, doch lässt sich die Lippenformung durch eine entsprechende Kennzeichnung mühelos einarbeiten.

Abb. 2



nach BUBMANN 2002, 739

An dieser Darstellung lässt sich die Bildungsweise der unterschiedlichen Vokalqualitäten gut ablesen; man braucht nur etwas Phantasie, um sich die Zungenbewegungen vorzustellen. So wird beispielsweise das [i:] dadurch erzeugt, dass der vordere Zungenabschnitt sehr hoch an den Gaumen tritt, ohne ihn allerdings zu berühren. Die Lippen sind dabei ungerundet. Bei gleicher Zungenstellung, aber gerundeten Lippen wird dagegen ein [y:] (wie in *fühlen*) erzeugt. Behält man die Lippenrundung bei, senkt aber den Zungenrücken ein wenig ab, wird ein [ø] (wie in *Töne*)

produziert. Zieht man dagegen von der [y:]-Position den Zungenrücken unter Beibehaltung der Lippenrundung nach hinten, erzeugt man ein [u:].

Man kann diese Vorgehensweise leicht um die noch ausstehenden Kombinationen von Zungenbewegungen und Lippenformungen ergänzen und kommt auf diese Weise zur Artikulation der übrigen Vokale des Deutschen. Dabei wird deutlich, dass die Artikulationsbewegungen ein Kontinuum fließender Übergänge bilden, die – allerdings nur bei der isolierten Produktion von Lauten (siehe aber unten 3.4) – an bestimmten Positionen einrasten. Und noch etwas wird dabei deutlich: Die allermeisten Vokale kommen gewissermaßen in zweifacher Ausführung vor, nämlich als kurze und als lange Vokale, wobei das jeweils lange Glied eines Paares in der IPA-Schrift durch einen hinter dem Zeichen für die Lautqualität stehenden Doppelpunkt gekennzeichnet ist.

Neben der **Qualität** unterscheiden sich die Vokale im Deutschen also auch in der **Quantität**. Allerdings ist der quantitative Unterschied bis auf eine Ausnahme jeweils mit einem qualitativen verbunden. Der kurz gesprochene Vokal eines Paares wird nämlich immer ein wenig offener artikuliert als der lange. Das wird auch in der Notierungskonvention der IPA-Schrift zum Ausdruck gebracht, denn die verwendeten Zeichen für die Vokalqualitäten weichen in jedem Paar geringfügig voneinander ab: [i]/[iː], [u]/[uː], [o]/[oː] usw.

Neben den „einfachen“ Vokalen (**Monophthongen**), deren Qualität während der ganzen Artikulation konstant bleibt, gibt es im Vokalismus eine Erscheinung, die mit den Affrikaten im konsonantischen Bereich zu vergleichen ist: die **Zwilaute** bzw. **Diphthonge**. Auch hier handelt es sich um kombinierte Laute, die aber mit gleitender Zungenstellung ineinander übergehen. Im Deutschen sind das [ai] (wie in *heiß*), [au] (wie in *Haus*) und [ɔɪ] (wie in *Eule*). Zu unterscheiden sind die Diphthonge von echten Vokalkombinationen wie [aɔ] in *Chaos*, bei denen die Vokale beim lauten Skandieren auf zwei Silben verteilt werden, also *Cha-os*, was bei *heiß*, *Haus* etc. nicht der Fall ist.

3.4 Laute im Kontext des Sprechens – Koartikulation

Bei der bisherigen artikulatorischen Beschreibung der Laute haben wir stillschweigend zwei aufeinander aufbauende Prämissen zugrunde gelegt, die bei näherer Betrachtung nicht haltbar sind:

- a) Laute lassen sich aus den sprachlichen Äußerungen problemlos als Einheiten isolieren.

- b) Die artikulatorischen Merkmale eines Lautes sind immer die gleichen, ganz egal, welcher Laut ihm in der sprachlichen Äußerung vorausgeht oder folgt.

Beide Annahmen entsprechen wohl auch dem Common-Sense-Verständnis von Sprache. Sie resultieren z. T. aus unserer Erfahrung mit der Schriftlichkeit, denn zumindest beim Maschinenschreiben reihen wir tatsächlich Einheit an Einheit, und weder der vorausgehende noch der nachfolgende Buchstabe hat irgendeinen Einfluss darauf, wie das „in der Mitte stehende“ Schriftzeichen konkret gestaltet wird.

Lautsprache funktioniert aber grundsätzlich anders; hier ist gewissermaßen „alles im Fluss“. So zeigen Röntgenaufnahmen der Artikulationsbewegungen, dass scharfe Bewegungseinschnitte, anhand derer man die Grenzen zwischen Laut x und Laut y genau lokalisieren könnte, bei der Produktion einer zusammenhängenden Äußerung meist nicht erkennbar sind; die Laute gehen vielmehr ineinander über. Und selbst die aus dem Schriftbild so vertrauten und fast naturgegeben erscheinenden Wortgrenzen sind auf dem Röntgenfilm nur dann klar nachweisbar, wenn ein Wort vom Sprecher bewusst aus dem Redefluss isoliert wird.

Aus diesem Befund folgt zwingend, dass auch die zweite oben formulierte Annahme reichlich brüchig ist, nach der ein Laut immer das gleiche artikulatorische Gesicht hat, unabhängig davon, in welcher lautlichen Umgebung er gesprochen wird. Tatsächlich beeinflussen sich die Laute durch ihre Aussprache im Redefluss gegenseitig; man nennt diese Erscheinung **Koartikulation**. Solche koartikulatorischen Prozesse sind physiologisch unvermeidbar, weil die Artikulationsorgane in ständiger Bewegung sind und dabei Merkmale nachfolgender Laute vorwegnehmen bzw. bereits ausgeführte Bewegungen Auswirkungen auf nachfolgende haben. Was damit genau gemeint ist, hat TERNES an einem einfachen Beispiel illustriert, das wir in dem folgenden Zitat wiedergeben wollen:

Wenn wir drei deutsche Wörter wie *Kiefer*, *Küfer*, *Kufe* aussprechen, haben wir zunächst den Eindruck, jedesmal denselben Laut [k] zu sprechen. Bei genauerer Untersuchung stellt sich jedoch heraus, dass das [k] jedesmal ein wenig anders artikuliert wird. Im Fall von *Kiefer* wird das [k] etwas weiter vorn im Mund artikuliert (prävelar) als das, was man als Durchschnittswert für [k] ansehen könnte (velar). Dies geschieht, weil sich das [k] im Wort *Kiefer* bereits auf die Position des unmittelbar folgenden *vorderen* Vokals [i:] einstellt und diese teilweise vorwegnimmt. Demgegenüber wird das [k] in *Kufe* etwas weiter hinten im Mund (postvelar) als der Durchschnittswert artikuliert, weil sich in diesem Fall das [k] bereits auf die Position des unmittelbar folgenden *hinteren* Vokals [u:] einstellt.

[...] Vergleichen wir nun das [k] in *Kiefer* mit demjenigen in *Küfer*, stellen wir fest, dass diese beiden [k]-Laute in Bezug auf die Artikulationsstelle in etwa die gleiche Position einnehmen, da [y:] ebenso ein vorderer Vokal ist wie [i:]. [...] Es besteht jedoch ein anderer Unterschied zwischen den [k]-Lauten in diesen beiden Wörtern. Das [k] in *Kiefer* wird mit gespreizten (ungerundeten) Lippen gesprochen, da das folgende [i:] ein gespreizter Vokal ist. [...] Im Worte *Küfer* hingegen wird das [k] mit gerundeten Lippen gesprochen, da es die Lippenrundung des unmittelbar folgenden gerundeten Vokals [y:] vorwegnimmt. (TERNES 1987, 36f.)

Wir können also Folgendes festhalten: Hinter dem Laut, den wir oben in unserer artikulatorischen Beschreibung als velaren Verschlusslaut [k] gekennzeichnet haben, verbirgt sich nicht nur eine Vielzahl lautlicher Varianten, die aus geschlechtsspezifischen, altersspezifischen und ähnlichen individuell gefärbten Ausspracheunterschieden resultieren, von denen bei jeder phonetischen Beschreibung von vornherein abstrahiert wird. Das [k] umfasst vielmehr eine ganze Reihe koartikulatorisch bedingter Lautvarianten, die – zumindest in ihrer groben Tendenz – im ersten Arbeitsschritt der phonetischen Analyse mit Hilfe einer so genannten **engen phonetischen Transkription** gewöhnlich auch erfasst werden. So würde ein Phonetiker die [k]-Laute in *Kiefer*, *Kufe* und *Küfer* mit einem jeweils anderen Symbol des IPA-Alphabets wiedergeben: den in *Kiefer* mit dem Symbol [k̟], um damit die leicht nach vorne verlagerte Artikulationsstelle zu kennzeichnen, den in *Kufe* mit [k̠] zur Bezeichnung der nach hinten verlagerten Artikulationsstelle und den in *Küfer* mit [k^w], um damit auf die Lippenrundung hinzuweisen.

Allerdings sieht man in der Praxis meist rasch von koartikulatorisch bedingter Lautvarianz ab. Das deswegen, weil sich die Koartikulation – wie oben bereits angedeutet – zwangsläufig aus den Gegebenheiten des menschlichen Artikulationsapparates und seiner neuronalen Steuerung ergibt und sie deswegen auch in jeder Sprache auf vergleichbare Weise in Erscheinung tritt (vgl. TERNES 1987, 40).

3.5 Die phonologische Abstraktion

Bei der phonologischen Betrachtungsweise der Laute haben wir es mit einem über verschiedene Stufen verlaufenden Abstraktionsprozess zu tun; einige davon haben wir bereits angesprochen. Letztlich läuft das Vorgehen darauf hinaus, diejenigen lautlichen Kategorien zu finden, die in einer Sprache **bedeutungsunterscheidende Funktion** haben. Je nachdem, ob

es sich um eine fremde, womöglich noch exotische, oder um die eigene, vertraute Sprache handelt, die man untersuchen will, wird man in der Praxis sehr unterschiedliche Wege einschlagen, um das gewünschte Ergebnis zu erzielen. Allerdings gibt es einige Arbeitsschritte, die in beiden Fällen durchgeführt werden müssen. Wir wollen uns im Wesentlichen auf die Beschreibung dieser Schritte beschränken.

3.5.1 Segmentierung

Ausgangspunkt der Analyse ist eine Sammlung lautsprachlicher Äußerungen derjenigen Sprache, an der die Untersuchung durchgeführt werden soll. Diese Äußerungen liegen zunächst als lautsprachliches Kontinuum vor. Der erste Arbeitsschritt besteht darin, dieses Kontinuum in voneinander abgegrenzte lautliche Einheiten zu zerlegen. Man nennt diesen Vorgang **Segmentierung**. Bei fremden (und vielleicht bislang noch nicht einmal untersuchten) Sprachen kann die Zerlegung erhebliche Schwierigkeiten bereiten, denn die Lautgrenzen sind – wie wir oben bereits gesehen haben – im Material selbst nicht klar vorgegeben.

In der Praxis stützt sich ein phonetisch entsprechend geschulter Wissenschaftler in erster Linie auf den Gehöreindruck, d. h. er geht auditiv-phonetisch vor, wobei natürlich akustische und artikulatorische Analysen zur Unterstützung herangezogen werden können. In einer dem analysierenden Wissenschaftler bekannten Sprache sind die Dinge natürlich erheblich einfacher, denn hier kann man sich auf sein Sprachwissen verlassen, auf dessen Grundlage sich die Segmentierung beinahe automatisch ergibt.

Die ermittelten lautlichen Segmente werden dann in einer möglichst engen phonetischen Transkription aufgezeichnet – möglichst eng deswegen, weil auch die kleinste lautliche Nuance im System der Lautkategorien einer Sprache von Bedeutung sein kann, wenn man einmal von den im oben skizzierten Sinne koartikulatorisch bedingten Lautmerkmalen absieht. Das Ergebnis dieses Arbeitsschritts sind also Ketten von Lauteinheiten, die man durch IPA-Symbole beschreibt, wobei jedem Symbol eine lautliche Einheit entspricht. Diese lautlichen Einheiten nennt man **Phone**.

In vielen Darstellungen werden die Phone als die in einer Äußerung konkret vorkommenden Lautsegmente definiert. Genau genommen entsprechen diejenigen Einheiten, die als Ergebnis der Segmentierung mit den IPA-Symbolen festgehalten werden, nicht singulären Lautereignissen, wie sie in einer konkreten Äußerung artikuliert werden. Es handelt

sich vielmehr schon um typisierte Laute, die – wie inzwischen klar sein dürfte – bereits eine Reihe von Abstraktionsstufen durchlaufen haben.

Dabei spiegelt sich der Grad der Abstraktion in der Enge bzw. Weite der phonetischen Transkription wider, und im Grunde wäre es sehr nützlich, wenn man die verschiedenen Abstraktionsgrade durch eine unterschiedliche Notierung – etwa durch unterschiedliche Klammerungen – auch als solche kenntlich machen würde (vgl. TERNES 1987, 50). Aber unabhängig davon, welche Abstraktionsstufe gerade erreicht ist – der entscheidende Schritt steht noch aus: Die Phone müssen noch zu bedeutungsunterscheidenden Lauteinheiten, zu **Phonemen**, zusammengefasst werden.

3.5.2 Minimalpaarbildung

Nehmen wir an, die Segmentierung hat u. a. die Wörter mit den Lautfolgen [maus] und [laus] ergeben, die beide natürlich schon weitgehend von den Erscheinungen der Koartikulation bereinigt sind. Insofern haben wir es bereits mit einer ziemlich weiten, sprich: abstrakten phonetischen Transkription zu tun.

Bei diesen Wörtern fällt auf, dass sie bis auf die Anfangssegmente [l] bzw. [m] identisch sind. Solche Wortpaare werden methodisch genutzt, um die Phone auf ihre bedeutungsunterscheidende Funktion hin zu überprüfen. Das Verfahren, das man auch **Substitution** nennt, ist auf den ersten Blick sehr trivial. Das Anfangssegment [m] wird durch [l] ersetzt, und es zeigt sich, dass sich die Bedeutung der Lautfolge dadurch verändert: Aus dem Wort *Maus* wird das Wort *Laus*. Man nennt solche „Wörter (oder Formen) verschiedener Bedeutung, die sich nur in einem Phon in derselben Position unterscheiden“ (TERNES 1987, 51), auch **Minimalpaare** und sagt, dass zwischen den betreffenden lautlichen Segmenten eine **phonologische Opposition** besteht. Letztlich führt das dazu, den entsprechenden Einheiten Phonemstatus zuzuerkennen. Das entspricht der klassischen Definition, nach der **Phoneme** die kleinsten bedeutungsunterscheidenden Einheiten des Sprachsystems sind.

Ganz so einfach, wie es zunächst den Anschein hat, sind die Dinge allerdings bei näherer Betrachtung doch nicht. Angenommen, als Ergebnis der Segmentierung liegen auch die Wörter mit den Lautfolgen [p^ham] und [bain] vor. Zunächst scheint es so, als hätten wir es wieder mit einem Minimalpaar zu tun, denn der Austausch von [p^h] durch [b] führt zu einem Bedeutungswechsel. Man müsste also den genannten Segmenten ebenfalls Phonemstatus zusprechen.

Dabei gibt es allerdings Folgendes zu bedenken: Unser letzter Abstraktionsschritt sollte dazu dienen, alle lautlichen Merkmale zu tilgen, die für die bedeutungsunterscheidende Funktion von Lauten keine Rolle spielen. Schon ganz intuitiv stellt man sich die Frage, ob das im Falle von [p^h] auch gelungen ist, denn die durch das hochgestellte ^h signalisierte Lautqualität „Behauchung“ mutet auf den ersten Blick gewissermaßen als überflüssige Beigabe zum einfachen [p]-Laut an, wie er in bestimmten lautlichen Kontexten, beispielsweise vor [s] in *mopsen*, tatsächlich gesprochen wird.

Allerdings darf man sich hier nicht durch die graphische Darstellung täuschen lassen, denn dass das Merkmal „Behauchung“ (anders als z. B. das Merkmal „Stimmhaftigkeit“ des Lautes [b]) dem eigentlichen Lautsymbol als diakritisches Zeichen nur beigefügt ist, beruht lediglich auf einer Konvention; man könnte für den behauchten und den unbehauchten [p]-Laut auch ein jeweils eigenes Symbol verwenden.

Um sich über den Status des Merkmals „Behauchung“ Gewissheit zu verschaffen, muss man prinzipiell die gleiche Analyseverfahren anwenden, wie oben vorgeführt. Man überprüft, ob [p^h] und [p] in phonologischer Opposition zueinander stehen oder nicht. Dass das nicht der Fall ist, zeigt sich schon, wenn man die Behauchung in Wörtern wie *Pein*, *Pass*, *Pilz* bewusst weglässt und die Lautfolge dann jeweils mit der normalen behauchten Variante konfrontiert: Ob mit oder ohne Behauchung gesprochen, die Lautketten transportieren keine unterschiedlichen Bedeutungen. Und auch sonst wird man keine Wörter finden, deren Bedeutungsunterschied allein durch den Wechsel von [p^h] und [p] bedingt ist. Wir können also festhalten, dass man im Deutschen auf der letzten Stufe der phonologischen Abstraktion von dem Merkmal „Behauchung“ absehen kann – anders als z. B. im Thai, wo behauchte und unbehauchte Verschlusslaute bedeutungsdifferenzierend sind.

Dieser Befund wirft natürlich auch ein anderes Licht auf das eingangs angeführte Paar *Pein/Bein*, anhand dessen wir den vermeintlichen Phonemstatus von [p^h] ermittelt hatten. Weil wir nämlich jetzt wissen, dass die Behauchung phonologisch irrelevant ist, kann man *Pein* auf der nun erreichten Abstraktionsstufe problemlos als [paɪn] transkribieren. Die Substitution mit dem ersten Segment aus [baɪn] ergibt dann die oben bereits konstatierte Bedeutungsänderung, sodass man als Phoneme letztlich /p/ und /b/ erhält. Wie man sieht, muss man bei phonologischen Analysen an verschiedenen Stellen ansetzen, um keine irreführenden Ergebnisse zu erzielen.

Sobald man sich auf der Phonemebene befindet, ändern sich auch die Notierungskonventionen. Während die Zeichen für Phone in eckigen Klammern notiert werden, stehen die Zeichen für Phoneme in Schrägstrichen. Damit wird auch der veränderte Status der Einheiten markiert, denn ein Phonem kann man „als solches nicht unmittelbar aussprechen. Was man jeweils ausspricht, ist nur *eine* konkrete *phonetische Realisierung* [...] eines Phonems“ (TERNES 1987, 41). Deswegen besteht zwischen [p] und /p/ trotz der Verwendung desselben IPA-Symbols auch keineswegs Identität: [p] ist genau wie [p^h] eine der phonetischen Realisierungen des Phonems /p/.

Man nennt solche Realisierungsvarianten auch **Allophone** eines Phonems. Diese Redeweise zeigt im Grunde nur einen Perspektivwechsel an: Sobald sich Phone durch die entsprechenden Analyseprozeduren als Realisierungen desselben Phonems entpuppt haben, können sie als dessen Allophone identifiziert werden. Dabei verbirgt sich streng genommen auch hinter diesen Allophonen eine Vielzahl von individuell gefärbten bzw. koartikulatorisch bedingten konkreten Lautvarianten, für die in der Phonologie leider keine eigene Bezeichnung zur Verfügung steht (zum Allophon [p^h] z. B. ein wegen des nachfolgenden gerundeten Vokals mit leichter Lippenrundung artikuliertes [p^h] in Wörtern wie *Püree*).

Mit Hilfe von Minimalpaarbildungen lässt sich das Phoneminventar des Deutschen (wenn auch manchmal auf Umwegen) bis auf einige Problemfälle relativ leicht herausarbeiten. Bei den Konsonanten beispielsweise ergeben sich aus den Paaren *Masse/Kasse* die Phoneme /m/ und /k/, aus *Tasse/Tasche* /s/ und /ʃ/, aus *Nadel/Nagel* bzw. *Dock/Bock* /d/, /g/ und /b/, aus *laufen/saufen* /l/ und /z/, aus *sinnen/singen* /n/ und /ŋ/, aus *hohl/wohl* /h/ und /v/, aus *Jugend/Tugend* /j/ und /t/, aus *Puder/Fuder* /p/ und /f/. Etwas komplizierter ist die phonologische Bewertung der Phone [ç] und [x].

Auf Detailprobleme im Bereich des Konsonantismus, wie etwa die phonologische Bewertung der Affrikaten (ein oder zwei Phoneme?), können wir hier nicht eingehen (vgl. dazu u. a. MEINHOLD/STOCK 1980). Bei den Vokalen lassen sich durch entsprechende Minimalpaare im Großen und Ganzen diejenigen Phontypen als phonologisch relevante Einheiten ermitteln, die wir bereits in 3.4 unter artikulatorischem Gesichtspunkt beschrieben haben. Problembereiche stellen hier u. a. der Schwa-Laut [ə] und der lange offene [ɛ:]-Laut dar.

3.5.3 Kombinatorische und freie Varianten

Im Fall von [ç] und [x] stoßen wir prinzipiell auf das gleiche Phänomen, das wir oben bei der ersten Analyse von [p^h] bereits kennen gelernt haben – die ersten Minimalpaare können trügerisch sein. Wenn wir beispielsweise *Nacht* mit *nackt* auf der einen und *weichen* mit *weiten* auf der anderen Seite kontrastieren und die entsprechenden Substitutionen ausführen, scheint der Phonemstatus sowohl von [ç] als auch von [x] erwiesen zu sein. Die Crux ist wieder, dass es kein Minimalpaar gibt, bei dem die in Frage stehenden Phone [ç] und [x] ihrerseits in Opposition zueinander stehen.

Der Grund liegt darin, dass das Vorkommen von [ç] und [x] an jeweils spezifische lautliche Kontexte gebunden ist: [ç] kommt nur nach den vorderen Vokalen und nach Konsonanten vor, [x] dagegen nur nach hinteren Vokalen. Das bedeutet, dass es keinen lautlichen Kontext gibt mit [ç] gibt, in dem auch [x] stehen könnte, und umgekehrt. Deswegen ist es logisch ausgeschlossen, dass mit den beiden Lauten Minimalpaare gebildet werden können, denn Minimalpaare dürfen sich ja – wie oben definiert – nur in einem in der gleichen Position stehenden Phon unterscheiden.

Phone, die in der skizzierten Weise **komplementär verteilt** sind, fasst man zu einem Phonem zusammen – in unserem Fall also [ç] und [x] zum Phonem /x/, wobei sich die Entscheidung zugunsten /x/ als Vertreter der Kategorie ziemlich komplizierten Detailüberlegungen verdankt, auf die wir hier nicht eingehen können (vgl. aber TERNES 1987, 103f.). Aus der umgekehrten Perspektive kann man [ç] und [x] wieder als Allophone des Phonems /x/ werten, wobei man ihren besonderen Eigenschaften entsprechend auch von **kombinatorischen** bzw. **stellungsbedingten Varianten** spricht.

Eigentlich handelt es sich auch bei [p] und [p^h] um kombinatorische Varianten, weil – wie wir in 3.3.3.2 schon ausgeführt haben – auch die Behauchung von Verschlusslauten im Deutschen umgebungsabhängig ist. Allerdings tritt sie manchmal auch in Stellungen auf, wo sie normalerweise nicht vorkommt, etwa wenn der Sprecher besonders akzentuiert sprechen will oder wenn er wütend schreit. Die umgebungsspezifische Realisierung von [ç] und [x] ist dagegen von solchen Faktoren völlig unabhängig.

Neben den stellungsbedingten Varianten gibt es als zweiten Haupttyp der Allophonie die **fakultativen** bzw. **freien Varianten**. Wie die Bezeichnung schon andeutet, ist diese Form der Varianz nicht an bestimmte lautliche Kontexte gebunden. Das typische Beispiel im Deutschen sind

die drei Realisierungsvarianten des Phonems /r/, Zäpfchen-, Zungenspitzen und Rachen-r. Man kann sie in allen Kontexten, in denen ein (konsonantischer) r-Laut überhaupt gesprochen werden kann, bedenkenlos gegeneinander austauschen, ohne dass das irgendeinen Einfluss auf die Bedeutungsübermittlung hat (siehe auch 3.2).

Allerdings lassen sich aus der Verwendung einer bestimmten Realisierungsvariante sprecherbezogene Merkmale wie die regionale Herkunft entnehmen, wobei „die geographische Verteilung [...] eher kleinräumig zu beschreiben“ ist (TERNES 1987, 84). So ist beispielsweise im Ruhrgebietsdeutschen das Rachen-r die vorwiegend verwendete Aussprachevariante.

3.5.4 Phoneme als Bündel distinktiver Merkmale

Bei den bisherigen Überlegungen war immer wieder die Rede von Lautmerkmalen, von denen man im Rahmen der phonologischen Abstraktion absehen kann. Verfolgt man diesen Ansatz konsequent, kann man Phoneme auch als Bündel phonologisch relevanter Merkmale ansehen:

Das Phonem [...] kann, wie ein Akkord in der Musik, in kleinere, gleichzeitig vorhandene Komponenten aufgespalten werden [...] Das französische Phonem *b* kann z. B. (in solchen Wortserien wie *bu*, *pu*, *vu*, *mu* etc.) durch die Phoneme *p*, *v*, *m* ersetzt werden; im Gegensatz zu *p* ist es stimmhaft, im Gegensatz zu *v* explosiv, im Gegensatz zu *m* oral (nicht-nasal) etc. Wenn wir auf diese Weise den differentiellen Wert des französischen Phonems *b* untersuchen, dann stellen wir seinen sprachlichen Inhalt fest: Stimmhaftigkeit, Explosivität, Oralität etc. (JAKOBSON 1974, 143)

Das Prinzip besteht also darin, dass man durch eine Gegenüberstellung genau diejenigen Merkmale herauszufinden versucht, die ein Phonem von allen anderen Phonemen unterscheiden. Dabei geht man möglichst schrittweise vor und kontrastiert nacheinander Phoneme, die nur in einem Merkmal voneinander abweichen. Auf ein deutsches Beispiel angewandt:

- /d/ unterscheidet sich von /t/ nur durch das Merkmal „Stimmhaftigkeit“,
- von /z/ nur durch das Merkmal „Verschluss“,
(/z/ ist ein Reibelaut bei ansonsten gleicher Bildungsweise)
- von /g/ bzw. /b/ nur durch das Merkmal „alveolar“,
(die Artikulationsstelle von /g/ ist velar, von /b/ bilabial, die übrigen Merkmale „Verschluss“ und „Stimmhaftigkeit“ stimmen mit /d/ überein).

Nach einer solchen Auffassung sind letztlich nicht mehr die Phoneme als Ganzheiten bedeutungsunterscheidend, sondern die **distinktiven Merkmale**. Dabei geht man davon aus, dass diese Merkmale universellen Charakter haben, also sprachübergreifend gelten. Jede Sprache trifft allerdings eine spezifische Auswahl, und es passiert häufig, dass ein Merkmal, das in der einen Sprache bedeutungsunterscheidend wirkt, in der anderen entweder gar nicht vorkommt oder dort lediglich artikulatorische Beigabe ist, wie etwa die Behauchung im Deutschen.

3.5.5 Phonotaktik

Genau wie sich Sprachen in ihren Phoneminventaren unterscheiden, variieren sie auch im Hinblick darauf, welche Regularitäten bei der Kombination von Phonemen zu größeren Einheiten wie Silbe, Morphem oder Wort gelten. Beschrieben werden diese Regularitäten im Rahmen der phonologischen Teildisziplin **Phonotaktik**; die Regeln selbst werden entsprechend **phonotaktische Regeln** genannt.

Die Basiseinheit für die Beschreibung phonotaktischer Regeln ist die **Silbe**, die im Gegensatz zum Morphem oder Wort keine Bedeutung trägt, sondern eine rein lautliche Einheit ist. Deswegen ist es auch eine verkürzte Redeweise, wenn man sagt, Wörter bestehen aus Silben, denn es ist genau genommen nur die lautliche Seite der Wörter, die aus einer bzw. mehreren Silben aufgebaut ist.

Silben haben eine interne Struktur; sie bestehen:

- a) aus einem obligatorischen **Silbenkern** bzw. **Nukleus**,
- b) einem den Kern fakultativ umgebenden **Silbenrand**, bei dem man je nach Position zwischen dem **Silbenkopf** (vor dem Kern) und der **Silbenkoda** (hinter dem Kern) unterscheidet.

Die einzelnen Positionen sind folgendermaßen besetzt: Der Kern wird in der Regel durch einen Vokal gebildet; Ausnahmen sind silbische Konsonanten wie [n] in [le:bn] oder [l] in [ʃʏsl] (*Schüssel*). Der Silbenrand besteht aus einem oder mehreren Konsonanten. Er ist insofern fakultativ, als eine Silbe auch nur aus einem einzigen Vokal bestehen kann (z. B. *Abend*), Kopf und Koda also unbesetzt sind.

In manchen Sprachen sind die Regeln der Silbenbildung ziemlich einfach. Im Japanischen z. B. hat jede Silbe die Form (K) V (N); dabei ist K ein Konsonant, V ein Vokal und N ein nasales Phonem; Klammern bedeuten, dass das Phonem in einer Silbe vorkommen kann, aber nicht muss. Die deutschen Regeln sind um einiges komplizierter, sie lassen Silben der

Struktur (K) (K) (K) V (K) (K) (K) zu. Eine solche deutsche Silbe ist beispielsweise das Wort *Strumpf*. (MILLER 1993, 104)

Bei der formalen Kennzeichnung der Silbenstruktur mit den Variablen K für Konsonant und V für Vokal ist allerdings zu bedenken, dass die Leerstellen nicht von beliebigen Vertretern ihrer Kategorie besetzt werden können. Für jede Sprache gilt nämlich, dass bestimmte Phoneme oder Phonemklassen in bestimmten Stellungen nicht vorkommen können. Solche **Distributionsbeschränkungen** werden nicht nur im Hinblick auf Silbenpositionen, sondern auch hinsichtlich der Stellungen im Wort bzw. Morphem oder auch relativ zu anderen Phonemen bestimmt, wobei sich die Bezugsgrößen teilweise überschneiden. Für das Deutsche gelten z. B. die folgenden Beschränkungen:

- [h] kommt nicht am Silben- bzw. Wortende vor;
- [ŋ] steht nicht im Wort- bzw. Morphemanlaut;
- am Wort- bzw. Morphemanfang sind u. a. die Kombination [p] + [t], [m] + [g] oder [r] + [t] ausgeschlossen;
- von den Konsonantenpaaren /p/ – /b/, /t/ – /d/, /k/ – /g/, /f/ – /v/, /s/ – /z/ kommen die jeweils stimmhaften am Silben- bzw. Wortende nicht vor, also /b/, /d/, /g/, /v/, /z/. Deswegen sprechen wir beispielsweise den letzten Laut von *Hund* auch stimmlos als [t], obwohl der Plural des Wortes, nämlich *Hunde*, mit [d] realisiert wird. Zwei verschiedene Wortformen ein und desselben Wortes werden also in einem lautlichen Segment unterschiedlich ausgesprochen, und zwar in Abhängigkeit von dessen Stellung.

Dass wir es hier mit einer sprachspezifischen Erscheinung zu tun haben, zeigt der Vergleich mit dem Englischen. Die genannten stimmhaften Verschluss- bzw. Reibelaute kommen hier auch am Wortende vor, was für viele deutsche Muttersprachler beim Erwerb der englischen Sprache ein Problem darstellt. Sie übertragen nämlich die muttersprachlichen Regularitäten unbewusst auf das Englische und sprechen Wörter wie *give*, *globe* etc. jeweils mit stimmlosem Endkonsonanten aus. Das ist im Übrigen ein schöner Beleg für die Wirksamkeit des phonematischen und phonotaktischen Filters auch bei der Sprachproduktion (vgl. 3.2).

3.5.6 Zum Status phonologischer Analysen

Wir wollen das letztgenannte Beispiel von Distributionsbeschränkungen im Deutschen nutzen, um abschließend noch einige Bemerkungen zum Status phonologischer Analysen zu machen. Gerade hier zeigt sich nämlich, dass phonologische Analysen grundsätzlich theorieabhängig sind

und dass man je nach Vorgehensweise und sprachtheoretischer Grund-
auffassung zu teilweise anderen Ergebnissen kommen kann als denen, die
wir in diesem Beitrag vorgestellt haben.

Der beschriebene Sachverhalt, nach dem die stimmhaften Phoneme
/b/, /d/, /g/, /v/, /z/ am Wort- bzw. Silbende nicht auftreten, sondern nur
ihre stimmlosen Pendanten /p/, /t/, /k/, /f/, /s/, wird in der einschlägigen
Forschungsliteratur auch als **Auslautverhärtung** bezeichnet. Die in der
Bezeichnung liegende Metaphorik macht genau genommen nur Sinn,
wenn man annimmt, dass einer in dieser Position gesprochenen verhärteten,
sprich: stimmlosen Einheit eine nichtverhärtete, sprich: stimmhafte
Einheit zugrunde liegt. Denn – um im Bild zu bleiben – verhärtet kann
sich ja nur etwas, was vorher in einem „weichen Aggregatzustand“ vor-
gelegen hat.

Tatsächlich geht man im Rahmen so genannter generativer Sprachmo-
delle von solchen dynamischen Prozessen aus. Sehr vereinfacht gesagt,
setzt man für ein Wort wie *Hund* eine **abstrakte phonologische Reprä-
sentation** /hund/ an, auf die eine **phonologische Regel** angewandt wird,
die besagt, dass /t/ im Silbenauslaut als [t] wie in [hunt] und im Silben-
anlaut als [d] wie in [hun-də] realisiert wird. (Vgl. dazu ausführlicher
GREWENDORF/ HAMM/STERNEFELD 1987, 76ff.)

Noch etwas anders bewertet man den nämlichen Sachverhalt in Kon-
zeptionen, die bereits in den 1930er Jahren von Linguisten der **Prager
Schule** entwickelt worden sind. Hier geht man davon aus, dass die bedeu-
tungsunterscheidende Kraft von /t/ und /d/ etc. am Wort- oder Silbende
suspendiert bzw. neutralisiert ist.

Daraus zieht man den Schluss, dass das Phonem, das in Wörtern wie
Hund vorkommt, weder /t/ noch /d/ ist, sondern eine übergeordnete, ge-
wissermaßen unentschiedene Einheit, die /t/ und /d/ integriert. Man nennt
sie **Archiphonem** und notiert sie abweichend von den sonstigen Konven-
tionen mit großen Buchstaben, also /T/.

Stellt man die genannten Positionen gegenüber, ergeben sich letztlich
drei unterschiedliche phonologische Bewertungen ein und desselben
Sachverhalts, wie man an den verschiedenen phonologischen Transkriptionen
unseres Beispielwortes *Hund* ablesen kann: /hunt/, /hund/ oder
/hunT/.

Diese Ergebnisse sind einerseits Ausdruck prinzipiell anderer Auf-
fassungen vom Gegenstand Sprache selbst; sie sind andererseits bedingt
durch unterschiedliche theoretische Ansätze und methodische Zugriffs-
weisen.

Literatur

- BUBMANN, HADUMOD 2002: Lexikon der Sprachwissenschaft. 3. Aufl. Stuttgart
- CRYSTAL, DAVID 1993: Die Cambridge Enzyklopädie der Sprache. Frankfurt/M. (engl. Original: The Cambridge Encyclopedia of Language. Cambridge usw. 1987)
- COULMAS, FLORIAN 1985: Reden ist Silber, Schreiben ist Gold. In: Zeitschrift für Literaturwissenschaft und Linguistik 15, H. 59, S. 94-112
- DUDEN Band 6: Aussprachewörterbuch. Wörterbuch der deutschen Standardaus-sprache. 4., neu bearb. u. erw. Aufl. Mannheim usw. 2000
- DUDEN Band 4: Grammatik der deutschen Gegenwartssprache, hg. v. der Duden-redaktion. 6., neu bearb. Aufl. Mannheim usw. 1998
- EISENBERG, PETER 1998: Grundriß der deutschen Grammatik. Bd. 1: Das Wort. Stuttgart
- GREWENDORF, GUNTHER/HAMM, FRITZ/STERNEFELD, WOLFGANG 1987: Sprachliches Wissen. Eine Einführung in moderne Theorien der grammatischen Beschreibung. Frankfurt/M.
- HAKKARAINEN, HEIKKI J. 1995: Phonetik des Deutschen. München
- HENGARTNER, THOMAS/NIEDERHAUSER, JÜRIG 1993: Phonetik, Phonologie und phonetische Transkription. Grundzüge, Begriffe, Methoden und Materialien. Aarau usw.
- JAKOBSON, ROMAN 1974: Form und Sinn. Sprachwissenschaftliche Betrachtungen. München
- MAAS, UTZ 1999: Phonologie. Opladen/Wiesbaden
- MEINHOLD, GOTTFRIED/STOCK, EBERHARD 1980: Phonologie der deutschen Ge-genwartssprache. Leipzig
- MILLER, GEORGE A. 1993: Wörter. Streifzüge durch die Psycholinguistik. Heidel-berg usw. (engl. Original: The Science of Words. New York 1991)
- PÉTURSSON, MAGNUS/NEPERT, JOACHIM 2002: Elementarbuch der Phonetik. 3., durchges. u. bearb. Aufl. Hamburg (1. Aufl. 1991)
- POMPINO-MARSCHALL, BERND 2003: Einführung in die Phonetik. 2., durchges. u. erw. Aufl. Berlin/New York
- RAMERS, KARL H. 2001: Einführung in die Phonologie. 2., unveränderte Aufl. München
- TERNES, ELMAR 1999: Einführung in die Phonologie. 2., verb. u. erw. Aufl. Darm-stadt (1. Aufl. 1987)

4 Morphologie: Die Lehre von den Bausteinen der Wörter

- 4.1 Was versteht man in der Linguistik unter Morphologie
- 4.2 Morphematische Strukturen von Wörtern
- 4.3 Wie erhält man kleinste sprachliche Einheiten?
 - 4.3.1 Segmentierungsverfahren
 - 4.3.2 Allomorphie
 - 4.3.3 Portemanteau-Morpheme
- 4.4 Morphologische Konstruktionen
 - 4.4.1 Affigierung
 - 4.4.2 Andere morphologische Konstruktionen
- 4.5 Klassifikation der Morpheme unter lexikalischen Gesichtspunkten

4.1 Was versteht man in der Linguistik unter Morphologie?

Die **Morphologie** beschäftigt sich mit den Erscheinungsformen, der Struktur und den Bauformen von Wörtern. Wörter können in **Morpheme** zerlegt werden. Das Morphem ist – wie in Kap. 1 bereits beschrieben – eine Einheit des Sprachsystems; mit Morphemen bezeichnet man die kleinsten bedeutungstragenden Einheiten der Sprache, d. h. Einheiten, die nicht weiter zerlegt werden können, ohne dass ihre Bedeutung verloren geht.

Nach SAUSSURES Zeichenmodell kommt ihnen eine Ausdrucks- und eine Inhaltsseite zu. Bei der Ausdrucksseite handelt es sich in der geschriebenen Sprache um eine graphische Einheit, die aus einem oder mehreren Buchstaben bestehen kann, in der mündlichen Sprache um eine Einheit aus einem oder mehreren Lauten.

Die meisten Wörter der deutschen Sprache lassen sich so, wie sie im tatsächlichen Sprachgebrauch vorkommen, in mehrere Morpheme aufgliedern, z. B.:

Er geht.

Die Verbform ist in die beiden Bestandteile *geh-* und *-t* zu zerlegen. Wir finden hier zwei von der Ausdrucks- und Inhaltsseite verschiedene Morpheme, wobei dem Morphem *-t* offensichtlich eine andere Art von „Inhalt“ zukommt als dem Morphem *geh-*. Der Ausdrucksseite, die hier aus der Graphemfolge bzw. dem Graphem <g> + <e> + <h> und <t> besteht, lässt sich im ersten Fall die Bedeutung 'sich [mit den Füßen] fortbewe-

gen', im zweiten Fall die grammatische „Bedeutung“ bzw. Funktion '3. Pers. Sing. Präs. Ind. Akt.' zuordnen.

Wenn gesagt wurde, dass die Morphologie die kleinsten bedeutungstragenden Einheiten der Sprache beschreibt, dann wird Bedeutung also in einem sehr weiten Sinne verstanden: einmal handelt es sich um **lexikalische Bedeutungen** (wie in *Haus*, *geh-* oder *Garten*), zum zweiten um **funktionale** (bzw. **grammatische**) **Bedeutungen** (Konjugationsmorpheme wie *-e*, *-en* oder *-t*; Deklinationsmorpheme wie das *-s* für den Genitiv oder das *-e* für den Plural von Nomen; Wortbildungs- oder Derivationsmorpheme wie *ver-*, *ent-* oder *-heit* und *-los*; Genaueres s. u.).

An dieser Stelle wird bereits deutlich, dass ein Morphem (besonders wenn es eine grammatische Funktion erfüllt) mehrere Bedeutungen haben kann, wie das Morphem *-e*, das u. a. folgende grammatische Funktionen vertritt: '1. Pers. Sing. Ind.' bei Verben und 'Pl.' bei Nomen.

Morpheme dürfen nicht mit Silben verwechselt werden, welche (als Silben) nur unter lautlichen Gesichtspunkten, nicht aber nach lexikalischen oder grammatischen Kriterien beschreibbar sind. Folgende Fälle können vorkommen:

- A Ein Morphem kann aus einer einzelnen Silbe bestehen (wie *leb-* in *lebenslos*).
- B Ein Morphem kann aus mehreren Silben bestehen (z. B. *Arbeit* = 1 Morphem, aber 2 Silben).
- C Eine Silbe kann mehrere Morpheme enthalten (vgl. *kannst*).
- D Morphemgrenzen stimmen oft nicht mit Silbengrenzen überein; vgl. *lau | fen* [Silbengrenze]; *lauf | en* [Morphemgrenze].

4.2 Morphematische Strukturen von Wörtern

Aufgrund ihrer morphematischen Struktur lassen sich Wörter in drei Klassen unterteilen:

1. **Simplizia** (einfache Wörter); sie enthalten ein **Kernmorphem*** (mit einer lexikalischen Bedeutung) und eventuell ein oder mehrere **grammatische Morpheme** (in der graphischen Darstellung steht „K“ für Kernmorphem, „F“ für Flexionsmorphem):

* Für Morpheme mit lexikalischer Bedeutung, die den Kern von Wörtern flektierbarer Wortarten bilden, finden sich in der Literatur, je nach theoretischem Ansatz, verschiedene Bezeichnungen mit je eigenen Definitionen, u. a.: Kern-, Basis-, Grund-, Stamm-, lexikalisches, Morphem. Vgl. hierzu z. B. FLEISCHER/BARZ 1992, 24f.; siehe auch unten.