

Grundriss der deutschen Grammatik

Band 1: Das Wort

Bearbeitet von
Peter Eisenberg

4., aktualisierte und überarbeitete Auflage 2013. Buch. XIV, 515 S. Hardcover

ISBN 978 3 476 02425 1

Format (B x L): 17 x 24,4 cm

Gewicht: 1084 g

[Weitere Fachgebiete > Literatur, Sprache > Angewandte Sprachwissenschaft > Studien zu einzelnen Sprachen & Sprachfamilien](#)

Zu [Inhaltsverzeichnis](#)

schnell und portofrei erhältlich bei

**beck-shop.de**
DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.

978-3-476-02425-1 Eisenberg, Grundriss der deutschen Grammatik/Band 1: Das Wort/
4., aktualisierte und überarbeitete Auflage
© 2013 Verlag J.B. Metzler (www.metzlerverlag.de)



J.B.METZLER

2. Die phonetische Basis

2.1 Phonetik und Sprachsignal

2.1.1 Übersicht: Phonetische Beschreibung von Lautereignissen

Artikulatorische, akustische und auditive Phonetik

Viele Schallereignisse, die uns täglich begegnen, deuten wir spontan und mit großer Sicherheit. Wir erkennen, ob es an der Tür klingelt, ob ein Auto anfährt, ein Glas herunterfällt, ein Klavier spielt oder ob jemand spricht. Sprachliche Äußerungen in gesprochener Sprache sind an Schallereignisse besonderer Art gebunden, die man als *Lautereignisse* bezeichnet. Wir erkennen ein Lautereignis als sprachlich im Allgemeinen auch dann, wenn aus irgendwelchen Gründen nicht zu verstehen ist, was geäußert wurde.

Verstehen wir eine Äußerung, dann haben wir das Lautereignis in geeigneter Weise verarbeitet. Das Lautereignis ist so strukturiert, dass es bei Äußerung einer sprachlichen Einheit – z. B. eines Satzes – erkennbar und verarbeitbar ist. Wie es strukturiert ist, sagt uns die Phonetik. Die Phonetik ermittelt und beschreibt die Struktur von Lautereignissen, wie sie bei gesprochenen Äußerungen von sprachlichen Einheiten entstehen.

Die Struktur eines Lautereignisses kann auf unterschiedliche Weise erfasst werden, und jeder dieser Weisen entspricht eine Teildisziplin der Phonetik. Verfolgen wir seinen Weg von der Erzeugung beim Sprecher bis zur Verarbeitung beim Hörer.

Die Erzeugung eines Lautereignisses beim Sprecher nennt man seine Artikulation, wobei als Artikulation häufig nur der äußere, an die Bewegung der Artikulationsorgane gebundene Teil des Produktionsprozesses gilt. Als solcher ist er Gegenstand der *artikulatorischen Phonetik*. In ihrer modernen Form beschäftigt sich die Phonetik aber auch damit, welche neurologischen und neurophysiologischen Prozesse der Artikulation zugrunde liegen und wie der Sprecher den Artikulationsprozess durch Mithören und Rückbindung taktiler Signale überwacht und steuert.

Die große praktische Bedeutung sowie die vergleichsweise lange Geschichte der artikulatorischen Phonetik beruhen darauf, dass Artikulationsvorgänge in ihren Grundzügen vergleichsweise leicht beobachtbar sind. Ein [m] wie in **Maus** wird mit geschlossenen Lippen gebildet, ein [f] wie in **Fisch** mit der Unterlippe an der oberen Zahnreihe. Beides sieht und fühlt man. Beim [n] (**neu**) sieht man kaum etwas, aber man fühlt, dass der vordere Zungenteil am Zahndamm hinter der oberen Zahnreihe anliegt. Natürlich sind nicht alle Artikulationsvorgänge auf einfache Weise und direkt beobachtbar, sie sind es aber im Vergleich zu den Hörvorgängen. An das Innenohr etwa kommt man nur schwer heran. Die Beschreibung von Sprachlauten auf der Basis von Hörvorgängen war lange Zeit hindurch nicht viel mehr als eine theoretische Möglichkeit.

Die artikulatorische Phonetik erforscht die Artikulation sprachlicher Einheiten insgesamt. Sie beschränkt sich nicht auf die Analyse von Einzellauten, sondern betrachtet ebenso Laute im Kontext und größere Einheiten wie Silben, Wortformen und Sätze. Und sie beschreibt den Artikulationsvorgang als Ganzen, mit allen daran beteiligten Organen (2.1.3).

Das durch Artikulationsvorgänge entstandene Lautereignis kann nun wie jedes andere Schallereignis physikalisch untersucht werden. Es wird damit zum Gegenstand der *akustischen Phonetik* (2.1.2).

Die akustische Phonetik macht etwa Aussagen darüber, welche physikalischen Eigenschaften bei ›demselben‹ Laut in verschiedenen Kontexten konstant bleiben. Beispielsweise stellt man bei genauem Hinhören fest, dass das [u] in Wörtern wie **Mut** und **Krug** gewisse Unterschiede aufweist. Diese Unterschiede kann man messtechnisch erfassen. Daneben weist das [u] aber in beiden Kontexten physikalische Gemeinsamkeiten – etwa im Frequenzgemisch – auf, die die akustische Phonetik eben als charakteristisch für den Vokal [u] erweist.

Neben den Eigenschaften von Lauten im Kontext misst die akustische Phonetik Tonhöhen- und Lautstärkeverläufe innerhalb größerer sprachlicher Einheiten wie Wörtern, Phrasen und Sätzen. Was bedeutet es physikalisch, wenn wir in **Ámeise** die erste und in **Amóbe** die zweite Silbe als betont hören? Was ist physikalisch der Unterschied zwischen **Sie kauft ein Áuto** und **Sie kauft éin Auto**? Die akustische Phonetik sagt uns auch, wie sich die physikalischen Eigenschaften von sprachlichen Einheiten in Abhängigkeit von der Sprechgeschwindigkeit ändern, und sie trennt allgemeine von den eigentlich strukturierenden Eigenschaften: Eine Kinderstimme hat im Allgemeinen eine höhere Grundfrequenz als eine Frauenstimme; Heiserkeit, Flüstern, Singen, lautes und leises Sprechen lassen sich physikalisch von normalem Sprechen leicht unterscheiden.

Bei all dem interessiert sich die Phonetik nicht nur für das fertige Lautereignis und den Artikulationsvorgang, sondern sie möchte wissen, wie beide zusammenhängen. Welche Artikulationsvorgänge führen zu bestimmten akustischen Eigenschaften? Gibt es nur einen artikulatorischen Weg, ein bestimmtes akustisches Ergebnis zu erzielen, d. h. verhalten sich alle Sprecher im Prinzip artikulatorisch gleich?

Zur Charakterisierung eines Lautereignisses als sprachlich gehört schließlich die Kenntnis seiner Verarbeitung durch das Gehör. Welche physikalischen Eigenschaften eines Lautereignisses der Sprecher als sprachlich bewertet, was er etwa als einen Laut, eine Silbe, die lautliche Form eines Wortes, als Akzentmuster oder Intonationsbogen hört, ist Gegenstand der *auditiven Phonetik*.

Dazu gehört zunächst das Verständnis der Arbeitsweise des Ohres und des Gehörsinnes selbst, auch unabhängig vom Sprachlichen. Welche Tonhöhen, Lautstärken, Tonhöhen- und Lautstärkedifferenzen hört der Mensch und wie hört er sie? Beispielsweise werden Töne als im gleichen Abstand – nämlich einer Oktave – zueinander gehört, wenn wir ihre Frequenz jeweils verdoppeln. Die Empfindung der Tonhöhe ist logarithmisch proportional der Frequenz. Ähnlich die Lautstärkeempfindung. Ihr physikalisches Korrelat, die Dichte der Schallenergie, muss exponentiell wachsen, damit der Eindruck einer linearen Lautstärkeerhöhung entsteht.

Ein für die Sprachwahrnehmung wichtiges und interessantes Gebiet der auditiven Phonetik ist das des kategorialen Hörens. In modernen Phonetiklabors ist es problemlos möglich, Lautereignisse künstlich zu synthetisieren. Einzellaute kann man dabei aus ihren physikalischen Bestandteilen so zusammensetzen, dass sich auch alle Übergänge zwischen verwandten Lauten herstellen lassen, etwa zwischen dem [s] in **reißen** und

dem [f] in **reifen**. Geht man nun in kleinen Schritten von einem echten [s] zu einem echten [f] über und bittet man eine Versuchsperson, zu sagen, was sie höre, so wird in der Regel entweder der eine oder der andere Laut identifiziert. Die Zone des Übergangs, d. h. der Unsicherheit, ist relativ klein und ändert sich wenig von Person zu Person. Welcher Laut gehört wird, ist andererseits aber wieder kontextabhängig. Bietet man vor dem Übergang viele [s] an, so wird früher als sonst ein [f] gehört, d. h. der Kontrast zum Ewiggleichen wird höher bewertet. Der Bereich des [f] dehnt sich aus. Auditive Untersuchungen dieser Art sind unerlässlich, wenn man verstehen möchte, welche physikalischen und damit artikulatorischen Eigenschaften von Lauten unter welchen Bedingungen funktional sind.

Wie wir sie dargestellt haben, erscheint die akustische Phonetik als Mittlerin zwischen der artikulatorischen einerseits und der auditiven andererseits. Physikalische Daten werden interpretiert in Hinsicht darauf, wie sie auditiv wahrgenommen werden und auf welchen artikulatorischen Vorgängen sie beruhen. Sprachwissenschaftlich interessiert letztlich aber die *direkte Beziehung* zwischen Artikulation und auditiver Wahrnehmung. Ein artikulatorischer Unterschied beliebiger Art kann linguistisch von Interesse sein, wenn er hörbar ist, und hörbare Unterschiede müssen auch artikulierbar sein. Die Regularitäten, denen die Lautstruktur folgt, sind an die Fähigkeiten der Artikulationsorgane genau so gebunden wie an die des Gehörssinnes.

So zeigt sich im Aufbau einer Silbe wie [ʃtrʊmpf] (**Strumpf**) eine Regularität, die wohl artikulatorisch fundiert ist. Die drei Laute in der Mitte, nämlich das [ʀ], [ʊ] und [m], haben Stimmtöne, die vier äußeren nicht. Sie sind stimmlos. Das kann man mit der Feststellung verallgemeinern, dass in jeder Silbe die Laute mit Stimmtönen zusammenstehen. Sie bilden einen Block. Als Erklärung drängt sich ein Ökonomiegesichtspunkt auf: Der Stimmtöne wird genau einmal pro Silbe aktiviert. Jede weitere Aktivierung würde zu einer weiteren Silbe führen.

Anders verhält es sich wahrscheinlich mit bestimmten Konsonanthäufungen, wie wir sie im Deutschen am Silbenanfang vorfinden. Als Folgen von drei Konsonanten kommen vor [ʃpl] (**Splint**), [ʃpr] (**Sprung**) und [ʃtr] (**Strumpf**). Warum gerade diese drei und nicht etwa auch [ʃtr] in einer Form wie ***Ftrich**? Wahrscheinlich fehlt diese Form aus guten Gründen. Es könnte sein, dass [f] von [ʃ] auditiv nicht genügend absticht, dass also der spezielle Kontext zu einer auditiven Beschränkung führt. Damit soll nicht behauptet werden, im ersten Beispiel spiele allein Artikulatorisches, im zweiten allein Auditives eine Rolle. Man kann jeweils gut den anderen Gesichtspunkt ins Spiel bringen. Eine Berufung auf artikulatorische oder auditive Beschränkungen allein reicht jedenfalls insgesamt nicht aus (dazu weiter 4.2).

Wenn wir Laute mit artikulatorischen Merkmalen wie »bilabial« (mit beiden Lippen gebildet) oder »apikal« (mit der Zungenspitze gebildet) beschreiben, dann meinen wir damit in der Regel auch etwas Auditives nach dem Muster »Ein Laut, der auf eine bestimmte Weise gebildet ist, wird auf eine darauf beziehbare Weise gehört.« An dieser Formulierung wird ein wesentlicher Unterschied zwischen Artikulatorischem und Auditivem deutlich, den wir noch einmal hervorheben wollen.

Eine auditive Phonetik ist einer artikulatorischen nicht insofern komplementär, als letztere Laute mit Produktionsmerkmalen beschreibe, erstere aber mit Perzeptionsmerkmalen. Eine auditive Phonetik, die Laute etwa mit Begriffen der Anatomie des Ohres beschreibt, hat es nie gegeben. Gäbe es sie, wäre sie nicht das Komplement zur artikulatorischen Phonetik, sondern sie könnte diese allenfalls ersetzen. Eine auditive Phonetik als Wahrnehmungslehre würden wir dennoch benötigen. Sprachliche Einheiten werden ge-

hört im Sinne von »wahrgenommen«. Es handelt sich um das Erkennen strukturierter Gebilde als der Formseite sprachlicher Einheiten.

Signalphonetik, Experimentalphonetik und Symbolphonetik

Artikulatorische, akustische und auditive Phonetik sind die klassischen Teilgebiete dieser Disziplin, wenn man sich am Weg des Signals von seiner Entstehung zu seiner Perception orientiert. Systematisch wird die Phonetik heute aber meist anders gegliedert. Einer Signalphonetik steht dann eine Experimentalphonetik gegenüber. Die *Signalphonetik* misst, was sich beim Sprechen als Signal messen lässt, von der Lippenbewegung bis zur Hirndurchblutung. Die *Experimentalphonetik* dagegen untersucht Zusammenhänge zwischen dem als Signal Gemessenen und dem, was wahrgenommen wird.

Diesen beiden steht als dritte Teildisziplin die *Symbolphonetik* gegenüber, von der bisher noch gar nicht die Rede war. Die Symbolphonetik entwickelt Verfahren, mit denen sprachliche Lautereignisse so geschrieben werden können, dass ihre wesentlichen Eigenschaften aus dem Geschriebenen rekonstruierbar sind. Das ideale symbolphonetische System ist zur graphischen Fixierung von Lautereignissen aus sämtlichen Sprachen geeignet.

Zielsetzungen dieser Art sind mit der Konzeption verschiedener phonetischer Alphabete verbunden, unter ihnen das Internationale Phonetische Alphabet IPA, das wir in unserer Grammatik so weit wie möglich verwenden.

Das IPA stellt Zeichen zur Verfügung, mit denen viel mehr artikulatorische und auditive Details darstellbar sind als man sie für die meisten Zwecke braucht. Die Kunst im Umgang mit dem phonetischen Alphabet besteht darin, dass man genau das hinschreibt, was man hört. Jede sprachliche Einheit erscheint dabei als eine Verkettung von Symbolen für Laute, Akzente, Pausen usw. Ein geschulter Ohrenphonetiker ist in der Lage, eine genaue Transkription für Einheiten aus ihm unbekanntem Sprachen anzufertigen. Genauso ist er aber in der Lage, eine »objektive«, d. h. nur am Lautereignis selbst orientierte Transkription von Einheiten aus Sprachen anzufertigen, die er kennt. Projektionen von der einen auf die andere Sprache vermeidet er, darauf kommt es vor allem an. Seine Ohren hören die Laute unabhängig von der Einzelsprache.

Weil die Arbeitsweise der Symbolphonetik sowohl von großem praktischen Interesse als auch grundlegend für ein Verständnis der Beziehung zwischen gesprochener und geschriebener Sprache ist, behandeln wir sie relativ ausführlich in einem eigenen Abschnitt (2.3). Abschnitt 2.2 sagt etwas über die Artikulation von Konsonanten und Vokalen, wobei schon die Verhältnisse im Deutschen im Vordergrund stehen. Zunächst geht es aber um die ganz allgemeine Darstellung einiger auditiver sowie akustischer Eigenschaften von Lauten einerseits (2.1.2) und des Artikulationsapparates andererseits (2.1.3).

Auch wenn unsere Ausführungen zur Phonetik wenig ausführlich sind, kann man sich fragen, ob sie überhaupt in eine einzelsprachliche Grammatik gehören. Eine Grammatik beschäftigt sich mit Form und Funktion der Einheiten einer Sprache, wobei die Beschreibung der lautlichen Form sprachlicher Einheiten in Hinsicht auf ihre Funktion Gegenstand der Phonologie ist. Darüber hinaus ist vieles, was wir zur Phonetik sagen, nicht spezifisch für das Deutsche, auch wenn fast alle illustrativen Beispiele dem Deutschen entnommen sind.

Gewisse phonetische Grundlagen sind dennoch für jede einzelsprachliche Grammatik unerlässlich. Schon weil die Grenze zwischen Phonetik und Phonologie umstritten ist und recht unterschiedlich gezogen wird, weil es Unklarheiten über den für die Phonologie relevanten Funktionsbegriff gibt und weil phonologische Begriffe in vielen Ansätzen

phonetisch fundiert sein sollen, ist das der Fall. Gerade in neueren phonologischen Theorien spielen phonetische Gesichtspunkte eine immer größere Rolle.

Wir liefern im vorliegenden Kapitel ein auf das Deutsche bezogenes phonetisches Minimum, wollen und können damit aber keinesfalls eine Einführung in die Phonetik ersetzen. Deshalb hier zusammenfassend einige Literaturhinweise.

Klassisch gewordene Einführungen sind von Essen 1953 und Ladefoged 1993 (1. Aufl. 1975; Schwerpunkt artikulatorische Phonetik). Mit weitreichendem Erklärungsanspruch bezüglich der Fundierung sprachlicher Kommunikation Tillmann/Mansell 1980. Eine ausgefeilte Darstellung insbesondere der Artikulation unter Berücksichtigung der Verhältnisse im Deutschen gibt Kohler 1995 (1. Aufl. 1977). Eine moderne, kurzgefasste Gesamtdarstellung findet sich in Pompino-Marschall 2009, mit Verbindung zur Phonologie R. Wiese 2011. Elementar fundierend mit vergleichsweise viel akustischer Phonetik ist Reetz 2003, mit vergleichsweise viel neurophysiologisch-auditiver Phonetik Pétursson/Neppert 2002.

2.1.2 Töne, Geräusche, Laute

Stimmton und musikalischer Ton

Die im Alltagsverständnis wichtigen Begriffe zur Beschreibung von *Tönen* sind Tonhöhe, Klangfarbe und Lautstärke. Sie lassen sich akustisch explizieren mit Hilfe von Grundbegriffen der Schwingungslehre. Der akustisch einfachste Ton, der Sinuston, hat eine bestimmte Frequenz, die die vom Ohr wahrgenommene Tonhöhe bestimmt. Die Tonhöhe ist proportional der Frequenz. Der (junge) Mensch hört Töne mit einer Frequenz von etwa 20 Hertz (Schwingungen pro Sekunde) bis 20.000 Hertz.

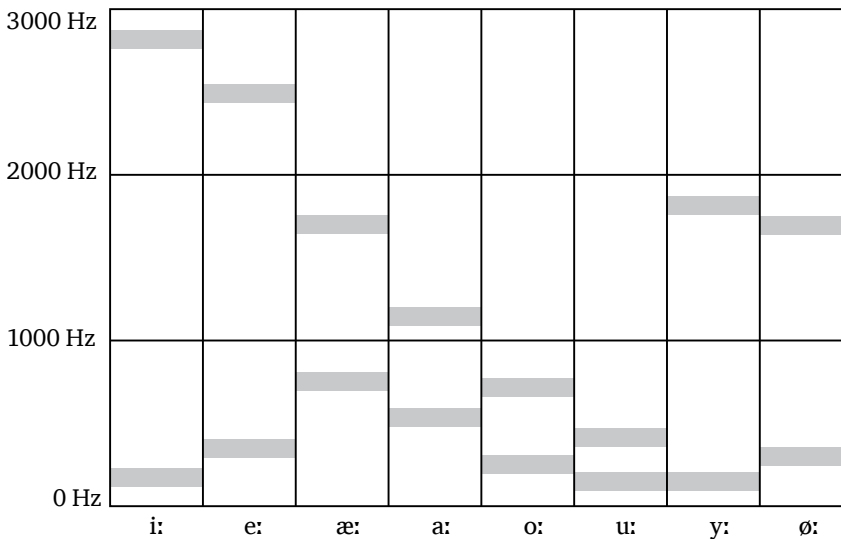
Die Lautstärke, mit der ein Sinuston wahrgenommen wird, ist abhängig von seiner Amplitude, dem Ausschlag der Schwingung. Direkt proportional der Amplitude ist der Schalldruck, wie er etwa auf der Membran eines Mikrophons entsteht. Dem Quadrat des Schalldrucks proportional ist die sog. Schallintensität oder Schalldichte, auf die die wahrgenommene Lautstärke bezogen ist. Wird eine Schallintensität I_1 als Lautstärke L_1 gehört, dann wird eine Schallintensität I_2 mit $\log I_2 = 2 \log I_1$ als doppelt so laut empfunden. Die Schallintensität muss also exponentiell steigen, damit der Eindruck einer linear ansteigenden Lautstärke entsteht. So ist das Ohr ganz gut gegen Lärmschäden geschützt. Die Lautstärke wird in dB (Dezibel) gemessen. Bei einem Ton von 1000 Hz entspricht 0 dB der Hörschwelle des Ohres. Flüstern hat 20 dB, normales Sprechen 40 dB, starker Verkehrslärm 80 dB und bei etwa 130 dB ist die Schmerzgrenze erreicht. Die ohrenbezogenen dB heißen auch Phon.

Reine Sinustöne kann man künstlich erzeugen, als natürliche Töne kommen sie nicht vor. Ein natürlicher Ton ist wie ein Sinuston periodisch, nur hat die Schwingung nicht die Sinusform. Jedoch ist jeder natürliche Ton aus Sinustönen aufgebaut, d. h. man kann seine Form rechnerisch und ihn selbst messtechnisch in Sinustöne unterschiedlicher Frequenz, seine Teiltöne, zerlegen. Der Teilton mit der niedrigsten Frequenz heißt Grundton, höhere Teiltöne heißen Obertöne. Die Frequenz der Obertöne beträgt ganzzahlige Vielfache der Frequenz des Grundtones. Die Stärke der einzelnen Teiltöne macht für das menschliche Ohr die Klangfarbe des resultierenden Tones aus. In der Akustik bezieht man sich mit ›Ton‹ meist auf reine Sinustöne, sonst spricht man von ›Klängen‹. Wir bleiben im Folgenden meist beim umgangssprachlichen ›Ton‹.

Bei musikalischen Tönen ist das Frequenzgemisch so, dass für das Ohr bestimmte höhere Frequenzen besonders in Erscheinung treten, das sind die Töne der sog. natürlichen Obertonreihe. Der tiefste Teilton bestimmt als Grundton die Tonhöhe für das Ohr. So liegt die Grundfrequenz des Kammertons a, den man mit einer Stimmgabel erzeugt, bei 440 Hz. Der erste Oberton eines musikalischen Tones liegt eine Oktave über dem Grundton. Physikalisch heißt das, er hat die doppelte Frequenz, Verhältnis 1:2 (beim a 880 Hz). Der zweite gehörte Oberton liegt eine Quinte über dem ersten, Frequenzverhältnis 2:3 (1320 Hz). Der dritte ist die Quarte über dem zweiten, Frequenzverhältnis 3:4, also 1760 Hz und damit zwei Oktaven über dem Grundton. So geht es bei den musikalischen Tönen weiter mit den ›Harmonischen‹ über den gesamten Bereich, wobei die Klangfarbe eben dadurch bestimmt ist, wie stark die einzelnen Harmonischen sind.

Auch bei gesprochenen Tönen findet eine sehr spezifische Filterung der höheren Frequenzen statt (zu den zahlreichen historischen Versuchen, die Sprechstimme durch musikalische Töne nachzubilden, Gessinger 1994). Vokale haben die Schallenergie in relativ schmalen Frequenzbändern konzentriert, den Formanten. Die Vokalqualität beruht auf dem Frequenzverhältnis der Formanten. Für das menschliche Ohr reichen die beiden ersten (die tiefsten) Formanten aus, um die Vokale einer Sprache zu identifizieren. Sehr natürlich klingen Vokale mit nur zwei Formanten allerdings nicht. In 1 sind die beiden ersten Formanten für eine Reihe von Vokalen des Deutschen zusammengestellt.

(1) Formanten



Es handelt sich bei den Angaben in 1 um Durchschnittswerte für die betonten Vokale in Wörtern wie **Biene**, **Leben**, **Nähe**, **Rabe**, **Woge**, **Bruder**, **Tüte** und **Flöte**. Die absoluten Angaben sind die für eine Männerstimme. Bei Frauenstimmen liegen die Werte des ersten Formanten um zwischen einer Terz (beim [i:]) und einer Oktave (beim [u:]) höher. Noch etwas höher liegen die Werte für Kinderstimmen. Allen gemeinsam ist natürlich das Frequenzverhältnis der Formanten, denn von ihm hängt ja die Vokalqualität für das Ohr ab.

Was hier als Ton von Sprachlauten und musikalischer Ton beschrieben wurde, darf nicht verwechselt werden mit dem sog. phonologischen Ton. Mit dem phonologischen

Ton ist ein Tonhöhenverlauf gemeint, der in Tonsprachen wie dem Chinesischen dazu dient, Wörter voneinander zu unterscheiden, die bis auf den Ton lautlich identisch sind. Auch um eine Verwechslung beider Begriffe von Ton zu vermeiden, spricht man vom die Klangfarbe eines Lautes bestimmenden Ton als vom Stimmton oder einfach von Stimme, entsprechend von stimmhaften Lauten.

Geräusche

Wir kommen zur zweiten großen Gruppe von Schallereignissen, den *Geräuschen*. Geräusche weisen nicht das für Töne charakteristische Frequenzgemisch auf, das allgemein als Klang und im Einzelnen als Klangfarbe wahrgenommen wird. Geräusche sind im Vergleich zu Tönen unspezifisch: Was kein Ton ist, ist ein Geräusch. Für die Beschreibung von Sprachlauten ist es sinnvoll, drei Typen von Geräuschen zu unterscheiden.

Der erste Typ sind die *Zischgeräusche*, wie wir sie bei den Reibelauten (Frikativen) vorfinden, beispielsweise dem [f] in **Waffe**, dem [s] in **Wasser**, dem [ʃ] in **Asche** und dem [ç] in **Sichel**. Diese Laute haben gemeinsam, dass der Luftstrom durch eine von irgendwelchen Artikulationsorganen gebildete Friktionsenge gepresst wird (2.1.3).

Physikalisch sind Zischgeräusche als breitbandig zu kennzeichnen. Die Schallenergie ist nicht wie bei den Tönen auf bestimmte Frequenzen oder schmale Frequenzbänder konzentriert, sondern auf breite Bereiche verteilt. Von bestimmten einzelnen Frequenzen als Charakteristikum eines derartigen Geräusches lässt sich nicht mehr sprechen, und folglich auch nicht davon, dass es sich um einen periodischen Vorgang handle. Das Frequenzgemisch ist als solches nur statistisch beschreibbar. Trotzdem gibt es natürlich physikalische Charakteristika von Zischgeräuschen, die das Ohr wahrnimmt. Ausschlaggebend ist, in welchen Frequenzbereichen ein Geräusch seine Energie konzentriert und wie das Verhältnis der Frequenzbänder zueinander ist.

Ein ganz anderer Typ von Geräusch entsteht als periodisch gegliedertes Schallereignis. Die elementaren Schallereignisse dieser Gliederung folgen so schnell aufeinander (ca. 20 Hz), dass für das Ohr nicht mehr das einzelne Ereignis zum Gegenstand der Wahrnehmung wird, wie es etwa bei einer Folge von Schlägen der Fall wäre. Die Ereignisfrequenz ist andererseits so niedrig und das Einzelereignis mit so viel Geräusch verbunden, dass ein tiefer Brummtton sozusagen nur als Nebenprodukt entsteht. In einem solchen Fall ist »das wahrgenommene Resultat ... ein konstantes Klangereignis mit der inneren Ereignisstruktur eines Trillers« (Tillmann/Mansell 1980: 40).

Auf einem *Trillergeräusch* beruht von den Lauten des Deutschen zweifelsfrei das explizit artikulierte Zungen-r, geschrieben [r]. Beim [r] hört man die einzelnen Schläge der Zunge als Bestandteil eines Kontinuums, begleitet von einem tiefen Brummtton. In der Regel wird auch das Zäpfchen-r (transkribiert als [ʀ]) hier angesiedelt, jedenfalls kann es bei expliziter Artikulation als Trillergeräusch realisiert werden. Die Grenze zum Zischgeräusch (Frikativ) bleibt aber fließend. Zu den Trillergeräuschen gehört weiter das Knarren, das wir am Kehlkopf erzeugen können; und ebenso ein Lippengeräusch, das entsteht, wenn wir die Lippen im austretenden Luftstrom rhythmisch gegeneinander schwingen lassen.

Der dritte Typ von Geräusch ist sowohl aperiodisch als auch nicht kontinuierlich, nennen wir ihn *Schlaggeräusch*. Die Schallenergie ist kurzzeitig auf breite Frequenzbänder verteilt. Der Einsatz des Geräusches ist scharf begrenzt, was aber nicht verhindert, dass Schlaggeräusche physikalisch und für das Ohr fein differenzierbar sind.

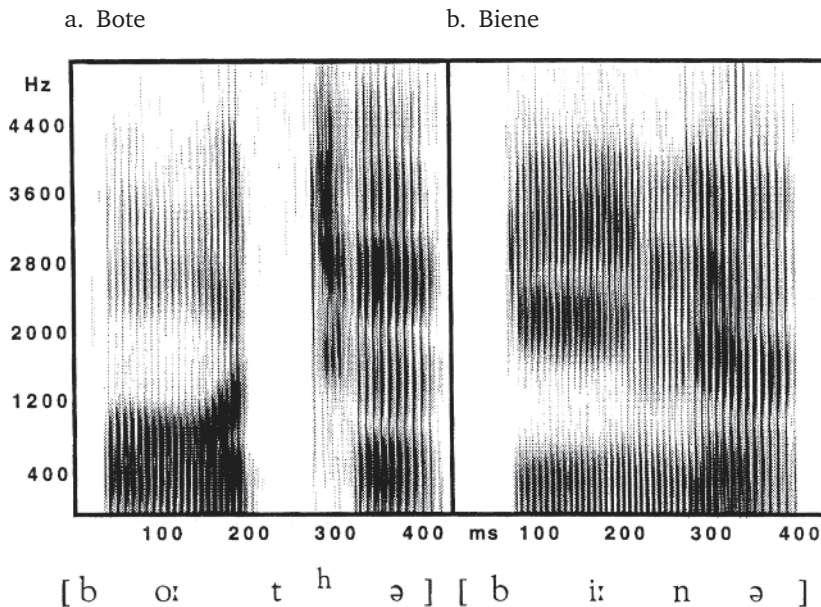
Sprachlaute mit den Charakteristika von Schlaggeräuschen sind die Plosive. Der Schlag entsteht, wenn ein Verschluss im Mundraum oder am Kehlkopf unter Druck ge-

setzt und übergangslos geöffnet wird. Beispiele aus dem Deutschen sind das [p] wie in **Pilz** (Verschluss mit beiden Lippen), das [t] wie in **Tuch** (Verschluss mit der Zunge am Zahndamm), das [k] wie in **Koch** (Verschluss mit dem hinteren Teil des Zungenrückens und Gaumen) sowie der sog. glottale Verschlusslaut [ʔ] vor Vokalen am Wortanfang (Verschluss der Stimmritze).

An der Grenze zwischen Schlag- und Zischgeräuschen liegt das Schnalzen. Es entsteht durch Unterdruckbildung im Mundraum und Öffnen eines vorderen Verschlusses, z. B. des Verschlusses durch Anlegen der Vorderzunge an den Zahndamm. Schnalzlaute gibt es ebenfalls in mehreren Sprachen. Jeder nennt hier das Zulu und das Hottentottische.

Frequenzgemisch und Energieverteilung von Sprachlauten und größeren sprachlichen Einheiten können als *Visible Speech* in Sonagrammen anschaulich dargestellt werden (2). Die technischen Hilfsmittel dazu stehen seit den 1940er Jahren mit dem sog. Sonagraphen zur Verfügung. Auf der Abszisse des Diagramms ist die Zeit in Millisekunden aufgetragen, auf der Ordinate die Frequenz in Hz. Schwärzungen zeigen an, wo die Schallenergie konzentriert ist. Eine dritte Dimension kommt durch den Schwärzungsgrad hinzu. Er ist proportional der Energiedichte.

(2) Sonagramme



Phonetiker können Sonagramme fast so lesen wie andere ein Buch. Aber auch ohne große Übung erkennt man schon eine ganze Menge, etwa die Formanten der Vokale [o:] und [i:], die Gemeinsamkeiten im Anlaut der beiden Formen mit dem [b], den plötzlichen Geräuscheinsatz bei Sprengung des Verschlusses des [t], die geringe Energiedichte beim [n] sowie die Übereinstimmung im Auslaut der beiden Formen.

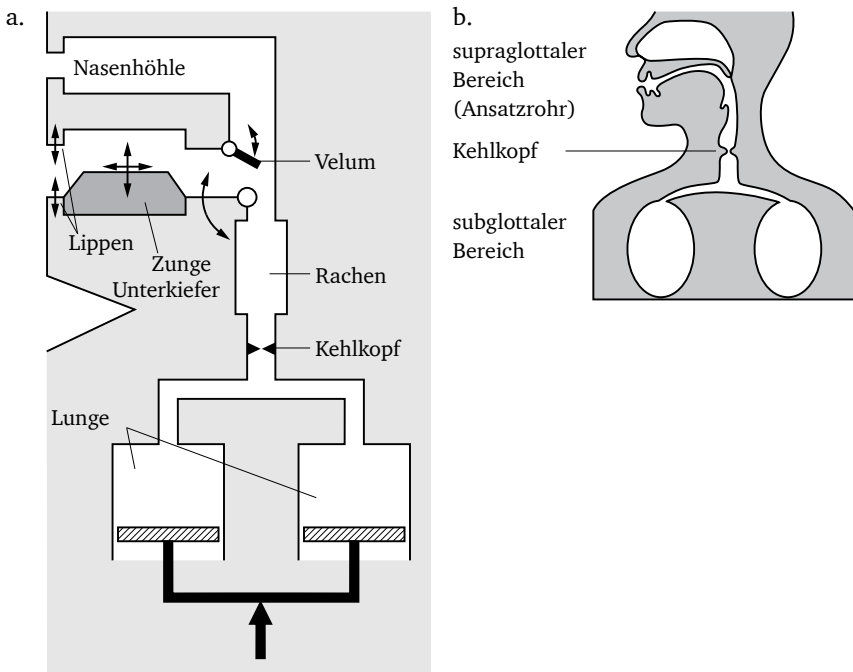
2.1.3 Artikulation

Der Artikulationsapparat

Die Artikulation der meisten Sprachlaute hat zwei Grundkomponenten. Einmal findet irgendwo im Mund- und Rachenraum eine Schallerzeugung statt, deren Resultat ein Ton oder ein Geräusch oder beides ist (Rohschall). Bei den Vokalen beispielsweise ist der Kehlkopf mit den Stimmbändern, die den Stimmton erzeugen, die primäre Schallquelle. Zum Zweiten wird der primär erzeugte Ton oder das Geräusch durch die Bewegung und Stellung der übrigen Artikulationsorgane verändert. Was als Laut hörbar wird, ist in der Regel etwas ganz anderes als der Rohschall. Für jeden einzelnen Laut muss deshalb das Verhalten des gesamten Artikulationsapparates im Blick bleiben.

Technisch gesprochen erfolgt die Artikulation von Lauten durch Manipulation von Luftströmen, in aller Regel des Luftstromes beim Ausatmen. Zur Aufzählung der Artikulationsorgane setzt man daher mindestens bei der Lunge (lat. *pulmo*) an, denn hier wird der Luftdruck für das Ausatmen erzeugt (pulmonaler Luftstrom). 1 zeigt grob schematisch, welche Wege der Luftstrom nehmen kann.

(1) Luftwege, funktionales Schema



Der für das Ausatmen und für die Artikulation erforderliche Luftdruck entsteht in der Lunge durch Bewegung des Zwerchfells und der Brustmuskulatur, die Lunge dient als direkter Energiespeicher für die Lautproduktion. Durch Öffnen des Kehlkopfes strömt die Luft zunächst in die Rachenhöhle und tritt dann entweder durch die Mundhöhle oder die Nasenhöhle nach außen. Der gesamte oberhalb des Kehlkopfes liegende Bereich (Rachen-, Mund- und Nasenhöhle) wird *Ansatzrohr* genannt. Die beiden Hauptwege im An-

satzrohr sind durch das Gaumensegel (Velum) getrennt, das als Luftweiche fungiert. Senken des Velums öffnet die Nasenhöhle, Heben des Velums schließt sie.

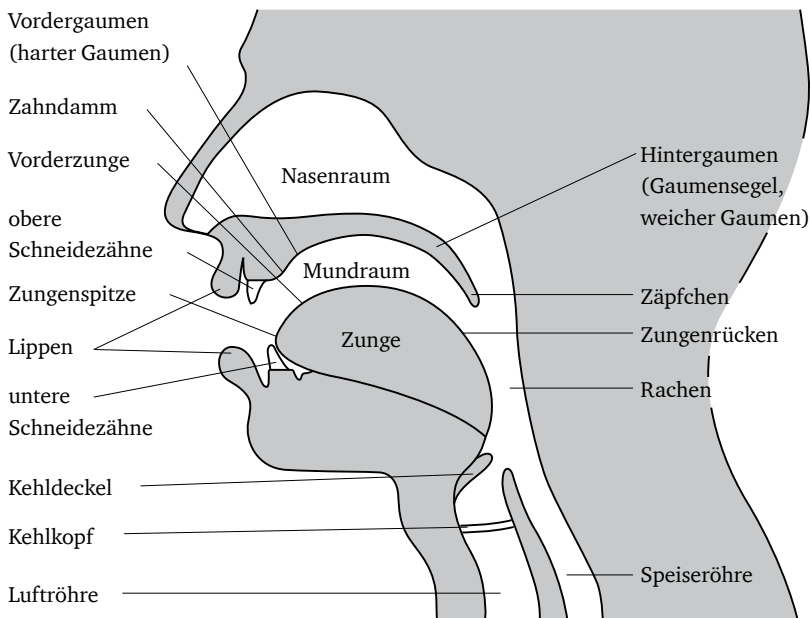
Für die Lautproduktion hat das Ansatzrohr zwei Grundfunktionen. Einmal dient es als Resonanzraum, also zur Ausgestaltung von Tönen und Geräuschen, zum Zweiten dient es der Geräuschproduktion selbst. Am wichtigsten ist dafür die Mundhöhle mit ihren beweglichen Organen Velum, Zunge, Lippen und ihrer Veränderbarkeit durch die Beweglichkeit des Unterkiefers insgesamt.

Die Schemata in 1 machen einen wichtigen Grund für Eigenart und Komplexität der Sprachlautproduktion deutlich, der uns veranlassen sollte, mit der Redeweise von ›den Artikulationsorganen‹ vorsichtig umzugehen. Die beteiligten Organe sind zum Sprechen da, aber sie sind nicht nur zum Sprechen da. Unter physiologischem Aspekt sind es einige nicht einmal in erster Linie. Zunge (trotz lingua!), Lippen, Zähne, Nase usw. haben als primäre Funktionen ganz offensichtlich dem Atmen, der Nahrungsaufnahme und vielleicht sogar der Lautproduktion unabhängig vom Sprechen zu dienen. Bei der Komplexität und Polyfunktionalität dieser Organe liegt einer der Gründe für die Schwierigkeit, die Sprachlautproduktion auf einfache und eindeutige Weise zu beschreiben. Würden wir mit einem spezialisierten Organ nur das artikulieren, was zur Differenzierung von 30 bis 60 Sprachlauten notwendig ist – wir kämen mit einer vergleichsweise einfachen Apparatur aus.

Die Artikulationsorgane

Zur Beschreibung der Hauptfunktionen einzelner Artikulationsorgane orientieren wir uns am sog. Sagittalschnitt in Schema 2. In 3 sind die wichtigsten deutschen und lateinisch/griechischen Bezeichnungen mit den daraus abgeleiteten Merkmalsbezeichnungen alphabetisch nach dem Deutschen aufgeführt (nach Duden 2009: 22).

(2) Artikulationsorgane

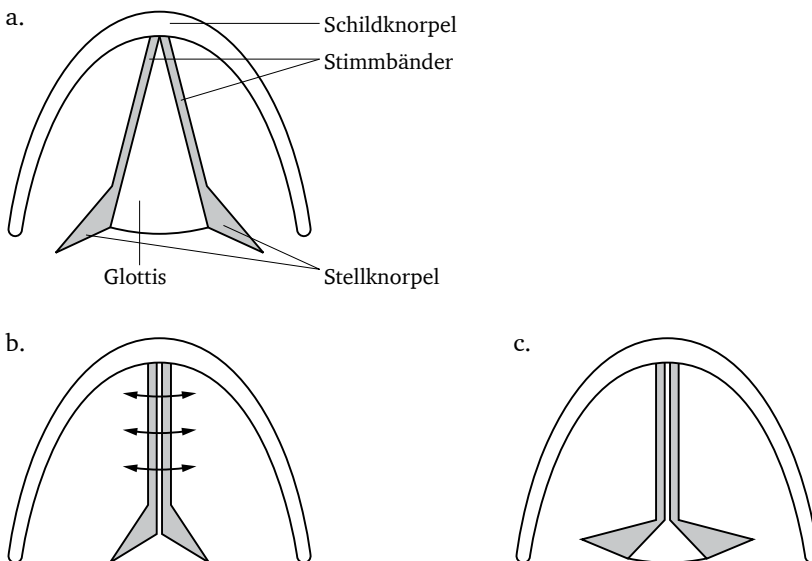


(3) Artikulationsorgane und artikulatorische Merkmale

Gaumensegel, weicher Gaumen Hintergaumen	velum (velar)
Kehlkopf	larynx (laryngal)
Lippe	labium (labial)
Mund	os (oral)
Nase	nasus (nasal)
Rachen	pharynx (pharyngal)
Stimmritze	glottis (glottal)
Vordergaumen, harter Gaumen	palatum (palatal)
Vorderzunge, Zungenkranz	corona (koronal)
Zahn	dens (dental)
Zahndamm	alveoli (alveolar)
Zäpfchen	uvula (uvular)
Zungenrücken	dorsum (dorsal)
Zungenspitze	apex (apikal)

Das komplizierteste der beteiligten Organe ist der *Kehlkopf* (der Larynx), der als Ventil das gesamte Ansatzrohr nach unten abschließt. Den zentralen Bereich des Kehlkopfes bilden die *Stimmbänder*, das sind zwei längliche, einander gegenüberstehende Muskelstränge mit komplexen Bewegungsmöglichkeiten. Die Stimmbänder sind vorne am Schildknorpel (bei Männern sichtbar als Adamsapfel) und hinten an zwei beweglichen Stellknorpeln angewachsen. Sie können gespannt werden und sie können aufeinander zu bewegt werden, so dass zwischen ihnen ein Spalt von veränderlicher Form und Größe entsteht, die *Stimmritze* oder *Glottis* (4).

(4) Kehlkopf



Beim normalen Atmen ist die Glottis offen, die Atemluft strömt ungehindert ein und aus (4a, schematische Sicht von oben). Zum Erzeugen eines Tones wird die Glottis geschlossen (4b) und im Rhythmus der Grundfrequenz des Tones geöffnet und wieder geschlossen. Das rhythmische Abreißen und Wiedereinsetzen des Luftstromes ist als Summton mit der Grundfrequenz hörbar.

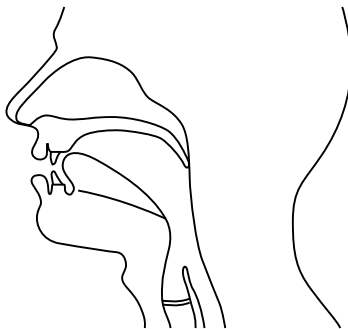
Der einzelne Öffnungs- und Schließvorgang ist physikalisch eine sog. Kippschwingung. Durch Schließen der Glottis nach dem Einatmen kann in der Lunge und Luftröhre ein erhöhter Luftdruck aufgebaut werden. Der Druck wird so eingestellt, dass sich die Glottis öffnet und die Luft ausströmt. Beim Ausströmen sinkt der Druck, die Glottis schließt sich und unterbricht den Luftstrom, der Öffnungsvorgang beginnt von neuem. Wichtig ist, im Bewusstsein zu halten, dass mit diesem Vorgang nur der Rohschall erzeugt wird. Man hört seine Grundfrequenz als Stimmtone, während die Klangfarbe des gehörten Tones durch die Filterwirkung der supraglottalen Lufträume mitbestimmt ist.

Zu den weiteren Modi der Lauterzeugung im Kehlkopf gehört das Flüstern (Glottis im vorderen Bereich geschlossen, im hinteren geöffnet, 4c; Steinberg 2008), das Falsett (Stimmbänder durch Zurückklappen der Stellknorpel sehr stark gedehnt) und die Geräuschproduktion beim glottalen Frikativ [h] (Glottis leicht geöffnet) sowie dem glottalen Plosiv [ʔ] (plötzliches Öffnen der Glottis, vgl. 2.2.1).

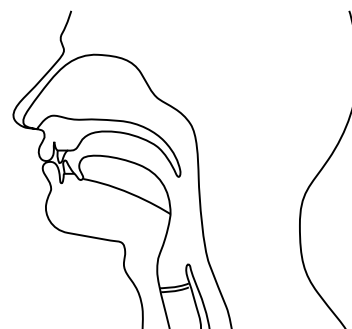
Die Glottis wird nach oben geschützt durch die *Epiglottis* (Kehldeckel), die sich beim Schlucken automatisch schließt. In manchen Sprachen, z. B. im Arabischen, gehört sie zu den aktiven Artikulatoren. Die darüber liegende Rachenhöhle (der *Pharynx*) dient als Resonanzraum. Größe und Form der Rachenhöhle sind durch Bewegung der Zunge und des Kehlkopfes variabel, ein aktiv artikulierendes Organ befindet sich in diesem Bereich aber nicht.

Der Pharynx wird nach oben durch den hinteren Teil des Zungenrückens (das *Dorsum*) und das Gaumensegel (auch weicher Gaumen, das *Velum*) begrenzt. Das Velum kann, wie schon erwähnt, gehoben und gesenkt werden (orale vs. nasale Laute). Bei der Produktion oraler Laute wird die Nasenhöhle ganz geschlossen, der gesamte Luftstrom geht durch die Mundhöhle. Bei den nasalten Lauten wird die Nasenhöhle geöffnet, die Mundhöhle wird aber nicht unbedingt am Velum verschlossen, sondern teilweise weiter vorn, beim [m] ganz vorn an den Lippen (5a: oraler Laut, hier [s]; 5b: nasaler Laut, hier [m]).

(5) a.



b.



Bei gehobenem Velum kann zwischen Zunge und Velum eine Enge variabler Größe gebildet werden. Ist die Öffnung so eng, dass beim Hindurchpressen des Luftstroms ein Zischgeräusch entsteht, dann ergibt sich ein Frikativ wie das [x] in **suchen**. Wird die Öffnung

zwischen Zunge und Velum ganz geschlossen und plötzlich geöffnet, so entsteht ein Plosiv wie das [k] in **Luke**.

Die Variabilität der Öffnung zwischen dem unteren, beweglichen Teil der Mundhöhle (Zunge, untere Zahnreihe, Unterlippe) und der oberen, weniger oder gar nicht beweglichen Begrenzung der Mundhöhle (harter Gaumen – das Palatum, Zahndamm – die Alveolen, obere Zahnreihe, Oberlippe) spielt für den ganzen prävelaren Bereich eine entscheidende Rolle. Eine große Öffnung führt zu Vokalen, eine kleine zu Frikativen und ein Verschluss mit plötzlicher Öffnung zu Plosiven. Wie dieser Bereich gegliedert werden soll und welche Bereiche für Engebildung und Verschluss man unterscheiden soll, ist nicht ohne weiteres klar (genauer 2.2.1).

Zusammenwirken der Artikulationsorgane

Unsere bisherigen Aussagen zur Artikulation betreffen einzelne Teile des Artikulationsapparates. An Beispielen von Einzellauten wurde gezeigt, welche Art von Ton oder Geräusch produziert werden kann. Eine Beschreibung dieser Art erfasst wesentliche Merkmale der sprachlichen Artikulation nicht. Auch wenn man von ihrer Ungenauigkeit absieht, ist sie aus mindestens zwei Gründen unrealistisch.

Der erste Grund: Sprachlaute werden nicht isoliert, sondern als Bestandteile größerer Einheiten artikuliert. Damit steht erst einmal die Frage, wo ein Lautsegment aufhört und wo das folgende anfängt. Dieser Frage gehen wir an anderer Stelle nach (2.3.1; 4.2).

Sieht man nun eine Silbe oder eine Wortform als Folge von Einzellauten, also Lautsegmenten an, dann stellt man fest, dass die Artikulation der Einzellaute nicht unabhängig voneinander erfolgt. Die Artikulationen benachbarter oder nahe beieinander liegender Laute sind aufeinander bezogen, wobei dieser Bezug von recht unterschiedlicher Art sein kann. Als ein Beispiel wurde früher schon erwähnt, dass Laute mit Stimmton innerhalb einer Silbe unmittelbar aufeinander folgen. Sie haben alle ein artikulatorisches Merkmal gemeinsam, sind artikulatorisch aneinander angeglichen.

Eine solche Angleichung kann auch dadurch hervorgerufen sein, dass ein Laut bestimmte Eigenschaften eines Nachbarlautes übernimmt und sich dadurch verändert. Man spricht dann von *Assimilation*. Das bekannteste Beispiel aus dem Deutschen ist die Nasalassimilation. Ein Wort wie **Bank** wird mit **n** geschrieben, hat aber die Lautfolge [baŋk]. Es gibt nun gute Gründe für die Annahme, dass die Schreibweise des Wortes kein Zufall ist, sondern dass das [ŋ] hier ›eigentlich‹ ein [n] ist. Das [n] wird alveolar, also vorn artikuliert. Das [ŋ], der velare Nasal, wird wie das ihm folgende [k] hinten (velar) artikuliert. Das [ŋ] wäre nach dieser Auffassung ein dem [k] assimiliertes [n] (weiter 4.3.2).

Die meisten Assimilationsvorgänge verändern einen Laut nicht so gravierend wie [n] zu [ŋ]. Das [t] in **Tusch** steht vor einem Vokal, der mit Lippenrundung artikuliert wird. Die Rundung des [ʊ] wird bei der Artikulation des [t] vorweggenommen, das [t] ist gerundet und wird nach dem IPA als [t̚] geschrieben. Ganz anders in **Tisch**. Hier folgt dem [t̚] ein [i], also ein definitiv ungerundeter Vokal. Rundung ist für das [t̚] eine marginale Eigenschaft. Beide Laute sind zweifelsfrei ein [t̚], aber sie sind doch so verschieden, dass man sie nicht austauschen darf. Man hört sonst nämlich etwas wie [tʁʊf] und [tʁɪf].

Assimilation hat ihre artikulatorische Basis in der *Koartikulation*. In **Tusch** sind [t̚] und [ʊ] bezüglich Lippenrundung koartikuliert. Lippenrundung ist an sich ein Merkmal des [ʊ]. Sein Auftauchen beim [t̚] besagt, dass beide Laute in gewisser Hinsicht gleichzeitig, eben koartikuliert sind.