Birgit Alber
birgit.alber@univr.it

Einführung in die Phonologie des Deutschen[1]

(3. Fassung, Sommersemester 2006/2007)

1. Einleitung

Die Phonologie beschäftigt sich mit der lautlichen Struktur der Sprache. Bevor wir jedoch damit beginnen, diese Struktur kennenzulernen, wollen wir uns zuerst ein paar allgemeinen Fragen zuwenden.

Wenn kleine Kinder sprechen lernen, dann gibt es etwa ab dem 2. Monat ein Stadium, in dem wir Erwachsenen zwar hören, dass das Kind anfängt, vor sich hinzuplappern, es fällt uns aber schwer, in dem Geplappere einzelne Sprachlaute auszumachen. Es gurrt und gurgelt, und schmatzt und schnalzt, und ab und zu glaubt man auch, einzelne Laute erkennen zu können, aber in dieser ersten Lallphase produziert das Kind eher noch Geräusche als Laute. Ein paar Monate nach dieser Phase, etwa ab dem 6. Monat, beginnt das Kind nun Lautsequenzen von sich zu geben, die wie *amma, mamma, papa* ... klingen. Das Kind will damit noch nicht seine Mutter oder seinen Vater rufen, [a] [m] und [p] gehören einfach zu den Lauten, die ein Kind als erstes aussprechen lernt. Es ist daher kein Zufall, dass die Namen, mit denen Kinder ihre Eltern bezeichnen, in vielen Sprachen der Welt genau diese Laute enthalten. Hier ist eine Liste von sehr verschiedenen Sprachen und das Wort, mit dem Kinder dieser Sprache ihre Mutter rufen:

(1)

Deutsch:	Mama, Mami
Italienisch:	mamma
Englisch:	mum, mummy
Polnisch:	mama
Estnisch:	mamma
Swahili:	mama
Maori:	mämä (ausgesprochen: maama)

Das Kind hat in dieser Phase damit begonnen, Laute voneinander zu unterscheiden. Es wird nun im Laufe seiner Entwicklung immer neue Laute dazulernen, bis es schließlich alle Laute seiner Sprache kennt und verwenden kann.

Warum ist aber die menschliche Sprache so beschaffen, dass wir Ketten von einzelnen Lauten bilden? Warum bleiben wir nicht bei der ersten Lallphase und verwenden als Bezeichnung für BAUM ein kräftiges Schmatzen? Im Grunde ist es doch viel anstrengender, sich Laut für Laut anzueignen. Immerhin brauchen kleine Kinder mehrere Monate, um ihre Zunge so weit "in den Griff" zu bekommen, dass sie Silbenketten bilden können.

Der Grund für die Verwendung von Lautsequenzen, im Gegensatz zu Geräuschen, liegt darin, dass auch die lautliche Struktur der Sprache, wie Sprache überhaupt, auf **diskreten Einheiten** beruht. So wie wir Wörter aus Morphemen zusammensetzen

---

[1] Wer dieses Material im Unterricht verwenden will, muss die Quelle angeben (z.B. als: Birgit Alber (2007*), Einführung in die Phonologie des Deutschen,* ms. Università di Verona.)

und Sätze aus syntaktischen Phrasen bauen, so bestehen auch lautliche Strukturen aus klar voneinander unterscheidbaren Einheiten.

Wie in der Morphologie und Syntax so beschäftigen sich auch Phonetik und Phonologie damit, wie diese Einheiten – Phone, Phoneme, phonologische Merkmale, Silben u.s.w. – beschaffen sind und nach welchen Regeln sie sich miteinander verbinden. Im 2. Kapitel dieser Einführung werden wir uns mit den Einheiten der Phonologie beschäftigen. Wir werden sehen, wie Laute klassifiziert werden und wie man Phone und Phoneme definieren kann.

Übung 1

Aus wievielen Lauten bestehen die folgenden Wörter? In welche Fällen entsprechen mehrere Buchstaben einem Laut? In welchen Fällen entsprechen mehrere Laute einem Buchstaben?

Schale
Zopf
Buch
Hexe
Mitte
Schuhe

Aber warum sind diskrete Einheiten so wichtig für die menschliche Sprache? Wenn menschliche Sprache nicht über diskrete Einheiten verfügen würde, dann müssten wir uns für jeden Begriff ein neues Geräusch merken. Wenn wir an die Einträge denken, die sogar ein kleines Wörterbuch hat, dann würde da eine ganz schöne Anzahl von Geräuschen zusammenkommen. Wenn wir unsere Wörter hingegen aus Lautsequenzen bilden, dann entlasten wir unser Gedächtnis um einiges. Eine Sprache wie Deutsch verfügt über etwa 40 Phoneme. Mit diesen Lauten können wir eine ziemlich große , im Prinzip unendliche Anzahl von Wörtern bilden, wir benötigen dazu aber nur eine endliche Anzahl von Lauteinheiten, nämlich 40 Phoneme.

Die Artikulation von Sprachlauten braucht große Präzision von Seiten der Sprechorgane und unser Ohr muss außerdem lernen, die einzelnen Laute voneinander zu unterscheiden. Die Hauptarbeit leistet aber eigentlich unser Gehirn. Denn obwohl wir oft undeutlich sprechen, obwohl oft ein Laut mit dem anderen verschmilzt, obwohl im Hintergrund oft viele Geräusche stören – dennoch gelingt es uns immer wieder, zu verstehen, was unsere Kommunikationspartner gesagt haben, welche Laute sie aneinandergereiht haben. Unsere Gehirn scheint irgendwie dafür bestimmt zu sein, ein Kontinuum von Lauten in Einzellaute zu unterteilen. Wir scheinen eine Prädisposition für die Analyse des Lautstroms in seine Einzelteile zu haben. Diese Tatsache zeigt uns auch, dass wir es bei der Zusammensetzung aus diskreten Einheiten mit einer sehr tiefen, universellen Charakteristik der menschlichen Sprache zu tun haben.

Wir haben bis jetzt nur von Lauten als Einheiten der phonologischen Struktur gesprochen. Es gibt aber auch andere Einheiten der lautlichen Struktur der Sprache, bei denen es sich nicht unbedingt um Einzellaute handelt. Die Rede ist hier von "größeren" Einheiten wie der Silbe, die ihrerseits aus Einzellauten besteht. Diese Einheiten werden im vierten Kapitel dieser Einführung ("Prosodische Phonologie") behandelt.

Wie in Syntax und Morphologie, so unterliegt auch die Aneinanderreihung von lautlichen Einheiten bestimmten **Regeln.** Nicht jede Sequenz von Lauten ist wohlgeformt. Eine Sequenz wie *sparf* könnte es im Deutschen geben, eine Sequenz wie *fpars* hingegen nicht. Dazu kommt noch, dass bestimmte Laute in gewissen Positionen **phonologischen Prozessen** unterliegen. Betrachten wir die folgenden Beispiele: [2]

(2) Das Wort <*Räte*> wird ausgesprochen als *Rä[t]e*
Das Wort <*Räder*> wird ausgesprochen als *Rä[d]er*

Das Wort <*Rat*> wird ausgesprochen als *Ra[t^h]*
Das Wort <*Rad*> wird ausgesprochen als *Ra[t^h]* !!

Wir sehen, obwohl der letzte Buchstabe des Wortes *Rad* ein <d> ist, wird der letzte Laut des Wortes als ein stimmloses, aspiriertes [t^h] ausgesprochen. Und das, obwohl wir im Plural von *Rad*, in *Räder*, ein stimmhaftes [d] haben. Die Aussprache von *Rad* ist also genau dieselbe wie die von *Rat*. Die Aussprache der beiden Wörter unterscheidet sich nur im Plural, nicht aber im Singular. Der Grund dafür ist ein phonologischer Prozess des Deutschen, den man **Auslautverhärtung** nennt. Wir werden uns diesen Prozess noch im Detail anschauen, im Augenblick soll nur gesagt werden, dass er dazu führen kann, dass stimmhafte Konsonanten am Ende des Wortes stimmlos werden.

Übung 2

Wie werden wohl die folgenden Wörter des Deutschen ausgesprochen, wenn man den Prozess der Auslautverhärtung berücksichtigt?

Tod
Lob
aktiv
Kind
lieb
sag!
Job

Phonologische Prozesse sind den Sprechern nicht bewusst. Sie werden von ihnen automatisch angewandt , und es ist für die Sprecher fast unmöglich, sie zu unterdrücken, wenn sie eine Fremdsprache sprechen. So ist es für deutsche Sprecher sehr schwierig, die Regel der Auslautverhärtung *nicht* anzuwenden, wenn sie z.B. Englisch sprechen. Da wird dann aus einem *jo[b]* leicht ein *Jo[p^h]*.

Interessanterweise haben die Laute, die von der Auslautverhärtung betroffen sind ([b, d, g, v ...], alle etwas gemeinsam. Es handelt sich hier immer um Laute, bei denen der Luftstrom, der bei ihrer Produktion entsteht, behindert wird. Entweder wird er ganz

[2] Grapheme (Buchstaben) schreibt man zwischen spitzen Klammern (<...>), Phone werden zwischen eckige Klammern gesetzt ([...]). Phoneme (zu dem Unterschied zwischen Phonen und Phonemen kommen wir noch!) schreibt man zwischen Schrägstrichen (/.../).

blockiert, wie bei [b, d, g] oder er wird so stark blockiert, dass ein Reibegeräusch entsteht (wie bei [v]). Diese Klasse von Lauten nennt man **Obstruenten.**

In jeder Sprache gibt es phonologische Prozesse und wir werden uns im 3. Kapitel mit den phonologischen Prozessen des Deutschen befassen.

Eine Frage, mit der sich Phonologen intensiv beschäftigen ist die nach den Charakteristiken von phonologischen Prozessen und Regeln. Wir werden uns im 4. Kapitel vor allem mit zwei Aspekten beschäftigen.

Erstens wollen wir uns in diesem Einführungstext immer wieder fragen, was an den phonologischen Prozessen und Regeln, die wir untersuchen, **sprachspezifisch** ist und was hingegen **universell** ist.

So wissen wir z.B., dass die Auslautverhärtung ein typischer Prozess des Deutschen ist, der in anderen Sprachen der Welt nicht auftaucht. Andererseits hat dieser Prozess auch universelle Charakteristiken. Er betrifft eine Klasse von Lauten (die Obstruenten) und es scheint in allen Sprachen der Welt so zu sein, dass Prozesse immer Klassen von Lauten betreffen und nicht eine Liste von zufällig zusammengewürfelten Einzellauten.

Ein zweiter Aspekt von universellen Regeln und Prinzipien, den wir vor allem im 4. Kapitel diskutieren werden, betrifft ihre **Verletzbarkeit**. Gelten universelle phonologische Prinzipien immer, in allen Sprachen der Welt? Wenn ein Prinzip universell ist, sollte es seiner Definition nach eigentlich immer und überall gelten. Seit etwa zehn Jahren gibt es in der Phonologie jedoch eine Theorie, die sogenannte **Optimalitätstheorie**, die einen ziemlich radikalen Standpunkt vertritt. Sie behauptet, dass alle phonologischen Beschränkungen universelle Gültigkeit haben. Wie sollen wir dann aber die Unterschiede zwischen den einzelnen Sprachen erklären? Die Hypothese der Optimalitätstheorie ist, dass alle phonologischen Beschränkungen universell sind, dass diese Beschränkungen jedoch in bestimmten Sprachen, in bestimmten Kontexten verletzt werden können. Ausserdem verwenden Sprachen verschiedene Strategien, um diese Beschränkungen zu erfüllen.

Schauen wir uns ein Beispiel an, damit diese Hypothese etwas klarer wird. Es gibt eine universelle Beschränkung, die besagt, dass eine Silbe mit einem Konsonanten beginnen sollte. Die Details dieser Beschränkung werden wir noch im 4. Kapitel genauer untersuchen, hier formulieren wir diese Beschränkung einfach folgendermaßen:

(3) Silbenanlautbeschränkung:
Silben müssen mit einem Konsonanten beginnen

Wir sehen, dass diese Beschränkung wahr sein muss, wenn wir uns überlegen, dass das Wort *Amerika* wohl immer als *A.me.ri.ka* silbifiziert werden wird und nie als *Am.er.ik.a*.

Das Deutsche verfolgt eine besondere Strategie, um diese Beschränkung zu befolgen. Wenn Wörter mit einem Vokal beginnen, dann wird vor diesen Vokal ein Konsonant eingefügt. Bei diesem Konsonanten handelt es sich um einen **glottalen Plosiv**. Das ist ein Konsonant, der gebildet wird, indem man die Öffnung zwischen den Stimmbändern kurz verschließt, und dann die Luft auf einmal herausströmen läßt. Versucht, leicht zu husten, dann produziert ihr genau diesen Laut. Das phonetische Symbol für den glottalen Plosiv ist [ʔ]. Man hört diesen Laut sehr deutlich in einem Satz wie dem folgenden, bei dem jedes Wort mit einem Vokal beginnt:

(4) [ʔ]Anton [ʔ]aß [ʔ]am [ʔ]Abend [ʔ]immer [ʔ]Auflauf

Wird die Silbenanlautbeschränkung im Deutschen immer befolgt? Nein, es gibt Fälle, in denen sie verletzt wird. Das geschieht zum Beispiel, wenn sich die vokalinitiale Silbe im Wortinneren befindet und nicht betont ist. Ein Beispiel dafür ist das Verb *sehen*. Die phonetische Transkription zeigt uns die Aussprache dieses Wortes:[3]

(5) [ze:.ən] <sehen>

Bei dem ersten Laut handelt es sich um ein stimmhaftes [z]. Dann folgt ein langes [e:] und anschließend ein weiterer, e-ähnlicher Laut, ein sogenanntes **schwa**. Beachtet, dass das <h> nicht ausgesprochen wird. Es dient nur dazu, uns zu zeigen, dass der vorangehende Vokal lang ist.

Die zweite Silbe dieses Wortes beginnt auch mit einem Vokal. Trotzdem wird hier kein glottaler Plosiv eingefügt. Die Silbenanlautbeschränkung kann also im Deutschen nur in bestimmten Fällen befolgt werden.

Im Italienischen ist die Sache wieder etwas anders. Auf den ersten Blick sieht es so aus, als würde die Silbenanlautbeschränkung im Italienischen keine besonders große Rolle spielen. Schließlich gibt es viele Wörter, die mit einem Vokal beginnen, wie *amore, agguato, o, e, inverno* u.s.w. Vor diese Wörter werden keine Konsonanten gesetzt, die ersten Silben beginnen also mit einem Vokal. Trotzdem sieht man den Einfluss der Silbenanlautbeschränkung auch im Italienischen. Wenn diese Wörter nämlich nach Wörtern mit einem finalen Konsonanten erscheinen, dann kann man etwas Auffallendes beobachten:

(6) co.n a.mo.re
del.l' in.ver.no
i.n ag.gua.to

Lest diese Beispiele langsam, wie bei einem Abzählreim (*in una conta*). Ihr werdet beobachten, dass die letzten Konsonanten des ersten Wortes mit dem ersten Vokal des zweiten Wortes silbifiziert werden. Das Italienische verwendet also eine andere Strategie als das Deutsche, um Silben mit einem Anfangskonsonanten zu versehen: es "zieht" den Konsonanten eines vorangehenden Wortes in die vokalinitiale Silbe hinüber. Durch diese Strategie kann es sogar zu Fällen von Ambiguität kommen, wie z.B. in den folgenden Sätzen:

(7) Ambiguität, die sich durch Resilbifizierung ergibt:

a. La spiaggia greca aveva u.n' a.re.na strepitosa
a'. La spiaggia greca aveva u.n a.re.na strepitosa

b. Il veterinario provò con u.n' a.ve.na diversa
b'. Il veterinario provò con u.na ve.na diversa

Wenn wir diese Sätze nur hören, aber nicht lesen, können wir nicht unterscheiden, ob es sich um die Bedeutung in a. oder a'., b. oder b'. handelt. Die Sätze klingen gleich, da das Italienische die Grenze zwischen *un* und dem darauffolgenden, vokalinitialen Word nicht markiert, sondern Resilbifizierung des letzten Konsonanten von *un* erlaubt.

---

[3] Die Silbengrenze bezeichne ich mit einem Punkt "."

Wir haben in diesen Beispielen drei Dinge gesehen. Erstens scheint es in der Tat eine universelle Silbenanlautbeschränkung zu geben, deren Einfluss sich sowohl im Deutschen als auch im Italienischen bemerkbar macht. Diese Beschränkung ist jedoch verletzbar: eine "schöne" Silbe sollte mit einem Konsonanten beginnen, aber im Deutschen ist das im Inneren des Wortes nicht immer möglich und im Italienischen ist das auch am Anfang des Wortes nicht immer möglich. Beide Sprachen versuchen jedoch – soweit möglich – die Silbenanlautbeschränkung zu befolgen. Dazu verfolgen Deutsch und Italienisch unterschiedliche Strategien: im Deutschen wird ein Konsonant eingesetzt, im Italienischen wird über Wortgrenzen hinweg silbifiziert.

Diese drei Charakteristiken von phonologischen Beschränkungen (Universalität, Verletzbarkeit, sprachspezifische Strategien, um sie zu erfüllen) werden wir im Kapitel über die Silbenstruktur noch genauer diskutieren.

Fassen wir nun kurz zusammen, was bisher gesagt wurde. Wir haben in dieser Einleitung einige Charakteristiken des Lautsystems der menschlichen Sprache gesehen, mit denen wir uns hier befassen werden:

- das Lautsystem der menschlichen Sprache besteht aus diskreten Einheiten. In Kapitel 2 werden wir sehen, wie diese Laut-Einheiten aussehen, welche Arten von Lauten es gibt und welche Laute im Deutschen verwendet werden. Wir werden in diesem Kapitel auch die Unterschiede zwischen den Einheiten **Phon** und **Phonem** diskutieren. In Kapitel 3.7 werden die kleinsten Einheiten der Phonologie, die phonologischen Merkmale besprochen. Kap. 4 diskutiert sehr ausführlich die Silbe, eine phonologische Einheit, die größer als ein Laut ist.

- das Lautsystem der menschlichen Sprache unterliegt regelhaften Prozessen. In Kapitel 3 werden wir uns vor allem mit den phonologischen Prozessen beschäftigen, die wir im Deutschen finden.

- das Lautsystem jeder menschlichen Sprache hat sowohl universelle als auch sprachspezifische Charakteristiken. Mit diesem Aspekt, und der Behandlung dieses Aspektes im Rahmen der Optimalitätstheorie werden wir uns in Kapitel 4 beschäftigen.

2. Diskrete Einheiten in der Phonologie

2.1 Phone – Exkurs zur Phonetik

Bevor wir die Einheiten vorstellen, mit denen sich die Phonologie befasst, müssen wir uns zuerst ein wenig mit den physikalischen Eigenschaften von Lauten im Allgemeinen beschäftigen.

Mit den physikalischen Charakteristiken der Sprachlaute beschäftigt sich das linguistische Teilgebiet der **Phonetik.** Die Phonetik untersucht, wie Laute von den Sprechorganen produziert werden (**artikulatorische Phonetik**), wie die akustischen Eigenschaften von Sprachlauten aussehen (**akustische Phonetik)** und wie Sprachlaute gehört werden (**auditive Phonetik**).

Im Gegensatz zur Phonetik beschäftigt sich die **Phonologie** mit dem **System**, in das die Sprachlaute eingebunden sind. Sie fragt z.B. danach, wie die

bedeutungsunterscheidenden Einheiten (die **Phoneme**, s. Kap. 2.2) einer Sprache organisiert sind, wieviele Phoneme es in einer Sprache gibt oder auch welche Phoneminventare es im Allgemeinen in den Sprachen der Welt gibt und welche nicht. Und sie beschäftigt sich damit, auf welche Art und Weise die Phoneme durch phonologische Prozesse manipuliert werden.

2.1.1 Akustische Phonetik

Wir werden uns im nächsten Abschnitt vor allem mit artikulatorischer Phonetik beschäftigen, da die Laute in der Phonologie zum größten Teil nach artikulatorischen Aspekten klassifiziert werden. Es wird also vor allem darum gehen, wie unsere Sprechorgane die einzelnen Laute produzieren.

Zur Illustration, wie die akustische Phonetik arbeitet, wollen wir uns jedoch ein Spektrogramm ansehen und versuchen zu verstehen, wie ein Laut physikalisch gesehen aussieht. Spektrogramme sind Darstellungen des physikalischen Sprachlautes.

Sprachlaute sind, wie andere Geräusche auch, Schallwellen, die gemessen werden können. Die Frequenz einer Welle gibt die Tonhöhe an, die Amplitude die Lautstärke. Sprachlaute - vor allem Vokale - sind nicht einfache, sondern komplexe periodische Wellen, d.h., sie bestehen aus mehreren überlappenden periodischen Wellen. Viele Laute haben deshalb mehrere Frequenzen. Vokale sind z.B. durch ein ganz bestimmtes Frequenzspektrum charakterisiert.

Wie sieht nun so ein Spektrogramm aus? Die x-Achse stellt die Zeit dar, die y-Achse die Frequenz der akustischen Welle. Die Intensität der Schwärzung gibt die Lautstärke des Lautes an.

Lo spetrogramma che dovrebbe apparire su questa pagina lo trovate sulla homepage di Peter Ladefoged

http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/course/chapter8/8.7.htm

oppure, in forma cartacea, alla fotocopisteria Rapida

Das Spektrogramm von Peter Ladefoged ist die Aufzeichnung der englischen Wörter *a bab, a dad* und *a gag*. Es zeigt uns einige interessante Dinge des physikalischen Sprachlautes:

- Vokale sind durch mehrere schwarze "Ringe" gekennzeichnet. Diese Ringe sind die charakteristischen Frequenzen des betreffenden Vokals. Die Frequenzen eines spezifischen Vokals werden auch als seine Formantenstruktur bezeichnet.
- Konsonanten wie [b], [d] und [g] sind gekennzeichnet durch - Stille! Wenn ihr genau hinschaut, so erscheint über den Konsonanten fast gar keine Schwärzung.
- worin unterscheiden sich dann die einzelnen Konsonanten? Wenn wir z.B. das erste [b] von *a bab* mit dem ersten [g] von *a gag* vergleichen, dann sehen wir, dass der Unterschied vor allem in der Formantenstruktur des vorhergehenden Vokals besteht: die zweiten und dritten Formanten des Vokals vor dem [g] bilden eine Art von Kurve, die sie vor [b] nicht bilden.

[b], [d] und [g] sind **Plosive** (sie werden auch **Verschlusslaute** genannt). Sie werden gebildet, indem die Mundhöhle an einer bestimmten Stelle geschlossen wird. Im Mund wird so ein (Luft)druck aufgebaut, der dann plötzlich gelöst wird, sobald die Blockierung gelöst wird. Beim Plosiv [b] zum Beispiel wird der Luftstrom bei den Lippen blockiert. Bei diesem Prozess werden keine Schallwellen produziert, deshalb hinterlassen Plosive auf dem Spektrogramm oft keine Spur. [b], [d] und [g] sind stimmhafte Laute, bei ihrer Produktion vibrieren die Stimmbänder. Diese Vibration kann man auch ganz leicht auf den Spektrogrammen sehen. Bei der Produktion der stimmlosen Plosive [p], [t] und [k] hingegen herrscht vollkommene Stille. Andere Konsonanten, wie z.B. [f] und [s] produzieren Geräusche, die man auf dem Spektrogramm sehen kann.

In so kurzen Sequenzen wie auf dem vorliegenden Spektrogramm lassen sich einzelne Segmente noch ziemlich gut voneinander unterscheiden. Wenn wir jedoch ganze Sätze aufzeichnen, dann ist es oft schwierig zu verstehen, wo ein Laut aufhört und wo der andere anfängt. Der Grund dafür ist, dass die Lautsequenz physikalisch gesehen ein Kontinuum bildet, ein Laut ist nicht immer klar von dem anderen abgegrenzt. Dennoch haben wir die Intuition, dass ein Wort aus einzelnen, klar voneinander unterscheidbaren Lauten besteht.

2.1.2 Artikulatorische Phonetik – Die Sprechwerkzeuge

In der Phonologie werden Laute traditionell nach artikulatorischen Gesichtspunkten klassifiziert. Wie werden nun aber Sprachlaute produziert?

Schauen wir uns zuerst die Sprechwerkzeuge an, jene Organe, mit denen wir die Laute unserer Muttersprache produzieren:

Sprechwerkzeuge:
- die **Lunge** (it.: *polmoni*): bei der Erzeugung von Lauten funktioniert die Lunge wie eine Pumpe, die den Luftstrom liefert. In den meisten Sprachen der Welt werden Laute mit einem Luftstrom gebildet, der von der Lunge über die Mundhöhle nach außen dringt (egressiver, pulmonarer Luftstrom). Es gibt aber auch in einigen Sprachen Laute, die anders gebildet werden. Ein Beispiel dafür sind die Klick-Laute in bestimmten afrikanischen Sprachen (z.B. Xhosa und Zulu). Einige dieser Laute

ähneln dem Zungenschnalzen eines Kutschers, der seine Pferde antreibt, oder dem Schmatzen, wenn man sich auf Entfernung ein Küsschen zuwirft.[4]

- **Lippen** (it.: *labbra*): die Lippen werden dazu verwendet, labiale Laute wie z.B. [p, b, f, m] zu produzieren

- **Zähne** (it.: *denti*): in der Region der Zähne werden Laute wie das Englische interdentale [θ] (wie im Wort *thin*) ausgesprochen. Auch Laute wie [d, t, s, z, n, l] können eine dentale Realisierung haben.

- **Alveólen** (it.: *alvéoli*): die Alveolen, auch *Zahndamm* genannt, befinden sich in der Mundhöhle gleich über den Zähnen. In manchen Sprachen werden die Laute [d, t, s, z, n, l] an den Alveolen produziert, man nennt sie also *alveolare Laute*. Für die Sprachen Deutsch, Italienisch und Englisch ist die Unterscheidung zwischen *dental* und *alveolar* nicht relevant. Ihr könnt also die Laute [d, t, s, z, n, l] sowohl *alveolar* als auch *dental* nennen.

- **Palatum** (harter Gaumen) (it.: *palato duro*): am harten Gaumen werden im Deutschen palatale Laute wie das [j] oder das [ç] (der sogenannte *ich-Laut*) produziert.

- **Velum** (weicher Gaumen) (it.: *palato molle*): in der Region des Velum werden die velaren Laute [k, g, ŋ] und das [x] (der sogenannte *ach-Laut*) produziert.

- **Uvula** (Zäpfchen) (it.: *ugola*): an der Uvula werden zwei r-Laute des Deutschen, der Vibrant [R] und der Frikativ [ʁ] produziert

- die **Zunge** (it.: *lingua*): die Zunge ist ein wichtiges Sprechorgan. Sie wird normalerweise in drei Zonen eingeteilt
    - Apex (Zungenspitze) (*apice*): dieser Teil der Zunge vibriert, wenn das apikale [r] des Italienischen produziert wird
    - Dorsum (Zungenrücken) (*dorso*): der Zungenrücken wird zur Produktion von palatalen, velaren und uvularen Konsonanten verwendet, die man dann alle zusammen *dorsale Konsonanten* nennt.
    - Zungenwurzel (*radice*): sie wird z.B. zur Produktion von pharyngalen Konsonanten verwendet, wie wir sie in den arabischen Sprachen finden.

- **Glottis** (*glottide*): ist eigentlich kein Organ, sondern der Raum zwischen den Stimmbändern
    - sie ist offen → bei stimmlosen Lauten, also wenn wir z.B. ein [p] produzieren
    - sie ist geschlossen → beim glottalen Plosiv [ʔ]
    - Stimmbänder sind eng beieinander und vibrieren → bei stimmhaften Lauten

Zustand der Glottis: der Zustand der Glottis, d.h., ob sie offen oder geschlossen ist, ob die Stimmbänder vibrieren oder nicht, ob die Stimmbänder gleich am Anfang der Lautproduktion oder erst später vibrieren - all diese Dinge charakterisieren die Qualität der Sprachlaute. Hier sollen nur einige besondere Charakteristiken der deutschen Sprachlaute in Bezug auf den Zustand der Glottis genannt werden.

---

[4] Auf der homepage des Phonetikers Peter Ladefoged kann man sich diese *clicks* anhören und sogar ein Röntgenbild von einem Sprecher sehen, der einen Klick produziert: http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/course/chapter6/6aiarstream.html

Stimmhaftigkeit: wir produzieren stimmhafte Laute [b, d, g, v, z, m, l] oder natürlich auch alle Vokale, indem wir die Stimmbänder zusammenziehen, die Glottis also praktisch schließen, und die Stimmbänder zum Vibrieren bringen. Wenn ihr nicht sicher seid, ob ein Laut stimmhaft oder stimmlos ist, dann legt einen Finger auf den Hals, auf der Höhe der Glottis. Bei der Produktion eines stimmhaften Lautes werdet ihr die Vibration spüren. Ihr könnt auch eure flache Hand auf den Kopf legen, auch dort spürt ihr die Vibrationen, die ein stimmhafter Laut verursacht.

Teilweise Stimmhaftigkeit: oft sind Konsonanten im Deutschen nicht durchgehend stimmhaft bzw. stimmlos. Wenn z.B. ein [b] am Wortanfang steht, wie im Wort *Baum,* so wird dieses [b] in seiner ersten Aussprachephase als stimmlos realisiert. Die Stimmbänder beginnen erst nach dieser ersten stimmlosen Phase zu vibrieren. Diese teilweise Stimmlosigkeit von Lauten wie [b, d, g, v] erweckt dann bei Nicht-Muttersprachlern oft den Eindruck, man habe [p, t, k, f] gesagt. Das kann besonders einem italienischen Muttersprachler passieren, denn im Italienischen werden die Laute [b, d, g, v] meist voll stimmhaft realisiert. Euch ist sicher schon aufgefallen, dass ein deutscher Muttersprachler, wenn er Italienisch spricht, oft die stimmhaften Laute zu "hart" ausspricht. Das Wort *bambino* klingt dann oft wie *pampino*. Jetzt wisst ihr, woran das liegt, ein deutscher Muttersprachler spricht den Laut [b] eben oft nicht als voll stimmhaft aus.

In der phonetischen Transkription werden wir teilweise Stimmhaftigkeit durch einen Kringel unter bzw. über den stimmhaften Konsonanten ausdrücken, wie in den folgenden Beispielen:

(8)	<böse>	[b̥]öse
	<doch>	[d̥]och
	<gut>	[g̊]ut
	<Wiese>	[y̥]iese

Aspiration: ein aspirierter Laut entsteht, wenn nach der Explosion des Konsonanten die Stimmbänder noch für eine kurze Zeit offen bleiben. Luft fließt dann für kurze Zeit ohne Behinderung durch die Glottis. Erst danach setzt wieder Stimmhaftigkeit (des darauffolgenden Vokals) ein. Im Standarddeutschen werden die stimmlosen Konsonanten [p, t, k] vor allem am Wortanfang, am Wortende oder vor betonten Vokalen stark aspiriert (s. Kap. 3.3). Das gilt allerdings nicht für die süddeutschen Varietäten, wo Aspiration weitgehend unbekannt ist. Aspiration eines Lautes wird durch ein hochgestelltes [^{h}] ausgedrückt, wie in den folgenden Beispielen:

(9)	<Park>	[p^{h}]ark
	<Ton>	[t^{h}]on
	<kam>	[k^{h}]am

Wenn man vergleicht, was wir hier über teilweise Stimmhaftigkeit und Aspiration gesagt haben, dann kann man festhalten, dass sich die Konsonanten [b, d, g] und [p, t, k] am Wortanfang also nicht so sehr in ihrer Stimmhaftigkeit unterscheiden, sondern dadurch, dass erstere teilweise stimmhaft und letztere aspiriert sind.
So unterscheiden sich die beiden Wörter *Bar* und *Paar* dadurch, dass der Anfangskonsonant des ersten Wortes teilweise stimmhaft, der des zweiten Wortes stimmlos und aspiriert ist:

(10) <Bar> vs. <Paar> = [b̥]ar vs. [p^{h}]aar

Übung 3

1. Sprecht alle Laute, die als Beispiele für die einzelnen Sprechwerkzeuge angegeben sind, laut nach. Versucht zu "spüren", wo diese Laute genau ausgesprochen werden
3. Sprecht nacheinander die Laute [p, t, k], dann [b, d, g], dann [m, n, ŋ] aus und versucht zu spüren, wie sich die Zunge bei der Produktion dieser Laute immer weiter nach hinten verschiebt.
3. Gibt es eurer Meinung nach im Deutschen *stimmhafte* aspirierte Konsonanten wie [b^h, d^h, g^h]? Versucht, diese Laute auszusprechen.

Falls ihr über die Antwort auf die dritte Frage dieser Übung ins Grübeln geraten sein solltet: im Deutschen gibt es keine stimmhaften aspirierten Konsonanten! Diesen Konsonantentyp gibt es aber in anderen Sprachen, z.B. in vielen Sprachen des indischen Subkontinents.

Wir benutzen jetzt die Diskussion zur Aspiration, um einen kleinen Exkurs in Richtung Phonologie zu unternehmen. Fragen wir uns zunächst, wie wichtig die Aspiration für die deutsche Sprache ist. Sie scheint nicht die gleiche Wichtigkeit zu besitzen wie z.B. der Unterschied zwischen den Lauten [t] und [k]. Wenn wir nämlich im Wort *Tante* den ersten Laut [t] durch ein [k] ersetzen, dann erhalten wir ein ganz neues Wort, das Wort *Kante*. Wenn wir jedoch im Wort *Park* das aspirierte [p^h] durch ein nicht aspiriertes [p] ersetzen, dann erhalten wir *kein* neues Wort! Unsere Aussprache klingt dann vielleicht etwas süddeutsch und nicht ganz standardmäßig, aber wer uns zuhört wird immer noch verstehen, dass wir von einem Ort in der Stadt mit vielen Bäumen sprechen, eben von einem *Park*. In der Phonologie sagt man, dass der Unterschied zwischen [t] und [k] **distinktiv** ist (man kann dadurch Wörter voneinander unterscheiden), während die Aspiration im Deutschen nicht distinktiv ist (sie kann keine Wortbedeutungen voneinander unterscheiden).

Es gibt allerdings Sprachen, in denen die Aspiration distinktiv ist. So gibt es z.B. im Hindi das folgende **Minimalpaar:**

(11) [p]al = 'sich um jemanden kümmern'
[p^h]al = 'Messerklinge'

Ein Minimalpaar ist ein Paar von Wörtern, das sich nur in einem distinktiven Merkmal unterscheidet, in diesem Fall in dem der Aspiration. Wenn wir *[p]al* ohne Aspiration aussprechen, so hat das Wort eine bestimmte Bedeutung, wenn wir es mit Aspiration aussprechen, so bedeutet es etwas ganz anderes. Die Aspiration kann somit im Hindi Wörter unterscheiden, sie ist distinktiv. Ein phonetisches Merkmal wie das der Aspiration kann also von Sprache zu Sprache unterschiedlich wichtig sein in dem Sinn, dass es in einer Sprache distinktiv ist und Bedeutungen unterscheiden kann, in einer anderen Sprache jedoch nicht.

2.1.3 Artikulatorische Phonetik - Die Klassifikation der Konsonanten

Konsonanten werden traditionell nach ihrem **Artikulationsort** (wo im Mund der Laut ausgeprochen wird) und nach ihrer **Artikulationsart** (wie er ausgesprochen wird) klassifiziert. Außerdem sagt man bei Plosiven und Frikativen noch dazu, ob sie stimmhaft oder stimmlos sind:

(12) [p] ist ein stimmloser bilabialer Plosiv
↓ ↓
Artikulationsort Artikulationsart

Es folgt hier eine Liste der Artikulationsorte, wie sie nach dem IPA, dem Internationalen Phonetischen Alphabet normalerweise angenommen wird. Ihr solltet diese Liste mit der Liste der Sprechwerkzeuge im letzten Kapitel vergleichen, die beiden Listen stimmen fast überein. Die Erklärung der Aussprache einzelner Laute findet ihr in der Diskussion des IPA, wo wir versuchen werden, die Konsonanten des Deutschen, Italienischen und des Englischen im IPA wiederzufinden.

(13) **Artikulationsorte** nach dem IPA:

bilabial: z.B. [p, b, m]
labiodental: z.B. [f, v]
dental: z.B. [t, d, s, r][5]
alveolar: z.B. [t, d, s, r]
postalveolar: z.B. [ʃ]
retroflex: z.B. [ʈ] in der sizilianischen Aussprache von *Trapani*, (s. unten)
palatal: z.B. [ç, j]
velar: z.B. [k, g, x, ŋ]
uvular: z.B. [R, ʁ]
pharyngal: z.B. [ʕ], in den arabischen Sprachen (s. unten)
glottal: z.B. [ʔ]

Was die **Artikulationsart** betrifft, so gibt es verschiedene Arten, einen Konsonanten zu produzieren: man kann den Lufstrom dabei an irgendeiner Stelle im Mund komplett blockieren, teilweise blockieren, kaum blockieren. Man kann die Luft durch den Mund entweichen lassen, oder durch die Nase.

Plosive (auch *Verschlusslaute* genannt, engl: *stops* oder *plosives*, it.: *occlusive*): zwei Sprechorgane blockieren den Luftstrom gänzlich. So wird er z.B. bei den labialen Lauten [p, b] von den Lippen blockiert, bei den dentalen Lauten [t, d] von der Zunge und den Zähnen, bei den velaren Lauten [k, g] von Zungenrücken und Velum. Beim glottalen Plosiv [ʔ] wird der Lufstrom durch das Schließen der Glottis blockiert.

Frikative (engl.: *fricatives,* it.: *fricative*): der Luftstrom wird teilweise blockiert, dabei entsteht ein reibendes Geräusch. z.B. [f, v], [s, z], [ʃ], [ç, x], [ʁ], [h].

Nasale (engl.: *nasals,* it.: *nasali*): der Luftstrom wird in der Mundhöhle blockiert, aber die Luft kann durch die Nase entweichen (bei gesenktem Velum). z.B. [m], [n], [ŋ],

---

[5] Wie schon gesagt werden diese Laute in manchen Sprachen eher dental, in anderen eher alveolar ausgesprochen.

Laterale *(*engl.: *laterals,* it.: *laterali)*: der Luftstrom wird in der Mundhöhle blockiert, aber die Luft kann seitlich entweichen. z.B. [l]

Vibranten (engl.: *trills* it: *vibranti*): ein bewegliches Sprechorgan vibriert: z.B. die Zungenspitze, bei [r], oder die Uvula, bei [R]

Approximanten *(*engl.: *approximants,* it.: *approssimanti)***:** der Luftstrom muss wie bei den Frikativen durch einen engen Kanal, aber es entsteht kein Reibegeräusch. z.B. [j]

Übung 4

Warum muss man bei Nasalen, Lateralen, Vibranten und Approximanten nicht angeben, ob sie stimmhaft oder stimmlos sind?

Plosive und Frikative erscheinen in vielen Sprachen sowohl in ihrer stimmhaften als auch in ihrer stimmlosen Form. So gibt es z.B. im Italienischen einen stimmhaften labialen Plosiv, das [b], und daneben gibt es auch den stimmlosen labialen Plosiv, das [p]. Nasale, Laterale, Vibranten und Approximanten hingegen erscheinen normalerweise immer als stimmhaft, es ist also überflüssig, außer Artikulationsart und -ort noch die Stimmhaftigkeit anzugeben.

Auf der folgenden Seite findet ihr das **Internationale Phonetische Alphabet (IPA),** das versucht, eine international gültige Transkriptionsweise aller Laute aller Sprachen anzubieten. Wir werden uns in unseren Transkriptionen nach diesem Alphabet richten. Zu allererst wollen wir uns die Tabelle der Konsonanten anschauen, und uns diejenigen Konsonanten heraussuchen, die in den Sprachen Deutsch, Italienisch und Englisch vorkommen. Wie schon gesagt, man klassifiziert einen Konsonanten nach seiner Artikulationsart und seinem Artikulationsort, ihr findet also im IPA in der 1. Spalte alle relevanten Artikulationsarten (*plosive, nasal, trill* u.s.w.) und in der 1. Zeile alle relevanten Artikulationsorte (*bilabial, labiodental, dental* u.s.w.).

Im Netz hat Peter Ladefoged ein IPA zur Verfügung gestellt, bei dem man die einzelnen Laute anklicken und dann ihre Aussprache hören kann: ihr findet das "sprechende" IPA unter
http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/

Schaut dort nach unter: → Vowels and Consonants
→ 1. Sounds and languages -The IPA chart sounds

IPA

Ihr könnt die Tabelle des Internationalen Phonetischen Alphabets an folgenden Stellen finden:

- im Internet, auf der Web-Seite der *International Phonetic Association:* http://www2.arts.gla.ac.uk/IPA/images/ipachart.gif

- in jedem Einführungsbuch zur Phonologie oder Linguistik (z.B. Hall 2000, Nespor 1999)

- in der *copisteria Rapida*, wo es eine Papierfassung dieser *dispensa* gibt.

Schauen wir uns die verschiedenen Gruppen von Lauten, geordnet nach Artikulationsarten an und überlegen uns, welche von ihnen es in den Sprachen Deutsch, Italienisch und Englisch gibt.

Plosive: im Deutschen finden wir die bilabialen Plosive [p, b], die alveolaren (oder dentalen) Plosive [t, d], die velaren Plosive [k, g] und den glottalen Plosiv [ʔ]. Falls ihr Schwierigkeiten habt, den glottalen Plosiv zu produzieren, so versucht, leicht zu husten. Ihr werdet merken, wie sich eure Glottis schließt. Nun versucht, das Wort *[ʔ]Anton* auszusprechen, indem ihr eure Glottis schließt, bevor ihr den initialen Vokal aussprecht. Auch wenn das seltsam erscheinen mag, der glottale Plosiv ist vom artikulatorischen Standpunkt aus ein Konsonant wie jeder andere auch. Im Deutschen wird er dazu verwendet, einen Anfangskonsonanten für Silben zu beschaffen, die mit Vokal beginnen, aber im Arabischen ist dieser Konsonant distinktiv und gleich wichtig wie z.B. ein [t]. Im Italienischen und Englischen gibt es auch bilabiale, alveolare und velare Plosive. Im Englischen, aber nicht im Italienischen, gibt es auch den glottalen Plosiv [ʔ]. Er ersetzt in bestimmten Varietäten des Englischen (z.B. Cockney, Scottish English) am Silbenanfang vor unbetontem Vokal oder am Wortende die "normalen" Plosive, wie z.B. in der Aussprache der folgenden Beispiele (Giegerich 1992):

(14)	<pit>	[pɪʔ]	<mat>	[maʔ]
	<Peter>	[piʔə]	<map>	[maʔ]
	<mighty>	[maɪʔɪ]	<mack>	[maʔ]

Weder im Deutschen noch im Italienischen oder Englischen gibt es retroflexe Plosive. Es gibt in diesen Sprachen überhaupt keine retroflexen Laute, mit Ausnahme der amerikanischen Aussprache des r-Lautes. Retroflexe Laute kann man allerdings in einigen sizilianischen Dialekten finden. Vielleicht habt ihr schon einmal gehört, wie Leute aus Trapani den Namen *Trapani* aussprechen? Sie produzieren am Wortanfang einen retroflexen Konsonantencluster. Verbreiteteter sind retroflexe Laute allerdings in einigen Sprachen Indiens. Wer sich dafür interessiert, wie retroflexe Laute gebildet werden, und wie sie sich anhören, der sollte die homepage von Peter Ladefoged (s. oben) besuchen.

Auch uvulare Laute sind im Deutschen, Italienischen und Englischen kaum präsent. Wir finden hier nur die r-Laute [R, ʁ] des Deutschen (s. unten), aber keine Plosive oder Nasale. Uvulare Plosive finden wir z.B. in einigen Eskimosprachen.

Nasale: Im Deutschen, Italienischen und Englischen finden wir den bilabialen Nasal [m], den alevolaren/dentalen Nasal [n] und den velaren Nasal [ŋ]. Das [ŋ], auch *angma* [ɛŋma] genannt, taucht z.B. in den folgenden Wörtern auf:

(15)	Dt.:	<Ring>	[Rɪŋ]
	It.:	<anche>	[aŋke]
	Engl.:	<sing>	[sɪŋ]

Im Italienischen gibt es außerdem noch den palatalen Nasal [ɲ], wie in *lasagne, gnocchi*.

Vibranten: das italienische "r" wird als ein alveolarer Vibrant [r] realisiert. Da bei der Produktion dieses Lautes die Zungenspitze vibriert, spricht man auch von einem apikalen "r". Im Deutschen hingegen finden wir den uvularen Vibranten [R], den man

besonders deutlich am Anfang eines Wortes hören kann, wie z.B. in *raus, rein, rasen*.

Taps und **Flaps:** der Tap/Flap [ɾ] ist ein Laut, bei dem die Zungenspitze nicht vibriert, sondern nur einmal kurz an die Alveolen schlägt. Diesen Laut finden wir im Amerikanischen Englisch in der Aussprache von [t] und [d] nach einem betonten Vokal (Giegerich 1992), wie z.B. in:

(16)	<city>	[pɪɾɪ]	<writing>	[raɪɾɪŋ]
	<pity>	[pɪɾɪ]	<riding>	[raɪɾɪŋ]

Frikative: Im Deutschen gibt es die labiodentalen Frikative [f, v], die alveolaren Frikative [s, z], die postalveolaren Frikative [ʃ, ʒ], den palatalen Frikativ [ç], den velaren Frikativ [x], den uvularen Frikativ [ʁ] und den glottalen Frikativ [h]. Der stimmhafte postalveolare Frikativ [ʒ] kommt im Deutschen nur in Fremdwörtern wie *Genie* [ʒe'ni:] und *Garage* [ga'Ra:ʒə] vor. Und auch bei der Aussprache dieser Wörter ersetzen viele deutsche Sprecher das stimmhafte [ʒ] durch ein stimmloses [ʃ] und sagen [ʃe'ni:] und [ga'Ra:ʃə]. Der stimmlose palatale Frikativ [ç] wird auch *ich-Laut* genannt, da er genauso ausgesprochen wird, wie der Konsonant im Wort *ich*. Der stimmlose velare Frikativ [x] hingegen wird *ach-Laut* genannt, weil er so wie der Konsonant im Wort *ach* ausgesprochen wird. Der uvulare Frikativ [ʁ] entspricht einer der Realisierung des r-Lautes im Deutschen. Wir finden diesen Laut vor allem im Wortinneren, z.B. in Wörtern wie *fahren*. Wir können ihn jedoch auch am Wortanfang hören. Wir werden die Realisierung der verschiedenen r-Laute in einem späteren Kapitel noch ausführlich diskutieren.
Das Italienische hat weniger Frikative als das Deutsche, ihm fehlen die Frikative [ʒ, ç, x] und [h]. Im Englischen finden wir den stimmlosen interdentalen Frikativ [θ] wie in *thin* und den stimmhaften interdentalen Frikativ [ð] wie in *the*.

Approximanten: unter den Approximanten finden wir den Laut [j], der dem Anfangskonsonanten im Wort *ja* entspricht.

Laterale: Unter *lateral appoximant* finden wir den einzigen Lateralen des Deutschen, das alveolare/dentale [l]. Im Italienischen finden wir außerdem noch einen palatalen Lateralen [ʎ], der den Buchstaben <gl> in Wörtern wie *aglio, meglio* entspricht. Laterale und r-Laute werden oft unter dem Begriff **Liquide** zusammengefasst.

Affrikaten: erscheinen nicht in der IPA-Konsonantentabelle, denn sie sind in einem gewissen Sinne Kombinationen von zwei Konsonanten. Im Deutschen haben wir Affrikaten, deren erster Teil aus einem Plosiv besteht, der zweite aus einem Frikativ, und zwar:
- die stimmlose labiale Affrikate: [pf] wie in <Pferd>
- die stimmlose alveolare Affrikate: [ts] wie in <Zeit>
- die stimmlose postalveolare Affrikate: [tʃ] wie in <Matsch>

Übung 5

1. "Bearbeitet" eure IPA-Tabelle: macht um alle Laute des Deutschen einen Kreis, malt die Konsonanten des Italienischen rot an, schattiert die Konsonanten des Englischen - verschafft euch einen Überblick über die Laute, die es in diesen drei Sprachen gibt.
2. Sind die folgenden Laute stimmhaft oder stimmlos? [ʃ, ç, x, h]
3. Macht eine kleine kontrastive Studie:
- welche Konsonanten , die es im Deutschen gibt, gibt es im Italienischen nicht?
- welche Konsonanten, die es im Italienischen gibt, gibt es im Deutschen nicht?
- gibt es im Italienischen auch Affrikaten? Welche?
- mit welchen Konsonanten werden also deutsche bzw. italienische Sprecher besondere Schwierigkeiten haben?
4. Schaut euch noch einmal die englischen Beispiele zu den glottalen Plosiven und den Taps/Flaps an. Welche Probleme könnten sich bei den Beispielen *mat, map, mack, writing, riding* für den Hörer ergeben?

Fassen wir hier noch einmal zusammen, was wir zu den Konsonanten des Deutschen und des Italienischen gesagt haben und versuchen wir, die größtenUnterschiede zwischen den beiden Sprachen festzuhalten. Im Deutschen gibt es zwei uvulare Laute [R, ʁ] und zwei glottale Laute [ʔ, h], während die uvulare und glottale Region im Italienischen nicht vertreten ist. Im Italienischen gibt es hingegen zwei palatale Laute , [ɲ, ʎ], die es im Deutschen nicht gibt. Im Italienischen fehlen außerdem die beiden Frikative [ç, x] des Deutschen. Unterschiede zwischen den beiden Sprachen gibt es auch in der Realisierung der r-Laute: das Italienische verwendet dazu den alveolaren Vibranten [r], das Deutsche hingegen den uvularen Vibranten [R] und den uvularen Frikativ [ʁ].

Affrikaten gibt es im Italienischen auch, es sind aber nicht dieselben wie im Deutschen. Im Italienischen finden wir die alveolare Affrikate [ts], wie in *mazzo,* [mats:o][6] , und, im Gegensatz zum Deutschen, auch ihr stimmhaftes Gegenstück, wie in *zona* [dzɔna]. Es gibt außerdem die postalveolare Affrikate [tʃ], wie in *cena* [tʃena] und ihr stimmhaftes Gegenstück [dʒ] wie in *maggio* [madʒ:o].

Ihr könnt euch nun vorstellen, mit welchen Lauten deutsche Sprecher besondere Schwierigkeiten haben und auch, wie sie diese Laute aussprechen werden. Probleme gibt es mit dem apikalen [r], das wahrscheinlich durch die uvularen r-Laute [R, ʁ] ersetzt werden wird. Auch [ɲ, ʎ] sind für einen deutschen Sprecher schwer auszusprechen. Hier hört man oft eine Aussprache, die sich nach der Schrift richtet, also *Lasa[gn]e* und *a[gl]io*. Problematisch sind auch die stimmhaften Affrikaten, die es im Deutschen nicht gibt, sie werden oft stimmlos ausgesprochen. Man hört dann also für die oben beschriebenen Wörter die Aussprachen [tso:na] und [matʃo].

Umgekehrt wird ein italienischer Sprecher Schwierigkeiten mit den r-Lauten des Deutschen haben, er wird sie wahrscheinlich apikal als [r] realisieren. Auch die

[6] Ein Doppelpunkt nach Konsonanten und auch nach Vokalen drückt aus, dass diese lang sind. Lange Konsonanten nennt man auch **Geminaten**.

Frikative [ç, x] können schwer auszusprechen sein, sie werden bei italienischen Sprechern manchmal durch [k] (z.B. *machen* als [maken]) ersetzt. Die glottalen Laute [ʔ, h] werden von italienischen Sprechern oft einfach ausgelassen. Das ist weniger schlimm beim glottalen Plosiv [ʔ], da dieser im Deutschen nicht distinktiv ist. Es kann aber zu Verständigungsschwierigkeiten beim [h] führen, wenn z.B. *Haus* wie *aus* ausgesprochen wird. Schwierigkeiten kann es auch mit der labialen Affrikate [pf] geben. Diese wird aber auch in vielen deutschen Varietäten zu [f] vereinfacht. In diesen Varietäten wird dann *Pferd* als *[f]erd* ausgesprochen.

Was die glottalen Verschlusslaute und die flaps/taps im Englischen betrifft, so gibt es hier ein ganz offensichtliches Problem. Wenn z.B. im Scottish English alle Plosive am Ende eines Wortes durch den glottalen Plosiv ersetzt werden, so weiß ein Hörer, wenn er *[maʔ]* hört, nicht mehr, ob der Sprecher nun *mat, map* oder *mack* meint. Und beim amerikanischen flapping kann es dazu kommen, dass man die beiden Wörter *writing* und *riding* nicht mehr unterscheiden kann! Einen phonologischen Prozess dieser Art nennt man **Neutralisierung**, da durch ihn Unterschiede zwischen distinktiven Lauten neutralisiert werden.

2.1.4 Artikulatorische Phonetik - Klassifikation der Vokale

Auf der IPA-Tabelle seht ihr in der Mitte rechts das sogenannte **Vokaltrapez.** Es ist eine schematische Darstellung der Artikulationsorte der Vokale. Es symbolisiert sozusagen den Raum in unserer Mundhöhle, in dem die Vokale artikuliert werden. Am besten versteht man das, wenn man es selbst ausprobiert. Sprecht zuerst den Vokal [i] aus, dann den Vokal [u]. Ihr werdet spüren, dass die Zunge bei beiden Vokalen hoch gehoben wird. Der Unterschied zwischen den beiden Vokalen besteht darin, dass die Zunge beim [i] nach vorne, beim [u] hingegegen nach hinten geschoben wird. Nun sprecht hintereinander die Vokale [i], [e] und [a] aus. Ihr solltet dabei merken, dass die Zunge bei der Aussprache dieser Vokale von oben nach unten wandert. Das [i] ist also ein hoher Vokal, das [e] ein mittlerer Vokal und das [a] ein tiefer Vokal. Nun versucht dasselbe mit der Lautsequenz [u, o, a]. Ihr werdet wieder merken, bei [u] steht die Zunge höher als bei [o] und bei [o] höher als bei [a]. Nun sprecht hintereinander die Laute [i, u, o, a, e] aus, und ihr habt ungefähr die Ecken des Vokaltrapezes mit der Zunge nachgezeichnet. Wenn ihr alle Ecken genau nachzeichnen möchtet, solltet ihr versuchen, für die untere rechte Ecke ein hinteres [ɒ] und für die untere linke Ecke ein vorderes [a] zu sprechen. Die Lautpaare [i] und [u] auf der einen Seite und [e] und [o] auf der anderen unterscheiden sich nicht nur dadurch, dass die Zunge vorne bzw. hinten steht, sondern auch dadurch, dass bei den Vokalen [u, o] die Lippen gerundet werden, bei den Vokalen [i, u] jedoch nicht. In dem Raum, der durch die Vokale [i, u, o, ɒ, a, e] eingegrenzt wird, können viele Vokale realisiert werden. Ihr könnt sie im Vokaltrapez des IPA sehen: dieses enthält alle Vokale aller Sprachen der Welt.

Wir werden nun im Folgenden versuchen, die Vokale des Deutschen zu beschreiben und sie mit denen des Italienischen zu vergleichen.

Die Vokale des Deutschen werden nach folgenden Merkmalen klassifiziert:
- nach der Höhe der Zungenposition: **hohe, tiefe, mittlere** Vokale
- nach der horizontalen Position der Zunge: **vordere, hintere, zentrale** Vokale
- nach der Rundung der Lippen: **gerundete** oder **ungerundete** Vokale
- danach, ob die Zunge bei der Produktion gespannt oder ungespannt ist: **gespannte** und **ungespannte** Vokale.

In den folgenden Tabellen, die das Vokaltrapez simulieren, sind die Vokale des Deutschen zuerst alleine und dann anhand von Beispielen dargestellt. Wir werden sie weiter unten im Einzelnen diskutieren.

(17) Die Vokale des Deutschen:

	vorne	zentral	hinten
hoch	gespannt: **[i:]** ungespannt: **[ɪ]** gerundet gespannt: **[y:]** gerundet ungespannt: **[ʏ]**		gespannt: **[u:]** ungespannt: **[ʊ]**
mittel	gespannt: **[e:]** ungespannt: **[ɛ]** **[ɛ:]** gerundet gespannt: **[ø:]** gerundet ungespannt: **[œ]**	schwa: **[ə]** vokalisiertes "r": **[ɐ]**	gespannt: **[o:]** ungespannt: **[ɔ]**
tief		gespannt: **[a:]**, ungespannt: **[a]**	

(18) Die Vokale des Deutschen: mit Beispielen

	vorne	zentral	hinten
hoch	gespannt: M**[i:]**te, <Miete> ungespannt: M**[ɪ]**tte, <Mitte> gerundet gesp.: H**[y:]**te, <Hüte> gerundet ungesp.: H**[ʏ]**tte, <Hütte>		gesp.: sp**[u:]**ken, <spuken> ungesp.: sp**[ʊ]**cken, <spucken>
mittel	gespannt: B**[e:]**t, <Beet> ungespannt: B**[ɛ]**tt, <Bett> **[ɛ:]**re, <Ähre> gerundet gesp.: H**[ø:]**le, <Höhle> gerundet ungesp.: H**[œ]**lle, <Hölle>	schwa: ruf**[ə]**n, <rufen> vok. "r": nu**[ɐ]**, <nur>	gesp.: Sch**[o:]**ten, <Schoten> ungesp.: Sch**[ɔ]**tten, <Schotten>
tief		gesp.: B**[a:]**n, <Bahn> ungesp.:B**[a]**nn, <Bann>	

Die ersten beiden Merkmale, nach denen die Vokale des Deutschen klassifiziert werden, sind einfach zu verstehen. Wie in allen Sprachen der Welt gibt es im Deutschen Vokale, die nach der Zungenhöhe unterschieden werden:

- hohe Vokale: [iː, ɪ, yː, ʏ, uː, ʊ]
- mittlere Vokale: [eː, ɛ, ɛː, øː, œ, ə, ɐ, oː, ɔ]
- tiefe Vokale [aː, a]

Außerdem gibt es Vokale, die nach der horizontalen Zungenposition unterschieden werden:

- vordere Vokale: [iː, ɪ, yː, ʏ, eː, ɛ, ɛː, øː, œ]
- zentrale Vokale: [ə, ɐ, aː, a]
- hintere Vokale: [uː, ʊ, oː, ɔ]

Ein wichtiges Merkmal im Deutschen ist außerdem das der Lippenrundung. In den meisten Sprachen der Welt sind die hinteren Vokale gerundet, die vorderen Vokale hingegen ungerundet. Es gibt dann aber auch Sprachen wie das Deutsche (und z.B. auch das Französische oder bestimmte lombardische Dialekte), in denen es die etwas ungewöhnlicheren gerundeten Vordervokale gibt:

- gerundete Hintervokale: [uː, ʊ, oː, ɔ]
- gerundete Vordervokale: [yː, ʏ, øː, œ]

Versucht, ein [iː] wie im Wort *Miete*, und danach ein [yː] wie im Wort *Hüte* auszusprechen. Ihr solltet dabei merken, dass die Zunge ungefähr in derselben Position verbleibt. Das einzige, was diese beiden Laute unterscheidet ist, dass das [iː] ungerundet ist, das [yː] jedoch gerundet. Nun wiederholt dieselbe Ausspracheübung mit [eː] wie in *Beet* und [øː] wie in *Höhle*. Wieder unterscheiden sich beide Laute nur durch die Lippenrundung. Nun wisst ihr, was ihr jemandem raten sollt, der die Umlaute <ü> und <ö> nicht aussprechen kann: zuerst ein [iː] bzw. ein [eː] aussprechen und dann die Lippen runden!

Es gibt auch einige Sprachen, in denen es ungerundete Hintervokale gibt. So wird z.B. im Japanischen der hohe hintere Vokal als ein ungerundetes [ɯ] ausgesprochen. Versucht selbst, ungerundete hintere Vokale zu artikulieren!

Ein weiteres Merkmal, das im Deutschen relevant ist, ist das der Gespanntheit. Ein gespannter Vokal unterscheidet sich von einem ungespannten Vokal dadurch, dass die Zunge bei seiner Produktion weiter oben und weiter gegen die äußere Grenze des Vokaltrapezes gestreckt steht als beim ungespannten Vokal. Vergleicht auf dem Vokaltrapez des IPA die Position des gespannten [i] und des ungespannten [ɪ], um zu verstehen, was hier gemeint ist.

- gespannte Vokale: [iː, yː, eː, øː, aː, uː, oː]
- ungespannte Vokale: [ɪ, ʏ, ɛ, ɛː, œ, a, ʊ, ɔ]

Man hört den Unterschied zwischen gespannten und ungespannten Vokalen am besten, wenn man Minimalpaare wie *Miete* und *Mitte, Hüte* und *Hütte, Beet* und *Bett, Höhle* und *Hölle, spuken* und *spucken, Schoten* und *Schotten* vergleicht. Bitte beachtet, dass sich die Wörter *Miete* und *Mitte* nicht durch die Länge des Konsonanen [t] unterscheiden. Im Unterschied zum Italienischen gibt es im Deutschen keine langen Konsonanten (keine Geminaten). Die Doppelkonsonanten,

die man in der Orthographie verwendet, zeigen nur an, dass der vorhergehende Konsonant kurz ist. Die Wörter *Miete* und *Mitte* werden also als [mi:tə] und [mɪtə] ausgesprochen. Beide Wörter enthalten ein [t] und unterscheiden sich nur durch die Gespanntheit bzw. Länge der beiden Vokale.

In den meisten Varietäten des Englischen finden wir einen ähnlichen Kontrast zwischen gespannten und ungespannten Vokalen, wie z.B. in den folgenden Wörtern (Giegerich 1992):

(19) <beat> , b[i:]t vs. <bit>, b[ɪ]t
<pool>, p[u:]l vs. <pull>, p[ʊ]ll

Im Deutschen korreliert die Gespanntheit der Vokale mit der Vokallänge, d.h. im nativen Wortschatz sind gespannte Vokale immer lang und ungespannte Vokale immer kurz. Die einzige Ausnahme bildet hier der Vokal [ɛ:]. Es handelt sich hier um einen ungespannten, mittleren Vokal, der jedoch lang ist. Es gibt Sprecher des Deutschen, die diesen Vokal in bestimmten Wörtern, die mit Umlaut <ä> geschrieben werden, verwenden (z.B. in *Ähre* oder *Bär*). Es gibt aber auch viele Sprecher, die diesen Vokal nie verwenden, und stattdessen ein langes [e:] aussprechen. Im nicht-nativen Wortschatz, d.h. in Fremdwörtern, ist die Korrelation zwischen Gespanntheit und Vokallänge nicht so streng. Hier können wir dann z.B. auch ungespannte Vokale finden, die kurz sind. Es handelt sich dabei immer um Vokale, die nicht betont sind, bei betonten Vokalen gilt wieder die Korrelation Länge-Gespanntheit:

(20) <Philosophie> Ph[i]l[o]s[o]ph[i:]
<Bibliothek> B[i]bl[io]th[e:]k

Es kann für einen Sprecher des Italienischen sehr schwierig sein, gespannte und ungespannte Vokale auseinanderzuhalten, da es diesen Unterschied im Italienischen höchstens für die mittleren Vokale gibt (s. unten). Es ist etwas leichter - aber immer noch schwierig - zu erkennen, ob ein Vokal lang oder kurz ist. Wenn man das festgestellt hat, dann kann man, jedenfalls für den nativen Wortschatz, Rückschlüsse auf die Gespanntheit anstellen: wenn man einen langen Vokal hört, dann sollte er auch gespannt sein, wenn man einen kurzen Vokal hört, dann ist er höchstwahrscheinlich ungespannt.

Was die Gespanntheit betrifft, so unterscheiden sich die tiefen Vokale [a:] und [a] von den anderen Vokalen dadurch, dass sich zwischen dem langen, gespannten [a:] und dem kurzen, ungespannten [a] kaum ein Unterschied in der Vokalqualität hören lässt. In der Tat ist der Status von [a] im Deutschen kontrovers. Manche Phonologen nehmen an, dass es auch beim [a] einen Unterschied in der Vokalqualität zwischen kurzem ungespanntem [a] (wie in *Bann*) und langem, gespanntem [ɑ:] (wie in *Bahn*) gibt. Wir werden mit Wiese (1996) annehmen, dass es im Deutschen nur eine Unterscheidung zwischen langem [a:] und kurzem [a] gibt, dass sich aber die beiden Laute nicht in ihrer Vokalqualität unterscheiden

Es gibt im Deutschen auch zwei mittlere zentrale Vokale, das sogenannte schwa [ə] und das vokalisierte "r" [ɐ]. Das schwa kommt im Deutschen nie als Wurzelvokal vor und kann nie betont werden. Wir finden es aber in den Vokalen aller Flexionssuffixe und in vielen Derivationssuffixen. Hier sind einige Beispiele:

(21)	Infinitivsuffix:	<lieben>	lieb[ə]n
	Personalendungen am Verb:	<liebe>	lieb[ə]
	Adjektivflexion:	<schöne>	schön[ə]
	Pluralflexion:	<Tage>	Tag[ə]
	Substantive auf -e:	<Liebe>	Lieb[ə]
	Präfix be-:	<begreifen>	b[ə]greifen

Das vokalisierte [ɐ] ist das Resultat eines phonologischen Prozesses, den wir weiter unten noch genauer besprechen werden. Um den Unterschied zwischen der Aussprache des schwa und des vokalisierten [ɐ] zu verstehen, solltet ihr einen Muttersprachler bitten, euch einige Minimalpaare wie die folgenden vorzusprechen:

(22)	<Liebe> vs. <lieber>	[li:bə] vs. [li:bɐ]
	<Ehe> vs. <eher>	[ʔe:ə] vs. [ʔe:ɐ]
	<schöne> vs. <schöner>	[ʃø:nə] vs. [ʃø:nɐ]

Wenn ihr euch vor allem die letzten beiden Beispiele, die beiden Adjektivformen *schöne* und *schöner,* vor Augen haltet, so werdet ihr plötzlich verstehen, warum es so schwierig ist, ein deutsches Diktat zu schreiben: der Unterschied in der Aussprache der beiden Vokale in der Flexion ist minimal, man muss also schon wissen, ob in der Nominalphrase *ein schön_ Hund* die Endung *-e* oder die Endung *-er* eingesetzt werden muss. Hören kann das ein Nicht-Muttersprachler oft nicht!

Übung 6

1. Versucht, anhand von folgenden Beispielen das Vokalsystem des Italienischen zu skizzieren:
butti, batte, *botte* (nel senso di *picchiare*), *botte* (un barile), *pesca* (frutto), *pesca* (attività per catturare pesci), *ride, rade, rode, rude*.
2. Welche Vokalklassen gibt es im Deutschen, die es im Italienischen nicht gibt? Welche Schwierigkeiten ergeben sich dadurch für Sprecher des Italienischen, wenn sie Deutsch lernen? Versucht, die Vokalklassen anzugeben, nicht nur die einzelnen Vokale aufzulisten.
3. Gibt es, eurer Meinung nach, im Italienischen lange und kurze Vokale, so wie im Deutschen?

Das Italienische verfügt über viel weniger Vokale als das Deutsche, wodurch sich einige Schwierigkeiten für italienische Sprecher ergeben, die Deutsch lernen. Auch im Italienischen gibt es hohe Vokale, wie [i] und [u], mittlere Vokale, wie [e, ɛ, o, ɔ] und den tiefen Vokal [a]. Aber im Italienischen gibt es keine gerundeten Vordervokale, wie sie die Vokale [y:, ʏ, ø:, œ] im Deutschen darstellen. Auch der Gespanntheitskontrast ist bei den italienischen Vokalen stark reduziert. Es gibt keinen Unterschied zwischen gespannten und ungespannten Vokalen was die hohen und die tiefen Vokale [i, u, a] betrifft. Bei den mittleren Vokalen gibt es diesen Kontrast allerdings auch im Italienischen. Man spricht hier oft von *vocali aperte* und *vocali chiuse*. Es gibt jedoch nur wenige Minimalpaare, an denen man ihn direkt beobachten kann. Für einige (aber nicht alle!) Sprecher des Italienischen gibt es einen Kontrast zwischen *p[ɛ]sca* (der Pfirsich) und *p[e]sca* (der Fischfang), für einige

(aber nicht alle!) Sprecher gibt es auch einen Kontrast zwischen *b[ɔ]tte* (die Schläge) und *b[o]tte* (das Fass). Die meisten Sprecher werden außerdem zwischen *[ɛ]* (<*è*>, 3. Person Sg. der Verbes *essere*) und [*e*] (<*e*>, koordinierende Konjunktion) und zwischen *[ɔ]* (<ho>, 1. Person Sg. des Verbes *avere*) und *[o]* (<o>, disjunktive Konjunktion) unterscheiden. Aber viele Minimalpaare dieser Art kann man für den Gespanntheitskontrast der mittleren Vokale nicht finden. Völlig fehlen im Italienischen die zentralen Vokale schwa und das vokalisierte [ɐ]. Das schwa kann man jedoch in einigen italienischen Dialekten wie z.B. dem Neapoletanischen wiederfinden. Auch hier taucht das schwa in der Endung auf, z.B. im letzten Vokal der neapolitischen Aussprache von *Napule*.
Die Antwort auf die Frage, ob das Italienische lange und kurze Vokale hat, findet ihr im Kapitel 4.1.4, wo die Struktur der deutschen und der italienischen Silbe verglichen wird.

Die folgende Tabelle stellt das italienische Vokalsystem noch einmal schematisch dar:

(23) Die Vokale des Italienischen:

	vorne	zentral	hinten
hoch	**[i]** <ride>		**[u]** <rude>
mittel	gespannt: **[e]** <e> ungespannt: **[ɛ]** <è>		gespannt: **[o]** <o> ungespannt: **[ɔ]** <ho>
tief		**[a]** <rade>	

Ihr solltet nun kurz einen Schritt zurücktreten und euch die Vokalsysteme des Deutschen und des Italienischen aus einiger Entfernung ansehen. Ist es nicht beeindruckend, wie symmetrisch beide Systeme sind? Jedem Vordervokal entspricht ein Hintervokal, jedem hohen Vokal ein mittlerer Vokal und im Deutschen hat jeder gespannte Vokal sein ungespanntes Gegenstück. Wir können uns ohne Probleme ein Vokalsystem vorstellen, das nicht so symmetrisch ist, das z.B. über die Vokale [i, ø:, ɔ, a] verfügt. Extrem unsymmetrische Vokalsysteme sind aber höchst selten. Im Allgemeinen tendieren Vokalsysteme dazu, symmetrisch zu sein. Sie nützen den Artikulationsraum, den unsere Mundhöhle bietet, auf die bestmögliche Weise aus und verteilen die Vokale symmetrisch auf diesen Raum.

2.2 Phoneme - die distinktiven Einheiten des Lautsystems

2.2.1 Was ist ein Phonem?

Wir haben schon im Kap. 2.1.2, als wir die Aspiration besprochen haben, erwähnt, dass bestimmte Laute bzw. Merkmale von Lauten wie die Aspiration distinktiv sein können. Wir haben auch von Minimalpaaren gesprochen, d.h. von Wörtern, die sich nur durch ein bestimmtes phonetisches Merkmal voneinander unterscheiden. All

diese Begriffe wollen wir nun in diesem Kapitel präzisieren, indem wir uns intensiv mit dem Begriff des **Phonems** beschäftigen.

Die Linguisten Trubetzkoy und Jakobson haben, neben anderen, das Phonem als die **minimale bedeutungsunterscheidende Lauteinheit** definiert. Was bedeutet das? Wir wissen, dass Laute an und für sich keine Bedeutung haben, die kleinste bedeutungstragende Einheit der Sprache ist das Morphem. Laute, oder auch bestimmte Merkmale von Lauten, können jedoch Bedeutungen voneinander unterscheiden. So kann z.B. das Merkmal der Gespanntheit die beiden Wörter *Miete* und *Mitte* unterscheiden.

In vielen phonologischen Arbeiten wird das Phonem noch als eine untrennbare Einheit gesehen. Man sagt dann z.B., dass es im Deutschen die beiden Phoneme /i:/ und /ɪ/ gibt.[7] Aus Bequemlichkeit werden auch wir diese Ausdrucksweise auch heute noch oft, aber wenn man ganz präzise sein will, dann sollte man eigentlich nur von **distinktiven Merkmalen** sprechen, die sich zusammen zu einem Phonem verbinden. Jakobson & Halle (1956: 20) haben diese Einsicht folgendermaßen formuliert:

(24) " The distinctive features are aligned into simultaneous bundles called phonemes"

Wir können also Phoneme als Bündel von gleichzeitig vorkommenden distinktiven Merkmalen sehen. Jakobson und Halle meinen damit, dass sich in jedem Phonem ein Bündel von Eigenschaften vereinigt und jede dieser Eigenschaften kann in der betreffenden Sprache Bedeutungen unterscheiden. So beinhaltet das Phonem /i:/ im Deutschen z.B. das Merkmal [+gespannt][8], das es vom Phonem /ɪ/ unterscheidet. Es enthält auch das Merkmal [-rund], das es vom Phonem /y:/ unterscheidet, das Merkmal [+hoch], das es von den Phonemen /e:/ und /a:/ unterscheidet und das Merkmal [-hinten], das es von /u:/ unterscheidet. Alle diese Merkmale sind im Deutschen distinktiv, denn es genügt eines dieser Merkmale, um Wörter mit verschiedener Bedeutung zu unterscheiden. Man kann sich ein Phonem wie /i:/ wie eine Schachtel vorstellen: in der Schachtel sind die Merkmale, die das Phonem kennzeichnen, in diesem Fall [+gespannt, -rund, +hoch, -hinten]. Auf der Schachtel steht: /i:/.

Auch im Italienischen ist /i/ ein Phonem, es kann Wörter wie z.B. *ride* und *rude* unterscheiden. Aber die distinktiven Merkmale, die dieses Phonem kennzeichnen, sind nicht dieselben wie im Deutschen. So ist z.B. das Merkmal [±rund] im Italienischen nicht distinktiv. Es stimmt zwar, dass die hinteren Vokale im Italienischen rund sind und die vorderen ungerundet, aber um vordere und hintere Vokale zu unterscheiden genügt das Merkmal [±hinten], wir brauchen kein Merkmal [±rund] dazu. Die Schachtel des Phonems /i/ wird also im Italienischen kein Merkmal [-rund] enthalten, weil dieses Merkmal im Italienischen nicht distinktiv ist.

Untersuchen wir noch ein Beispiel, indem wir die labialen Plosive im Deutschen und im Hindi vergleichen. Im Hindi ist /p^h/ ein Phonem, es kann Bedeutungen unterscheiden, wie wir schon in Kap. 2.1.2 anhand der Beispiele [p]al, 'sich um jemanden kümmern' , und [p^h]al, 'Messerklinge', gesehen hatten. Das Phonem /p^h/ besteht aus einer ganzen Reihe von distinktiven Merkmalen, das, jedes für sich,

---

[7] Wie schon erwähnt schreiben wir Phoneme zwischen Schrägstrichen /.../, um sie von den Phonen zu unterscheiden, die wir zwischen eckigen Klammern [...] schreiben.

[8] Eine genaue Definition der distinktiven Merkmale des Deutschen findet sich in Kap. 3.7, in dem der Begriff des phonologischen Merkmals im Detail erläutert wird.

Bedeutungen unterscheiden kann. So unterscheidet das Merkmal [- stimmhaft] /p^h/ von /b^h/, denn je nachdem, ob ein Wort /p^h/ oder /b^h/ enthält, kann sich seine Bedeutung ändern. Das Merkmal [LABIAL][9] wiederum unterscheidet /p^h/ vom Phonem /t^h/, das nicht labial sondern dental ist. Das Merkmal [+aspiriert] schließlich unterscheidet /p^h/ von /p/. Auch dieses Merkmal kann im Hindi, wie wir gesehen haben, Bedeutungen unterscheiden. Die Phonemschachtel des /p^h/ enthält also im Hindi das distinktive Merkmal [±aspiriert].

Im Deutschen ist die Situation etwas anders. Auch im Deutschen gibt es ein aspiriertes [p^h], wie z.B. in dem Wort *Pein*, [p^haɪn]. Es gibt auch das Minimalpaar *Pein,* [p^haɪn] vs. *Bein* [b̥aɪn], das sich genau im ersten Laut des Wortes unterscheidet. Bedeutet das nun, dass die Aspiration im Deutschen distinktiv ist? Nein, und zwar aus zwei Gründen. Erstens gibt es kein Minimalpaar, das sich *nur* in der Aspiration eines Lautes unterscheidet. Im gerade genannten Beispiel unterscheiden sich [p^h] und [b̥] nicht nur durch das Merkmal [±aspiriert], sondern auch durch das Merkmal [±stimmhaft], auch wenn das [b̥] nur als teilweise stimmhaft realisiert wird. Das bedeutet zugleich, dass es keine zwei Wörter im Deutschen gibt, bei denen die Anwesenheit bzw. Abwesenheit der Aspiration die Bedeutung verändert. Es gibt keine zwei Wörter wie z.B. [p^haɪn] und [paɪn], bei denen das erste *Schmerzen* bedeutet und das zweite vielleicht *Freude*. Die Aspiration hat also im Deutschen keine distinktive, d.h. bedeutungsunterscheidende Kraft. Der zweite Grund dafür, dass wir [p^h] nicht als Phonem werten können ist, dass die Aspiration im Deutschen **vorhersagbar** (it.: *predicibile*, engl.: *predictable*) ist. Wir werden die Kontexte, in denen Aspiration vorkommt, noch in einem späteren Kapitel genau diskutieren, aber inzwischen können wir festhalten, dass stimmlose Plosive im Standarddeutschen auf jeden Fall am Wortanfang, vor betontem Vokal aspiriert werden, wie in den folgenden Beispielen:

(25) <Pein> [p^haɪn]
<Tonne> [t^hɔnə]
<kann> [k^han]

Im Wortinneren hingegen, besonders vor unbetontem Vokal, werden die stimmlosen Plosive nicht aspiriert:

(26) <Rappe> [Rapə]
<rate> [ra:tə]
<wecke> [v̥ɛkə]

Wenn die Verteilung eines Lautmerkmales vorhersagbar ist, dann bedeutet das zugleich, dass es nicht distinktiv sein kann. Denn wenn ein Merkmal vorhersagbar ist, dann hängt sein Auftreten vom phonologischen Kontext ("Anfang des Wortes", "Ende des Wortes", "vor betontem Vokal" u.s.w.) und eben gerade nicht von der Bedeutung des Wortes ab. Mit anderen Worten, man könnte auch sagen, dass die Phoneme jenen Teil des Lautsystems darstellen, der nicht vorhersagbar ist. Viele Phonologen nennen die Phonemsequenzen, die sich zu Wörtern zusammenschließen, auch **das Lexikon**, da die Phonemsequenzen der einzelnen Wörter, wie die Lexeme eines Wörterbuchs, von den Sprechern auswendig gelernt werden müssen und nicht vorhersagbar sind.

---

[9] Wenn es euch interessiert, warum [LABIAL] nicht als [±labial] geschrieben wird, dann solltet ihr euch Kapitel 3.7 anschauen.

Aber kehren wir zu den stimmlosen Plosiven des Deutschen zurück. Wie sollen wir dann die Phoneme nennen, durch die sich das Minimalpaar *Pein* vs. *Bein* unterscheidet? Man nimmt im Allgemeinen an, dass es im Deutschen zwei Phoneme gibt, /p/ und /b/, und dass sich diese beiden Phoneme durch das Merkmal [±stimmhaft] unterscheiden. Und das ist auch genau das Merkmal, das das Minimalpaar *Pein* vs. *Bein* unterscheidet. Wir können dann noch hinzufügen, dass das Phonem /p/ in bestimmten Kontexten (z.B. am Wortanfang vor betontem Vokal) als aspiriertes Phon [p^{h}] realisiert wird. In anderen Kontexten (z.B. im Wortinneren vor unbetontem Vokal) wird es hingegen ganz "normal" als stimmloses Phon [p] realisiert. Das Phonem /b/ hingegen wird am Wortanfang nur teilweise stimmhaft als [b̥] realisiert. Im Wortinneren, besonders zwischen Vokalen, wird das Phonem /b/ hingegen als vollkommen stimmhaftes Phon [b] realisiert.

Die Beziehung zwischen den Phonemen und ihren phonetischen Realisierungen wird oft als ein Modell mit zwei Ebenen dargestellt. Auf der Ebene der **zugrundeliegenden Repräsentation** (it.: *livello soggiacente*, engl.: *underlying representation, UR*) stellen wir die Phoneme dar, auf der Ebene der **Oberflächenrepräsentation** (it.: *livello di superficie*, engl.: *surface representation, SR*) die phonetischen Realisierungen dieser Phoneme. Für die oben diskutieren Beispiele würde das die folgende Darstellung ergeben:

(27) <Pein> vs. <Bein>:

/p/ein /b/ein ← zugrundeliegende Repräsentation (Ebene der Phoneme)

↓ ↓

[p^{h}]ein [b̥]ein ← Oberflächenrepräsentation (Ebene der Phone)

Die Wörter *Pein* und *Bein* enthalten die Phoneme /p/, bzw. /b/, die sich durch das Merkmal [±stimmhaft] unterscheiden. Da sich diese beiden Phoneme am Wortanfang befinden, wird das Phonem /p/ als ein aspiriertes Phon realisiert, während das Phonem /b/ als ein teilweise stimmhaftes Phon realisiert wird.

(28) <Rappe> und <Rabe>

Ra/p/e Ra/b/e ← zugrundeliegende Repräsentation (Ebene der Phoneme)

↓ ↓

Ra[p]e Ra[b]e ← Oberflächenrepräsentation (Ebene der Phone)

Die Wörter *Rappe* und *Rabe* enthalten ebenfalls die Phoneme /p/ bzw. /b/. In diesem Fall befinden sich diese Phoneme aber im Wortinneren, vor einem unbetonten Vokal. Deshalb wird das Phonem /p/ ohne Aspiration realisiert und /b/ wird als ein vollkommen stimmhaftes Phon [b] realisiert.

Diese Darstellung auf zwei Ebenen zeigt uns zugleich, dass das Phonem etwas Abstraktes ist. Wir können Phoneme nicht aussprechen, wir sprechen nur ihre konkrete Realisierung, die Phone aus.

Fassen wir noch einmal zusammen: Phoneme sind Bündel von distinktiven Merkmalen. Wir können annehmen, dass es eine bestimmte, begrenzte Menge von distinktiven Merkmalen gibt, aus der sich die Sprachen bedienen. Anders gesagt, die Menge an distinktiven Merkmalen ist universal gegeben und jede Sprache verwendet eine Teilmenge davon. Wir können diese Selektion von distinktiven Merkmalen von Seiten der einzelnen Sprachen schematisch folgendermaßen darstellen:

(29) Schematische Darstellung der Selektion von distinktiven Merkmalen:

	→	M_1		→	M_1			M_1
	→	M_2			M_2		→	M_2
L_1		M_3	L_2	→	M_3	L_3		M_3
	→	M_4		→	M_4		→	M_4
	→	M_5			M_5			M_5
	→	M_6		→	M_6			M_6
		....			....			....
		M_n			M_n			M_n

Wenn wir annehmen, dass der menschlichen Sprache im Allgemeinen n distinktive Merkmale zur Verfügung stehen, dann können wir uns vorstellen, dass die Sprache L_1 z.B. davon die Menge { M_1,M_2, M_4, M_5, M_6} auswählt, L_2 wählt die Menge { M_1, M_3, M_4, M_6} und L_3 die Teilmenge, {M_2, M_4}.

Um festzustellen, ob ein Merkmal in einer Sprache distinktiv ist, müssen wir herausfinden, ob es Bedeutungen unterscheiden kann, d.h., ob es Minimalpaare gibt, die sich nur in diesem Merkmal unterscheiden. Wir können außerdem versuchen herauszufinden, ob das Merkmal vorhersagbar ist. Wenn der Kontext, in dem es vorkommt, vorhersagbar ist, dann kann es sich nicht um ein distinktives Merkmal handeln. In Kapitel 3 werden wir die wichtigsten phonologischen Prozesse des Deutschen diskutieren. Dort werden wir also auch herausfinden, welche Laute das (voraussagbare) Resultat eines solchen Prozesses sind und welche hingegen Phoneme sind.

2.2.2 Allophone - zwei (oder mehrere) Realisierungen eines Phonems

Übung 7

Untersucht die Distribution der Laute [l] und [r] im Koreanischen (Daten nach Kenstowicz 1994:83):

mul	'Wasser'	mu.re	'am Wasser'
mul.ka.ma	'ein Platz für Wasser'		
mal	'Pferd'	ma.re	'am Pferd'
mal.ka.ma	'ein Platz für ein Pferd'		
pal	'Fuß'	pa.ri	'vom Fuß'
səul	'Seoul'	ru.pi	'Rubin'
il.kop	'sieben'	ra.tio	'Radio'

1. In welchem Kontext finden wir den Laut [l], in welchem Kontext den Laut [r]? Überlegt euch vor allem, in welcher Silbenposition die jeweiligen Laute stehen.
2. Ist die Distribution von [l] voraussagbar? Ist die Distribution von [r] voraussagbar?
3. Ist [l] also ein Phonem? Ist [r] ein Phonem?
4. Gibt es einen Kontext, in dem sowohl [l] als auch [r] vorkommen können?
5. Wie wird ein koreanischer Sprecher wohl das deutsche Wort *Lehrer* aussprechen?

Wenn ihr die koreanischen Daten sorgfältig untersucht habt, dann habt ihr sicher herausgefunden, dass wir den Laut [l] immer am Ende einer Silbe finden, den Laut [r] immer am Silbenanfang. Das ist besonders auffallend bei Wörtern, die das gleiche Lexem enthalten. So enthalten beispielsweise sowohl *mul* als auch *mu.re* das Lexem WASSER, aber im ersten Fall finden wir den Laut [l], der immer am Silbenende auftaucht, im zweiten Fall den Laut [r], weil am Silbenanfang nur [r] stehen kann. Wir müssen schließen, dass die Distribution der Liquide [l] und [r] vollkommen voraussagbar ist: [l] steht immer am Silbenanfang, [r] immer am Silbenende. Die beiden Laute [l] und [r] können also keine Phoneme sein. Es gibt außerdem keinen Kontext, in dem sowohl [l] als auch [r] stehen können, denn Silbenanfang und Silbenende schließen sich gegenseitig als Kontexte aus: was am Silbenanfang steht kann nicht zugleich am Silbenende stehen. Man spricht in diesen Fällen von **komplementärer Distribution**: die Verteilung dieser beiden Laute ist komplementär, wo der eine steht kann nie auch der andere stehen. Und was die letzte Frage betrifft: ein koreanischer Sprecher wird das Wort *Lehrer* wahrscheinlich irgendwie wie *Rerel* aussprechen, indem er den ersten silbeninitialen Liquid durch [r] ersetzt und den letzten, silbenfinalen durch [l].

Wenn [l] und [r] im Koreanischen keine Phoneme sind, was sind sie denn dann? Laute, die in komplementärer Distribution zueinander stehen und die man als Realisierungen ein und desselben Phonems betrachten kann, nennt man **Allophone.**

Es gibt also im Koreanischen ein Phonem, nennen wir es ganz allgemein 'Liquid', das, je nach Kontext, zwei Allophone realisieren kann, entweder [l] oder [r]:

(30) /Liquid/ ← zugrundeliegende Repräsentation

[l] [r] ← Oberflächenrepräsentation (Ebene der (Allo)phone)

Kontext: Silbenende — Kontext: Silbenanfang

Schauen wir uns noch einmal die Beispiele *mul* und *mure* genau an, um noch einen Vorteil zu erkennen, den ein Modell bietet, das eine zugrundeliegende phonematische Ebene und eine phonetische Oberflächenrepräsentation annimmt. Wir können das Wort *mure* in zwei Morpheme zerlegen, *mur-e*, von denen das erste der Wurzel mit der Bedeutung WASSER entspricht, das zweite einer Lokalangabe mit der Bedeutung NAHE BEI/AM. Das würde aber bedeuten, dass das Lexem WASSER, wenn es allein durch die Wurzel ausgedrückt wird, die Form *mul* hat, wenn es aber durch ein lokales Suffix ergänzt wird, die Form *mur*. Wie ist dieses Lexem wohl in unserem mentalen Lexikon abgespeichert? Mit beiden Formen? Mit einer von beiden? Wenn wir ein Modell mit zwei Ebenen annehmen, dann können wir einfach sagen, dass wir in unserem mentalen Lexikon die zugrundeliegende Form, also die Phonemsequenz abspeichern. Wir würden uns also für das Lexem WASSER im Koreanischen ungefähr die Phonemsequenz /m/, /u/ und /Liquid/ merken. Vielleicht merken wir uns auch die Phonemsequenz /mul/ oder die Phonemsequenz /mur/, d.h., vielleicht entscheiden wir, dass /l/ bzw. /r/ das zugrundeliegende Element ist, statt ein unbestimmtes Phonem /Liquid/. Aber was wichtig ist, ist, dass wir uns nicht zwei Formen für das Lexem WASSER merken müssen, sondern nur eine. Das spart Speicherplatz im Gehirn. Die konkrete Realisierung des Phonems /Liquid/ (oder, wenn ihr wollt, /l/, oder /r/) hingegen leiten wir von einer Regel ab, die wir auch in unserem Gehirn abgespeichert haben und die uns sagt, dass wir am Silbenanfang

immer [r] finden müssen, am Silbenende immer [l]. Wir müssen uns also nicht für jedes Lexem, das einen Liquid enthält, zwei Formen, eine mit [l] und eine mit [r] merken. Wir merken uns nur jeweils eine Form des Lexems und wir merken uns die Regel, die uns dabei hilft, die Liquide korrekt zu realisieren.

Nun schauen wir uns ein Beispiel für allophonische Realisierung eines Phonems im Deutschen an, die **Assimilation der dorsalen Frikative (ich- und ach-Laut)**

Übung 8

Lasst euch die folgenden Beispiele von einem deutschen Muttersprachler vorsprechen und versucht herauszufinden, in welchen Wörtern der palatale Frikativ [ç](*ich-Laut*) und in welchen der velare Frikativ [x] (*ach-Laut*) zu hören ist.

dich	Dach	Kuhchen (kleine Kuh)
Tuch	Tschechow	fauchen
Tücher	Milch	Pfauchen (kleiner Pfau)
Loch	manch	Chemie
löchern	Kuchen	China

1. Schreibt neben jedes Wort, ob es ein [ç] oder ein [x] enthält.
2. Macht eine Liste der Laute, die vor [ç] auftreten und eine Liste der Laute, die vor [x] auftreten. Gebt auch an, ob sich vor den beiden Frikativen jeweils eine Morphemgrenze befindet oder nicht
3. Beschreibt die Distribution von [ç] oder ein [x]:
[ç] tritt in folgendem Kontext auf:_____
[x] tritt in folgendem Kontext auf:_____
4. Sind [ç] und [x] zwei Phone oder zwei Phoneme?
5. Falls [ç] und [x] Allophone sind, welches ist dann das entsprechende Phonem? /ç/ oder /x/?

Wenn ihr gut zugehört habt, dann solltet ihr den ich-Laut [ç] in folgenden Wörtern gehört haben: *dich, Tücher, löchern, Tschechow, Milch, manch, Kuhchen, Pfauchen*. Wahrscheinlich habt ihr ihn auch als ersten Laut in den Wörtern *Chemie, China* gehört. Falls der Muttersprachler, den ihr befragt habt, aus dem Süden des deutschen Sprachraumes kommt, dann hat er wahrscheinlich *[k]emie* und *[k]ina* gesagt. Den ach-Laut [x] solltet ihr in den Wörtern *Tuch, Loch, Dach, Kuchen, fauchen* gehört haben.

Die Liste der Laute, die vor dem ich-Laut auftauchen sieht demnach folgendermaßen aus: [ɪ, y:, œ, ɛ, l, n, u:+Morphemgrenze, aʊ+Morphemgrenze, Wortanfang]. Mit *[u:]+Morphemgrenze* und *[aʊ]+Morphemgrenze* sind die letzten beiden Beispiele gemeint, *Kuh-chen* und *Pfau-chen*, bei denen der ich-Laut genau an der Morphemgrenze, am Anfang des Diminutivmorphems *-chen* zu stehen kommt. Mit *Wortanfang* sind die beiden Beispiele *Chemie* und *China* gemeint, in denen der ich-Laut am Wortanfang steht. Eigentlich handelt es sich in diesen beiden Fällen auch um eine Morphemgrenze, denn ein Wortanfang ist auch immer zugleich der Anfang eines Morphems. Wir können nun versuchen, diese Liste von Kontexten in einer etwas allgemeineren Formulierung zusammenzufassen. Wenn wir die Liste der

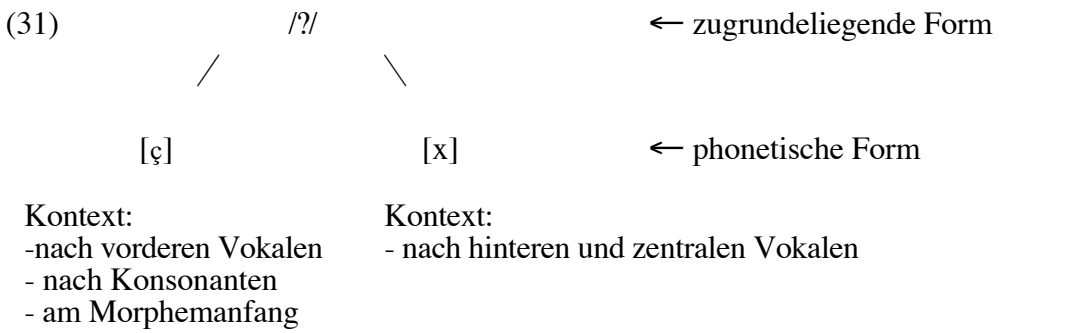
Vokale betrachten, nach denen der ich-Laut auftaucht, so können wir feststellen, dass es sich bei diesen Vokalen immer um vordere Vokale handelt. Wir können also den Kontext für die Distribution von [ç] folgendermaßen formulieren: Der ich-Laut [ç] erscheint nach vorderen Vokalen, nach Konsonanten und am Anfang eines Morphems.

Den ach-Laut finden wir in den obigen Beispielen nach den Lauten [u:, ɔ, a, u:, aʊ]. Wenn wir hier sagen sollen, zu welcher Klasse diese Vokale gehören, dann fällt uns auf, dass es sich dabei um nicht-vordere, also zentrale oder hintere Vokale handelt. Wir stellen also fest, dass der ach-Laut [x] im Kontext hinter zentralen oder hinteren Vokalen auftritt. Kehren wir auch noch einen Augenblick zu den Beispielen *Kuh-chen* und *Pfau-chen* zurück. Da wir in diesen Beispielen einen ich-Laut finden, müssen wir schließen, dass dieser ich-Laut auf die Morphemgrenze des Suffixes *-chen* zurückzuführen ist, und nicht auf die Laute [u:] und [aʊ], die dem Frikativ vorausgehen, denn hinter den nicht-vorderen Vokalen [u:] und [aʊ] würden wir uns sonst einen ach-Laut [x] erwarten. Das sehen wir in den sehr ähnlichen Wörtern *Kuchen* und *fauchen* in denen wir den ach-Laut finden, denn vor dem Frikativ befinden sich direkt die hinteren Vokale [u:, aʊ], jedoch keine Morphemgrenze.

Wir haben nun die beiden Kontexte gefunden, in denen ich-Laut und ach-Laut vorkommen. Es handelt sich bei der Verteilung dieser Laute ganz klar um komplementäre Distribution, denn es gibt keinen Kontext, in dem wir sowohl ich-Laut als auch ach-Laut finden können. Wenn die beiden Laute komplementär verteilt sind, dann kann es sich bei ihnen nicht um Phoneme handeln. Ich- und ach-Laut sind zwei Phone, deren Verteilung absolut voraussagbar und komplementär verteilt ist, also handelt es sich bei ihnen um zwei Allophone desselben Phonems.

Welchen Namen können wir nun aber diesem Phonem geben? Ist das zugrundeliegende Phonem /ç/, das dann die zwei Realisierungen [ç] und [x] aufweist oder ist das Phonem /x/, wiederum mit den Realisierungen [ç] und [x]?

(31)



Es hat in der phonologischen Literatur viel Diskussion dazu gegeben (s. z.B. Wurzel 1981, Hall 1992, Ramers&Vater 1995, Wiese 1996, Hall 2000), ob im Fall der Assimilation der dorsalen Frikative /ç/ oder /x/ als zugrundeliegend anzunehmen sei Wir haben hier nicht die Möglichkeit, im Detail die Gründe für die eine oder die andere Lösung nachzuvollziehen, deshalb werden wir nur eines der Argumente ausführlich diskutieren, da es sich um ein typisches Argument für Probleme dieser Art handelt. Die Argumentation verläuft folgendermaßen: als zugrundeliegendes Phonem müsse man /ç/ annehmen, da die Realisierung [ç] keinen homogenen Distributionskontext aufweise. Was bedeutet das? Wir können sehen, dass die Kontexte, in denen [ç] vorkommt, sehr verschieden (nicht homogen) sind. Es handelt

sich eigentlich um drei Kontexte: nach vorderen Vokalen, nach Konsonanten und am Morphemanfang.

Nehmen wir nun zuerst an, dass das zugrundeliegende Phonem /ç/ ist. Dann ist es ziemlich einfach, eine Regel zu formulieren, nach der das Phonem /ç/ nach [+hinten] Vokalen als [x] realisiert wird. Zu den [+hinten] Vokalen werden im Deutschen sowohl die hinteren Vokale als auch die zentralen Vokale [a, a:] gezählt, da letztere sich in phonologischen Prozessen wie hintere Vokale verhalten (s. Kap. 3.7). In allen anderen Fällen, die wir nicht weiter spezifizieren müssen, wird das Phonem /ç/ dann einfach als [ç] realisiert. Wenn wir aber umgekehrt annehmen, dass das zugrundeliegende Phonem /x/ ist, so müssten wir sagen, dass das Phonem /x/ in drei ganz verschiedenen Kontexten (nach vorderen Vokalen, nach Konsonanten und am Morphemanfang) als [ç] realisiert wird. Diese Regel wäre also weitaus komplizierter und weniger elegant als die erste. Da wir davon ausgehen müssen, dass unser Gehirn einfache Regeln bevorzugt, werden wir die erste Hypothese wählen, und annehmen, dass es sich bei dem zugrundeliegenden Phonem um /ç/ handelt, wie im Folgenden dargestellt:

(32)

/ç/ ← zugrundeliegende Form

/ \

[ç] [x] ← phonetische Form

Kontext:	Kontext:
-nach vorderen Vokalen	- nach hinteren und zentralen Vokalen
- nach Konsonanten	
- am Morphemanfang	

Beachtet, dass wir eine Argumentation dieser Art bei unserem koreanischen Beispiel nicht anwenden könnten, da dort beide Kontexte (Silbenanfang vs. Silbenende) gleich homogen sind.

Treten wir nun wieder einen Schritt zurück und überlegen wir uns, was wir hier eigentlich beschrieben haben. Wir haben die Distribution des ich-Lautes und des ach-Lautes beschrieben. Irgendwo tauchte dann auch der Name dieses phonologischen Prozesses auf: Assimilation der dorsalen Frikative. Warum sprechen wir in diesem Fall von einer **Assimilation**? Nun, schauen wir uns noch einmal an, was hier geschieht: ein palatales Phonem /ç/ wird zu einem velaren Laut [x], wenn es sich nach einem [+hinten] Vokal befindet. Wir erinnern uns vielleicht, dass das Palatum weiter vorne in der Mundhöhle angesiedelt ist als das Velum. Wenn wir sagen "ein palataler Laut wird zu einem velaren Laut", dann bedeutet das, dass der Laut nach hinten verschoben wird. Nun sollte es uns eigentlich nicht überraschen, dass ein Laut nach hinten verschoben wird, wenn er nach einem [+hinten] Vokal steht. Es handelt sich hier um einen klaren Fall von Assimilation, bei dem der Vokal den auf ihn folgenden Frikativ beeinflusst. Ein hinterer Vokal wie z.B. [u:] wird also den Frikativ [ç] sozusagen "nach hinten ziehen", bis er zu einem velaren [x] wird. Beachtet, dass diese einfache Erklärung nicht möglich ist, wenn wir von einem Phonem /x/ ausgehen. In diesem Fall könnten wir auch sagen, dass das /x/, wenn es nach [-hinten] Vokalen zu stehen kommt, "nach vorne gezogen wird" und als [ç] realisiert wird. Aber es wird viel schwieriger zu erklären, warum /x/ auch nach Konsonanten und am Morphemanfang nach vorne verschoben wird.

In Kap. 3.1 werden wir die Diskussion um ich- und ach-Laut wieder aufnehmen und diesen Prozess in einer phonologischen Regel ausdrücken.

2.2.3 Freie Variation - Nicht-komplementäre Realisierung eines Phonems

Wir haben bis jetzt von allophonischer Verteilung von Phonen gesprochen, also von Fällen, in denen zwei oder mehrere Phone als Realisierungen ein und desselben Phonems zu werten sind und eine komplementäre Distribution haben. Es scheint aber auch Fälle zu geben, in denen die Verteilung bestimmter Phone nicht komplementär, sondern frei ist. Das bedeutet, dass ein Phonem durch mehrere Phone realisiert werden kann, die aber im selben Kontext auftreten können.

Es ist in der Linguistik eine viel diskutierte Frage, ob es echte freie Variation in den Sprachen der Welt überhaupt gibt. Oft hängt das Vorkommen bzw. Nichtvorkommen eines Lautes in einem bestimmten Kontext von außersprachlichen Faktoren wie Dialektzugehörigkeit des Sprechers, Sprechgeschwindigkeit u.ä. ab.

Im Deutschen wird die Distribution der aspirierten Laute am Wortende manchmal als Beispiel für freie Lautvariation angeführt (Ramers&Vater 1995):

(33) to[t] oder to[t^h]

Am Wortende muss ein stimmloser Plosiv nicht aspiriert werden, man hört ihn in diesem Kontext aber auch oft mit Aspiration. Das Auftreten der Aspiration ist in diesem Kontext frei.

Das Phonem /t/ hat also in diesem Fall zwei Realisierungen, [t] und [t^h], die jedoch in ihrer Distribution nicht komplementär sondern frei sind. Wir können das in unseren Transkriptionen ausdrücken, indem wir die Aspiration in diesem Kontext in Klammern setzen: to[$t^{(h)}$]. Wie wir in einem späteren Kapitel sehen werden, ist die Aspiration in anderen Kontexten, wie z.B. dem Wortanfang nicht frei, sondern absolut vorhersagbar, also allophonisch.

Auch in Bezug auf die beiden r-Laute des Deutschen, den uvularen Vibranten [R] und den uvularen Frikativ [ʁ], wird oft von freier Variation gesprochen. So können wir z.B. am Silbenanfang manchmal den einen, manchmal den anderen Laut hören, wie in den folgenden Beispielen:

(34)	[R]aus	oder	[ʁ]aus
	[R] iese	oder	[ʁ]iese
	fah[R] en	oder	fah[ʁ]en
	hö[R] en	oder	hö[ʁ]en

Vielleicht finden wir am Anfang des Wortes etwas öfter den Vibranten [R] und im Wortinneren häufiger den Frikativ [ʁ], vielleicht ist das [R] charakteristisch für ein langsameres Sprechtempo, aber keines dieser Beispiele klingt seltsam oder gar falsch, egal welcher der beiden r-Laute realisiert wurde.

Im Bereich der r-Laute gibt es neben dieser freien Verteilung von [R] und [ʁ] auch noch eine allophonische Distribution, die in Teil 3 unter dem Kapitel zur r-Vokalisierung ausführlich diskutiert wird.

Der alveolare Vibrant [r], der typische r-Laut des Italienischen, ist im Deutschen nicht mehr als eine freie Variante neben [R] und [ʁ] zu werten. Dieser Laut war bis zum

Zweiten Weltkrieg charakteristisch für die Bühnenaussprache. In der Tat kann man ihn noch sehr gut in Chansons aus den zwanziger und dreißiger Jahren von Marlene Dietrich oder Lotte Lenya hören. Auch Adolf Hitler, der sich offensichtlich der Bühnenaussprache bediente, verwendete in seinen Reden den alveolaren Vibranten. Heutzutage findet man das [r] aber nur mehr in einigen bayrischen Dialekten, in der Standardsprache wird es nicht mehr verwendet und die meisten Sprecher des Deutschen sind auch nicht fähig, es zu produzieren.

Fassen wir nun kurz den Inhalt von Kap. 3 zusammen. Wir haben in diesem Abschnitt über die diskreten Einheiten gesprochen, die in der Phonologie eine Rolle spielen. Zuerst wurden die Laute (Phone) im Allgemeinen, ihre Produktionsweise und ihre Klassifikation im Bereich der artikulatorischen Phonetik besprochen. Dann haben wir uns den Phonemen zugewandt, den kleinsten bedeutungsdifferenzierenden Einheiten der Sprache. Wir haben gesehen, dass ein Phonem eine oder mehrere Realisierungen haben kann und dass diese Realisierung allophonisch (also mit komplementärer Verteilung) oder auch frei sein kann. Die freie Distribution von verschiedenen Realisierungen eines Phonems ist in der Linguistik kontrovers. Wir haben einige Beispiele gebracht, von denen angenommen wird, dass sie im Deutschen durch freie Variation charakterisiert sind (uvularer Vibrant vs. uvularer Frikativ, Aspiration am Wortende). Außerdem haben wir in diesem Kapitel gelernt, wie man allophonische Distribution beschreibt. Wir wenden uns jetzt noch weiteren phonologischen Prozessen des Deutschen zu und versuchen, diese anhand von phonologischen Regeln präzise zu beschreiben.

3. Phonologische Prozesse und Regeln

In diesem Kapitel werden die wichtigsten phonologischen Prozesse des Deutschen vorgestellt. Wir werden die Daten untersuchen, die für diese Prozesse relevant sind und Regeln formulieren, um zu beschreiben, was in dem jeweiligen phonologischen Prozess passiert. Dazu werden wir das oben beschriebene Modell verwenden, das davon ausgeht, dass es eine zugrundeliegende Form gibt, die die Sequenz der unvorhersagbaren Elemente, die Phoneme enthält, und dass es eine Oberflächenform gibt, die uns zeigt, wie die einzelnen Phoneme realisiert werden. Eine Theorie, die mit (zugrundeliegenden) Phonemen und Phonen (an der Oberfläche) arbeitet, kann erklären, warum die Struktur der Sprache im Grunde ziemlich ökonomisch ist. Wir müssen uns von einem Wort nicht alle phonetischen Details merken, sondern nur die Phoneme, aus denen es gebildet ist - und die phonologischen Regeln, die für die ganze Sprache gelten. Wenn wir die phonologischen Regeln auf die Phonemsequenzen anwenden, erhalten wir schließlich die phonetische Oberflächenform, die nun tatsächlich ausgesprochen wird. Wir müssen uns also eigentlich gar nicht so viel merken. Wir müssen uns z.B. nicht merken, dass das Wort *Tuch* den ach-Laut enthält, *Tücher* aber den ich-Laut. Wir merken uns einfach, dass sowohl *Tuch* als auch *Tücher* das Phonem /ç/ enthalten. Und wir merken uns die Regel der dorsalen Assimilation der Frikative, die im Fall von *Tuch* das Phonem /ç/ in einen ach-Laut [x] umwandelt. Diese Annahme, dass wir nicht jedes Wort mit all seinen phonetischen Details abspeichern müssen bedeutet einen enormen Gewinn an Speicherplatz für das Gehirn. So wird überhaupt erst erklärbar, warum Kinder innerhalb von wenigen Jahren eine riesige Menge an Wörtern lernen können. Sie müssen nämlich nur die Phonemsequenzen der Wörter und die phonologischen Regeln ihrer Muttersprache lernen, die konkrete Aussprache ergibt sich dann, indem man die Regeln auf die einzelnen Wörter anwendet.

3.1. Assimilation der dorsalen Frikative (ich- und ach-Laut)

Kehren wir hier noch kurz zum Prozess zurück, der die Distribution von ich- und ach-Laut bestimmt. Wir hatten in Kap. 2.2.2 gesehen, dass bei diesem Prozess das Phonem /ç/ nach [+hinten] Vokalen als ach-Laut [x] realisiert wird. Um die Diskussion dieses typischen phonologischen Prozesses des Deutschen abzuschließen, wollen wir nun noch zeigen, wie man ihn in einer phonologischen Regel darstellen kann.

Eine phonologische Regel hat die Struktur

(35) A → B / C _ D

In dieser Formel steht "A" für den input (meistens ein Phonem) und "B" für den output, d.h. das Resultat des phonologischen Prozesses. Im Fall der Assimilation der dorsalen Frikative wäre der input also das Phonem /ç/ und der output das Phon [x]. Der Strich "_" ist ein Platzhalterstrich und "C" und "D" stehen für den Kontext, in dem sich das Phonem befinden muss, damit der phonologische Prozess stattfindet. "C" steht für die Laute, Morphemgrenzen u.s.w., die vor dem Phonem stehen, "D" für das, was hinter dem Phonem steht. Auf dem Platzhalterstrich müsst ihr euch das Phonem selbst vorstellen. Die Regel für die Assimilation der dorsalen Frikative würde also folgendermaßen lauten:

(36) Assimilation der dorsalen Frikative
/ç/ → [x] / [+hinten] _

Wir können die Regel so lesen: das Phonem /ç/ wird als Phon [x] realisiert im Kontext nach [+hinten] (also hinteren und zentralen) Lauten. Warum haben wir nichts für die Variable "D" geschrieben? In diesem Fall war der Kontext *nach* dem Phonem nicht relevant: es ist egal was nach /ç/ kommt. Um zu wissen, wie dieses Phonem realisiert wird, müssen wir wissen, was vor ihm steht, nicht was nach ihm steht. Müssen wir nun auch noch eine Regel schreiben, in der angegeben wird, dass /ç/ zu [ç] wird? Nein, das müssen wir nicht. /ç/ → [ç] ist der sogenannte *elsewhere-case*. Wir nehmen an, dass in allen Fällen, die nicht unter unsere Regel fallen (den *elsewhere*-Fällen), gar nichts passiert und das Phonem keine Veränderung erfährt.

Überlegen wir uns noch, wie die Regel aussehen würde, wenn wir von einem Phonem /x/ ausgingen:

(37) Assimilation der dorsalen Frikative: Hypothese mit Phonem /x/

$$/x/ \rightarrow [ç] \,/\, \begin{matrix}[\text{-hinten}] \\ [\text{Konsonant}] \\ [_{\mu}\end{matrix} \quad _$$

Diese Regel liest sich so: das Phonem /x/ wird als Phon [ç] realisiert im Kontext nach vorderen Lauten, Konsonanten und dem Morphemanfang. Für den Anfang bzw. das Ende einer morphologischen Einheit schreibt man oft eine sich öffnende eckige Klammer "[" (für den Anfang der Einheit) bzw. eine sich schließende eckige Klammer "]" (für das Ende einer Einheit). Man gibt außerdem den Namen der jeweiligen Einheit mit griechischen Buchstaben an, z.B. μ für *Morphem,* ω für *Wort,* σ für *Silbe* u.s.w. Man sieht nun auf Anhieb, dass diese Regel mit ihren drei Kontextangaben komplizierter ist, als die Regel, die wir zuerst formuliert hatten. Es

ist außerdem nicht ersichtlich, was diese drei Kontexte miteinander zu tun haben sollten. Wie in anderen Dingen im Leben, so sollten wir auch in der Linguistik versuchen, unsere Darstellungen so einfach und elegant wie möglich zu halten. Wir werden uns also freuen, wenn es uns gelingt, eine einfache Regel wie in (36) zu formulieren und werden hingegen skeptisch die Augenbrauen runzeln, wenn wir eine Formalisierung wie in (37) finden.

3.2 Vokalisierung des /r/[10]

Übung 9
Lasst euch die folgenden Beispiele von einem deutschen Muttersprachler vorsprechen und versucht herauszufinden, in welchen Wörtern die r-Laute [R] und [ʁ] zu hören sind und wo wir hingegen das vokalisierte [ɐ] hören.

Tür	bitter	rein
Tor	Lehrer	Rolle
mehr	weiter	rufen
Herr	Wetter	Türen
hörte	Seher	Tore
gestorben	sehr	fahren
Hirte	hörst	hören
wurde	Sport	mehrere
zer (wie in "zer-reißen")	Wirt	Lehrerin
ver (wie in "ver-gessen")		

1. Um welche Laute handelt es sich hier? Nennt Artikultationsort und -art (für die Konsonanten) bzw. Zungenposition (für den Vokal):
[R] ist ....
[ʁ] ist ...
[ɐ] ist ...
2. Schreibt neben jedes Wort, ob das /r/, das es enthält, als [R] bzw. [ʁ] oder aber als [ɐ] ausgesprochen wird.
3. Beschreibt in Worten, in welchen Kontexten das /r/ vokalisiert wird (also wann es als [ɐ] realisiert wird). Falls ihr nicht herausfindet, wann das passiert, dann schaut euch noch einmal genau das Beispiel zum Koreanischen im letzten Kapitel an.
4. Konstruiert eine phonologische Regel (Regel der r-Vokalisierung), die das ausdrückt, was ihr unter 3. mit Worten beschrieben habt.
5. Fasst zusammen: welche Realsierungen hat das Phonem /r/ im Deutschen? Welche dieser Realisierungen müssen wir als allophonisch werten und welche nicht?

Wenn ihr euch noch einmal die IPA-Tabelle anschaut, dann werdet ihr sehen, dass es sich bei dem Laut [R] um einen uvularen Vibranten handelt. Der Konsonant [ʁ] ist auch ein uvularer Laut, aber ein Frikativ. Das [ɐ] ist kein Konsonant, sondern ein

[10] Für das Phonem, das den r-Lauten zugrunde liegt, wird in der Phonologie des Deutschen traditionell /r/ geschrieben. Das ist etwas seltsam, da das Phon [r] (der alveolare Vibrant) im Standarddeutschen als Laut praktisch nicht mehr vorkommt. Wir werden in diesem Text jedoch die traditionelle Schreibweise beibehalten.

zentraler Vokal. Sein Artikulationsort ist ähnlich wie der des schwa, aber etwas tiefer (s. Kap. 2.1.4, wo schwa und [ɐ] verglichen werden).
Ihr solltet in den obigen Beispielen folgende r-Laute gehört haben: In der ersten Spalte finden wir Beispiele, bei denen die meisten Sprecher ein vokalisiertes [ɐ] verwenden; es gibt auch Sprecher, die nach kurzen Vokalen, wie in den Beispielen *Herr, Hirte, wurde* kein vokalisiertes [ɐ] sondern [R] oder [ʁ] verwenden; die Präfixe *zer-* und *ver-* werden meist als [tsɛɐ] oder [tsɐ] bzw. [fɛɐ] oder [fɐ], aber jedenfalls mit vokalisiertem [ɐ] ausgesprochen. Die Beispiele in der zweiten Spalte werden auch mit vokalisiertem [ɐ] ausgesprochen, außer - für manche Sprecher - die Beispiele *Sport* und *Wirt*, die einen kurzen Vokal enthalten. Wir können außerdem feststellen, dass in den Beispielen *bitter, Lehrer, weiter, Wetter, Seher* das schwa vor dem finalen /r/ wegfällt. Diese Wörter werden also als [b̥ɪtɐ, le:Rɐ/le:ʁɐ, v̥aɪtɐ, v̥ɛtɐ, ze:ɐ] ausgesprochen. Beachtet, dass demnach die Wörter *Seher* und *sehr* beide als [ze:ɐ] ausgesprochen werden. Der einzige Unterschied zwischen den beiden ist, dass *Seher* aus zwei Silben besteht, während *sehr* einsilbig ist. Die Beispiele in der dritten Spalte enthalten entweder ein [R] oder ein [ʁ]. Die Distribution der beiden Laute ist in diesen Beispielen frei, außer vielleicht die Vorliebe des [R]s für die wortinitiale Position.

In welchen Kontexten wird also das Phonem /r/ vokalisiert? Lassen wir zuerst einmal die Sprecher, die nach kurzen Vokalen das /r/ nicht vokalisieren, beiseite. Für alle anderen Sprecher können wir folgendes beobachten: Am Ende eines Wortes wird ein /r/ auf jeden Fall vokalisiert. Das sehen wir an den Beispielen *Tür, Tor, mehr, Herr, bitter, Lehrer, weiter, Wetter, Seher, sehr*. "Ende des Wortes" genügt aber noch nicht als Kontext für die r-Vokalisierung, denn das Phonem /r/ wird auch in den Beispielen *hörte, gestorben, Hirten, wurde, zer-, ver-* vokalisiert. In diesen Wörtern bzw. Morphemen steht das /r/ nicht am Ende eines Wortes, es steht jedoch am Ende einer anderen linguistischen Einheit, der Silbe. Sollen wir nun also sagen, dass das /r/ "am Ende eines Wortes und am Ende einer Silbe" vokalisiert wird? Nun, diese Regel würde ziemlich genau die angegebenen Beispiele umfassen, aber sie ist nicht besonders elegant. Wie wäre es mit "das Phonem /r/ wird am Ende einer Silbe vokalisiert"? Diese Regel ist viel kürzer und sie umfasst sowohl Beispiele wie *gestorben* als auch Beispiele wie *Tür*, denn das Ende eines Wortes ist immer auch das Ende einer Silbe. Wir haben jetzt aber noch die Beispiele *hörst, Sport* und *Wirt* nicht berücksichtigt. In diesen Beispielen steht das /r/ schon irgendwie am Ende der Silbe, aber wenn man ganz genau hinschaut, dann stehen zwischen dem /r/ und dem Silbenende noch einer oder mehrere Konsonanten. Ich illustriere das am Beispiel *hörst:*

(38) . h ø: ɐ s t .

Wenn man Silbenanfang und Silbenende mit zwei Punkten ausdrückt, dann sieht man, dass sich zwischen dem vokalisierten [ɐ] und dem Silbenende "." die beiden Konsonanten [s] und [t] befinden. Wir können also unsere Regel noch etwas präzisieren: "das Phonem /r/ wird am Silbenende vokalisiert, auch wenn zwischen dem /r/ und dem Silbenende noch einer oder mehrere Konsonanten stehen". Wenn wir nun die Beispiele in der dritten Spalte betrachten, dann können wir sehen, dass unsere Regel stimmen muss, denn in diesen Beispielen steht das /r/ immer am Silbenanfang und folglich wird es nicht vokalisiert.

Wenden wir uns nun noch den Sprechern zu, die das /r/ nach kurzen Vokalen nicht vokalisieren. Für sie gilt unsere Regel nur bedingt, sie muss etwas eingeschränkt werden, z.B. so: "das Phonem /r/ wird nach langen Vokalen am Silbenende

vokalisiert, auch wenn zwischen dem /r/ und dem Silbenende noch einer oder mehrere Konsonanten stehen".

Wir versuchen nun, die Distribution des vokalisierten [ɐ], die wir soeben in Worten ausgedrückt haben, in einer formalen Regel darzustellen. Als input nehmen wir das Phonem /r/, als output geben wir das Resultat des phonologischen Prozesses, das [ɐ] an:

(39) /r/ → [ɐ]

Das Phonem /r/ wird durch das Phon [ɐ] realisiert. Nun müssen wir noch angeben, in welchem Kontext das geschieht. Wir hatten gesagt, "am Ende der Silbe". Den Kontext "Ende der Silbe", können wir durch eine sich schließende eckige Klammer und den griechischen Buchstaben *sigma* ausdrücken: $]_{\sigma}$. Wo steht nun der Platzhalterstrich, der die Position des Phonems angibt? Vor dem Zeichen für das Silbenende oder nach ihm? Wenn wir uns ein Beispiel wie *hörte* [hø:ɐ.tə] anschauen, dann können wir feststellen, dass sich das vokalisierte [ɐ] *vor* dem Silbenende, das durch einen Punkt ausgedrückt wird, befindet. Der Platzhalterstrich steht also vor dem Symbol für das Silbenende:

(40) /r/ → [ɐ]/ __ $]_{\sigma}$

Nun fehlt nur noch ein wichtiger Teil unserer Regel: wir müssen ausdrücken, dass dieser phonologische Prozess auch dann stattfindet, wenn zwischen dem /r/ und dem Silbenende einer oder mehrere Konsonanten stehen. Wir können das darstellen, indem wir zwischen den Platzhalterstrich und das Symbol für das Silbenende das Zeichen "(C)" einfügen. "C" steht für "Konsonant" und die Klammer bedeutet, dass dieser Konsonnt optional auftritt. Es *muss* also kein Konsonant zwischen dem /r/ und dem Silbenende stehen, aber es *kann* ein Konsonant dort stehen. Unsere Regel wird am Ende also folgendermaßen aussehen:

(41) Regel der r-Vokalisierung:
/r/ → [ɐ]/ __ (C) $]_{\sigma}$

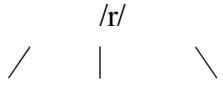
Diese Regel können wir lesen als: "Das Phonem /r/ wird in folgendem Kontext als [ɐ] realisiert: wenn es vor dem Silbenende steht, wobei zwischen /r/ und dem Silbenende ein optionaler Konsonant stehen kann."

Wie müssen wir diese Regel modifizieren, wenn sie auch für diejenigen Sprecher gelten soll, die das /r/ nur nach langen Vokalen vokalisieren? Wir müssen den Kontext "nach langem Vokal" in unsere Regel integrieren:

(42) Regel der r-Vokalisierung: (für einige Sprecher)
/r/ → [ɐ]/ V: __ (C) $]_{\sigma}$

Zusammenfassend können wir über die Distribution der r-Laute im Deutschen folgendes sagen: Es gibt im Deutschen ein Phonem /r/, das drei Realisierungen hat.

(43)



← zugrundeliegende Form

[R] [ʁ] [ɐ] ← phonetische Form

Kontext:	Kontext:
- am Silbenanfang	- am Silbenende (vor optionalen Konsonanten)

Zwei dieser Realisierungen, [R] und [ʁ], kommen im Kontext des Silbenanfangs vor, sind aber untereinander durch freie Variation charakterisiert. Die dritte Realisierung, [ɐ] steht in komplementärer Distribution zu den ersten beiden Realisierungen und kommt am Silbenende vor.

3.3 Nasalassimilation und g-Tilgung

Übung 10

Untersucht die Distribution von [n] und [ŋ] in den folgenden Beispielen:

Hand	[hant]	Bank	[b̥aŋk$^{(h)}$]
ganz	[g̊ants]	sink!	[zɪŋk$^{(h)}$]
Kind	[k^{h}ɪnt$^{(h)}$]	denken	[d̥ɛŋkən]
Wand	[v̥ant$^{(h)}$]	Tango	[t^{h}aŋgo]
Hans	[hans]	Ungar	[ʔuŋgaʁ]
Wunsch	[v̥ʊnʃ]	Ingo	[ʔɪŋgo]
Hanf	[hanf]		
Senf	[sɛnf]		

1. Bestimmt die Artikulationsart und den Artikulationsort von [n] und [ŋ] nach dem IPA.
2. Gebt an, in welchen Kontexten wir [n] und [ŋ] finden.
3. Handelt es sich hier um komplementäre Distribution? Sind [n] und [ŋ] Phoneme oder Allophone?
4. Beschreibt den Prozess, den ihr beobachtet habt, in Worten.
5. Beschreibt den Prozess, den ihr beobachtet habt, mithilfe einer phonologischen Regel, indem ihr davon ausgeht, dass das Phonem /n/ in bestimmten Kontexten als Phon [ŋ] realisiert wird.
6. Was könnte die Funktion dieses Prozesses sein? Warum findet er statt?
7. Wir können [n] vor Frikativen finden, [ŋ] aber nur vor Plosiven. Was könnte der Grund für diese Einschränkung sein?

Bei den beiden Nasalen, die wir in den obigen Beispielen vorfinden, handelt es sich um den alveolaren Nasal [n] und den velaren Nasal [ŋ] (auch *angma* [ɛŋma] genannt). Wenn wir die Beispiele in den beiden Spalten vergleichen, dann können wir beobachten, dass wir [n] vor den alveolaren Plosiven und Frikativen [t, s, ʃ] und vor dem labiodentalen Frikativ [f] finden.[11] Der velare Nasal [ŋ] hingegen tritt nur vor den velaren Plosive [k, g] auf. Bei der Distribution der beiden Laute muss es sich um komplementäre Distribution handeln, denn wir finden nie ein [n] vor einem velaren Laut und nie ein [ŋ] vor einem nicht-velaren Laut. Die Nasale [n] und [ŋ] sind also aller Wahrscheinlichkeit nach Allophone desselben zugrundeliegenden Phonems. Wir gehen davon aus, dass das zugrundeliegende Phonem /n/, und nicht /ŋ/ ist, da wir [n] in einer größeren Anzahl von Kontexten (vor alveolaren Plosiven und Frikativen, vor labiodentalen Frikativen) finden als [ŋ] (nur vor velaren Plosiven).[12] Wir können diesen phonologischen Prozess also folgendermaßen formulieren: das Phonem /n/ wird vor den velaren Plosiven [g] und [k] als velarer Nasal [ŋ] realisiert. Als phonologische Regel können wir den Prozess so darstellen:

---

[11] Vor [f] wird [n], bei schneller Aussprache, auch oft zum labiodentalen Nasal [ɱ] assimiliert ([haɱf, sɛɱf]). Wir ignorieren diese Aussprache hier.

[12] s. Wiese (1996: 218) für eine ausführlichere Diskussion des anzunehmenden zugrundeliegenden Phonems.

(44) Nasalassimilation:
/n/ → [ŋ] / __ [g, k]

Was könnte der Grund für diesen Prozess sein? Wir sehen hier, dass ein Nasal vor einem velaren Laut zu einem velaren Laut wird. In anderen Worten, der Artikulationsort des Nasals und des darauffolgenden Plosivs gleichen sich an. Es handelt sich hier um einen typischen Fall von **Assimilation**, in diesem Fall von Assimilation des Artikulationsortes. Wir haben einen ähnlichen Fall schon in Kap. 2.2.2 beschrieben, wo wir über die Distribution von ich- und ach-Laut gesprochen haben. Assimilierte Strukturen kommen der Faulheit des Sprechers entgegen: er muss die Zunge nicht, wie in [nk], von den Zähnen zum Velum bewegen, sondern er kann sie gleich hinten, beim Velum positionieren. Assimilation bedeutet also eine Ersparnis an Zungenbewegungen. Was gut für den Sprecher ist, ist oft aber nicht gut für den Hörer. Zu viel Assimilation macht es für den Hörer schwierig, die einzelnen Segmente auseinanderzuhalten. In einem Sprachsystem gibt es deshalb zwar oft Assimilationsprozesse, ihnen sind aber auch immer gewisse Grenzen gesetzt.

Warum findet dieser Prozess nur vor Plosiven, nicht aber vor anderen Konsonanten statt? Nun, wenn wir uns das IPA anschauen, dann sehen wir, dass der einzige andere velare Konsonant, den es im Deutschen gibt, der ach-Laut [x] ist. Der ach-Laut ist also der einzige Konsonant, vor dem im Prinzip ein /n/ zu einem velaren [ŋ] werden könnte. Wir wissen jedoch, dass [x] ein Allophon von [ç] ist, das nur nach hinteren und zentralen Vokalen auftaucht. Nach Konsonanten, also auch Nasalen, finden wir immer den ich-Laut [ç]. In anderen Worten, den Kontext [nx], der Assimilation auslösen könnte, werden wir im Deutschen nie finden, da wir [x] nie nach einem Konsonanten finden. Wir wissen also eigentlich gar nicht, ob die Nasalassimilation auf Plosive beschränkt ist, da es einfach keine Kontexte gibt, in denen wir beobachten könnten, was mit [n] vor einem Nicht-Plosiv passiert.

Wir können also eigentlich unsere Regel noch etwas allgemeiner formulieren, indem wir sagen, dass Nasalassimilation im Allgemeinen vor velaren Konsonanten stattfindet. Da der velare Frikativ [x] nie nach [n] vorkommt, kann er auch kein Problem für diese Formulierung sein. Velare Konsonanten sind durch das Merkmal [DORSAL] charakterisiert (s. Kap. 3.7). Wir könnten also z.B. die Regel folgendermaßen formulieren:

(45) Nasalassimilation: allgemeiner formuliert
/n/ → [DORSAL] / __ [DORSAL]

Nun sieht man sehr schön, dass es sich hier um Assimilation handeln muss: ein Nasal wird dorsal in einem dorsalen Kontext.

Bei dem gerade beschriebenen Prozess handelt es sich um eine sogenannte **regressive Assimilation,** da ein Segment ein ihm vorausgehendes Segment beeinflusst:

(46) Regressive Assimilation
n ← k

Es gibt im Deutschen aber auch einen Prozess der **progressiven Nasalassimilation.** Dieser Prozess betrifft die sogenannten **silbischen Nasale**, d.h. Nasale, die den Nukleus einer Silbe bilden. Silbische Nasale entstehen, wenn in einer Flexionsendung das schwa wegfällt und der Nasal die Position des

Silbennukleus besetzt (s. Kap. 4.1). Das passiert eigentlich immer bei einem normalen Sprechtempo. Nur wenn wir langsam und deutlich sprechen, wird das schwa wirklich ausgesprochen. Wir kennzeichnen silbische Konsonanten mit einem vertikalen Strich unter dem IPA-Zeichen, wie im folgenden Beispiel:

(47) <reden>:	Langsame Aussprache	normale bis schnelle Aussprache
	[Re:.dən]	[Re:.dn̩]

Wenn nun der Nasal silbisch ist, dann untergeht er nach einem Plosiv einer progressiven Assimilation, d.h., er passt sich an den vorhergehenden Laut an, wie die folgenden Beispiele zeigen:

(48)	Langsame Aussprache	normale bis schnelle Aussprache
loben	[lo:bən]	[lo:bm̩]
geben	[ge:bən]	[ge:bm̩]
raten	[ra:tən]	[ra:tn̩]
laden	[la:dən]	[la:dn̩]
legen	[le:gən]	[le:gŋ̍]
fragen	[fra:gən]	[fra:gŋ̍]

Wir sehen, dass ein silbischer Nasal nach einem labialen [b] zu einem labialen Nasal [m] und nach einem velaren [g] zu einem velaren Nasal [ŋ] wird. Nach einem dentalen Plosiv wie [t, d] bleibt der Nasal natürlich ein dentales [n].

Bei dieser Assimilation handelt es sich um eine progressive Assimilation, denn der erste Laut, der Plosiv, beeinflusst den auf ihn folgenden Laut, also den Nasal:

(49) Progressive Assimilation
Plosiv → silbischer Nasal

Übung 11

Handelt es sich bei der Assimilation der dorsalen Frikative (ich-/ach-Laut) um eine regressive oder eine progressive Assimilation?

Kehren wir nun zum ersten Assimilationsphänomen, das wir besprochen haben, zurück, zur Assimilation von [n] vor einem velaren Plosiv. Dieser Prozess ist eng mit dem Prozess der sogenannten **g-Tilgung** verbunden. Wenn der velare Nasal vor einem [g] steht, dann gibt es nämlich einige Kontexte, in denen das [g] wegfällt, wie z.B. in

(50) <sing!> [zɪŋ]

Wenn wir uns dieses Beispiel genau anschauen, dann sollte uns eigentlich eine Sache besonders verblüffen: wir haben in *[zɪŋ]* ein assimilierstes [ŋ], obwohl der Auslöser der Assimilation, das [g], gar nicht mehr da ist! Dieses Phänomen nennt man in der Phonologie **Opazität.** In unserem Fall ist die Struktur *[zɪŋ]* opak, weil der Prozess der Nasalassimilation stattgefunden haben muss (immerhin sehen wir ein velares [ŋ]),

aber die Struktur an der Oberfläche so aussieht, als hätte er nicht stattfinden können (wir sehen den Auslöser der Assimilation, das [g], nicht mehr).

Diese verzwickten Fälle löst eine regelbasierte Phonologie mit einer geordneten Reihenfolge von Regeln. Nasalassimilation und g-Tilgung können wir beispielsweise folgendermaßen analysieren:
- zuerst findet ein Prozess der Nasalassimilation statt
- dann findet der Prozess der g-Tilgung statt

In unserem Ebenenmodell können wir das so darstellen:

(51) Ableitung von <sing>:

/zɪng/	← zugrundeliegende Repräsentation (Ebene der Phoneme)
↓	← Nasalassimilation
[zɪŋg]	← Zwischenebene
↓	← g-Tilgung
[zɪŋ]	← Oberflächenrepräsentation (Ebene der Phone)

Wir können uns das Ebenenmodell wie eine Maschine vorstellen. Ganz oben werfen wir die zugrundeliegende, phonematische Form in die Maschine. Diese Form erfährt dann den Prozess der Nasalassimilation, der eine Zwischenform schafft, [zɪŋg]. Beachtet, dass wir diese Zwischenform eigentlich nie sehen, denn das Wort <sing> wird nie als [zɪŋg] ausgesprochen. Auf diese Form wird nun die g-Tilgung angewandt und so erhalten wir am Ende die Oberflächenform [zɪŋ].

Die genaue Umgebung, in der der Prozess der g-Tilgung stattfindet, ist ziemlich kompliziert (s. Wiese 1996:224), aber wir können für unsere Zwecke hier festhalten, dass ein [g] nach einem Nasal getilgt wird, wenn es sich in folgenden Kontexten befindet:
- vor einem schwa [ə] oder vokalisierten [ɐ]
- am Ende eines Morphems, aber nur, wenn danach kein betonter Vokal kommt

So finden wir g-Tilgung also in den folgenden Wörtern, bei denen [g] vor schwa oder vokalisiertem [ɐ] steht:

(52)	Dinge	[d̥íŋə]
	Engel	[ʔéŋəl]
	Inge	[ʔíŋə]
	Finger	[fíŋɐ]

[g] fällt jedoch nicht weg, wenn nach ihm ein "voller" Vokal steht:

(53)	Ingo	[ʔíŋgo]
	Ungarn	[ʔúŋgaɐn]
	Singular	[zíŋgula:ɐ]

In den folgenden Fällen steht [g] am Ende eines Morphems. Das [g] fällt weg, wenn auf die Morphemgrenze ein unbetonter Vokal folgt:

(54)	läng-lich	[léŋlɪç]

Beding-ung	[b̥ədíŋʊŋ]
Diphtong	[d̥ɪftɔ́ŋ]

Es bleibt erhalten, wenn auf die Morphemgrenze ein betonter Vokal folgt:

(55)	fing-íer-en	[fiŋgí:ʁən]
	diphtong-íer-en	[d̥ɪftɔŋgí:ʁən]

Übung 12

Welche Form würden wir für das Wort /dɪng/ erhalten, wenn wir die Ordnung der beiden Regeln "Nasalassimilation" und "g-Tilgung" invertieren würden, d.h., wenn wir zuerst die g-Tilgung anwenden würden und dann die Nasalassimilation? Zeichnet ein Ebenenmodell dieser Ableitung.

Beginnen wir mit der zugrundeliegenden Form /dɪng/. Wenn wir auf diese Form die Regel der g-Tilgung anwenden, dann passiert - nichts! Denn wir hatten gesagt, dass das [g] nach einem velaren Nasal getilgt wird. In der zugrundeliegenden Form haben wir aber keinen velaren Nasal. Nun wenden wir die Regel der Nasalassimilation an. Wir erhalten die Form [dɪŋg]. Was unsere beiden Regel betrifft, so könnten wir hier aufhören. Wir wissen aber, dass es im Deutschen auch noch die Regel der Auslautverhärtung gibt, die stimmhafte Plosive am Ende eines Wortes zu stimmlosen Plosiven macht. Wenn wir diese Regel auf die Form [dɪŋg] anwenden, so erhalten wir [dɪŋk]. Der erste Konsonant des Wortes ist außerdem nicht vollkommen stimmhaft, also wird die endgültige Oberflächenform [d̥ɪŋk] sein.

Was hier wie ein Spiel mit der Anordnung der verschiedenen Regel aussieht, besitzt linguistische Realität. Es gibt norddeutsche Sprecher, die für /dɪng/ nicht [d̥ɪŋ], sondern [d̥ɪŋk] sagen. Diese Aussprache ist typisch nicht allein für bestimmte Dialekte, sondern im Allgemeinen für gewisse norddeutsche Umgangssprachen. Wir haben hier eine einfache und elegante Art gezeigt, wie man diesen Unterschied zwischen norddeutscher und standarddeutscher Aussprache erklären kann: die Sprecher der beiden Aussprachevarianten besitzen die gleichen Regeln, aber die Anordnung dieser Regeln ist in den beiden Varietäten verschieden.

Wir können die "norddeutsche" Ableitung von /dɪng/ schematisch so darstellen:

(56) Ableitung von <Ding>: norddeutsche Varietäten

/dɪng/	← zugrundeliegende Repräsentation (Ebene der Phoneme)
↓	← g-Tilgung (nicht anwendbar)
[dɪng]	← Zwischenebene
↓	← Nasalassimilation
[dɪŋg]	← Zwischenebene
↓	← Auslautverhärtung, teilweise Stimmhaftigkeit
[d̥ɪŋk]	← Oberflächenrepräsentation (Ebene der Phone)

3.4 Aspiration der Plosive

Wir haben schon in vorhergehenden Kapiteln einige Mal bemerkt, dass die stimmlosen Plosive [p, t, k] im Deutschen oft aspiriert werden. Wir wollen uns nun genau anschauen, in welchen Kontexten das geschieht (s. Hall 1992, Ramers&Vater 1995, Wiese 1996 für eine ausführliche Diskussion).

Wir halten zuerst fest, dass nur jene stimmlosen Plosive aspiriert werden, die alleine am Silbenanfang stehen. Wenn auf den stimmlosen Plosiv noch ein weiterer Konsonant folgt oder ihm ein anderer Konsonant vorausgeht, dann finden wir keine Aspiration, wie die folgenden Beispiele zeigen:

(57) Keine Aspiration wenn der Plosiv nicht alleine am Silbenanfang steht

Traube	[tRaʊbə]
Platz	[plats]
stehen	[ʃte:ən]
sparen	[ʃpa:ʁən]

Wir finden eine starke Aspiration des Plosivs, wenn er vor einem betonten Vokal steht:

(58) Aspiration vor betonten Vokalen

Tanz	[t^{h}ánts]
Pech	[p^{h}éç]
getan	[g̊ət^{h}á:n]
Natur	[nat^{h}ú:ɐ]

Am Wortanfang finden wir oft Aspiration, auch wenn der erste Vokal nicht betont ist:

(59)

Talent	[t^{h}alɛ́nt$^{(h)}$]
Pedant	[p^{h}edánt$^{(h)}$]
Kamel	[k^{h}amé:l]

Am Wortende ist Aspiration optional (s. Kap. 2.2.3)

(60)

tot	[t^{h}o:t$^{(h)}$]
lieb	[li:p$^{(h)}$]
lag	[la:k$^{(h)}$]

Zusammenfassend können wir also sagen, dass die stimmlosen Plosive in folgenden Kontexten aspiriert werden:
- vor betonten Vokalen (stark)
- am Wortanfang (weniger stark)
- optional am Wortende

Wie wir schon erwähnt hatten, gibt es die Aspiration in vielen süddeutschen Varietäten überhaupt nicht. Im Englischen werden stimmlose Plosive auf sehr ähnliche Weise wie im Deutschen aspiriert. Im Niederländischen, einer Sprache, die dem Deutschen sehr ähnlich ist, werden die stimmlosen Plosive wiederum nicht aspiriert. Wir hatten außerdem beobachtet, dass die Aspiration im Deutschen - im Gegensatz zu Sprachen wie dem Hindi - nicht distinktiv ist.

3.5 Auslautverhärtung

Übung 13

Untersucht die Alternation der Stimmhaftigkeit in den folgenden Beispielen:

Lobes, Lob	[lo:bəs]	[lo:p$^{(h)}$]
Rades, Rad	[Ra:dəs]	[Ra:t$^{(h)}$]
Tages, Tag	[t^{h}a:gəs]	[t^{h}a:k$^{(h)}$]
aktive, aktiv	[ʔakti:və]	[[ʔakt^{h}i:f]
ewige, ew'ge	[ʔe:vɪgə]	[ʔe:fgə]
Rose, Röschen	[Ro:zə]	[Rø:sçən]
fragen, fragte	[fRa:gən]	[fRa:ktə]
jagen, Jagd	[ja:gən]	[ja:kt]
loben, lobst	[lo:bən]	[lo:pst]

1. Gebt an, welche Laute hier stimmlos werden.
2. Beschreibt in Worten den Kontext, in dem die Laute stimmlos werden.
3. Schreibt eine phonologische Regel, die diesen Prozeses beschreibt.

Wir können an diesen Beispielen beobachten, dass die Phoneme /b, d, g, v, z/ in bestimmten Kontexten als stimmlose Phone [p, t, k, f, s] realisiert werden. Bei diesen Phonemen handelt es sich um die Klasse der Plosive und Frikative. Zusammen bilden sie die Klasse der Obstruenten, die so genannt werden, weil der Luftstrom bei der Produktion dieser Laute teilweise oder ganz behindert wird. Obstruenten haben auch noch eine weitere gemeinsame Charakteristik: in vielen Sprachen der Welt, wie auch im Deutschen, Italienischen und Englischen, kommen sie immer "in Paaren" vor, d.h., es gibt von den meisten Obstruenten sowohl die stimmhafte als auch die stimmlose Version. So gibt es zum [p] seine stimmhafte "Schwester", das [b], zum [f] das [v] u.s.w. Dadurch unterscheiden sich die Obstruenten von den sogenannten **Sonoranten**, die normalerweise nur in ihrer stimmhaften Version auftauchen. Zu den Sonoranten gehören die Nasale, die Vibranten, die Laterale, die Approximanten und die Vokale. Von diesen Lauten sagt man, dass sie spontan stimmhaft sind. Zum Unterschied zwischen Obstruenten und Sonoranten kommen wir im Kap. 3.7 zurück, wo über distinktive Merkmale gesprochen wird.

In welchem Kontext werden nun die Obstruenten stimmlos? Wenn wir die Beispiele *Lob, Rad, Tag, aktiv* betrachten, so könnten wir meinen, dass die Auslautverhärtung das Ende des Wortes betrifft. Aber schon beim Beispiel *ewige, ew'ge* werden wir eines besseren belehrt. Das Wort *ew'ge* stellt eine Zusammenziehung des Wortes *ewige* dar, wie wir sie in lyrischen Texten finden können. Die Zusammenziehung dient dazu, das Wort um eine Silbe zu verkürzen, damit es in die metrische Struktur des Gedichtes passt. Im Wort *ewige* ist der Obstruent [v] stimmhaft, in der Zusammenziehung *ew'ge* hingegen wird er zu einem stimmlosen [f]. Wir müssen uns also fragen, worin der Unterschied zwischen *ewige* und *ew'ge* besteht. Wenn wir genau hinsehen, dann können wir beobachten, dass das [v] in *ewige* am Anfang einer Silbe steht, das [f] in *ew'ge* hingegen am Ende einer Silbe. Wir müssen also schließen, dass die Auslautverhärtung am Ende einer Silbe stattfindet. Es kann auch nicht sein, dass das /v/ am Ende eines Morphems stimmlos wird, denn wir sehen, dass das /v/ weder in *ewig-e* noch in *ew'g-e* am Ende eines Morphems steht. Wir

können nun die anderen Beispiele untersuchen und kontrollieren, ob unsere Hypothese "Auslautverhärtung findet am Silbenende statt" auch für diese Beispiele zutrifft.

Im Beispiel *Röschen* befindet sich das Phonem /z/ auch am Silbenende und wird deshalb als stimmloses [s] realisiert. Dasselbe gilt für das Beispiel *fragte*, wo /g/ am Silbenende als [k] realisiert wird. Im Beispiel *Jagd* sehen wir, dass der stimmhafte Obstruent /g/ nicht am absoluten Silbenende stehen muss, es kann sich zwischen ihm und dem Silbenende auch noch ein Konsonant befinden. Dieser Konsonant wird natürlich auch als stimmlos realisiert, in diesem Fall als ein stimmloses [t]. Eine ähnliche Situation finden wir in *lobst* vor. Das stimmhafte /b/ wird stimmlos, auch wenn sich zwischen dem /b/ und dem Silbenende die beiden Konsonanten [st] befinden.

Wir können also unsere Regel für die Auslautverhärtung folgendermaßen formulieren: Obstruenten werden am Silbenende stimmlos, auch wenn zwischen dem Obstruenten und dem Silbenende noch einer oder mehrere Konsonanten stehen.

Wir versuchen jetzt, diese Beobachtung in einer phonologischen Regel zu formulieren. Beginnen wir damit, dass wir den *input* angeben. Wir könnten z.B. einfach alle betroffenen Phoneme auflisten:

(61) /b, d, g, v, z/ →

Wir würden so jedoch eine Generalisierung übersehen und zwar, dass es sich hier nicht um eine zufällige Liste von Phonemen handelt, sondern um Phoneme, die alle etwas gemeinsam haben. Es handelt sich um eine Klasse von Lauten, die Klasse der Obstruenten. Wir können also für den input den Namen der Klasse einsetzen:[13]

(62) Obstruent →

Was geschieht nun mit den Obstruenten? Sie werden stimmlos:

(63) Obstruent → [-stimmhaft]

Jetzt müssen wir noch den Kontext angeben, in denen dieser Prozess stattfindet, "am Silbenende, auch wenn vor dem Silbenende noch einer oder mehrere Konsonanten stehen". Wir hatten schon beim Prozess der r-Vokalisierung gesehen, dass sich ein Kontext dieser Art folgendermaßen darstellen lässt:

(64) Auslautverhärtung
Obstruent → [-stimmhaft]/ __ (C) $]_{\sigma}$

Wir können nun noch einige Überlegungen zur Formulierung dieser Regel anstellen. Warum haben wir z.B. nicht als input "stimmhafte Obstruenten" angegeben?

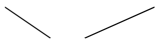
(65) Auslautverhärtung ?
Obstruent → [-stimmhaft]/ __ (C) $]_{\sigma}$
[+stimmhaft]

In den Beispielen, die wir untersucht haben, waren es ja immer stimmhafte Plosive und Frikative, die stimmlos wurden. Andererseits können wir uns jedoch auch fragen,

[13] Eine Darstellung des input in Form von distinktiven Merkmalen werden wir im Kap. 3.7 vornehmen.

ob die Regel ohne den Zusatz [+stimmhaft] für die Obstruenten nicht doch funktioniert. Ohne diesen Zusatz betrifft sie nicht nur stimmhafte, sondern auch stimmlose Obstruenten, wie z.B. ein /p/. Wir würden also sagen, dass ein /p/ am Ende einer Silbe stimmlos wird. Das ist nicht problematisch, da ein /p/ sowieso schon stimmlos ist. Da die Regel also auch ohne den Zusatz [+stimmhaft] funktioniert, werden wir diesen Zusatz weglassen, da eine Regel immer so einfach wie möglich sein sollte. Wir können die Regel in einem gewissen Sinn auch als allgemeines Prinzip formulieren: Obstruenten werden am Silbenende als stimmlos realisiert, egal ob sie im input stimmhaft oder stimmlos sind. Bei diesem phonologischen Prozess handelt es sich um eine **Neutralisierung**, da in einer bestimmten Position (dem Silbenende) der Wert für [±stimmhaft] zu [-stimmhaft] neutralisiert wird. Wir können das in unserem Ebenenmodell am Beispiel der Phoneme /p/ und /b/ folgendermaßen darstellen:

(66) /p/ /b/ ← zugrundeliegende Form



[p] ← phonetische Form

Kontext:
Ende der Silbe

Die Auslautverhärtung ist einer der wichtigsten phonologischen Prozesse des Deutschen. Wie wir schon in Kap. 1 gesagt hatten, verwenden deutsche Sprecher diesen Prozess auch unbewusst, wenn sie andere Sprachen, wie z.B. Englisch sprechen. Durch die Auslautverhärtung kann es dann oft zu echten Verständnisschwierigkeiten kommen, wenn z.B. die finalen Konsonanten in den Wörtern *lice* und *lies* und *laugh* und *love* immer als stimmlos ausgesprochen werden, obwohl sie im Englischen einmal stimmlos, einmal stimmhaft sind. Die Auslautverhärtung gibt es aber nicht nur im Deutschen, sondern auch in anderen Sprachen, wie z.B. im Katalanischen und im Russischen.

Wir wollen jetzt noch ein kleines Gedankenexperiment durchführen, um uns in den Schritten der linguistischen Analyse zu üben. Wir hatten gesagt, dass es sich bei dem Prozess der Auslautverhärtung um einen Prozess handelt, bei dem stimmhafte Laute stimmlos werden. Das ist auch sicher richtig, unsere Daten stimmen mit der Formulierung dieser Regel überein. Nun nehmen wir aber einmal an, dass es genau umgekehrt ist: stimmlose Laute werden zwischen zwei Vokalen stimmhaft. Wenn wir das erste Beipiel aus unserer Datenliste nehmen

(67) Lobes, Lob [lo:bəs] [lo:p$^{(h)}$]

so könnten wir zum Beispiel annehmen, dass das zugrundeliegende Phonem in diesem Fall /p/ ist, wie in [lo:p$^{(h)}$]. Dieses Phonem wird dann stimmhaft, wenn es sich zwischen zwei Vokalen befindet, wie in [lo:bəs].

Was hindert uns daran, so eine Regel der "Verstimmlichung" zwischen Vokalen anzunehmen? Was sagt uns, dass die Regel der Auslautverhärtung richtig ist, jene der "Verstimmlichung" aber falsch? Im Allgemeinen ist es eine gute Idee, sich solche alternative Erklärungen für ein linguistisches Phänomen vorzustellen und dann zu versuchen, eine der beiden Erklärungen zu falsifizieren.

Welche Regel ist also falsch, welche ist richtig? Wenn wir weiterhin auf die Daten starren, die wir bis jetzt verwendet haben, werden wir nicht weiterkommen, für sie

funktionieren beide Regeln. Doch wir hatten in Kap. 1 noch ein paar Beispiele gesehen, die hier interessant sein könnten:

(68) Räte, Rat [Rɛ:tə] [Ra:t$^{(h)}$]

An diesen Beispielen sehen wir, dass unsere Regel der "Verstimmlichung" zwischen Vokalen nicht richtig sein kann, denn dann müsste auch das [t] in *Räte* zu einem [d] werden. Aber die Tatsache, dass wir stimmlose Obstruenten zwischen Vokalen finden können, beweist uns, dass es keinen allgemeinen Prozess gibt, der Obstruenten in dieser Position stimmhaft macht. Wir können also mit gutem Gewissen die Auslautverhärtung als richtige Beschreibung des Phänomens betrachten.

Ich finde übrigens den Begriff *Auslautverhärtung* äußerst vage. Was ist mit "Auslaut" gemeint? Der letzte Laut des Wortes? Des Morphems? Der Silbe? Und was genau bedeutet "Verhärtung"? Im Englischen nennt man diesen Prozess *final devoicing*, im Italienischen *desonorizzazione finale,* was der Sache schon etwas näher kommt.

3.6. Die Phoneme /s/ und /z/

Übung 14

Schaut euch die folgenden Daten zur Distribution von [s] und [z] genau an und beantwortet dann die Fragen.

a.	reißen ~ reisen	['ʁaɪsən] ~ ['ʁaɪzən]
	biegsam	[bi:kzam]
	Stöpsel	[ʃtœpsəl]
b.	See	['ze:]
	so	['zo:]
	sind	['zɪnt]
c.	Stall	['ʃtal]
	Spiel	['ʃpi:l]
	Stein	['ʃtaɪn]
	Smoking	['smokɪŋ]
	Slalom	['sla:lɔm]
d.	Gräser, Gras	['g̊rɛ:.zɐ]['g̊ra:s]
	Rose, Röschen	['ʁo:.zə], ['ʁø:s.çən]

1. Handelt es sich bei [s] und [z] um zwei Phoneme oder Allophone eines einzigen Phonems? Schaut euch zur Beantwortung dieser Frage besonders die Daten in a. gut an.
2. Beschreibt die Distribution von [s] und [z] in Worten. In welchen Kontexten kommt immer nur [s] vor, in welchen Kontexten kommt immer nur [z] vor?
3. Beschreibt den Prozess, der wortinitiale Phone [z] generiert, mit Hilfe einer phonologischen Regel nach dem Schema A→ B/ C_ D
4. Wie ist das im Italienischen? Sucht nach italienischen Wörtern, in denen das <s> in den gleichen Kontexten steht wie wir sie hier im Deutschen haben. Beschreibt, was dort passiert.

Wenn ihr diese Übung in allen ihren Teilen beantwortet habt, dann solltet ihr Folgendes herausgefunden haben.
Die Beispiele in (a) zeigen uns, dass /s/ und /z/ im Deutschen Phoneme sind, denn sie können Minimalpaare wie *reißen* und *reisen* unterscheiden. Auch die Wörter *biegsam* und *Stöpsel* weisen auf einen bedeutungsdifferenzierenden Kontrast hin: in beiden Fällen steht vor dem [z], bzw. [s] ein stimmloser Konsonant und dahinter ein Vokal, der Kontext ist in beiden Fällen also sehr ähnlich, doch im ersten Beispiel haben wir einen stimmhaften, im zweiten Beispiel einen stimmlosen Sibilanten.
Die Beispiele in (b) zeigen uns, dass die Phoneme /s/ und /z/ am Wortanfang vor Vokal jedoch immer als [z] realisiert werden. Wir können dieses Phänomen mit einer phonologischen Regel folgendermaßen ausdrücken:

(69) /s/ → [z]/ [$_{\omega}$ __ V

Diese Regel besagt, dass das Phonem /s/ am Anfang eines Wortes und vor einem Vokal als stimmhaftes [z] realisiert wird. Die offene eckige Klammer mit der Etikette "ω" steht für den Wortanfang, der Platzhalterstrich "__" zeigt uns den Kontext des betreffenden Phonems an und "V" steht für "Vokal". Müssen wir hier noch eine zweite phonologische Regel angeben, die besagt, dass das Phonem /z/ in demselben Kontext auch als [z] realisiert wird, also:

(70) /z/ → [z]/ [$_{\omega}$ __ V

Nein, das ist natürlich nicht notwendig, denn die Realisierung von /z/ als [z] ist der "Normalfall" (im Englischen sagt man, der *default*-Fall). Im Normalfall nehmen wir an, dass ein Phonem ohne phonetische Veränderungen realisiert wird.

Wie bei der Auslautverhärtung, so handelt es sich auch bei diesem phonologischen Prozess um eine Neutralisierung: am Wortanfang vor Vokal werden die beiden Phoneme /s/ und /z/ zu einem einzigen Laut [z] neutralisiert, d.h., sie werden beide mit demselben Laut [z] realisiert:

(71) Neutralisierung der Phoneme /s/ und /z/ zu [z] am Wortanfang vor Vokal:
/s/ ↘
[z]
/z/ ↗

Das bedeutet, dass wir von einem Wort wie ['ze:] gar nicht sagen können, ob es zugrundeliegend das Phonem /s/ oder das Phonem /z/ hat. Beide Phoneme würden auf jeden Fall als [z] realisiert werden.

Die Beispiele in (c) zeigen uns einen weiteren Neutralisierungsprozess. Wir sehen, dass /s/ und /z/ am Wortanfang vor Konsonant als [ʃ] realisiert werden. Das gilt jedenfalls für native Wörter wie *Stall, Spiel, Stein*. Bei Fremdwörtern finden wir in diesem Kontext auch oft die Realisierung [s], wie z.B. in *Smoking* und *Slalom*. Wir können also für den nativen Wortschatz folgende Regel aufstellen:

(72) /s/, /z/ → [ʃ]/ [$_{\omega}$ __ C

Die Regel besagt, dass die Phoneme /s/ und /z/ am Wortanfang vor einem Konsonanten ("C") als [ʃ] realisiert werden. Da es sich um einen Neutralisierungsprozess handelt, wissen wir auch in diesem Fall bei einem Wort wie z.B. *Spiel* nicht, ob zugrundeliegend ein /s/ oder /z/ vorliegt. Es könnte

zugrundeliegend sogar ein [ʃ] vorliegen, auch wenn die Orthographie uns eher an ein /s/ oder /z/ denken lässt. Was aber auch immer das zugrundeliegende Phonem sein mag, alle drei Phoneme können in diesem Kontext im nativen Wortschatz nur als [ʃ] realisiert werden:

(73) Neutralisierung der Phoneme /s/, /z/, /ʃ/ zu [ʃ] am Wortanfang vor Konsonant:
/s/ ↘
/ʃ/ → [ʃ]
/z/ ↗

Die Beispiele in (d) zeigen uns einen Neutralisierungsprozess, den wir schon kennen, die Auslautverhärtung. Das Phonem /z/ wird am Silbenende als stimmloses [s] realisiert.

Vergleichen wir nun die Realisierung von /s/ und /z/ im Deutschen mit der Distribution derselben Laute im Italienischen. Sind [s] und [z] im Italienischen bedeutungsunterscheidende Einheiten, also Phoneme? Bevor wir diese Frage beantworten, müssen wir vorausschicken, dass sich die Realisierung der beiden Laute je nach regionaler Varietät ziemlich unterscheidet. So kann man davon ausgehen, dass /s/ und /z/ im Toskanischen zwei Phoneme sind, wie uns das folgende Minimalpaar zeigt (Canepari 1979):

(74) chie[s]e ~ chie[z]e *fragte (passato remoto), Kirchen (pl.)*

Allerdings gibt es nicht viele Minimalpaare mit /s/ und /z/ . Da das Toskanische oft auch als Maßstab für die italienische Standardaussprache benutzt wird, gehen die meisten Lehrbücher davon aus, dass das Italienische /s/ und /z/ als Phoneme kontrastiert.

In den norditalienischen Varietäten, wie sie z.B. im Veneto gesprochen werden, ist es jedoch so, dass im Wortinneren, zwischen Vokalen, immer ein stimmhaftes [z] erscheint. Betrachtet die folgenden "norditalienischen" Aussprachen:

(75) ca[z]a
ro[z]a
co[z]a
me[z]e

Wenn allerdings vor dem Sibilanten im Wortinneren ein Konsonant steht, dann ist dieser in den Varietäten des Veneto stimmlos:[14]

(76) sen[s]o
sal[s]a

Wenn hinter dem Sibilanten ein konsonant steht, dann können wir einen Assimilationsprozess beobachten: wenn der auf ihn folgende Konsonant stimmlos ist, dann ist auch der Sibilant stimmlos, wenn der auf ihn folgende Konsonant stimmhaft ist, dann ist der Sibilant stimmhaft:

---

[14] In den Varietäten des Trentino ist allerdings ein alveolarer Sibilant, der in der Wortmitte auf einen (stimmhaften) Konsonanten folgt oft stimmhaft. So hört man im Trentino oft die Aussprache *sen[z]o*.

(77) re[s]to
lo[s]co
bi[z]bocce
Le[z]bo
O[z]lo

Wie sieht es nun am Wortanfang aus? Sehen wir uns Beispiele an, in denen auf den Sibilanten ein Vokal folgt:

(78) [s]enza
[s]olo
[s]ale

In allen diesen Fällen ist der betreffende Laut stimmlos.
Und wenn auf den Sibilanten ein Konsonant folgt?

(79) [s]torto
[s]pendere
[s]carto
[z]drammatizzarte
[z]berla
[z]gangherato
[z]lovacchia

Wir sehen hier einen Assimilationsprozess: der Sibilant ist stimmlos, wenn ein stimmloser Konsonant folgt und stimmhaft, wenn ein stimmhafter Konsonant folgt.

Es gibt auch italienische Varietäten, in denen wir in diesem Kontext einen ähnlichen Prozess wie im Deutschen finden. So wird der Sibilant z.B. in den Dialekten der Campania vor einem Konsonanten zu [ʃ] und wir finden Aussprachen wie *[ʃ]porco, [ʃ]catola*.

Und wie sieht es am Wortende aus? Nun, das ist nicht so leicht zu sagen, da es im Italienischen kaum Wörter mit wortfinalen Konsonanten gibt. Aber in den wenigen Wörtern, wo der Sibilant am Ende des Wortes steht, wird er eindeutig stimmlos ausgesprochen:

(80) lapi[s]
ga[s]
rebu[s]

3.7 Phonologische/Phonetische Merkmale

Wir haben jetzt die wichtigsten phonologischen Prozesse des Deutschen besprochen und sie in vielen Fällen mit ähnlichen Prozessen im Italienischen verglichen. Bevor wir uns einem neuen Thema zuwenden, wollen wir uns noch mit dem Begriff des phonetischen, bzw. phonologischen Merkmals beschäftigen.

Der Begriff des phonologischen oder phonetischen Merkmals ist in den einzelnen Kapiteln dieses Buches immer wieder aufgetaucht. In Kap. 2 hatten wir im Zusammenhang mit dem Begriff des Phonems schon von phonologischen Merkmalen gesprochen und betont, dass Phoneme als Bündel von distinktiven Merkmalen zu verstehen sind. Im Kap. 3 sind phonologische Merkmale dann wieder bei der Formulierung der phonologischen Regeln aufgetaucht. In diesem Kapitel

wollen wir nun den Begriff des phonologischen bzw. phonetischen Merkmals präzisieren und eine Liste der distinktiven Merkmale des Deutschen vorstellen.

Die Merkmale (it: *tratti*, engl.: *features*) sind die "Atome", aus denen Phone und Phoneme zusammengesetzt werden. Sie sind somit die kleinsten phonologischen bzw. phonetischen Einheiten, die es gibt. Phonologische/phonetische Merkmale verbinden sich zu Segmenten, Segmente verbinden sich zu Silben (s. Kap. 4) und Silben verbinden sich dann zu weiteren, größeren phonologischen Einheiten, über die wir noch in den nächsten Kapiteln sprechen werden.

Nehmen wir zum Beispiel das Phon [p]. Es kann mithilfe der folgenden phonetischen Merkmale charakterisiert werden

$[p^h]$ = [+ konsonantisch]
[-sonorantisch][15]
[- stimmhaft]
[LABIAL]
[+aspiriert]
u.s.w.

Von diesen phonetischen Merkmalen sind aber nur einige, nicht alle, distinktiv, es können also nur einige dieser Merkmale zu den phonologischen Merkmalen gezählt werden. So sind z.B. [±konsonantisch], [±sonorantisch], [±stimmhaft] und [LABIAL] im Deutschen distinktive, also phonologische Merkmale, [±aspiriert] ist jedoch nicht distinktiv und somit nur ein phonetisches Merkmal.

Distinktive Merkmale können dazu benutzt werden, um Phoneme minimal zu unterscheiden. So unterscheiden sich die beiden Phoneme /m/ und /b/ im Merkmal [±sonorantisch][16], /b/ und /p/ genau im Merkmal [±stimmhaft], /p/ und /t/ im Merkmal [LABIAL].[17]

Außerdem charakterisieren Merkmale **natürliche Klassen** von Lauten, also all jene Mengen von Lauten, auf die phonologische Regeln angewandt werden oder die das Resultat von phonologischen Regeln sind. Natürliche Klassen sind Mengen von Lauten, die sich in phonologischen Prozessen gleich verhalten. So kann man zum Beispiel das Resultat der Auslautverhärtung ganz einfach beschreiben, indem man sagt: alle Obstruenten bekommen das Merkmal [-stimmhaft].

(81) Auslautverhärtung
Obstruent → [-stimmhaft]/ __ (C) $]_\sigma$

Wenn wir keine Merkmale hätten, müssten wir eine ganze Reihe von Regeln schreiben:

(82)
/b/	→	[p] /	___ (C)]	$_\sigma$
/d/	→	[t] /	___ (C)]	$_\sigma$
/g/	→	[k] /	___ (C)]	$_\sigma$
/v/	→	[f] /	___ (C)]	$_\sigma$
/z/	→	[s] /	___ (C)]	$_\sigma$

[15] [-sonorantisch] sind alle Laute, die keine Sonoranten sind, dieses Merkmal charakterisiert also die Klasse der Obstruenten (s. unten).
[16] Und im Merkmal [±nasal].
[17] Und im Merkmal [KORONAL].

Eine Regelansammlung dieser Art würde aber die Generalisierung übersehen, dass in allen diesen fünf Regeln dasselbe passiert: ein stimmhafter Konsonant einer bestimmten Klasse von Lauten (den Obstruenten) verändert ein einziges seiner distinktiven Merkmale, das Merkmal [±stimmhaft]. Es wäre höchst verwunderlich, wenn wir plötzlich einen phonologischen Prozess finden würden, bei dem sich das Resultat nicht durch ein gemeinsames Merkmal charakterisieren ließe und die *output*-Laute also keine natürliche Klasse bilden würden, wie im folgenden fiktiven Beispiel:

(83) Unmöglicher phonologischer Prozess: der *output* des Prozesses bildet keine natürliche Klasse:

/b/	→	[d] /	___ (C)] $_{\sigma}$
/d/	→	[t] /	___ (C)] $_{\sigma}$
/g/	→	[m] /	___ (C)] $_{\sigma}$
/v/	→	[a] /	___ (C)] $_{\sigma}$
/z/	→	[s] /	___ (C)] $_{\sigma}$

Genauso unvorstellbar ist es, dass der *input* des phonologischen Prozesses keine natürliche Klasse bildet, also nicht durch gemeinsame Merkmale charakterisierbar ist, wie z.B. im folgenden hypothetischen Beispiel:

(84) Unmöglicher phonologischer Prozess: der *input* des Prozesses bildet keine natürliche Klasse:

/t/	→	[p] /	___ (C)] $_{\sigma}$
/d/	→	[b] /	___ (C)] $_{\sigma}$
/x/	→	[f] /	___ (C)] $_{\sigma}$
/n/	→	[m] /	___ (C)] $_{\sigma}$

In diesem Beispiel könnte man allenfalls den *output* durch ein gemeinsames Merkmal charakterisieren, da alle *input*-Laute das Merkmal [LABIAL] bekommen, aber der *input* ist eine zufällige Ansammlung von Lauten, ohne ein Merkmal, das sie alle zusammen charakterisieren würde.

Wir werden uns nun jene distinktiven Merkmale anschauen, die im Deutschen eine Rolle spielen. Die Beschreibung der einzelnen Merkmale folgt der in Hall (2000):

Allgemeine Merkmale:

[± konsonantisch]: dieses Merkmal unterteilt die Phoneme in zwei große natürliche Klassen: Vokale und Konsonanten. Alle Laute, die [+konsonantisch] sind (also alle Konsonanten), sind durch eine Verengung des Artikulationsraumes zwischen Larynx und Lippen gekennzeichnet. So wird z.B. ein [p] produziert, indem der Artikulationsraum bei den Lippen geschlossen wird. Laute, die [- konsonantisch] sind (also Vokale), werden ohne Verrengung produziert. Wenn wir z.B. ein [a] produzieren, dann strömt die Luft ohne Verengung durch den Artikulationsraum.

[± sonorantisch]: Dieses Merkmal unterscheidet Obstruenten von Sonoranten. Obstruenten (also Frikative und Plosive) sind [- sonorantisch]. Sonoranten sind [+sonorantisch]. Zur Gruppe der Sonoranten gehören Nasale, Liquide (also r- und l-Laute) und Approximanten (im Deutschen [j]).
Der große Unterschied zwischen Sonoranten und Nicht-Sonoranten besteht darin, dass Sonoranten spontan stimmhaft sind, Nicht-Sonoranten jedoch nicht. Das bedeutet, dass Sonoranten normalerweise immer stimmhaft sind, während es bei den Obstruenten die typischen Paare von stimmhaften und stimmlosen Konsonanten gibt.

Ein Sonorant wie [m], z.B., ist normalerweise stimmhaft. Ein Nicht-Sonorant wie [b] hingegen hat normalerweise eine stimmlose "Schwester", nämlich [p].[18]

Merkmale, die (vor allem) zur Klassifizierung von Konsonanten verwendet werden:

[± stimmhaft]: dieses Merkmal trennt stimmhafte von stimmlosen Lauten. Es unterscheidet also Laute wie /b, d, m, a/ von Lauten wie /p, t, s/.

Merkmale, die die Artikulationsart (der Konsonanten) betreffen:

[± kontinuierlich]: dieses Merkmal ist wichtig, weil es Plosive von Frikativen unterscheidet. Bei Lauten mit dem Merkmal [- kontinuierlich] kommt es bei der Produktion zu einem kompletten Verschluss in der Mundhöhle, es handelt sich dabei also um die Plosive. Frikative hingegen sind [+kontinuierlich], weil bei ihrer Produktion der Luftstrom kontinuierlich durch den Artikulationsraum fließt. Nasale werden normalerweise auch als [- kontinuierlich] klassifiziert, da sie in phonologischen Regeln oft eine natürliche Klasse mit den Plosiven bilden. Außerdem werden auch Nasale mit einem kompletten Verschluss produziert.

[± nasal]: dieses Merkmal unterscheidet nasale Laute von nicht-nasalen (oralen) Lauten.

[± lateral]: dieses Merkmal unterscheidet die Laterale von allen anderen Lauten.

Merkmale, die den Artikulationsort (der Konsonanten) betreffen:

[LABIAL]: dieses Merkmal charakterisiert die labialen Laute wie z.B. /p, b, f, v, m/. Dieses Merkmal ist ein sogenanntes **privatives Merkmal**, das heißt, ein Laut hat entweder das Merkmal [LABIAL] oder er hat es nicht, aber es gibt nicht [+labiale] Laute und [-labiale] Laute. Man nimmt an, dass [LABIAL] ein privatives Merkmal ist, da sich die Laute, die dieses Merkmal besitzen, wie eine natürliche Klasse verhalten, die Merkmale, die es nicht besitzen, bilden jedoch keine natürliche Klasse. Das bedeutet, dass es keine natürliche Klasse mit dem Merkmal [-labial] gibt. Da es keine natürliche Klasse dieser Art gibt, kann es auch das Merkmal [-labial] nicht geben. Mit anderen Worten, labiale Laute wie /p, b, f/ u.s.w. sind manchmal als Gruppe das Ziel eines phonologischen Prozesses (bilden also eine natürliche Klasse). Es gibt aber keinen phonologischen Prozess, der Nicht-Labiale als Gruppe betrifft.

[KORONAL]: dieses Merkmal charakterisiert im Deutschen dentale/alveolare Laute [d, t, s, z, n, l] und postalveolare Laute [ʃ]. Diese Laute werden auch als "koronal" bezeichnet, weil bei ihrer Artikulation der mittlere Teil der Zunge, die *corona* beteiligt ist.
[KORONAL] ist auch ein privatives Merkmal. Entweder besitzt ein Laut dieses Merkmal, oder er besitzt es nicht.

---

[18] Es gibt Sprachen, in denen stimmlose Sonoranten vorkommen, aber es handelt sich dabei meist um allophonische Prozesse, die den Sonoranten in einem bestimmten Kontext stimmlos machen. Es gibt auch Sprachen, die z.B. nur ein [p], aber kein [b] haben. Wenn aber einer von den beiden Lauten im Lautinventar einer Sprache fehlt, dann ist das immer das [b]. Diese Tatsache zeigt uns, dass der "normale", unmarkierte Plosiv das stimmlose [p] sein muss, nicht das stimmhafte [b]. Denn in allen Sprachen finden wir [p], aber nicht in allen finden wir auch [b]. Die unmarkierte Charakteristik von Sonoranten ist also die Stimmhaftigkeit, die unmarkierte Charakteristik von Obstruenten hingegen ist die Stimmlosigkeit.

[± anterior]: dieses Merkmal kann im Deutschen dazu verwendet werden, um die alveolaren Sibilanten [s, z] von dem postalveolaren Sibilanten [ʃ] zu unterscheiden. Dieses Merkmal ist notwendig, da alle diese Sibilanten das Merkmal [KORONAL] besitzen.

[DORSAL]: zu den dorsalen Lauten gehören die palatalen, velaren und uvularen Konsonanten des Deutschen, also [ç, x, k, g, ŋ, R, ʁ]. Diese Laute werden "dorsal" genannt, weil sie unter Beteiligung des *dorsum,* also des hinteren Teils der Zunge, artikuliert werden.
Auch [DORSAL] ist ein privatives Merkmal.

Merkmale, die (vor allem) zur Klassifizierung von Vokalen verwendet werden:

[± rund]: wir verwenden dieses Merkmal, um gerundete Vokale wie /y, œ, u, o/ zu beschreiben.

[± hoch]: wir verwenden dieses Merkmal, um hohe Vokale wie /i, y, u/ von allen anderen Vokalen zu unterscheiden

[± tief]: wir verwenden dieses Merkmal, um tiefe Vokale wie /a/ zu kennzeichnen. N.B.: mittlere Vokale wie /e, o, œ/ sind durch die Merkmale [- hoch] und [-tief] gekennzeichnet

[± gespannt]: dieses Merkmal unterscheidet gespannte Vokale wie /i:, u:, o:/ von ungespannten Vokalen wie /ɪ, ʊ, ɔ/.

[± hinten]: wir verwenden dieses Merkmal, um hintere Vokale wie /u:, o:, ɔ/ von allen anderen Vokalen zu unterscheiden. Man nimmt an, dass auch die tiefen Vokale /a:, a/ das Merkmal [+hinten] besitzen, da sie sich in bestimmten phonologischen Prozessen wie z.B. der Assimilation der dorsalen Frikative wie hintere Vokale verhalten.

Übung 15
Gebt mindestens ein distinktives Merkmal an, das die folgenden Minimalpaare unterscheidet:

reißen ~ reisen	mehre ~ Möhre	mein ~ Bein
Bann ~ dann	sie ~ See	Land ~ Rand
nackt ~ Nacht	Sand ~ sind	
Sünde ~ Sunde	Pein ~ Bein	
Miete ~ Mitte	Pein ~ fein	
kein ~ dein	Reck ~ Rock	

Übung 16

Wir haben die Regel der Auslautverhärtung fogendermaßen formuliert:

Obstruent → [-stimmhaft] / ___ (C)] $_{\sigma}$

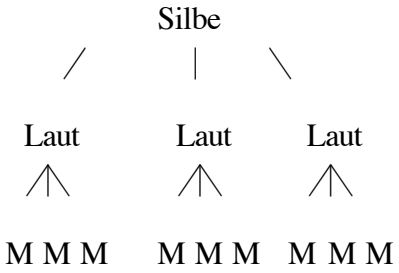
Überlegt euch nun, wie man diese Formulierung verbessern könnte. Wie kann man die natürliche Klasse der Segmente, die von dieser Regel betroffen sind, mit Hilfe von Merkmalen ausdrücken? Drückt den *input* der Regel durch distinktive Merkmale aus.

4. Prosodische Phonologie und ihre Regeln

4.1 Die Silbe

In der Phonologie werden Merkmale zu Lauten gebündelt, aber Laute und Merkmale sind nicht die einzigen phonologischen Einheiten. Laute werden wiederum zu einer größeren Einheit verbunden, der Silbe:

(85)



Die Sprecher einer Sprache haben eine gewisse Intuition, was Silben betrifft. Sie wissen meistens ziemlich genau, wo eine Silbe anfängt und wo sie aufhört, welche Silben in ihrer Sprache wohlgeformt sind und welche hingegen nicht möglich sind. So weiß ein Sprecher des Deutschen, dass *sparf* eine mögliche Silbe des Deutschen sein könnte, *sprf* jedoch nicht.

Diese Tatsachen sollte uns eigentlich ziemlich verwundern, denn es ist praktisch unmöglich, die Silbe einer Sprache zu messen. In einem Spektrogramm können wir also keine Silbengrenzen erkennen. Existiert ein sprachliches Element, das man nicht messen kann, überhaupt? Diese Unsichtbarkeit der Silbe hat einige Linguisten dazu geführt, zu behaupten, dass es die Kategorie der Silbe eigentlich nicht gibt. Wir werden jedoch gleich sehen, dass die Silbe eine große Rolle in unserem Sprachbewusstsein spielt. Wenn die Silbe also auch physikalisch gesehen unsichtbar ist, so scheint sie doch eine mentale Kategorie unserer Sprachkompetenz zu sein.

Dass Sprecher die Kategorie der Silbe besitzen, können wir z.B. beobachten, wenn Kinder einen Abzählreim aufsagen:

(86) Ich-und-du-Mül-lers-Kuh-Mül-lers-E-sel-das-bist-du

Bei jeder Silbe zeigt das Kind auf einen anderen Spieler. Es verwendet also, auch wenn es sich dessen nicht bewusst ist, die Kategorie der Silbe zum Zählen der Mitspieler.

Übung 17

Im Italienischen scheinen Abzählreime etwas anders zu funktionieren. Zerlegt die folgenden Abzählreime in ihre rhythmischen Einheiten und versucht zu beschreiben, wie diese Einheiten aussehen. Ihr könnt auch einen Abzählreim aus eurer Kindheit verwenden:

Uno, due, tre
chi non scappa c'è

Ambarabà, cicì, cocò,
tre galline sul comò,
che facevano l' amore,
con la figlia del dottore,
il dottore s' ammalò,
ambarabà, cicì, cocò

Obwohl die Sprecher einer Sprache meistens ziemlich genau wissen, wo die Silbengrenzen in einem Wort liegen, so gibt es doch einige Zweifelsfälle. Im Deutschen ist es zum Beispiel nicht völlig klar, wie das folgende Wort in Silben zerlegt wird:

(87) <Müller> [mʏlɐ]
a. [mʏl.lɐ] oder
b. [mʏ.lɐ] ?

Wir wissen, dass das Wort "Müller" im Deutschen mit nur einem [l] ausgesprochen wird. Es sollte eigentlich wie in b. getrennt werden. Die Sprecher sind sich allerdings nicht ganz sicher. Einige ziehen es vor, das Wort wie in a. in zwei Silben zu zerlegen, wobei das [l] sowohl die erste Silbe schließt als auch die zweite beginnt. Das hängt vielleicht auch damit zusammen, dass die orthographische Norm die Silbentrennung <Mül.ler> vorschreibt.

Die Phonologen sprechen in diesen Fällen von **ambisilbischen Konsonanten.** Sie gehen davon aus, dass das Wort nur ein Segment [l] enthält, das aber in einem gewissen Sinne sowohl zur ersten als auch zur zweiten Silbe gehört. Man könnte sagen, dass die Silbengrenze in die Mitte des Segments fällt. Ambisilbische Konsonanten tauchen immer dann auf, wenn der Vokal einer Silbe betont und kurz ist und kein anderer Konsonant die Silbe schließt.

Auch in Fällen wie den folgenden ist die Intuition der Sprecher nicht völlig eindeutig:

(88) Fen.ster oder Fens.ter ?

Gei.stes oder Geis.tes ?

Die orthographische Norm schreibt die Trennung *Fen.ster* bzw. *Gei.stes* vor, aber auch in diesen Fällen ist die Intuition der Sprecher nicht ganz klar.

Weitere Evidenz für die Existenz der Silbe können wir darin finden, dass bestimmte phonologische Prozesse als Domäne ihrer Anwendung die Silbe haben. So haben wir z.B. gesehen, dass die r-Vokalisierung und die Auslautverhärtung am Ende einer Silbe stattfinden. Wenn es die Kategorie der Silbe nicht gäbe, wie sollte dann der Anwendungskontext dieser Prozesse beschrieben werden?

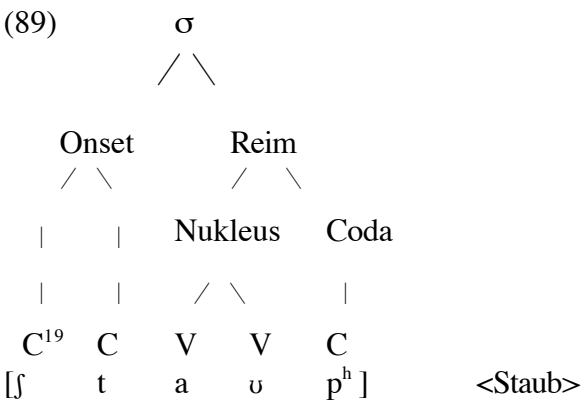
Zusammenfassend können wir sagen, dass es, auch wenn wir Silben physikalisch nicht messen können, doch einige Evidenz für die Existenz der Silbe gibt. Die Silbe scheint Teil unserer sprachlichen Kompetenz zu sein, denn wir haben eine gewisse Intuition, was diese Kategorie betrifft: wir wissen normalerweise, wo eine Silbe anfängt und wo sie aufhört, wie eine wohlgeformte Silbe in unserer Muttersprache aussieht, wir verwenden sie spontan in Abzählreimen. Außerdem können wir beobachten, dass die Silbe den Kontext für viele phonologische Prozesse ausmacht.

Die Silbe selbst besteht aus gewissen Untereinheiten. Diese innere Struktur der Silbe werden wir uns im nächsten Kapitel genauer anschauen.

4.1.1 Die Struktur der Silbe

Eine Silbe ist nicht einfach eine Sequenz von Segmenten. Sie hat eine innere Struktur. Es gibt in der Phonologie mehrere Vorschläge dazu, wie die Struktur der Silbe genau aussehen soll. Wir werden das folgende Modell verwenden:

(89)



Jede Silbe einer natürlichen Sprache besteht aus mindestens einem **Nukleus** (oder Silbengipfel), bei dem es sich meistens um einen Vokal handelt. In allen Sprachen der Welt gibt es außerdem Silben, die einen **Onset** (oder Silbenanlaut) besitzen, d.h. Silben, in denen dem Nukleus einer oder mehrere Konsonanten vorangehen. In manchen Sprachen gibt es außerdem auch Silben, die eine **Coda** (oder Silbenauslaut) besitzen, d.h., in denen auf den Nukleus noch einer oder mehrere Konsonanten folgen.

[19] "C" steht für "Konsonant" und "V" für "Vokal"

Übung 18

Transkribiert folgende Wörter und gebt ihnen eine Silbenstruktur:

Kraft, Nase, Onkel, Obst

Wie kommen aber eigentlich die Linguisten dazu, anzunehmen, dass die Struktur der Silbe genau so und nicht anders aussieht? Warum nimmt man an, dass es eine Kategorie wie den Onset gibt, oder dass Nukleus und Coda zusammen eine Einheit bilden, die wir hier mit *Reim* bezeichnet haben? Im Folgenden werden wir kurz einige Argumente diskutieren, die zu der oben abgebildeten Struktur geführt haben.

Warum nehmen wir an, dass es eine Kategorie *Onset* gibt? Ein gutes Argument für die Existenz der Kategorie *Onset* ist die Tatsache, dass es bestimmte Regeln gibt, die sich genau auf diese Kategorie beziehen. Wenn zum Beispiel im Italienischen ein Onset aus drei Konsonanten besteht, dann muss der erste dieser Konsonanten immer ein [s] sein. Wir können also eine Silbe mit einem Onset wie *spr* in *spre.co* haben, aber es gibt keine Silbe die etwa mit *fpr* beginnt.

Es gibt auch etwas lustigere Evidenz für die Existenz des Onset. In **Versprechern** ist es oft so, dass die Sprecher verschiedene Laute miteinander vertauschen, wie im folgenden Beispiel:[20]

(90) mit dem Zinger feigen <= mit dem Finger zeigen

In diesem Beispiel wurden die beiden Onsets der Wörer *Finger* und *zeigen* vertauscht. Es scheint so zu sein, dass bei Versprechern dieser Art meist Untereinheiten der Silbe miteinander vertauscht werden, also Onsets mit Onsets, Nuklei mit Nuklei, Codas mit Codas, Reime mit Reimen, aber nicht zum Beispiel Onset und Nukleus zusammen, denn diese beiden Elemente bilden keine Einheit!

Warum nehmen wir eigentlich an, dass es eine Katgorie *Reim* gibt, die aus Nukleus und Coda besteht? Ein gutes Argument für die Existenz der Kategorie *Reim* ist die Tatsache, dass diese Kategorie in der Lyrik verwendet wird: der Reim eines Gedichtes besteht meistens aus den Einheiten Nukleus und Coda, wie wir im folgenden Gedicht von Goethe sehen können:

(91) Wer reitet so spät durch Nacht und W**ind**
Es ist der Vater mit seinem K**ind**
Er hat den Knaben wohl in dem **Arm**,
Er faßt ihn sicher, er hält ihn w**arm**.

[20] Aus der Versprechersammlung von Helen Leuninger (http://web.uni-frankfurt.de/fb10/LSLeuninger/pl/pl-neb.htm)

Mein Sohn, was birgst du so bang dein Ges**icht**?
Siehst Vater, du den Erlkönig n**icht**!
Den Erlenkönig mit Kron' und Schw**eif**?
Mein Sohn, es ist ein Nebelstr**eif**.

(Die ersten beiden Strophen des Gedichtes *Erlkönig,* Johann Wolfgang von Goethe, 1749-1832)

Wir sehen hier, dass in zwei aufeinander folgenden Zeilen immer der Nukleus und die Coda der letzten Silbe (*ind, arm, icht, eif*), also der Reim, gleich bleiben.

Für die Existenz der Kategorie *Nukleus* spricht schon alleine die Tatsache, dass diese Einheit die einzige Untereinheit der Silbe ist, die immer realisiert werden muss. Im Deutschen kann der Nukleus durch einen (kurzen oder langen) Vokal, einen Diphthong, oder einen silbischen Konsonanten realisiert werden:

(92)	a. Kurzer Vokal:	W[a].che	
	b. Langer Vokal:	L[i:].be	
	c. Diphthong:	L[aʊ].be	
	d. Silbischer Konsonant:	war.[tn̩]	<warten>
		ne.[bm̩]	<neben>
		Fle.[gl̩]	<Flegel>

Silbische Konsonanten wie in den Beispielen d. werden im Deutschen vor allem bei schneller Aussprache verwendet (die schnelle Aussprache stellt den Normalfall dar, nicht die Ausnahme!). Im Beispiel *ne.[bm̩]* wurde außerdem der Nasal an den vorangehenden Plosiv assimiliert (s. Kap. 3.3). Die schwas werden in diesen Wörtern nur bei sehr sorgfältiger, langsamer Aussprache gesprochen. Dann klingen diese Beispiele folgendermaßen:

(93) war.[tən]
ne.[bən]
Fle.[gəl]

Auch für die Silbencoda gibt es Restriktionen, die uns zeigen, dass es eine Kategorie dieser Art innerhalb der Silbe geben muss. So können z.B. im Italienischen nur bestimmte Konsonanten in der Coda stehen. Ein Nasal in der Coda ist möglich, ein Plosiv hingegen nicht:

(94) a. ba[n].da
bi[m].bo

b. *ba[p].da
*bi[t].bo

Außer den Restriktionen in Bezug auf die Silbencoda gibt es außerdem noch Evidenz aus dem Bereich der **Geheimsprachen** und **Sprachspiele** für die Existenz dieser Kategorie. Schauen wir uns das folgende Gedicht von Joachim Ringelnatz an:

(95) Gedicht in Bi-Sprache

Ibich habibebi dibich,
Lobittebi, sobi liebib.

Habist aubich dubi mibich
Liebib? Neibin, vebirgibib.

Nabih obidebir febirn,
Gobitt seibi dibir gubit.
Meibin Hebirz habit gebirn
Abin dibir gebirubiht.

Joachim Ringelnatz (1883-1934)

Dieses Liebesgedicht ist in der sogenannten Bi-Sprache geschrieben, einer Geheimsprache, wie sie Kinder oft zum Spiel verwenden. Es gibt im Deutschen viele verschiedene Geheimsprachen dieser Art, die man unter den Name Hühnersprache, Löffelsprache, Gabelsprache, Erbsensprache oder Räubersprache finden kann. Wie funktioniert diese Sprache? Hier ist ein Tipp: die ersten Zeilen lauten in der "Übersetzung" folgendermaßen:

(96) Ich habe dich,
Lotte, so lieb

Wie werden die Wörter in dieser Sprache gebildet? Wenn man die erste Zeile "zerlegt", dann bekommt man:

(97) I-bi-ch ha-bi-be-bi di-bi-ch
Lo-bi-tte-bi, so-bi lie-bi-b

Schauen wir uns das Wort *di-bi-ch=dich* genauer an. Es wurde hier die Sequenz *-bi-* hinter den Nukleus und vor die Coda gesetzt. Wenn wir uns die anderen Wörter anschauen, so sehen, wir, dass dort dasselbe passiert: immer kommt die Sequenz *-bi-* genau hinter dem Nukleus und vor der Coda. Am Beispiel *He-bi-rz=Herz* sieht man auch, dass das *-bi-* nicht etwa vor den letzten Konsonanten des Wortes gesetzt wird, sondern vor alle finale Konsonanten, also vor die Coda.

Nun, wie lautet der Rest des Gedichtes?

Dieses Beispiel aus dem Bereich der Geheimsprachen zeigt uns, dass die finalen Konsonanten einer Silbe eine Einheit bilden, die Coda. Denn diese Einheit bestimmt, an welche Stelle die Sequenz *-bi-* gesetzt wird.

Übung 19

In den folgenden vier Zeilen habe ich den Anfang von Ringelnatz' Gedicht in die Hühnersprache übersetzt. Die Hühnersprache ist etwas komplizierter als die Bi-Sprache. Versucht herauskriegen, wie sie funktioniert! Für welche Unterheiten der Silbe liefert uns die Hühnersprache Evidenz?

Ihichlefich hahalefa behelefe dihichlefich
Loholefo ttehelefe soholefo liehieblefieb

Hahastlefast auhauchlefauch duhulefu mihichlefich
liehieblefieb? Neiheinlefein, veherlefer gihiblefib

Übung 20

Der folgende Vers ist ein sogenannter Schüttelreim:

1."Die Boxer aus der Meisterklasse
2. boxten sich zu Kleistermasse
3. Und aus dem ganzen Massenkleister
4. erhebt sich stolz der Klassenmeister"

a) Transkribiert das unterstrichene Wort in Zeile 1

b) Zerlegt dieses Wort in Silben und gebt die Silbenstruktur der einzelnen Silben an (N.B. <ss>, ausgesprochen als [s] ist ein ambisilbisches Segment, d.h. es gehört sowohl zur Coda der vorangehenden Silbe, als auch zum Onset der folgenden)

c) Erklärt das Muster, nach dem die ersten beiden Zeilen (1 und 2) des Schüttelreims gebildet werden. Verwendet dazu die Terminologie der Silbenstruktur.

d) Erklärt, wie das Rezept für den Schüttelreim nach Zeile 2 weitergeht, d.h., wie Zeile 3 aus Zeile 1 und 2 abgeleitet wird. Ist eure Erklärung morphologischer oder phonologischer Art?

4.1.2 Die Sonoritätshierarchie

Wir haben in den letzten beiden Kapiteln gesehen, dass es verschiedene Argumente für die Existenz der Silbe und ihre Untereinheiten Onset, Nukleus, Coda und Reim gibt. Ein weiterer Grund, an die Existenz der Silbe zu glauben, ist, dass es ein Prinzip gibt, das in allen Sprachen der Welt eine große Rolle bei der inneren Organisation der Silbe spielt. In den Sprachen der Welt sind Silben meistens so organisiert, dass die sonorsten Segmente im Zentrum (also im Nukleus) stehen, während die **Sonorität** zum Silbenrand hin abnimmt. Was ist hier mit Sonorität gemeint?[21]

Viele Linguisten nehmen an, dass sich Segmente auf einer Sonoritätsskala anordnen lassen. Wie dieses Skala genau aussieht ist noch nicht klar. Wir werden hier ein Modell nehmen, das von einigen Phonologen für das Deutsche vorgeschlagen wird (Wiese 1988, Vater 1996, Ramers&Vater 1995, Ramers 2001):

(98) **Sonoritätsskala:**
Plosive < Frikative < Nasale < l < r < Vokale

("<" bedeutet: "weniger sonor als)

Nach dieser Skala sind Plosive wenige sonor als Frikative, Frikative sind weniger sonor als Nasale, diese sind wiederum weniger sonor als [l] und [l] ist weniger sonor als [r]. Vokale sind die sonorsten Elemente.

---

[21] Der Begriff *Sonorität* darf nicht mit dem Begriff *Stimmhaftigkeit* (auf italienisch: *sonorità*) verwechselt werden!

Nun kann man beobachten, dass in den Sprachen generell Silben so konstruiert sind, dass die Sonorität bis zum Nukleus hin ständig ansteigt und nach dem Nukleus wiederum bis zum Ende der Silbe hin sinkt. Vor dem Nukleus steht also nie ein sonoreres Element vor einem weniger sonoren Element, nach dem Nukleus ist es genau umgekehrt. Wir formulieren diese Beobachtung im Sonoritätsprinzip:

(99) **Sonoritätsprinzip:**
Die Sonorität nimmt bis zum Nukleus stetig zu und fällt nach dem Nukleus stetig ab

Wir können das am Beispiel von deutschen Silben sehen:

(100) a. Kraft k < r < a > f > t
b. Kerl k < ɛ > r > l
c. Knie K < n < i:

Im ersten Beispiel sehen wir, dass die Elemente im Onset nach dem Sonoritätsprinzip organisiert sind: der Plosiv /k/ ist weniger sonor als /r/, deshalb muss ein /k/ im Onset immer vor einem /r/ stehen, eine Silbe, die mit /rk/ beginnt, ist im Deutschen nicht möglich. In der Coda ist es genau umgekehrt, die Sonorität fällt nach dem Nukleus ständig ab. Deshalb ist die Sequenz /ft/ erlaubt, bei der der sonorere Frikativ vor dem weniger sonoren Plosiv steht. Im zweiten Beispiel, *Kerl*, sehen wir, dass /r/ sonorer als /l/ sein muss, denn /r/ steht in der Coda näher am Nukleus als /l/. Das letzte Beispiel schließlich zeigt uns, dass Plosive weniger sonor als Nasale sind, denn das /k/ steht im Onset vor dem /n/.

Die Sonoritätshierarchie ist nicht unbedingt für alle Sprachen gleich und sie wird auch nicht immer in allen Sprachen befolgt. So gibt es z.B. im Russischen Wörter, bei denen das sonorere /r/ im Onset vor dem weniger sonoren /t/ steht:

(101) rta 'Mund, Genitiv Sg.'

Auch im Deutschen gibt es einige Lautsequenzen, die nicht die Sonoritätshierarchie befolgen. So kann z.B. im Onset vor einem Plosiv noch ein [ʃ] oder ein [s] stehen:

(102) Strumpf [ʃ]trumpf
Skala [s]kala

Sibilanten sind allerdings die einzigen Segmente, die in diesem Kontext stehen können. Wir werden in den nächsten Abschnitten noch sehen, welche Segmente sich im Deutschen in der Coda über die Sonoritätshierarchie hinwegsetzen können.

Trotzdem es Segmentsequenzen gibt, die sich nicht an die Sonoritätshierarchie halten, so kann man doch sagen, dass die Sonoritätshierarchie von den meisten Sprachen wenigstens größtenteils befolgt wird. Es handelt sich also um ein universelles Prinzip.

4.1.3 Universelle Tendenzen: unmarkierte Onsets und Codas

Außer der Befolgung der Sonoritätshierarchie gibt es auch noch andere universelle Tendenzen, die die Struktur der Silbe bestimmen. Schauen wir uns einige Beispiele an, die Hall (2000: 212) entnommen sind:

(103) [vzdwuʂ] 'entlang' Polnisch

	[fspla.nout]	'aufflammen'	Tschechisch
	[vpt͡skvni]	'ich schäle'	Georgisch
(104)	[a.lo.ha]	'Liebe'	Hawaiianisch
	[wa.hi.ne]	'Frau'	Samoanisch
	[du:.xo]	'Hochzeit'	Iraqw

In den ersten drei Beispielen sehen wir Wörter aus den Sprachen Polnisch, Tschechisch und Georgisch, in denen sehr viele Konsonanten im Onset der Silbe erlaubt sind. Nehmen wir z.B. das tschechische Beispiel *fspla.nout:* der Onset der ersten Silbe enthält vier Konsonanten! Die letzten drei Beispiele hingegen stammen aus Sprachen, in denen höchstens ein Konsonant im Onset stehen darf. Wenn wir nun sehr viele Sprachen untersuchen, dann können wir beobachten dass es einerseits Sprachen gibt, in denen nur ein Konsonant im Onset erlaubt ist, und andererseits Sprachen, die mehrere Konsonanten im Onset erlauben. Es gibt aber keine Sprachen, in denen obligatorisch mehrere Konsonanten im Onset stehen müssen! Das bedeutet zugleich, dass jede Sprache, die Silben mit einem komplexen Onset besitzt, auch Silben mit einem einfachen Onset hat. Wir können diese Beobachtung dahingehend interpretieren, dass es anscheinend ein Prinzip gibt, das komplexe Onsets verbietet. Manche Sprachen, wie z.B. Hawaiianisch, Samoanisch und Iraqw befolgen dieses Prinzip, andere Sprachen, wie Polnisch, Tschechisch, Georgisch, befolgen es nicht. In der Linguistik spricht man in diesen Fällen von **markierten** und **unmarkierten** Strukturen. Ein einfacher Onset ist die unmarkierte Struktur, denn wir finden sie in allen Sprachen der Welt, auch in denen, die komplexe Onsets erlauben. Ein komplexer Onset ist die markierte Struktur, denn wir finden sie nicht in allen Sprachen, sondern nur in Sprachen wie Polnisch, Tschechisch, Georgisch.

In Bezug auf den Onset können wir eine weitere Beobachtung machen. Alle Sprachen der Welt erlauben Silben mit einem Onset. In einigen Sprachen gibt es auch Silben ohne Onset (z.B. im Hawaiianischen, wie uns die erste Silbe in *a.lo.ha* zeigt). In anderen Sprachen hingegen sind Silben ohne Onset verboten. Es gibt aber keine Sprache der Welt, in denen Silben *mit* einem Onset nicht möglich sind. Wir können also hier wieder sagen: die unmarkierte Silbe ist eine Silbe, die einen Onset besitzt; die markierte Silbe ist die Silbe ohne Onset.

Es ist nicht ganz einfach, den Begriff *Markiertheit* zu erklären. Unter unmarkierten Strukturen versteht man in der Linguistik Strukturen, die im Normalfall vorkommen. Markierte Strukturen kommen hingegen nur in besonderen Kontexten vor. Unter unmarkierter Struktur verstehen wir hier eine Struktur, die in allen Sprachen der Welt möglich ist und in manchen Sprachen der Welt sogar ausdrücklich verlangt wird. Eine unmarkierte Silbe ist also somit eine wohlgeformte Silbe in allen Sprachen der Welt und manche Sprachen erlauben nur Silben dieser Art.

Wir fassen die beiden Beobachtungen zur unmarkierten Silbenstruktur in der Onset-Beschränkung zusammen (s. auch das Silbenanlautgesetz in Hall 2000: 213):

(105) Onset-Beschränkung, Teil I: Unmarkierte Silben haben einen Onset
Onset-Beschränkung, Teil II: Unmarkierte Silben haben nur einen Konsonanten im Onset

Wie verhält es sich nun mit der Coda? Welche universelle Tendenzen können wir hier beobachten? Schauen wir uns wieder ein paar Beispiele an (s. Hall 2000: 214 und die dort angegebenen Quellen):

(106)	[o:psts]	'Obsts, Genitiv'	Deutsch

	[t͡sɛhɛjəlɛwtxʷs]	'ihre Kirche'	Upriver Halkomelem
	[χosgwd͡ʒ]	'ich befahl'	Svan
(107)	[a.lo.ha]	'Liebe'	Hawaiianisch
	[wa.hi.ne]	'Frau'	Samoanisch
	[a.pá]	'Dach'	Ngiti

Die ersten drei Beispiele stammen aus Sprachen, in denen ziemlich lange Codas möglich sind. So haben wir im Beispiel aus dem Deutschen eine Coda, die aus vier Konsonanten besteht.[22] Im Svan gibt es sogar Silben, die fünf Konsonanten in der Coda haben können (die letzten beiden Konsonanten sind allerdings eigentlich nur ein Segment, eine Affrikate). Dann gibt es aber auch Sprachen, in denen überhaupt keine Silbencodas erlaubt sind. Die letzten drei Beispiele stammen aus Sprachen dieser Art: keine Silbe hat einen Konsonanten in der Coda.

Wir können also hier wieder eine universelle Tendenz feststellen. In allen Sprachen der Welt sind Silben ohne Codas erlaubt. In manchen Sprachen der Welt (wie im Hawaiianischen, Samoanischen und Ngiti), sind nur Silben ohne Coda erlaubt. Es gibt aber keine Sprache der Welt, in der Silben ohne Codas verboten sind.

Die unmarkierte Silbe scheint also eine Silbe zu sein, die nicht auf einen Konsonanten endet. Markierte Silben hingegen haben einen Konsonanten in der Coda.

Es gibt außerdem Sprachen, wie z.B. das Italienische, in denen nur ein Konsonant in der Coda stehen darf. Wir können also wie beim Onset-Prinzip beobachten, dass komplexe Codas markierter als einfache Codas sind.

Wir können diese Beobachtungen wiederum in zwei Beschränkungen formulieren:[23]

(108) Coda- Beschränkung, Teil I: Unmarkierte Silben haben keine Coda
Coda- Beschränkung, Teil II: Codas mit nur einem Konsonanten sind unmarkierter als Codas mit mehreren Konsonanten

Wenn wir nun Onset- und Coda-Beschränkung zusammen betrachten, dann müssen wir den Schluss ziehen, dass die universell unmarkierte Silbe eine CV-Silbe ist, also eine Silbe, die nur aus einem Konsonanten und einem Vokal besteht. In der Tat ist diese Silbe die einzige Silbe, die in allen Sprachen der Welt vorkommt!

Es gibt eine weitere Beobachtung, die darauf hinweist, dass es sich bei der CV-Silbe um die universell am wenigsten markierte Silbe handelt. Wenn Kinder anfangen, die ersten Wörter von sich zu geben, dann verkürzen sie diese Wörter gerne auf eine CV-Silbe. *Buch* wird dann zu *bu* und *Elefant* wird zu *fa*. Erst nach diesem CV-Stadium kommt dann ein weiteres Stadium, in dem Kinder auch Silbencodas zulassen. Und

---

[22] Wenn wir die Sprachen, aus denen diese Beispiele stammen, allerdings etwas genauer untersuchen, dann sehen wir, dass extrem komplexe Codas oft nur am Wortende auftauchen können. Das gilt auch für das Deutsche. Eine Coda, die aus der Sequenz [psts] besteht, finden wir zwar in wortfinalen Codas, nicht aber im Inneren eines Wortes. Dasselbe gilt für komplexe Onsets: am Anfang eines Wortes finden wir oft sehr komplexe Onsets, die im Inneren des Wortes nicht erlaubt sind.

[23] Für eine ausführliche Diskussion dieser universellen Prinzipien s. auch Jakobson (1962) und Prince&Smolensky (1993/2004)

erst nach diesem zweiten Stadium kommt ein drittes Stadium, in dem Kinder schließlich auch komplexe Onsets und Codas aussprechen:

(109) Stadien im Spracherwerb:

1. Stadium:	CV	*fa*
2. Stadium:	CVC	*fan*
3. Stadium:	CCVCC	*fant*

Wenn man es sich genau überlegt, dann ist dieser Spracherwerb in Stadien ziemlich überraschend, denn eigentlich können die Kinder die meisten Buchstaben schon aussprechen. Warum sagt das Kind *fa*, wenn es doch das *n* und das *t* der Coda schon aussprechen kann? Der Grund liegt darin, dass das Kind in seinen phonologischen Entwicklungsstadien generell mit unmarkierten Strukturen beginnt und erst nach und nach die markierteren Strukturen erwirbt.

Ein letzter Hinweis darauf, dass eine wohlgeformte Silbe einen Onset aber keine Coda besitzen sollte, bildet der Mechanismus der Silbifizierung. Wenn wir eine Sequenz von Lauten vor uns haben, dann gibt es im Prinzip mehrere Möglichkeiten, wie diese Sequenz silbifiziert werden könnte. Nehmen wir z.B.

(110) Amerika: a. A.me.ri.ka
b. *Am.er.ik.a

Es ist klar, dass ein Wort wie *Amerika* immer wie in a. silbifiziert werden wird und nie wie in b. Das heißt, die Silbifizierung wird immer so vor sich gehen, dass jede Silbe, wenn möglich, einen Onset hat, aber keine Coda.

Der Mechanismus der Silbifizierung ist allerdings noch etwas komplizierter. Generell wird eine Lautsequenz so silbifiziert, dass der Onset maximiert wird. Das bedeutet, dass die Konsonanten, soweit möglich, alle einem Onset zugeordnet werden. Nur wenn ein Konsonant im Onset keine wohlgeformte Struktur mehr ergeben würde, wird er einer Coda zugeordnet. Wir sehen das im folgenden Beispiel:

(111) Andreas: a. *A.ndre.as
b. An.dre.as
c. *And.re.as
d. *Andr.e.as

In a. wurde versucht, die Konsonanten /n, d/ und /r/ alle in den Onset der zweiten Silbe zu stecken. Das ist nicht möglich, denn die Sequenz /ndr/ ergibt keinen wohlgeformten Onset, da die Sonoritätshierarchie verletzt werden würde. Deshalb muss *n* in der Coda bleiben. Die anderen beiden Konsonanten können jedoch dem Onset zugeordnet werden, wie Beispiel b. zeigt. Beispiel c. und d. sind nicht wohlgeformt, da hier der Onset nicht maximiert wurde. In Beispiel c. wäre die erste Silbe *And* durchaus eine wohlgeformte Silbe, doch diese Silbifizierung wird nicht gewählt, da das *d* auch im Onset der darauffolgenden Silbe stehen kann (und also dort stehen muss).

4.1.4 Die deutsche und die italienische Silbe und ihre Restriktionen

In diesem Kapitel wollen wir uns die Silbenstruktur des Deutschen und die Silbenstruktur des Italienischen anschauen und sie miteinander vergleichen. Wie sieht der deutsche Onset aus, wie sieht der italienische Onset aus? Gibt es Ähnlichkeiten oder Unterschiede was die Segmente betrifft, die in den Untereinheiten

der Silbe auftreten können? Am Ende des Kapitels wollen wir uns einem Problem zuwenden, das jedem, der Deutsch und Italienisch spricht, sofort auffällt: deutsche Wörter sind kürzer als italienische Wörter.[24] Wir werden sehen, dass unsere kontrastiven Beobachtungen zur Silbenstruktur uns helfen können, diesen Unterschied zu erklären.

Beginnen wir mit dem Silbenonset im Deutschen. Der Onset der Silbe des Deutschen besteht aus höchstens drei, fast immer aber aus mindestens einem Konsonanten.

Wenn wir im Onset drei Konsonanten haben, dann ist der erste immer ein [s] oder ein [ʃ]. Solche maximale Onsets finden wir in der ersten Silbe der folgenden Beispiele:

(112)	[ʃ]trumpf	<Strumpf>
	[ʃ]prache	<Sprache>
	[ʃ]plitter	<Splitter>
	[ʃ]pray oder [s]pray	<Spray>

Das Graphem <s> wird am Wortanfang vor einem Konsonanten meistens als [ʃ] ausgesprochen. Nur in einigen Fremdwörtern wie *Spray, Skat, Slalom, Smoking* wird es als [s] ausgesprochen.

Da der dritte Konsonant in einem Onset immer nur ein [s] oder ein [ʃ] sein kann, haben einige Phonologen angenommen, dass der deutsche Onset eigentlich höchstens aus zwei Konsonanten bestehen kann, und dass es sich bei dem [s] oder [ʃ] um ein sogenanntes **extrasilbisches Segment** handelt (s. Wiese 1996). Es gibt auch noch einen weiteren Grund, um das anzunehmen: wenn das [s] oder [ʃ] wirklich Teil des Onset wäre, dann würde es die Sonoritätshierarchie verletzen, da es als Frikativ vor einem Plosiv stehen kann.

Es gibt noch einige andere Beispiele, in denen es so aussieht, als gäbe es drei Konsonanten im Onset:

(113)	[pfl]aume	<Pflaume>
	[pfr]opfen	<Pfropfen>
	[tsv]eig	<Zweig>
	[tsv]anzig	<zwanzig>

Wenn man sich diese Wörter allerdings genauer ansieht, dann sieht man, dass es sich in allen Fällen um Sequenzen handelt, in denen am Wortanfang eine Affrikate [pf] oder [ts] steht. Die Tatsache, dass Affrikaten hier in einem Kontext stehen, in dem normalerweise nur ein einzelner Konsonant vorkommt (also im Onset vor einem anderen Konsonanten) ist ein gutes Argument dafür, dass es sich bei Affrikaten im Grunde um ein Einzelsegment handelt, nicht um zwei unterschiedliche Segmente.

Wenn es in der zugrundeliegenden Form am Anfang einer Silbe keinen Konsonanten gibt, dann wird ein glottaler Plosiv [ʔ] eingesetzt. So bekommt auch eine Silbe, die

---

[24] Ich meine hier morphologisch einfache Wörter. Natürlich hat das Deutsche auch ziemlich lange Wörter, wenn man morphologisch komplexe Wörter in Betracht zieht, da das Deutsche sehr lange und komplexe Komposita zulässt.

sonst mit einen Vokal beginnen würde einen Onset. Wir hören den glottalen Plosiv besonders gut in einem Satz, in dem alle Wörter mit einem Vokal beginnen:

(114) [ʔ]Anton [ʔ]aß [ʔ]am [ʔ]Abend [ʔ]immer [ʔ]Auflauf

Aber nicht nur am Wortanfang finden wir den glottalen Plosiv. Er wird von den meisten Sprechern auch vor eine vokalinitiale Wurzel gesetzt:

(115) ver-[ʔ]arbeiten
ver-[ʔ]achten
be-[ʔ]eilen
[ʔ]auf-[ʔ]essen

Sprecher des Italienischen müssen bei der Aussprache dieser Wörter besonders aufpassen. Das Italienische hat keinen glottalen Plosiv, deshalb neigen Sprecher des Italienischen dazu, in diesen Wörtern die Coda des Präfixes als Onset für die Wurzelsilbe zu verwenden. Diese Aussprache klingt aber für ein deutsches Ohr ziemlich seltsam:

(116) Italienische Aussprache vokalinitialer Wurzeln:
*ve.rar.bei.ten
*ve.rach.ten
*au.fes.sen

Es gibt noch einen Kontext, in dem wir den glottalen Plosiv finden. Auch im Inneren eines Morphems kann es vorkommen, dass eine Silbe mit einem Vokal beginnt. Das geschieht dann, wenn zwei Vokale aufeinander treffen und einen sogenannten **Hiatus** bilden:

(117) The.[ʔ]á.ter
cha.[ʔ]ó.tisch
Be.[ʔ]á.te
in.di.vi.du.[ʔ]éll

Süddeutsche Sprecher produzieren in diesem Kontext eigentlich selten einen glottalen Plosiv, aber wer nicht aus dem Süden kommt, wird in diesen Wörtern wahrscheinlich diesen epenthetischen Konsonanten einsetzen.

Beachtet, dass in den letzten Beispielen der Vokal, vor dem der glottale Verschluss steht, immer betont ist. Es gibt nämlich einen Kontext, in dem morphemintern normalerweise kein glottaler Plosiv produziert wird, auch wenn die Silbe mit einem Vokal beginnt. Das geschieht, wenn der silbeninitiale Vokal nicht betont ist:

(118)	[zé:.ən]	*[ze:.ʔən]	sehen
	[ʔé:.ə]	*[ʔe:.ʔə]	Ehe
	[bé:.a]	*[bé:.ʔa]	Bea
	[ká:ɔs]	*[ká:ʔɔs]	Chaos[25]

Der glottale Plosiv steht also bei vokalinitialen Silben in folgenden Kontexten:

---

[25] Es gibt Sprecher, die auch in *Chaos* einen glottalen Verschluss verwenden. Es könnte sein, dass auf der zweiten Silbe dieses Wortes, *-os-*, ein (schwächerer) Nebenakzent liegt, und dass einige Sprecher glottale Plosive auch vor Vokale setzen, die einen Nebenakzent tragen.

- am Wortanfang
- am Anfang einer Wurzel
- morphemintern, wenn der darauffolgende Vokal betont ist

Wir sehen, auch wenn es so aussieht, als gäbe es im Deutschen Silben ohne Onset, so müssen wir doch schließen, dass das in den meisten Fällen nicht so ist, da in diesen Fällen im Onset ein glottaler Plosiv steht.

Der Onset des Deutschen richtet sich im Großen und Ganzen nach der Sonoritätshierarchie. Allerdings finden wir nicht alle Sequenzen, die nach der Sonoritätshierarchie erlaubt wären. Die folgenden Sequenzen müssten eigentlich möglich sein, es gibt aber keine Onsets dieser Art im Deutschen:

(119) *kfa
*mra
*tla

Wir haben schon weiter oben gesehen, dass der Silbennukleus im Deutschen aus einem kurzen oder langen Vokal, einem Diphthong oder einem silbischen Konsonanten bestehen kann:

(120)	a. Kurzer Vokal:	W[a].che	Wache
	b. Langer Vokal:	L[i:].be	Liebe
	c. Diphthong:	L[aʊ].be	Laube
	d. Silbischer Konsonant:	war.[tn̩]	warten
		ne.[bm̩]	neben
		Fle.[gl̩]	Flegel

In der Coda stehen im Deutschen im Allgemeinen maximal zwei Konsonanten, deren Abfolge sich wieder nach den Prinzipien der Sonoritätshierarchie richtet:

(121) Werk
halb
Hand
Helm

Es gibt allerdings Wörter, in denen stehen nach dem Nukleus mehr als zwei Konsonanten (s. Wiese 1996: 48):

(122) Obst
Markt
Herbst
Werft

Es ist interessant, dass der dritte und vierte Konsonant in diesen Fällen immer ein [s] oder ein [t] ist. Es können also nur alveolare Konsonanten als dritte und vierte Konsonanten in der Coda stehen. Da die Art der Segmente, die nach dem zweiten Konsonanten in der Coda stehen können so beschränkt ist, wurde auch hier vorgeschlagen, dass diese alveolaren Segmente als extrasilbisch zu werten sind.

Was die Affrikaten betrifft, so können wir auch in der Coda eine interessante Beobachtung machen. Schauen wir uns die folgenden Beispiele an:

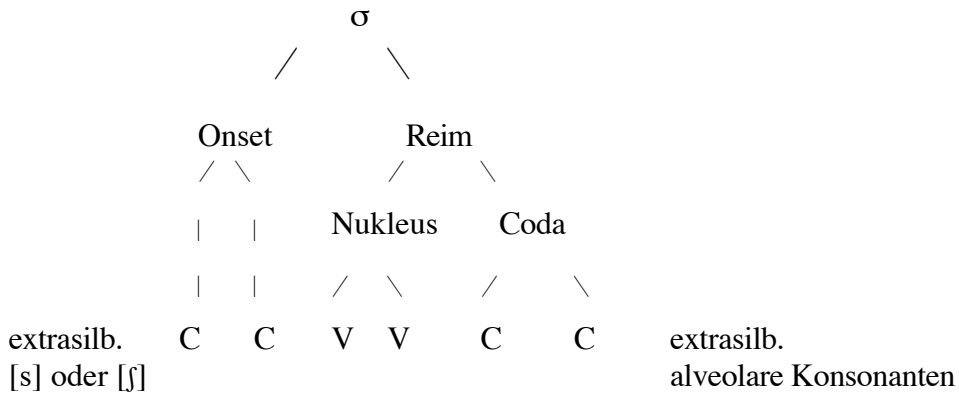
(123) Pro[pf] Pfropf

Ko[pf]	Kopf
Wi[ts]	Witz
Sa[ts]	Satz

Die Affrikaten [pf] und [ts] bestehen aus einem Plosiv-Teil ([p], bzw. [t]) und einem Frikativ-Teil ([f], bzw. [s]). Wenn diese beiden Teile als je ein Segment zählen würden, dann würden [pf] und [ts] in der Coda die Sonoritätshierarchie verletzen. Da wir jedoch diese Affrikaten, wie in den obigen Beispielen, in der Silbencoda finden, müssen wir annehmen, dass Affrikaten nur als ein Segment zählen. Wir haben also ein weiteres Argument für den monosegmentalen Status der Affrikaten gefunden.

Fassen wir noch einmal die gerade genannten Generalisierungen zur deutschen Silbe zusammen. Die maximale deutsche Silbe kann man schematisch folgendermaßen darstellen:

(124) Die maximale deutsche Silbe:



Die maximale Silbe im Deutschen besteht also aus zwei Onset-Konsonanten, einem langen Vokal oder Diphthong im Nukleus und zwei Coda-Konsonanten. Dazu können noch extrasilbische Sibilanten am Silbenanfang oder alveolare Konsonanten am Silbenende dazukommen.

Die minimale Silbe im Deutschen hat mindestens einen Nukleus und (außer in einigen wenigen Fällen wie die zweite Silbe in *sehen*) eigentlich auch immer einen Onset, der mindestens aus einem epenthetischen glottalen Plosiv besteht.

Wenden wir uns nun der italienischen Silbe zu. Wir wollen versuchen, die Charakteristiken der italienischen Silbenstruktur anhand einer Übung selbst zu erarbeiten. Wir können uns dabei an der Beschreibung der deutschen Silbenstruktur orientieren. Die italienischen Daten und Generalisierungen stammen weitgehend aus Nespor (1993).

Übung 21

1. Als erstes wollen wir herausfinden, wie der Onset der italienischen Silbe aussieht. Sehen wir uns dazu die folgenden Daten an:

(125)

[pr]anzo	[str]ano
[kl]ima	[spr]uzzo
[tr]eno	[skr]ivere

[gl]icine [zgr]adevole
a.[tl]e.ta

Beantwortet nun die folgenden Fragen:
a. Wieviele Konsonanten können im Italienischen maximal im Onset stehen?
b. Muss mindestens ein Konsonant im Onset stehen?
c. Welche Konsonanten können bei einem maximalen Onset in der ersten Position stehen, welche in der zweiten, welche in der dritten?
d. Welche Gemeinsamkeiten bzw. Unterschiede gibt es zum Onset des Deutschen?
e. Befolgen die Onsets des Italienischen die Sonoritätshierarchie?

2. Wenden wir uns nun dem Nukleus zu:

(126) [a].ra.bo
[au].la
p[ie].no

a. Welche Segmente können im Italienischen die Nukleusposition besetzen?
b. Welche Unterschiede zum Deutschen lassen sich hier feststellen?

3. Und nun zur Coda:

(127) pa[r].te fa[t].to
a[l].to le[g].go
ba[n].da ma[s].sa
ba[m].bi.no a[v].vi.so

pa[s].ta oder pa.[s]ta?

a. Wieviele Konsonanten können im Italienischen maximal in der Coda stehen?
b. Welche Art von Konsonanten kann in der Coda stehen?
c. Welche Unterschiede gibt es zwischen der italienischen und der deutschen Silbencoda?

Wenn ihr diese Übung in allen ihren Teilen ausgearbeitet habt, dann solltet ihr die folgenden Dinge beobachtet haben. Der italienische Onset gleicht in großen Zügen dem deutschen Onset. Auch im Italienischen haben wir maximal zwei Konsonanten, die die Onset-Position besetzen können, plus einen dritten Konsonanten, der aber dann ein /s/ oder /z/ sein muss. Diese Sibilanten könnten genauso wie im Deutschen als extrasilbisch gewertet werden. Das erste Element des maximalen Onset ist also immer ein /s/ oder /z/. Das zweite Element kann ein beliebiger Konsonant sein. Das dritte Element – und hier unterscheidet sich das Italienische wieder vom Deutschen – ist immer ein Liquid, also ein /r/ oder ein /l/. Es gibt noch einen Unterschied zum Deutschen. Während wir im Deutschen vor vokalinitiale Silben meistens einen glottalen Plosiv setzen, tun wir das im Italienischen nicht. Im Italienischen gibt es also häufiger Silben ohne Onset als im Deutschen.
Schließlich können wir noch beobachten, dass es im Italienischen Onsets gibt, die es im Deutschen nicht gibt, wie z.B. den Onset [tl] im Wort *a.tle.ta*.

Der Nukleus des Italienischen ist dem des Deutschen ziemlich ähnlich. In beiden Sprachen kann er aus einem einfachen Vokal oder auch aus einem Diphthong bestehen. Im Unterschied zum Deutschen können allerdings im Nukleus einer italienischen Silbe keine silbischen Konsonanten stehen.

In der Coda der beiden Sprachen finden wir wahrscheinlich die größten Unterschiede. Im Italienischen kommt in der Coda immer nur höchstens ein Konsonant vor. Dieser Konsonant kann außerdem nur ein Sonorant (also ein Nasal oder ein Liquid), die erste Hälfte einer Geminate oder ein [s] sein. Im Deutschen hingegen kann die Coda aus mehreren Konsonanten bestehen.

Was das [s] betrifft, so müssen wir seine Silbenposition in wortinternen Konsonantensequenzen noch etwas ausführlicher besprechen. Betrachet die folgenden Hypothesen zur Silbifizierung:

(128) As.co.li oder A.sco.li?
pas.ta oder pa.sta?
a.spro oder as.pro?

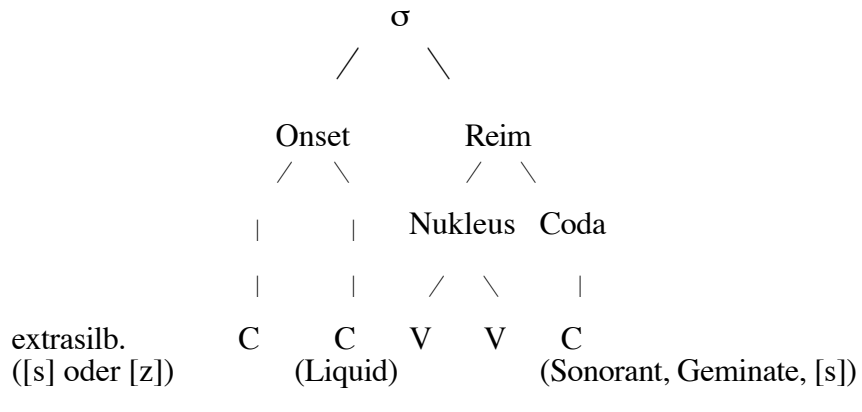
Es ist in diesen Fällen nicht ganz klar, ob das [s] zur Coda der vorhergenden Silbe oder zum Onset der nächsten Silbe gezählt werden soll. In der Schule lernt man, dass man diese Wörter als *A.sco.li* bzw. *pa.sta, a.spro* trennen soll. Aber es gibt gute phonologische Argumente dafür, dass die richtige Silbentrennung eigentlich *As.co.li* bzw. *pas.ta, as.pro* ist (s. Nespor 1993: 176). Wir können nämlich beobachten, dass der Vokal vor dem [s] immer kurz ist. Wenn die Trennung jedoch *A.sco.li, pa.sta, a.spro* wäre, dann müsste der Vokal gelängt werden, denn Vokale in offenen,[26] betonten Silben werden im Italienischen (vor allem in der vorletzten Silbe) immer gelängt. Wir können das in folgenden Beispielen sehen, in denen die betonten Vokale alle in offenen Silben stehen und deshalb gelängt werden:

(129) cá.sa
á.ra.bo
prá.to

Diese Beispiele zeigen uns außerdem, dass Vokallänge im Italienischen vom Akzent abhängt. Sie ist somit vorhersagbar und kann also nicht distinktiv sein. Hier unterscheidet sich das Italienische vom Deutschen, in dem die Vokallänge, in Verbindung mit dem Merkmal [±gespannt], Minimalpaare wie *Miete* und *Mitte* unterscheiden kann.

Wir können nun die Struktur der italienischen maximalen Silbe zusammenfassend folgendermaßen darstellen:

(130) Die maximale italienische Silbe:



[26] **Offene Silben** sind Silben, die auf einen Vokal enden (also nicht durch einen Konsonanten "geschlossen" werden).

Wir wenden uns hier zum Abschluss noch einem interessanten Problem zu, das die Länge der Wörter des Italienischen und des Deutschen betrifft. Wer die beiden Sprachen auch nur oberflächlich kennt, bemerkt bald, dass das Italienische eigentlich ziemlich lange Wörter hat, während deutsche Wörter meistens aus einer einzigen Silbe bestehen. Wir sprechen hier nur von morphologisch einfachen Wörtern, denn im Deutschen kann man natürlich sehr lange morphologisch komplexe Wörter bilden, wenn man z.B. an Komposita wie *Donaudampfschifffahrt* denkt. Doch es ist gar nicht so einfach, mehrsilbige deutsche Wörter zu finden, die keine Komposita sind und keine Suffixe oder Präfixe enthalten. Im Bereich der Fremdwörter gibt es natürlich auch mehrsilbige Wörter, wie z.B. *Universität, individuell* u.s.w., aber die mehrsilbigen Wörter, die nicht eindeutige Fremdwörter sind, sind eher selten. Es handelt sich um so Wörter wie *Arbeit, Ameise*, oder auch Wörter, die auf *–er* und *–el* enden, wie *Engel, Muster* oder ähnliches. Die meisten typisch deutschen Wörter bestehen jedoch aus nicht mehr als einer Silbe. Im Italienischen ist es umgekehrt: das typische italienische Wort besteht aus mehr als einer Silbe. Wie kann man sich den Unterschied zwischen diesen beiden Sprachen erklären?

Ein Grund für die Kürze der deutschen Wörter könnte in der Silbenstruktur liegen. Nehmen wir einmal an, dass es, um einen Begriff in einem Wort auszudrücken, im Durchschnitt sieben Segmente braucht. Im Deutschen können diese sieben Segmente ohne weiteres in eine einzige Silbe gezwängt werden, wenn wir z.B. eine Silbe nehmen, die ein oder mehrere extrasilbische Segmente hat. Im Italienischen brauchen wir für die gleichen sieben Segmente hingegen mindestens zwei Silben, da in einer italienischen Silbe maximal fünf Segmente Platz haben.

Ein weiterer Grund für die "Mehrsilbigkeit" des Italienischen liegt sicherlich darin, dass italienische (lexikalische) Wörter immer mit einem Vokal enden müssen. Ein einsilbiges Wort kann diese Beschränkung nur erfüllen, wenn die Silbe eine (C)CV-Struktur hat. In einer Struktur dieser Art kann man aber sehr wenige Segmente unterbringen (höchstens vier, wenn man die extrasilbischen Elemente dazuzählt). Es handelt sich hier also um eine nicht sehr "informative" Silbe.

4.1.5. Verletzbare universelle Beschränkungen - eine Mini-Einführung in die Optimalitätstheorie

Wir haben nun gesehen, wie die deutsche und die italienische Silbe aussehen. In diesem Kapitel wollen wir uns fragen, wie sich die Silben dieser Sprachen in Bezug auf die universellen Onset- und Coda-Beschränkungen verhalten.

Auf den ersten Blick scheint die Antwort auf diese Frage banal zu sein: beide Sprachen scheinen diese Beschränkungen nicht zu befolgen.

Im Deutschen haben wir Beispiele wie das folgende:

(131) [ze:.ən] <sehen>

In diesem Beispiel hat die zweite Silbe keinen Onset, sie verstößt also gegen die Onset-Beschränkung (1. Teil) und sie hat zugleich eine Coda, verstößt also gegen die Coda-Beschränkung (1. Teil).

Im Italienischen haben wir Beispiele wie:

(132) am.ba.scia.ta

wo die erste Silbe keinen Onset, wohl aber eine Coda hat.

Doch auch wenn die beiden universellen Beschränkungen in vielen Wörtern verletzt werden, so können wir doch sehen, dass sie auch in diesen beiden Sprachen eine gewisse Rolle spielen.

Wir werden sehen, dass beide Sprachen, wann immer möglich, die beiden Beschränkungen befolgen, und dass sie dazu bestimmte Strategien benützen.

Beginnen wir mit dem Deutschen. Wir haben im Deutschen Wörter, die mit einem Vokal beginnen. Wenn dieser Vokal am Wortanfang steht, dann wird vor ihm ein glottaler Plosiv eingesetzt. Das Wort *Esel* wird also folgendermaßen realisiert:

(133) /e:səl/ → [ʔe:səl]

Der glottale Plosiv sorgt dafür, dass die erste Silbe einen Onset bekommt. Wortintern kann der glottale Plosiv nur vor einen Vokal gesetzt werden, wenn dieser betont ist. Im Beispiel *[ze:.ən]* kann also kein glottaler Plosiv vor das schwa gesetzt werden, da dieses nicht betont ist.[27]

Prinzipiell könnte man sich für das Wort */e:səl/* verschiedene Strategien der Realisierung vorstellen:

---

[27] Wer an einer detaillierten Analyse dieses Prozesses interessiert ist, kann sich meinen Aufsatz dazu anschauen: Alber B. (2001). "Regional Variation and Edges: Glottal Stop Epenthesis and Dissimilation in Standard and Southern Varieties of German". *Zeitschrift für Sprachwissenschaft* Vol. 20.1, S. 3-41.

(134) /e:səl/ → a. [ʔe:səl]
b. *[səl]
c. *[e:səl]
/pak e:səl/ → d. *[pa.k e:.səl] <Packesel>

Wir können, wie in a., einen Konsonanten für die vokalinitiale erste Silbe einfügen. Dies ist die Strategie, die schließlich angewandt wird. Wir können uns aber auch vorstellen, dass man vielleicht den ersten Vokal streicht, wie in b. Durch diese Streichung erreichen wir, dass nur eine Silbe mit Onset übrigbleibt. Wir könnten auch die zugrundeliegende Form so lassen, wie sie ist, und damit die Onset-Beschränkung verletzen, wie in c. Schließlich könnten wir, wenn dem Wort *Esel* noch ein anderes Wort vorangeht, den Weg der Resilbifizierung wählen, wie in d. Das bedeutet, dass wir den Konsonanten des ersten Wortes nehmen und ihn als Onset für die erste Silbe des zweiten Wortes verwenden. Doch auch diese Strategie wird im Deutschen nicht gewählt.[28]

Jede dieser Strategien hat bestimmte Vorteile und bestimmte Nachteile. Die erste Strategie verletzt eine Beschränkung (auf engl.: ***constraint***), die wir *Füge nichts hinzu!* nennen könnten. Sie wird dadurch verletzt, dass wir den glottalen Plosiv - ein Segment, das in der zugrundeliegenden Form nicht da war - hinzufügen. Die zweite Strategie verletzt die Beschränkung *Streiche nichts!*, die dafür sorgt, dass alle Segmente, die in der zugrundeliegenden Form da sind, auch in der phonetischen Form realisiert werden. Die dritte Strategie befolgt diese beiden letzten Beschränkungen, aber sie verletzt die Onset-Beschränkung. Die letzte Strategie verletzt eine Beschränkung, die wir *Keine Resilbifizierung!* nennen könnten:

(135) a. [ʔe:səl] verletzt *Füge nichts hinzu!*
b. *[səl] verletzt *Streiche nichts!*
c. *[e:səl] verletzt *Onset-Beschränkung*
d. *[pa.k e:.səl] verletzt *Keine Resilbifizierung!*

Alle diese Beschränkungen sind von Phonologen schon in Bezug auf andere phonologische Prozesse in anderen Sprachen beobachtet worden. Wir werden im Folgenden die Namen für diese Beschränkungen verwenden, die wir auch in der linguistischen Literatur finden könnnen, und zwar:

(136) a. *Füge nichts hinzu!* = DEP
b. *Streiche nichts!* = MAX
c. *Onset-Beschränkung* = ONSET
d. *Keine Resilbifizierung!* = ALIGN (WORT/WURZEL, SILBE, L)

Der Name für die Beschränkung DEP steht für engl. *dependency* , da die Beschränkung besagt, dass jedes Element des output (der phonetischen Oberflächenform) ein ihm entsprechendes Element im input (der phonematischen, zugrundeliegenden Form) haben soll:

[28] Die Strategie der Resilbifizierung können wir – unter bestimmten Umständen – auch im Deutschen finden, z.B. wenn das vorangehende Wort ein Funktionswort ist, die erste Silbe des lexikalischen Wortes unbetont ist und das Sprechtempo relativ schnell ist. So können wir z.B. die Sequenz *in Amerika* ohne glottalen Verschluss und mit Resilbifizierung realisieren: *i.n A.me.ri.ka*

(137) DEP: jedem Element im output entspricht ein Element im input[29]

Diese Beschränkung ist in [ʔe:səl] verletzt, da dem ersten Element im output, dem glottalen Verschluss, kein input-Element entspricht:

(138) / e:$_1$ s$_2$ ə$_3$ l$_4$ / ← input
| | | | |
[ʔ$_1$ e:$_2$ s$_3$ ə$_4$ l$_5$] ← output

Die Beschränkung MAX (für engl.: *maximality*) besagt, dass es für jedes input-Element ein entsprechendes output-Element geben muss:

(139) MAX: jedem Element im input entspricht ein Element im output

Wenn wir ein input-Element tilgen, dann wird dieser *constraint* verletzt:

(140) / e:$_1$ s$_2$ ə$_3$ l$_4$ / ← input
| | | |
[s$_1$ ə$_2$ l$_3$] ← output

DEP und MAX sind sogenannte **Treuebeschränkungen** (engl.: *faithfulness constraints*). Sie sorgen dafür, dass der input (die zugrundeliegende Form) gleich aussieht wie der output (die phonetische Realisierung). Auch bei diesen Beschränkungen handelt es sich um universelle Beschränkungen, die man in allen Sprachen der Welt finden kann, die aber manchmal verletzt werden können.

Unsere Onset-Beschränkung (1. Teil) können wir auch als *constraint* definieren:

(141) ONSET: jede Silbe hat einen Onset

Diese Beschränkung ist natürlich durch die erste Silbe in [e:səl] verletzt. ONSET ist eine **Markiertheitsbeschränkung**, die universell unmarkierte Strukturen begünstigt. Ein Großteil der phonologischen Prozesse lässt sich als Kampf zwischen Treuebeschränkungen und Markiertheitsbeschränkungen darstellen. Wenn Treuebeschränkungen dominant sind, dann bleibt die zugrundeliegende Form so wie sie ist, wenn Markiertheitsbeschränkungen wichtiger sind, finden phonologische Prozesse statt.

Der *constraint* ALIGN (WURZEL/WORT, SILBE, L) gehört zur Familie der ***alignment constraints*** die möchten, dass die Ränder bestimmter morphologischer und prosodischer Kategorien übereinstimmen. Den hier relevanten *alignment constraint* können wir folgendermaßen definieren:

(142) ALIGN (WURZEL/WORT, SILBE, L): der linke Rand eines Wortes oder einer Wurzel stimmt mit dem linken Rand einer Silbe überein

In der Silbifizierung

(143) [pa.k] [e:.səl]

---

[29] Für die Definition dieses und der folgenden *constraints* s. Prince&Smolensky (1993/2004), McCarthy&Prince (1993) und McCarthy&Prince (1995).

beginnt die zweite Silbe vor dem Laut [k] und endet nach dem [e:]. Der Wort-bzw. Wurzelgrenze zwischen [k] und [e:] entspricht also keine Silbengrenze. Der constraint ALIGN ist verletzt.

Für das Deutsche können wir sagen, dass eine dieser Beschränkungen auf Kosten der anderen Beschränkungen verletzt werden kann. In einem gewissen Sinne ist sie "schwächer" als die anderen Beschränkungen. Es handelt sich dabei um die Beschränkung DEP. Indem wir im output einen glottalen Plosiv hinzufügen, verletzen wir diese Beschränkung. Dadurch, dass wir diese Beschränkung verletzen, können wir jedoch die anderen Beschränkungen erfüllen. Wir fügen ein Segment – den glottalen Verschluss – ein und können dadurch die Onset-Beschränkung erfüllen, ohne MAX und ALIGN zu verletzen.

Wir können diese vier Beschränkungen nach ihrer Wichtigkeit in einer Hierarchie anordnen:

(144) ONSET, MAX, ALIGN
|
DEP

DEP wird von den drei anderen Beschränkungen dominiert. Das führt dazu, dass diese Beschränkung verletzt wird, um die drei anderen Beschränkung zu erfüllen.

Wir können die Interaktion zwischen den verschiedenen Beschränkungen auch in einem sogenannten *constraint tableau* veranschaulichen:

Tableau 1 Bewertung des input /pak e:.səl/ (Packesel):

/pak e:.səl/	ONSET	MAX	ALIGN	DEP	
a. pak. e:səl	*!				2. Silbe ohne Onset
b. pak. səl		*!			Tilgung
c. pa.k e:.səl			*!		Resilbifizierung
☞ d. pak. ʔe:.səl				*	Epenthese

In der 1. Spalte dieser Tabelle steht der input des Wortes, das die Beschränkungen bewerten, /pak e:.səl/. Darunter folgen die *Kandidaten*, d.h., alle möglichen Realisierungen, die man sich für den input vorstellen kann. Kandidat a. ist eine Realisierung, bei der sich in Bezug auf den input nichts verändert hat. In Kandidat b. wurde der Vokal [e:] getilgt. Kandidat c. ist der Resilbifizierungskandidat, er unterscheidet sich von Kandidat a. dadurch, dass hier die Silbengrenze gleich hinter dem ersten Vokal liegt. Dadurch bekommt die zweite Silbe den Onset [k]. Kandidat c. zeichnet sich dadurch aus, dass ein Konsonant, der glottale Plosiv [ʔ] eingefügt wird. Diesen Prozess des Einfügens von Lauten nennt man in der Phonologie auch **Epenthese.** Dieser letzte Kandidat wird am Ende als optimaler Kandidat ausgewählt. Er ist der Gewinner, denn er wird das Rennen gegen die anderen Kandidaten gewinnen. Deshalb kennzeichnen wir ihn mit der kleinen Hand ☞. Die erste Hürde, die die Kandidaten überwinden müssen, ist der *constraint* ONSET. Hier bleibt Kandidat a. auf der Strecke, denn er verletzt diese Beschränkung, weil die zweite Silbe keinen Onset hat. Verletzungen geben wir mit einem Stern * an. Das Ausrufezeichen nach dem Stern bedeutet, dass die Verletzung in diesem Fall fatal war, der Kandidat scheidet aus dem Rennen aus, denn alle anderen Kandidaten waren besser als er, was den constraint ONSET betrifft. Es bleiben also nun Kandidat b, c. und d. übrig, die sich dem zweiten *constraint*, MAX, stellen. An diesem Punkt scheidet Kandidat b. aus, denn dieser Kandidat verletzt MAX, weil der input-Vokal

[e:] getilgt wurde. Es bleiben nun Kandidat c. und d. übrig. Kandidat c. verlässt uns nach der Begegnung mit der Beschränkung ALIGN. Dieser Kandidat verletzt ALIGN, da der linke Rand des Wortes *Esel* nicht mit dem linken Rand einer Silbe übereinstimmt. Nach dieser Runde bleibt nur Kandidat d. übrig. Dieser Kandidat verletzt den *constraint* DEP, doch da DEP am tiefsten in der Hierarchie der Beschränkungen steht, fällt diese Verletzung nicht mehr ins Gewicht. Alle Zellen der Tabelle, deren Inhalt für die Bewertung unwichtig ist, sind grau schattiert. Das gilt auch für die Zelle, in der die DEP-Verletzung von Kandidat d. steht: es ist egal, was in dieser Zelle steht, ob Kandidat d. eine DEP-Verletzung, zwei DEP-Verletzungen oder gar keine DEP-Verletzung aufweist, er gewinnt auf jeden Fall.

Die Trennungslinien zwischen Spalten der ONSET-, MAX- und ALIGN-Beschränkung sind gestrichelt, denn es ist nicht relevant, welche dieser Beschränkungen die jeweils andere dominiert. Wir können eigentlich auch gar nicht sagen, ob ONSET, MAX und ALIGN in einer Dominanzrelation zueinander stehen. Wichtig ist nur, dass alle drei Beschränkungen den *constraint* DEP dominieren.

Wie sieht es nun im Italienischen aus? Versucht, die Analyse des Italienischen selbst in der folgenden Übung zu erarbeiten:

Übung 22

1. Wir verwenden in dieser Übung die Beispiele *amore* und *con amore*. Erstellt nun mit diesen Beispielen eine Liste von möglichen outputs, wie wir sie im letzten Beispiel (*Esel*) für das Deutsche gesehen haben. Wie könnten *amore* und *con amore* prinzipiell realisiert werden? Ihr solltet alle vier Möglichkeiten anführen: Einfügen eines Segmentes, Streichen eines Segmentes, keine Änderung und Resilbifizierung.

	amore	*con amore*
a.	_____	_____
b.	_____	_____
c.	_____	_____
d.	_____	_____

Welche dieser Realisierungen gibt es nun im Italienischen wirklich, welche sind nicht möglich?

2. Gebt nun an, welche Beschränkungen in diesen vier Realisierungen verletzt werden:

a. _____
b. _____
c. _____
d. _____

Welche Beschränkungen können also im Italienischen verletzt werden, welche nicht?

3. Versucht nun, die Hierarchie der Beschränkungen für das Italienische zu bauen. In dieser Hierarchie werden die Beschränkungen, die verletzt werden können, von den Beschränkungen, die nicht verletzt werden können, dominiert. Stellt diese Hierarchie auf die gleiche Art und Weise wie im Deutschen dar.

4. Die Hierarchie des Italienischen ist allerdings etwas komplizierter als die im Deutschen. Schaut euch noch mal die beiden Strategien an, die im Italienischen möglich sind, und die Beschränkungen, die durch sie verletzt werden. Lässt sich zwischen diesen beiden Beschränkungen noch eine Subhierarchie feststellen? Welche Strategie wird in *con amore* bevorzugt? Welche von den beiden relevanten Beschränkungen wird also eher verletzt?

5. Zeichnet zwei *constraint tableaux*, eines für den input *con amore,* eines für den input *amore*. Beschreibt die Auswahl des Gewinners in diesen *tableaux.*

Wenn ihr diese Übung in all ihren Schritten gelöst habt, dann solltet ihr zu dem Schluss gelangt sein, dass im Italienischen die beiden Beschränkungen MAX und DEP am höchsten stehen. ONSET und ALIGN hingegen müssen von diesen beiden Beschränkungen dominiert werden, denn sie werden in den outputs *amore* und *co.n a.mo.re* verletzt. Wenn wir versuchen, eine Hierarchie zwischen ONSET und ALIGN herzustellen, dann müssen wir uns den output *co.n a.mo.re* anschauen. In diesem output wird die Beschränkung ALIGN verletzt, während ONSET befolgt wird. ONSET muss also ALIGN dominieren. Wir erhalten also für das Italienische die folgende Hierarchie:

(145)
MAX, DEP
|
ONSET
|
ALIGN

Die beiden Beschränkungen MAX und DEP stehen am höchsten in der Hierarchie, denn sie werden immer befolgt. ONSET und ALIGN stehen tiefer, denn sie werden in einigen Fällen verletzt. Wenn man aber die Wahl zwischen diesen beiden Beschränkungen hat, wie im Fall von *con amore*. dann entscheiden wir uns dafür, die Onset-Beschränkung auf Kosten von ALIGN zu befolgen. Wir resilbifizieren also die beiden Wörter als *co.n a.mo.re*, anstatt das zweite Wort ohne Onset zu lassen (wie in *con amore*).

Wir können nun die Auswahl des gewinnenden Kandidaten durch die Beschränkungen wieder in einem *constraint tableau* darstellen. Erstellen wir zuerst das *tableau* für den input *con amore:*

Tableau 2: Bewertung des input /con amore/:[30]

/con amore/	MAX	DEP	ONSET	ALIGN	
a. con. a.mo.re			*!		2. Silbe ohne Onset
b. con. mo.re	*!				Tilgung
☞ c. co.n a.mo.re				*	Resilbifizierung
d. con. ʔa.mo.re		*!			Epenthese

Wir sehen in der ersten Spalte dieselben Kandidaten, die wir schon im Deutschen hatten. Die Hierarchie der Beschränkungen ist aber im Italienischen eine andere als im Deutschen, deshalb wird auch ein anderer Gewinner unter den Kandidaten ausgewählt. Bei der Bewertung durch MAX scheidet Kandidat b. aus, da ihm der

[30] Um die Diskussion nicht unnötig zu komplizieren verzichte ich hier auf die phonetische Transkription der italienischen Beispiele.

input-Vokal [a] fehlt. Bei der Bewertung durch DEP verlieren wir Kandidat d., da er durch die Epenthese eines glottalen Verschlusslautes diesen *constraint* verletzt. Kandidat d. kann also, im Gegensatz zum Deutschen, hier nicht gewinnen. Kandidat a. scheidet aus, weil er im Vergleich zu Kandidat c. schlechter ist: er verletzt nämlich den höheren *constraint* ONSET, währen Kandidat c. nur den tieferen *constraint* ALIGN verletzt. Es gewinnt bei dieser Bewertung also Kandidat c., der Kandidat mit Resilbifizierung.

Schauen wir uns jetzt noch die Bewertung des input /amore/ an:

Tableau 3: Bewertung des input /amore/:

/amore/	MAX	DEP	ONSET	ALIGN	
☞ a. a.mo.re			*		2. Silbe ohne Onset
b. mo.re	*!				Tilgung
c. ?a.mo.re		*!			Epenthese

Wir haben hier wieder dieselbe Hierarchie von Beschränkungen, die für das Italienische gilt. Allerdings fehlt in der 1. Spalte ein Kandidat, der Kandidat mit Resilbifizierung. Ihn gibt es nicht, da der input nur aus einem Wort besteht, es steht kein Wort davor, von dem sich *amore* einen Konsonanten "ausleihen" könnte. Was passiert in diesem Fall? Nun, der Tilgungskandidat b. und der Epenthesekandidat c. scheiden wieder aus, da sie MAX bzw. DEP verletzen. Doch Kandidat a., der Kandidat ohne Onset, kommt jetzt zum Zug, da er von den möglichen Kandidaten der einzige ist, der die höher stehenden Beschränkungen nicht verletzt. Die Tatsache, dass er ONSET verletzt, ist also nicht so wichtig, da es keinen Kandidaten gibt, der besser wäre.

Was haben wir nun eigentlich in diesem Vergleich zwischen italienischen und deutschen Onset-Beschränkungen gesehen? Wir haben gesehen, dass die Beschränkungen, die die Silbenstruktur bestimmen, als universell angesehen werden können. Was bedeutet universell? Es bedeutet, dass wir ihren Einfluss in den meisten, vielleicht in allen Sprachen der Welt finden. Es bedeutet aber nicht, dass diese Beschränkungen immer und überall gelten. Sie können auch verletzt werden, aber nur, wenn es dadurch möglich ist, wichtigere Beschränkungen zu erfüllen.

Die Beschränkungen, die die Silbenstruktur bestimmen, sind universell, aber die Hierarchie ihrer Wichtigkeit ist sprachspezifisch. So ist z.B. die Onset-Beschränkung sowohl im Deutschen als auch im Italienischen ein wichtiges Prinzip. Aber im Deutschen steht sie höher in der Hierarchie als im Italienischen, sie wird im Deutschen also eher befolgt werden, als im Italienischen. Die **Optimalitätshierachie** (Prince&Smolensky 1993/2004) hat diese Ideen in den letzten 10 Jahren ausgearbeitet:

- die Idee, dass die Wohlgeformtheit von sprachlichen Strukturen durch universelle Beschränkungen (*constraints*) bestimmt wird;
- dass diese universellen Beschränkungen verletzbar sind;
- dass diese Beschränkungen in jeder Sprache in einer sprachspezifischen Hierarchie angeordnet sind.

Übung 23

Kennt ihr die Romane von Andrea Camilleri, in denen der commissario Montalbano die Hauptrolle spielt? Im commissariato von Montalbano gibt es den Polizisten Catarella, der manchmal eine etwas eigenartige Sprache spricht. So sagt er z.B. "Dottore! È arrivato un facchisi!" Er meint damit, dass ein *Fax* angekommen ist.

1. Transkribiert die Wörter *fax* und *facchisi*
2. Was hat Catarella mit dem Wort *fax* gemacht? Wie sieht die Silbenstruktur des Wortes *fax* aus, wie sieht die Silbenstruktur von *facchisi* aus?
3. Welche Silbenstruktubeschränkungen verletzen die Wörter *fax* und *facchisi*, welche befolgen sie?
4. Wie sieht also die Grammatik Catarellas aus, welche Beschränkungen stehen hoch in der Hierarchie, welche stehen tief?
5. Ist Catarellas Grammatik mit der des Standarditalienischen identisch?

4.2 Metrische Phonologie - Der Akzent

4.2.1 Füße - die Einheiten der Metrik

Im letzten Kapitel haben wir über die Silbe als prosodische Einheit gesprochen. Aber auch Silben sind noch längst nicht die größten prosodischen Einheiten. Wir wollen uns in diesem Kapitel noch mit der prosodischen Einheit des **Fußes** auseinandersetzen, der in einer Hierarchie der prosodischen Kategorien die nächstgrößere Kategorie nach der Silbe ist.

Wie in viele Sprachen der Welt so besitzen auch die Wörter des Deutschen eine alternierende rhythmische Struktur. Wir können zunächst feststellen, dass alle Wörter des Deutschen eine Silbe besitzen, die stärker als alle anderen Silben betont wird. Diese Silbe trägt den sogenannten **Hauptakzent.** So trägt z.B. im folgenden Wort die letzte Silbe den Hauptakzent:

(146) En.zy.klo.pä.**díe**

Wenn man nun genau hinhört, dann hört man auch noch, dass die erste Silbe dieses Wortes stärker betont ist als die zweite und die dritte wiederum stärker als die vierte. Das sind die sogenannten **Nebenakzente**:

(147) **È**n.zy.**klò**.pä.**díe**

Der Hauptakzent wird in der Phonologie mit einem **akuten Akzent (= ´)** transkribiert, der Nebenakzent mit einem **Gravis-Akzent (= `)**. In manchen Transkriptionen findet ihr aber auch vor der betonten Silbe einen hochgesetzten Strich für den Hauptakzent und einen tiefer gesetzten Strich für den Nebenakzent:

(148) ˌEn.zy.ˌklo.pä.ˈdie

Die Verteilung von Haupt- und Nebenakzenten folgt im Deutschen, wie auch in vielen anderen Sprachen, einem mehr oder weniger alternierenden Muster: eine Silbe betont – eine Silbe auslassen – eine Silbe betont – usw. Diese rhythmischen Muster

können jedoch von Sprache zu Sprache sehr verschieden sein. So kann z.B. der Hauptakzent auf die erste oder aber auf die letzte Silbe fallen, der erste Nebenakzent kann auf die erste oder auf die zweite Silbe fallen u.s.w.

Sehen wir uns zum Vergleich das Akzentmuster des Finnischen an, das auch alternierend ist, sich aber ziemlich von dem des Deutschen unterscheidet. Im Finnischen fällt der Hauptakzent auf die erste Silbe, Nebenakzente fallen meistens auf alle ungeraden Silben danach:

(149)	ó.	pet.	tè.	le.	mà.	na.	ni	'as something I have been learning'
	hm	ta	**hm**	ta	**hm**	ta	ta	
	1	2	3	4	5	6	7	

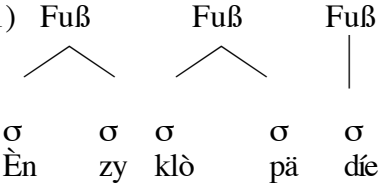
Allerdings: wenn sich im Wort eine **schwere Silbe** mit einem komplexen Reim (d.h. mit einem langen Vokal, einem Diphthong oder einem Vokal und einem darauffolgenden Konsonanten) befindet, dann zieht diese den Nebenakzent an und der streng alternierende Rhythmus wird unterbrochen:

(150)	vá.	**lis**.	tu.	**màt**.	to.	mi.	**àn**.	ne	'your uneducated, part.pl.'
		|		|			|	
		schwer		schwer			schwer		
	hm	ta	ta	**hm**	ta	ta	**hm**	ta	
	1	2	3	4	5	6	7	8	

In diesem Beispiel fallen die Akzente nicht mehr nur auf alle ungeraden Silben, sondern auch auf Silben in gerader Position (z.B. die 4. Silbe), wenn diese schwer sind. Man könnte das finnnische Akzentmuster so beschreiben: die erste Silbe trägt immer den Hauptakzent; danach folgen die Nebenakzente mit einem alternierenden Muster, außer wenn eine schwere Silbe auftaucht, die dann den Akzent anzieht.

Die meisten Phonologen sind sich heutzutage ziemlich einig, dass man Akzentmuster am besten beschreibt, indem man eine oder mehrere Silben zu einer Einheit zusammenschließt und in einen **Fuß** packt:

(151)



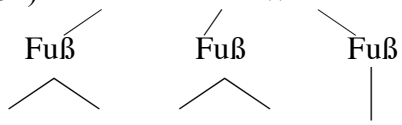
Wie diese Füße aber genau aussehen, ist eine schwierige Frage, die wir hier leider nicht ausführlich diskutieren können.[31] Bilden immer zwei Silben einen Fuß oder manchmal nur eine? Ist eine Silbe, die alleine einen Fuß bildet (wie z.B. die Silbe in der obigen Darstellung) immer auch einen schwere Silbe?

Segmente verbinden sich zu Silben, Silben verbinden sich zu Füßen und Füße verbinden sich am Ende zu einem prosodischen Wort, das ungefähr dem grammatikalischen Wort gleichkommt (die schwierigen Fälle, wo das nicht so ist, interessieren uns hier nicht). Wir haben also am Ende für ein Wort die folgende **prosodische Struktur**:

---

[31] Einen Überblick über die Ideen zur Struktur des metrischen Fußes findet ihr im Lexikoneintrag *Foot*, den ich für die *Encyclopedia of Language and Linguistics* geschrieben habe (Alber 2006).

(152) Prosodisches Wort



σ	σ	σ	σ	σ
V C	CCV	CCV	CV	CV:
[ɛ n	tsy	klo	pɛ	di:]

4.2.3 Kurzer Kommentar zum deutschen Akzent

Wir wollen uns in diesem Kapitel zuerst mit der Akzentstruktur in **nativen Wörtern** des Deutschen und dann mit der Akzentstruktur in **Fremdwörtern** beschäftigen. Denn native und entlehnte Wörter im Deutschen haben eine unterschiedlich Struktur, die das Akzentmuster in beiden Fällen als verschieden erscheinen lässt, auch wenn eine einheitliche Beschreibung der beiden Bereiche vielleicht möglich ist.

Wenden wir uns zunächst dem nativen Wortschatz zu. Bei den nativen Wörtern des Deutschen muss man wiederum zwischen **morphologisch einfachen** und **morphologisch komplexen** Wörtern unterscheiden. Native Wörter, die morphologisch einfach sind, die also nur aus einem Morphem (einer Wurzel) bestehen, sind im Deutschen sehr kurz. Meistens besteht die Wurzel eines Wortes nur aus einer einzigen Silbe. Folglich fällt der Akzent auf eben diese Silbe. In den wenigen nativen Wörtern, die aus mehr als einer Silbe bestehen, fällt der Hauptakzent auf die erste dieser Silben:

(153) Árbeit nicht: *Arbéit
Ántwort nicht: *Antwórt
Úrlaub nicht: *Urláub

Bei den morphologisch komplexen Wörtern müssen wir verschiedene Untergruppen unterscheiden.
In den Komposita fällt der stärkere Akzent – im Gegensatz zum Italienischen –auf das erste Element des Kompositums (das Bestimmungswort):

(154) Komposita: Akzent fällt auf den ersten Teil:
Léhrstùhl
Hóchschùle
Kíndergàrten
réchtfèrtigen
frűhstücken

Das zweite Element des Kompositums (der Kopf) trägt auch einen Akzent, der aber schwächer als der des ersten Elementes ist.

Verben mit trennbaren Präfixen haben eine Akzentstruktur, die sehr derjenigen der Komposita ähnelt, d.h. ein stärkerer Akzent fällt auf das Präfix und eine schwächerer Nebenakzent fällt auf die Verbwurzel:

(155) trennbare Präfixe:
áuf-hàlten
úm-dènken
űber-sètzen 'traghettare'

Bei untrennbaren Präfixen fällt hingegen der stärkste Akzent auf die Verbwurzel. Das Präfix selbst ist oft unbetont, besonders wenn es eine Silbe enthält, die als leicht gewertet werden kann:

(156) untrennbare Präfixe:
be-tréiben
zer-bréchen
er-láuben
ü̱ber-sétzen 'tradurre'

Es gibt dann noch einige Beispiele von Wörtern, die wie Komposita aussehen und also eigentlich eine Kompositabetonung haben sollten, die aber dennoch auf dem zweiten Element betont werden:

(157) Komposita ohne Kompositaakzent:
frohlócken
vollbríngen

Es könnte sein, dass es sich hier um exozentrische Komposita handelt, deren morphologische Struktur nicht mehr transparent ist.

Zum nativen Wortschatz ist noch zu sagen, dass das schwa [ə] und das vokalisierte [ɐ] nie einen Akzent tragen. Ein Wort wie das folgende kann also nur auf der ersten Silbe betont werden, auch wenn sich dadurch keine alternierende rhythmische Struktur ergibt:

(158) die trockeneren = die tróck[ə]n[ə]r[ə]n

Sehen wir uns nun den Akzent in den Fremdwörtern an. Fremdwörter sind, im Unterschied zu nativen Wörtern, oft mehrsilbig. Bei ihnen können wir also leichter eine alternierende rhythmische Struktur beobachten als in nativen Wörtern.

Bei den Fremdwörtern, die in die deutsche Sprache integriert wurden, finden wir den Hauptakzent auf einer der letzten drei Silben.

(159) a. Hauptakzent auf der letzten Silbe (Ultima): Ù.ni.vèr.si.t´ät
Èn.zy.klò.pä.díe

b. Hauptakzent auf der vorletzten Silbe (Pänultima): Pä.da.gó.gik
Mos.kí.to

c. Hauptakzent auf der vorvorletzten Silbe (Antepänultima): Pe.tró.le.um
Mó.ni.tor

Es ist nur schwer vorherzusagen, auf welche der letzten drei Silben der Hauptakzent jeweils fällt. So gibt es Wörter, die eine sehr ähnliche, oder sogar eine identische Struktur haben, in denen der Hauptakzent jedoch nicht auf dieselbe Silbe fällt:

(160) Té.nor 'contenuto, significato'
Te.nór 'tenore'

Bei *Té.nor* und *Te.nór* handelt es sich um zwei verschiedenen Wörter, mit unterschiedlicher Bedeutung, die aber dieselbe lautliche Struktur haben. Dennoch fällt der Akzent im ersten Fall auf die Pänultima, im zweiten auf die Ultima.

Es gibt auch Wörter, die zwei verschiedene Betonungsmöglichkeiten haben, wie z.B.:

(161) Kí.mo.no
Ki.mó.no

Dieses Wort kann im Deutschen sowohl auf der Pänultima als auch auf der Antepänultima betont werden.

Schließlich gibt es auch regionale Variation, was die Betonung bestimmter Wörter betrifft. So wird das Wort *Mathematik* im norddeutschen Raum eher auf der Ultima betont, im süddeutschen Raum eher auf der Pänultima:

(162) Mà.the.ma.tík (norddeutsche Varietäten)
Mà.the.má.tik (süddeutsche Varietäten)

Alle diese Variationen zeigen uns, dass der deutsche Fremdwortakzent nicht absolut vorhersagbar ist. Wir können jedoch zwei Regelmäßigkeiten festhalten:

- der Akzent fällt immer auf eine der letzten drei Silben
- wenn der Akzent auf die Antepänultima fällt, so scheint die Pänultima in den allermeisten Fällen eine leichte Silbe zu sein.

Wir können diesen zweiten Punkt an den Beispielen verifizieren, die wir bisher genannt haben. In keinem der Beispiele mit Antepänultimabetonung ist die Pänultima schwer, enthält also einen langen Vokal, einen Diphthong oder wird durch einen Konsonanten geschlossen.

Die Fremdwörter des Deutschen kommen zu einem großen Teil aus dem Griechischen und dem Lateinischen. Es wird uns also nicht verwundern, dass die Akzentstruktur, die wir gerade beschrieben haben, der Akzentstruktur des Italienischen sehr ähnlich ist. Auch im Italienischen fällt der Hauptakzent immer auf eine der letzten drei Silben, oft auf genau dieselbe Silbe, auf der sie auch im entsprechenden deutschen Fremdwort liegt:

(163)	sy.ste.má.tisch	si.ste.má.ti.co
	En.zy.klo.pä.díe	en.ci.clo.pe.dí.a
	u.ni.ver.sál	u.ni.ver.sá.le

Es wäre ein interessantes Forschungsprojekt festzustellen, ob auch im Italienischen die Antepänultima nur betont sein kann, wenn die Pänultima leicht (also im Italienischen: eine Silbe ohne Coda) ist.

5. Laute und Schrift

Im Deutschen gibt die Orthographie die Laute der Sprache ziemlich genau, aber doch nicht perfekt wieder. So gibt es z.B. Sequenzen von **Graphemen** (Zeichen der Schrift), denen nur ein Phonem entspricht:

<sch> [ʃ] 'schlafen' 3 Grapheme, ein Phonem

<ch> [x] oder [ç] 'ich, ach' 2 Grapheme, ein Phonem, 2 Allophone

Ein weiterer Fall ähnlicher Art sind die Grapheme <f> und <v>: sie drücken im Deutschen immer denselben Laut [f] aus:

<für> ['fy:ʁ] 2 Grapheme, ein Phonem
<viel> ['fi:l]

Umgekehrt gibt es Fälle, in denen wir zwei Laute haben, aber dieser Unterschied wird in der Orthographie nicht ausgedrückt. So gibt es z.B. kein orthographisches Zeichen für das stimmhaft [z]. Wir schreiben sowohl für [z] als auch für [s] ein <s>. Das "scharfe" <ß> (auch ess-zett genannt) entspricht allerdings immer einem stimmlosen [s].

<See> ['ze:] zwei Phoneme, ein Graphem
<das> ['das]

Auch die Grapheme <z> und <x> druecken in Wahrheit zwei Laute aus:

<zahm> ['tsa:m] zwei Laute, ein Graphem
<Hexe> ['hɛksə]

Ein weiteres Beispiel für die Nicht-Entsprechung zwischen Phonemen und Graphemen sind die Doppelkonsonanten. Im Deutschen gibt es keine Geminaten, d.h., es gibt keine Konsonanten, die phonetisch gesehen lang sind. In der Schrift gibt es aber sehr viele Doppelkonsonanten. Diese orthographischen Doppelkonsonanten drücken nicht die Länge der Konsonanten aus, sondern die Kürze des vorhergehenden Vokals:

<Mitte> ['mɪtə]
<Mappe> ['mapə]
<Ecke> ['ɛkə]
<Spitze> [ʃpɪtsə][32]

Auch bei den Diphthongen ist die Orthographie etwas eigenartig. So schreibt man <eu>, wie z.B. in <euch>, aber dieser Diphthong setzt sich nicht aus den Lauten [e] und [u] zusammen, sondern aus [ɔ] und [ɪ].

Übung 24
Wie wird in der deutschen Orthographie die Vokallänge wiedergegeben? Untersucht dazu folgende Daten:
dir
viel
ihr
See

Die Beispiele in dieser Übung zeigen uns, dass es im Deutschen verschiedene Möglichkeiten gibt, um die Vokallänge zu kennzeichnen. Manchmal, wie im Wort

[32] Komischerweise gibt es in der deutschen Orthographie keine Doppelkonsonanten <kk> und <zz>. Stattdessen wird <ck> und <tz> geschrieben.

dir, das ein langes [i:] enthält, wird die Vokallänge gar nicht angegeben. In anderen Wörtern, wie z.B. in *viel,* sehen wir, dass ein langes [i:] durch die Buchstabensequenz <ie> gekennzeichnet wird. In anderen Wörtern wiederum, wie in *ihr,* steht ein sogennantes **Dehnungs-h** nach dem Vokal und zeigt uns, dass dieser lang ist. Und schließlich gibt es noch Beispiele, in denen für einen langen Vokal einfach zwei Vokalgrapheme geschrieben werden, wie in *See*.

Die deutsche Orthographie wurde von Mönchen im Mittelalter entwickelt, auf der Basis der lateinischen Schrift. Ende des 19. Jahrhunderts und Anfang des 20. Jahrhunderts gab es dann einige Reformen in der Orthographie und deren Festlegung in einer Norm. Einige orthographische Eigenheiten wurden eliminiert, aber andere (siehe oben) sind geblieben. Die Orthographie ist eben auch das Resultat einer historischen Entwicklung. Wir können sie nicht zu einer perfekten Übereinstimmung mit den Phonemen bringen, ohne die Sprachgemeinschaft zu verstören. Stellt euch nur vor, wie das wäre, wenn es plötzlich ein Gesetz gäbe, das befiehlt, dass man alle Texte in phonetischer Transkription schreiben muss. Allerdings sind die deutsche und die italienische Rechtschreibung doch noch relativ nahe an der Lautstruktur der Wörter, im Vergleich z.B. zu der englischen Orthographie.

Appendix

Kommentar zur Transkription des Deutschen:

Die beste Grundlage für gutes Transkribieren ist gutes Hinhören. Wenn aber eine Stelle kommt, an der man den eigenen Ohren nicht recht traut, dann können vielleicht diese Hinweise weiterhelfen:

- phonetische Transkription kommt *immer* zwischen eckige Klammern ([...])! Phoneme kommen zwischen Schrägstriche (/ .... /) und orthographische Zeichen kommen zwischen spitze Klammern (< ... >). Der Hauptakzent wird mit einem Apostroph vor der betonten Silbe gekennzeichnet (z.B. ['t^{h}a:l]), oder mit einem akuten Akzent "´" (z.B. [t^{h}á:l]).

- Achtung! Passt auf die Klein- und Großschreibung auf! Große Buchstaben bedeuten im phonetischen Alphabet etwas anderes als Kleinbuchstaben! z.B.
[R] = uvularer Vibrant
[r] = apikaler Vibrant

- Vokalqualität und Vokalquantität: manchen Studenten fällt es schwer, die Vollvokale korrekt zu transkribieren. Vielleicht hilft euch dabei die Generalisierung, über die wir schon im Unterricht gesprochen haben:
außer in Fremdwörtern besteht im Deutschen folgende Korrelation: gespannte Vokale sind lang, ungespannte Vokale sind kurz. Wenn ich also ein Wort wie <Düne> vor mir habe, und ich weiß nicht, wie ich den Wurzelvokal transkribieren soll, dann versuche ich zuerst einmal festzustellen, ob dieser Vokal kurz oder lang ist. Er ist eindeutig lang (sonst hätte ich ja das Wort <dünne>). Deshalb transkribiere ich [y:] (und in <dünne> transkribiere ich <ü> als [ʏ]). Und vergesst nicht den Doppelpunkt für die Vokallänge! Es gibt eine einzige Ausnahme zu der oben genannten Generalisierung: das ungespannte [ɛ] kann manchmal auch lang sein (also [ɛ:]), wie z.B. in "Ähre")

- das schwa (=[ə]) in den Flexionsendungen: in den typischen deutschen Flexionsendungen wird immer ein schwa gesprochen, nicht ein [ɛ] oder gar ein [e]. (Oder aber es wird überhaupt kein Vokal, sondern ein silbisches [n̩], [m̩] oder [l̩] gesprochen). Ein schwa taucht auch auf bei Nomen, die auf <e> enden, auch wenn es sich hier nicht um ein Flexionsmorphem handelt. Also <Mühe> = [my:**ə**], nicht [my:**e**] oder [my:ɛ]

- ein typischer Fehler: Verwechseln von Orthographie und phonetischen Zeichen. Im Deutschen schreiben wir <v>, aber dieser Buchstabe entspricht meist der Aussprache [f], wird also mit dem phonetischen Zeichen [f] transkribiert. z.B.

<ver-> in <veranlassen> = [fɛɐ]

Wer sich hier nicht sicher fühlt, sollte das IPA noch einmal durchgehen und die Zeichen heraussuchen, die nicht den orthographischen Zeichen des Deutschen entsprechen.

- vergesst die glottalen Plosive (= [ʔ]) nicht! glottale Plosive werden mindestens in folgenden Kontexten gesetzt:

- vor Vokalen am Wortanfang
- vor Vokalen am Anfang einer Wurzel
(z.B. [ʔ]er[ʔ]arbeiten)

Die meisten Leute machen außerdem noch einen im Wortinnern, wenn die darauffolgende Silbe den Hauptakzent trägt:

z.B. The[ʔ]áter

Einige wenige (?) machen auch einen glottalen Plosiv, wenn die darauffolgende Silbe einen Nebenakzent trägt:

z.B. Oze[ʔ]ànographíe

N.B.:glottale Plosive werden im Aussprache-Duden nicht transkribiert. Da das aber eine wichtige Charakteristik des Deutschen ist, will ich sie in eurer Transkription sehen

- r-Laute am Silbenanfang:

- [r] (= apikaler Vibrant = "Zungenspitzen-r"): gibt es im Deutschen so gut wie gar nicht, außer in einigen Dialekten und in der Bühnenaussprache der dreißiger Jahre (s. Hitler).

- [R] (= uvularer Vibrant): kommt selten vor, meist auch eher bei Dialektsprechern, oder höchstens am Wortanfang.

- [ʁ] (= uvularer Frikativ): ist wahrscheinlich das gebräuchlichste "r" im Deutschen (bitte das [ʁ] mit Bauch nach rechts!)

am Silbenende: am Silbenende hat das Phonem /r/ ein vokalisiertes Allophon [ɐ], d.h., am Silbenende wird jedes /r/ zu [ɐ]. <wer> wird also als [v̥e:ɐ] transkribiert, nicht als [v̥e:ʁ], [v̥e:R], oder gar [v̥e:r]! Der einzige Fall, in denen manche Sprecher am Silbenende das /r/ nicht vokalisieren ist, wenn es sich nach einem kurzen Vokal befindet. Dort dürft ihr auch ein konsonantisches [ʁ] transkribieren, wie z.B. in <wird>, transkribiert als [v̥ɪʁt$^{(h)}$].

N.B.: der Aussprache-Duden ist ziemlich schlampig bei der Transkription der <r's>, in diesem Fall kann man sich nicht auf ihn verlassen.

- das "stumme h" heißt deshalb so, weil es nicht ausgesprochen wird. Die typische Situation für jemanden, der Deutsch lernen will, sieht so aus: er weiß nicht, ob man die <h's>, die da immer zwischen Vokalen *geschrieben* stehen, ausspricht, oder nicht. Also fragt er einen Muttersprachler. Der sagt ihm das Wort schön langsam vor, und macht dabei ein ganz deutlich zu hörendes [h]. Also macht das der Deutschlernende auch, und fällt deshalb durch seine lächerliche Aussprache auf. Wenn man Deutsch lernt, sollte man wissen:
- beim ganz langsamen Sprechen (z.B. Diktieren) wird in Wörtern wie <Mühe> oft ein [h] gesprochen
- bei der normalen Aussprache fällt das [h] vollkommen weg. Also:

<Mühe> = [my:ə] nicht [my:hə]

- wie transkribiert man Diphthonge? Das Deutsche hat drei Diphthonge, die als <ei> (auch <ai>), <au> und <eu> *geschrieben* werden. Aber wie werden sie *ausgesprochen*? Da scheiden sich die Geister. Die Qualität des ersten Vokals ist ziemlich klar: [a] für die ersten beiden Diphthonge, und [ɔ] (nicht [o]!) für den dritten. Den zweiten Teil des Diphthongs transkribiert jeder anders. Ich führe hier ein paar Möglichkeiten an:

<ei>, <ai> (z.B. <ein>):	[aɪ], [ai], [aj]
<au> (z.B. <Haus>):	[aʊ], [au]
<eu> (z.B. <euch>):	[ɔɪ], [ɔi], [ɔj]

Transkriptionsbeispiele:

<Tat>	['tʰa:t(ʰ)]
<sieht>	['zi: t(ʰ)]
<wo>	['v̥o:]
<müsst>	['mʏst(ʰ)]
<Helm>	['hɛlm]
<könnt>	['kʰœnt(ʰ)]
<Zimmer>	['tsɪmɐ]
<Sonne>	['zɔnə]
<böse>	['b̥ø:zə]
<Höhe>	['hø:ə]
<welcher>	['v̥ɛlçɐ]
<Honig>	['ho:nɪç], ['ho:nɪk] (süddt.)
<Tochter>	['tʰɔxtɐ]
<Töchter>	['tʰœçtɐ]
<Paar>	['pʰa:ɐ]
<Bar>	['b̥a:ɐ]
<mehrere>	['me:ʀəʀə], ['me:ʁəʁə]
<mehr>	['me:ɐ]
<Dächer>	['d̥ɛçɐ]
<sterben>	['ʃtɛɐbən], ['ʃtɛɐbm̩]
<König>	['kʰø:nɪç]
<Sänger>	['zɛŋɐ]
<Tier>	['tʰi:ɐ]
<Söhne>	['zø:nə]
<sinken>	['zɪŋkən], ['zɪŋkŋ̍]
<Hase>	['ha:zə]
<Wasser>	['v̥asɐ]
<Strich>	['ʃtʀɪç], ['ʃtʁɪç]
<Wiese>	['v̥i:zə]
<Wissen>	['v̥ɪsən], ['v̥ɪsn̩]
<Ding>	['d̥ɪŋ]
<verehren>	['fɛɐʔe:ʀən], ['fɛɐʔe:ʁən]
<erste>	['ʔɛɐstə]
<Wind>	['v̥ɪnt(ʰ)]
<Häuser>	['hɔɪzɐ]
<Häuschen>	['hɔɪsçən]

<Arbeit>	['ʔaʁbaɪt$^{(h)}$]
<Tuch>	['t^{h}u:x]
<sehr>	['ze:ʁ]
<brav>	['b̥ʁa:f], ['b̥Ra:f]
<weitere>	['v̥aɪtəʁə], ['v̥aɪtəRə]
<süchtig>	['zʏçtɪç]
<Möhre>	['mø:ʁə], ['mø:Rə]
<doch>	['d̥ɔx]
<Hunger>	['hʊŋɐ]
<zum>	['tsʊm]

Ausgewählte Bibliographie:

Alber, Birgit (2006). Eintrag *Foot* in: Keith Brown (Hg.). *Encyclopedia of Language and Linguistics*, Second Edition, Amsterdam, Elsevier, 539-546.

Di Meola, Claudio (2003). *La linguistica tedesca. Un' introduzione con esercizi e bibliografia ragionata*. Roma, Bulzoni.

DUDEN, *Aussprachewörterbuch: Wörterbuch der deutschen Standardaussprache. 2004*. Bearbeitet von Max Mangold in Zusammenarbeit mit der Dudenredaktion. (Der Duden 6). Dudenverlag, Mannheim/Wien/Zürich.

Eisenberg, Peter (1998). *Grundriß der deutschen Grammatik. Band 1: Das Wort*. Verlag J.B. Metzler, Stuttgart, Weimar.

Fromkin, Victoria (2000) (Hrsg.), *Linguistics. An Introduction to Linguistic Theory*. Blackwell, Oxford.

Giegerich, Heinz (1992), *English Phonology. An Introduction*. Cambridge University Press, Cambridge.

Goldsmith, John A. (1995) (Hrsg.). *The Handbook of Phonological Theory*. Blackwell, Oxford

Graffi, Giorgio & Sergio Scalise (2002). *Le lingue e il linguaggio. Introduzione alla linguistica*. Bologna, Il Mulino.

Großes Wörterbuch der deutschen Aussprache (1982). Herausgegeben von dem Kollektiv Eva-Maria Krech, Eduard Kurka, Helmut Stelzig, Eberhard Stock, Ursula Stötzer und Rudi Teske. VEB Bibliographisches Institut, Leipzig.

Gussenhoven, Carlos & Haike Jacobs (1998). *Understanding Phonology*. Arnold, London.

Hall, Tracy Alan (1992). *Syllable structure and syllable-related processes in German* (= Linguistische Arbeiten 276). Max Niemeyer, Tübingen.

Hall, Tracy Alan (2000). *Phonologie. Eine Einführung*. Berlin/New York, de Gruyter.

Jakobson, Roman (1962). *Selected writings I: phonological studies*. Mouton, The Hague.

Jakobson, Roman&Morris Halle (1956). *Fundamentals of Language*. Mouton, The Hague, Paris.

Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Blackwell, Oxford.

Kohler, K.J. (1977). *Einführung in die Phonetik des Deutschen*. Schmidt, Berlin.

Ladefoged, Peter (1993). *A Course In Phonetics*. Harcourt Brace College Publishers, Fort Worth et al.

McCarthy, J. & Prince, A. (1993). 'Generalized alignment' In Booij, G. & van Marle, J. (eds.) *Yearbook of morphology 1993*. Dordrecht: Kluwer. 79-153.

McCarthy, J. & A. Prince (1995). "Faithfulness and reduplicative identity." *University of Massachusetts Occasional Papers in Linguistics* **18***: Papers in Optimality Theory*, edited by J.Beckman & S. Urbanczyk & L. Walsh, 249-384. Graduate Linguistic Student Association, Amherst.

Nespor, Marina (1999). *Fonologia*. Bologna, Il Mulino.
Pinker, Steven (1994). *The Language Instinct*. William Morrow and Company, New York. [Dt. Übersetzung (1996): *Der Sprachinstinkt*. Kindler, München.]
Prince, Alan & Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms. Rutgers University, New Brunswick, and University of Colorado, Boulder. Veröffentlicht (2004) als *Optimality Theory. Constraint Interaction in Generative Grammar,* Blackwell, Oxford.
Ramers, Karl-Heinz & Heinz Vater (1995). *Einführung in die Phonologie*. (Klage 16), Gabel-Verlag, Hürth.
Ramers, Karl-Heinz (1998). *Einführung in die Phonologie*. W. Fink Verlag, München.
Spencer, Andrew (1996). *Phonology*. Blackwell, Oxford.
Trask, R. L. (1996). *A Dictionary of Phonetics and Phonology*. Routledge, London and New York.
Vater, Heinz (1996). *Einführung in die Sprachwissenschaft*. 2. Auflage, UTB, Fink-Verlag, München.
Wiese, Richard (1996). *The Phonology of German*. Oxford University Press, Oxford.
Wurzel, W.-U. (1981). "Phonologie: Segmentale Struktur", In: Heidolph, K.E. et al. (eds.). *Grundzüge einer deutschen Grammatik*. Akademie Verlag, Berlin. 898-990.