

SpeechTrainer

Version 2.0 (JAVA)



Handbuch

© Copyright Bernd J. Kröger 2000 – 2011. Alle Rechte vorbehalten.

SpeechTrainer ist eine eingetragene Wortmarke beim Deutschen Patent- und Markenamt.

Lizenzwerb für *SpeechTrainer* (siehe auch Kap. 1.3):

Birgit Kröger und Prof. Dr. Bernd J. Kröger

Buschstr. 1e

B-4728 Hergenrath

<http://www.speechtrainer.de>

Achtung:

Dieses Handbuch ist Bestandteil eines Lizenzproduktes.

SpeechTrainer

Version 2.0 (JAVA)

Handbuch

Inhalt

1 Einführung.....	2
2 Installation und Setup	4
3 Listeneinträge auswählen.....	5
4 Eigene Wörter eingeben.....	7
5 Einzellaute darstellen.....	9
6 Die Artikulationsorgane.....	10
7 Artikulationsorgane bewegen.....	11
8 Erstellen einer eigenen Liste.....	16
9 Vertonen einer Liste.....	18
10 Synchronisieren eines Listen- eintrages.....	20
11 Zusammenstellen einer eigenen Liste.....	23
12 Nachträgliches Ändern eines Listeneintrages.....	24
13 Überblick: Menüleiste und Programmfenster.....	26
14 Die Ebene des Artikulations- Modells.....	30
15 Die Text-nach-Laut- Konversion.....	31
16 Artikulatorische Targets und Bewegungen.....	33
17 Grundlagen der artikulation- rischen und akustischen Phonetik.....	36

18 Artikulatorische und aku- stische Lautmerkmale.....	44
19 Phonetische Transkription.....	48

Achtung:

**Dieses Handbuch ist Bestandteil eines Li-
zenzproduktes.**

SpeechTrainer ist eine eingetragene Wortmarke beim
Deutschen Patent- und Markenamt.

Lizenzwerb für *SpeechTrainer* (siehe auch Kap. 1.3):

www.speechtrainer.de

info@speechtrainer.de

Birgit Kröger und Prof. Dr. Bernd J. Kröger

Buschstr. 1e

B-4728 Hergenrath

1 Einführung

1.1 Funktionen des Programms *SpeechTrainer*

Ziel des Programms *SpeechTrainer* ist die Veranschaulichung (Animation / Visualisierung) der Artikulation beim Sprechen von Einzelwörtern und kurzen Sätzen (Veranschaulichung von Sprechbewegungen).

SpeechTrainer ist zur Unterstützung der Arbeit von Logopäden, Medizinern, Sprachheilpädagogen, Sprechern, Sprachlehrern mit ihren Patienten bzw. Schülern gedacht. Das Programm soll dem Erlernen bzw. dem Verbessern der Aussprache von Einzelwörtern und kurzen Sätzen dienen.

SpeechTrainer stellt vier grundlegende Hauptfunktionen zur Verfügung:

- **Listeneinträge auswählen** ("File > open list", Kapitel 3): 1. Laden einer (eigenen oder mitgelieferten) Liste. - 2. Markieren eines Listeneintrages, d.h. eines in der Liste vorhandenen Wortes oder kurzen Satzes. - 3. Ausführung: Visualisierung der Artikulation des Listeneintrages ("play animation").
- **Eigene Wörter eingeben** ("View > sounds", Kapitel 4): 1. Texteingabe eines beliebigen Wortes - 2. Automatische Erstellung der Lautung dieses Wortes - 3. Ausführung: Visualisierung der Artikulation des Wortes ("play animation").
- **Einzellaute darstellen** ("View > sounds", Kapitel 5): 1. Auswahl eines Lautes (Vokal oder Konsonant) - 2. Visualisierung der Artikulation dieses Lautes ("play animation").
- **Einzelne Artikulationsorgane bewegen** ("View > articulators" Kapitel 6 und 7): 1. Isolierte Bewegung von einem (oder mehreren) Artikulationsorgan(en) aufgrund phonetischer Kategorien bei der Lautbildung. - 2. Visualisierung der (isolierten) artikulatorischen Bewegung ("play animation").

Dem fortgeschrittenen Benutzer von *SpeechTrainer* stehen weitere Programmfunktionen zur Verfügung:

- **Erstellen einer eigenen Liste** ("File > new list", Kapitel 8): Der Benutzer kann eigene Wörter orthographisch (Schreibung) und / oder in Lautschrift (Lautung) eingeben, darstellen und in einer eigenen Liste abspeichern. Die Wörter bzw. kurzen Sätze werden auch als Listeneinträge bezeichnet. Die entstandene Liste ist zunächst „stumm“.
- **Vertonen eines Listeneintrages** ("View > list > add new word to list", Kapitel 9): Der Benutzer kann nun zu einem in der Liste angeklickten Wort bzw. kurzen Satz (Listeneintrag) dieses Wort bzw. diesen Satz selbst einsprechen und aufzeichnen. Diese Aufzeichnung (Sprachsignal als Audiodatei) ist dann dem entsprechenden Listeneintrag zuord-

net. (Jedem Listeneintrag wird eine Audiodatei zugeordnet).

- **Synchronisieren eines Listeneintrages** (Kapitel 10): Die aufgrund der Schreibung bzw. Lautung generierten Sprechbewegungen werden von einem Artikulationsmodell erzeugt und sind zunächst sprecherunabhängig. Deshalb müssen die Sprechbewegungen mit jeder Audiodatei synchronisiert werden. *SpeechTrainer* stellt Werkzeuge für die Synchronisation von Sprachsignal und Artikulation bereit.
- **Zusammenstellen einer eigenen Liste** ("file > open list", "file > append list", Kapitel 11): Es können anhand der Einträge der mitgelieferten Listen eigene Listen zusammengestellt werden (z.B. patientenspezifische / schülerspezifische Listen mit unterschiedlichem Schwierigkeitsgrad). Es entfällt dann die Arbeit der Vertonung und Synchronisation, da viele mitgelieferte Listen bereits vertont und synchronisiert sind. Diese Funktionalität kann aber auch für selbst erzeugte Listen genutzt werden.
- **Nachträgliches Ändern eines Listeneintrages** (Kapitel 12): Alle Listeneinträge – auch die der mitgelieferten Listen – können nachträglich nach dem Geschmack des Benutzers eigenmächtig geändert werden. Die Änderung umfasst dabei alle Ebenen – beginnend mit der Schreibung über die Lautung bis hin zur Synchronisation und zu Details der Artikulatorbewegungen bei der Animation.

In Kapitel 13 wird dann noch einmal ein zusammenfassender Überblick über das Menü von *SpeechTrainer* und über alle auftretenden Formulare gegeben. Im Kapitel 14 wird ein zusammenfassender Überblick über das Hauptfenster und die Unterfenster gegeben.

Zum effektiven Umgang mit der Text-nach-Laut-Konversion (automatische Umwandlung von Schreibung in Lautung) wird diese in Kap. 15 detailliert beschrieben.

Dem fortgeschrittenen Benutzer von *SpeechTrainer* wird dann in Kap. 16 eine Einführung in das Modell zur Generierung der Artikulationsbewegungen (Label und Targets) gegeben, da eine erfolgreiche Synchronisation erst nach Verstehen der Funktion der für jeden Laut generierten Label durchgeführt werden kann.

Da ein gewisses Verständnis phonetischer Grundlagen für die Erstellung der Lautung eines Listeneintrages aber auch für die Synchronisation von selbst eingegebenen Wörtern wichtig ist, wird dem fortgeschrittenen Benutzer von *SpeechTrainer* in Kap. 17 bis 19 eine kompakte Einführung in die artikulatorische und akustische Phonetik, in artikulatorische und akustische Lautmerkmale und in die phonetische Transkription vermittelt.

1.2 Audiovisuelle Darstellung / Animation

Die Visualisierung der Artikulation erfolgt über eine **Animation** der Artikulationsbewegungen eines Modellsprechers. Basis dieser Animation sind Querschnittsbilder des Kopfes durch die Medianebene (Symmetrieebene des Menschen), wie sie bei einer Magnet-Resonanz-Tomographie (MRT, englisch: MRI) oder Röntgenaufnahme entstehen. Der Modellsprecher von *SpeechTrainer* wurde anhand von MRT-Daten generiert. Synchron zur Visualisierung der Artikulation kann das Wort auch hörbar gemacht werden (Audioausgabe). *SpeechTrainer* ermöglicht somit eine audiovisuelle Darstellung von Sprechbewegungen einzelner Wörter bzw. kurzer Sätze.

1.3 Nutzungsrechte zum Programm *SpeechTrainer*

Der Benutzer von *SpeechTrainer* benötigt eine *auf seinen Namen lautende* gültige Nutzungslizenz. Zum Erwerb einer solchen Lizenz wenden Sie sich bitte an:

info@speechtrainer.de

<http://www.speechtrainer.de>

**Birgit Kröger und Prof. Dr. Bernd J. Kröger,
Buschstr. 1e, B-4728 Hergenrath.**

Wir machen darauf aufmerksam, dass die Nutzung von *SpeechTrainer* ohne gültige und auf Ihren Namen lautende Nutzungslizenz nicht erlaubt ist.

Dieses Handbuch ist Bestandteil der Nutzungslizenz.

Sie ermöglichen durch Einhaltung der Lizenzbedingungen die Weiterentwicklung des Programms und die Aufrechterhaltung aller Serviceleistungen rund um *SpeechTrainer*.

Wir hoffen, dass Sie *SpeechTrainer* erfolgreich und gewinnbringend einsetzen können und wünschen Ihnen nun beim Lesen dieses Handbuches und beim Arbeiten mit *SpeechTrainer* viel Spaß.

Hergenrath, im Juli 2011

Birgit Kröger und Bernd J. Kröger

2 Installation und Setup

2.1 Minimale Systemanforderungen

Zum Betrieb von *SpeechTrainer* benötigen Sie einen PC mit den folgenden Komponenten:

- Aktueller Desktop-Rechner oder aktuelles vollwertiges Notebook
- Windows XP oder höher
- Java Echtzeit Umgebung (Java Runtime Environment, JRE) von Sun Microsystems; diese ist frei im Internet erhältlich (z.B. auch zum Betrieb von open office), siehe: <http://www.java.com>. Auf vielen Rechnern ist JAVA aber bereits installiert.
- Audiokarte und Kopfhörer oder Lautsprecher
- USB-Headset mit qualitativ hochwertigem Mikrofon, z.B. Sennheiser Headset (Nur erforderlich, wenn die Erstellung eigener Listen mit *selbst gesprochener* Audioausgabe gewünscht wird.)

Anmerkung: Auf Netbooks oder Tablets kann es wegen der geringen Leistungsfähigkeit dieser Rechner zum „Ruckeln“ der Animationen kommen.

Das Programm ist JAVA-basiert und ist somit auch auf APPLE-Rechnern lauffähig.

2.2 Installation von *SpeechTrainer*

Kopieren Sie die Datei „Speechtrainer20.jar“ aus dem Internet herunter (www.speechtrainer.de > Software) herunter in ein Verzeichnis Ihrer Wahl auf Ihrem Rechner. Doppelklicken Sie dann „SpeechTrainer20.jar“. Es öffnet sich dann ein Dialogfenster zur Eingabe des Lizenzcodes für das Programm *SpeechTrainer*.

Lizenzierung: Geben Sie in diesem Dialogfenster die von uns erhaltenen Namen und anschließend die Lizenznummer ein. (Die Daten hierzu erhalten Sie per email über info@speechtrainer.de)

Anmerkung: Es reicht nicht die Eingabe der Lizenznummer. Auch Name und evtl. weitere Zusätze müssen - wie in der email aufgeführt - detailliert eingegeben werden. Am Ende jeder Zeile darf kein weiteres Leerzeichen eingegeben werden.

Klicken Sie anschließend auf das Schaltfeld „Lizenzieren“. Der Lizenzeditor verschwindet und das **Hauptfenster** des Programms *SpeechTrainer* wird automatisch geöffnet.

Klicken Sie nun im (aktuell sichtbaren) Listenfenster auf die Schaltfläche „Play Animation“ an. Sie erhalten eine artikulatorische Animation, d.h. eine Visualisierung der Sprechbewegungen des aktuellen Listeneintrages „Bitte“. Sie schließen die Anwendung *SpeechTrainer* durch den Menübefehl „File > Quit“.

2.3 Installation von mitgelieferten Listen auf Ihrem PC

Beim Start von *SpeechTrainer* erscheint im Listenfeld nur eine kurze, fest im Programm eingebaute **Liste**. Diese Liste enthält nur fünf Einträge: „Bitte“ „öffnen“ „Sie“ „eine“ „Liste“. Sie können jedes Wort (jeden Intrag der Liste) anklicken und durch Anklicken von „Play Animation“ die zugehörige Animation aktivieren. Die Animationen aller Einträge dieser Liste sind allerdings ohne Audioausgabe (ohne Ton).

Umfangreichere Listen und insbesondere auch Listen mit Audioausgabe finden Sie im Internet (www.speechtrainer.de > download > vertonte Wortlisten).

Anmerkung: Die unter SpeechTrainer 1.0 erstellten Listen können mit *speziellen Konverterprogrammen* in das SpeechTrainer-2.0-Format konvertiert werden (siehe www.speechtrainer.de).

2.4 Deinstallation von *SpeechTrainer*

Löschen Sie einfach die Datei „speechtrainer.jar“ auf Ihrem Rechner.

Es ist keine weitergehende Installation des Programms über ihr Betriebssystem erfolgt. Deshalb finden Sie das Programm auch nicht in ihrem Programm-Menü. Sie können das Programm also immer auch nur aus dem Ordner aufrufen, in das Sie das Programm hineinkopiert haben.

Anmerkung: die von Ihnen mit dem Programm erzeugten Daten-Dateien müssen Sie ebenfalls „per Hand“ löschen.

3 Listeneinträge auswählen

Grundsätzlich muss keine Liste geladen werden. Sie können immer auch direkt das gerade benötigte Wort per Hand eintippen und dann die Artikulation (zumindest stumm) visualisieren (siehe Kap. 4). In der Therapie- bzw. Unterrichtssituation kann es aber sinnvoll sein, auf bereits vorbereitete Wörter und damit auf bereits vorab angelegte Listen zurückzugreifen.

Je nach Patient oder Schüler können Sie aber auch eigene Listen erstellen (Kap. 8 bis Kap. 11). Zunächst können Sie aber auch auf die mitgelieferten Listen zurückgreifen. Sie müssen sich dann während der Arbeit mit dem Schüler oder Patienten nicht mit phonetischer Lautschrift, nicht mit der Erstellung von Audiodateien und nicht mit deren Synchronisation beschäftigen.

In diesem Kapitel wird das Laden einer mitgelieferten (d.h. im Netz abrufbaren) Liste und das Arbeiten mit der Liste erläutert.

3.1 Starten des Programms *SpeechTrainer*

Nach erfolgreicher Installation von *SpeechTrainer* (Kap. 2.2) starten Sie das Programm *SpeechTrainer* durch Anklicken von „SpeechTrainer20.jar“. Es öffnet sich sofort das **Hauptfenster** des Programms.

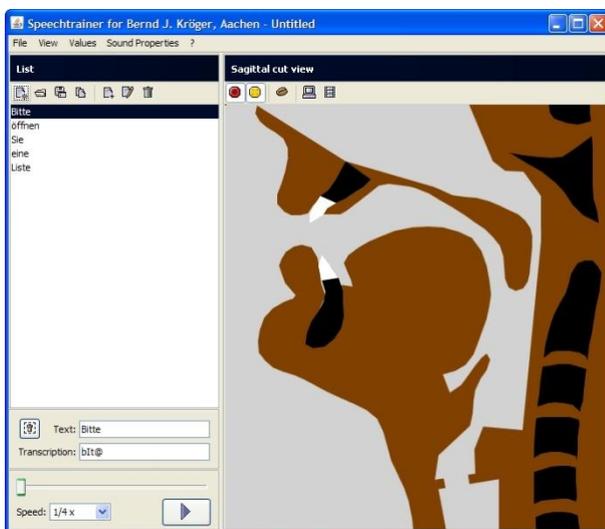


Abbildung 1. Das Hauptfenster von *SpeechTrainer*.

3.2 Die Oberfläche des Hauptfensters von *SpeechTrainer*

Das direkt nach dem Programmstart sichtbare **Hauptfenster** von *SpeechTrainer* untergliedert sich unterhalb der **Menüleiste** (File, View, Values, ...) in vier Bereiche (auch **Funktionsfenster** oder **Unterfenster**, Kap. 11).

- **Listenfenster** („list“, links): Darstellung der Listeneinträge (weißer Bereich) der aktuell geladene

Liste. Die Schaltfläche „Play Animation“ befindet sich ebenfalls im Listenfenster.

- **Gesprochen-Fenster** („transcription“, links unten): Lautung (Lautschrift) des **aktuellen Eintrages**.
- **Geschrieben-Fenster** („text“, links unten): (Orthographische) Schreibung des aktuellen Eintrages.
- **Animationsfenster** („sagittal cut view“, rechts): Darstellung der Artikulationsorgane und Animation der Sprechbewegungen.

Im Listenfenster aktivieren Sie durch Anklicken von „Play Animation“ die artikulatorische Animation zum **aktuellen Eintrag** der Liste (das blau unterlegte Wort; z.B. das Wort „Bitte“). Durch Anklicken einzelner Einträge im weißen Bereich des Listenfensters können Sie jeden Eintrag der Liste anwählen (blauer Hintergrundstreifen) und durch „Enter“ oder Anklicken des „Play“-Button können Sie die visuelle Animation starten.

Anmerkung: Bei „schwachen“ Netbooks (niedrige Rechenleistung) kann es sein, dass die Animation „ruckelt“. Dies ist aber auf gängigen Desktop-Rechnern oder vollwertigen Notebooks nicht der Fall.

Klicken Sie nun auf den Eintrag „Sie“ im Listenfenster. Dieser Eintrag wird nun zum **aktuellen Eintrag** (blauer Streifen). Der **aktuelle Eintrag** erscheint ebenfalls orthographisch im Geschrieben-Fenster und mit seiner Lautschrift im Gesprochen-Fenster. Durch Anklicken von „Play Animation“ aktivieren Sie nun die artikulatorische Animation zum aktuellen Eintrag (z.B. „Sie“).

Anmerkung: Das **Listenfenster** ist nur eines von **drei Funktionsfenstern** (**Unterfenstern** des Hauptfensters) von *SpeechTrainer*. Die beiden anderen Unterfenster können Sie durch Anklicken von „View > Sounds“ (**Lautfenster**) oder von „View > Articulators“ (**Artikulatoren-Fenster**) aktivieren. Alle drei Unterfenster erscheinen an dieser Stelle des Hauptfensters (linke Seite des Hauptfensters). Andere **Dialogfenster** erhalten Sie, indem Sie in der Menüleiste z.B. auf „Values > Sounds > Table“ klicken. Bitte stellen Sie das Hauptfenster von *SpeechTrainer* aber jetzt wieder auf das Listenfenster ein (Menü: „View > List“).

3.3 Öffnen einer mitgelieferten Liste

Beim Start von *SpeechTrainer* erscheint im Listenfeld nur eine kurze, fest im Programm eingebaute **Default Liste**. Diese Liste enthält nur die fünf Einträge „Bitte“, „öffnen“, „Sie“, „eine“, „Liste“. Weitere Listen und Listen mit Audioausgabe befinden sich im Netz unter www.speechtrainer.de > downloads“. Bitte kopieren Sie diese Listen zunächst aus dem Netz in einen Ordner auf Ihrem Rechner. Zum Öffnen einer kopierten Liste gehen Sie nun im **Menü** von *SpeechTrainer* auf: „File > open list“. Wechseln Sie dann zu dem Ordner, in den Sie die Listen aus dem Netz kopiert haben. Doppelklicken Sie nun z.B. auf die Datei „Listen\ Start\ Liste_01.st“. Sie haben nun diese Liste geladen und geöffnet. Die Einträge dieser Liste erscheinen nun im Listenfeld („Panne, Tanne, Kanne“).

Anmerkung: Es sollten nur die mit *SpeechTrainer* mitgelieferten bzw. von Ihnen selbst erstellten Listendateien über „Datei > öffnen“ geöffnet werden. Andere *.st-Dateien können von *SpeechTrainer* nicht sinnvoll gelesen werden.

3.4 Markieren eines Listeneintrages

Ein **Listeneintrag** ist ein in einer Liste vorhandener Eintrag, d.h. ein Wort oder ein kurzer Satz. Die Einträge der geöffneten Liste (z.B. „Liste_01.st“) erscheinen im Listenfeld von *SpeechTrainer*. Klicken Sie nun im Listenfenster auf den Eintrag „Tanne“. Dieser Eintrag ist nun der **aktuelle Eintrag** und erscheint auch im Geschrieben- und Gesprochen-Fenster.

Anmerkung: Die in *SpeechTrainer* geladene bzw. geöffnete Liste wird gelegentlich auch als **aktuelle Liste** bezeichnet. *SpeechTrainer* kann jeweils nur eine Liste laden und öffnen.

3.5 Ausführung: Visualisierung des Listeneintrages

Klicken Sie nun im Listenfenster auf die Schaltfläche „Play Animation“. Sie starten damit die artikulatorische Animation des aktuellen Eintrages. Sie sehen und hören nun die Animation des aktuellen Listeneintrages.

Achtung: Bitte warten Sie immer das Ende der Animation ab, bevor Sie eine weitere Aktion (Maus oder Tastatur) in *SpeechTrainer* ausführen. Beim Unterbrechen einer laufenden Animation kann es zu Abstürzen kommen.

Falls Ihnen die Animation zu langsam oder zu schnell abläuft, können Sie links unter unter „speed“ die aktuelle Ablaufgeschwindigkeit der Animation ändern.

Sie können die Animation aber auch mittels des **Schiebereglers** vorwärts oder auch rückwärts ablaufen lassen und an jeder beliebigen Stelle anhalten.

Anmerkung: Nicht jede Liste ist auch mit Audiodaten verknüpft. Beim Fehlen von Audiodateien (unvertonte und nicht synchronisierte Listeneinträge) bleibt die Animation stumm.

3.6 Darstellung der Animation ändern

Im rechten Teil des Hauptfensters („Sagittal cut view“) können Sie durch Klicken auf die Schaltfläche „Switch color theme“ die Grundfarben der Animation ändern.

Durch Anklicken der Schaltflächen „Show Phonation Points“ bzw. „Show Noise Points“ kann die Animation der Phonation bzw. die Animation des Reibegeräusches an unterschiedlichen Artikulationsstellen ein- bzw. ausgeschaltet werden.

Anmerkung: Lesen Sie hierzu auch Kap. 13.

3.7 Ändern der Darstellungsgröße des Animationsfensters

Wenn Sie die Darstellungsgröße des Animationsfensters ändern wollen, gehen Sie mit der Maus auf den **unteren Rand** des Programmfensters von *SpeechTrainer*, so dass sich das Aussehen des Mauszeiger zu einem Doppelpfeil wandelt. Ziehen Sie bei gedrückter linker Maustaste nun diese Ecke des Programmfensters auf die gewünschte Größe.

3.8 Einen anderen Listeneintrag wählen und ausführen

Nach Ablauf der Animation klicken Sie auf einen anderen Listeneintrag im Listenfenster (z.B. „Panne“). Dieses Wort wird nun zum aktuellen Listeneintrag (blauer Streifen) und erscheint auch im Geschrieben- und Gesprochen-Fenster. Klicken sie nun auf „Play Animation“. Sie sehen nun die Animation zum neuen aktuellen Eintrages.

3.9 Die mitgelieferten Listen von *SpeechTrainer*: Vertonte und stumme Listen

Da das Programm *SpeechTrainer* in der Version 2.0 erst seit kurzer Zeit existiert, ist die Anzahl der mitgelieferten Listen noch relativ begrenzt. Alle Listen finden Sie unter:

<http://www.speechtrainer.de>

Anmerkung: Die Liste „Listen\ Start\ Liste_01.st“ ist stumm. Laden Sie nun eine weitere Liste, nämlich die Liste „Listen\ Start\ Liste_02.st“. Es erscheinen die Listeneinträge „Banane, Tomate, Kiwi“. Diese Liste ist vertont. Wenn Sie die Animation starten („Play animation“), können Sie nun gleichzeitig das Wort in der Audioausgabe hören.

Anmerkung: Um die Stimme in der richtigen Tonhöhe zu hören, müssen Sie die Wiedergabegeschwindigkeit der Animation im Fenster „speed“ auf *normal* „1x“ stellen.

3.10 Beenden von *SpeechTrainer*

Nach Abschluss Ihrer Arbeiten mit einer Liste können Sie das Programm beenden. Klicken Sie dazu Menü „List > Quit“ an. Das Hauptfenster von *SpeechTrainer* wird nun geschlossen und die Ausführung der Anwendung (des Programms) *SpeechTrainer* wird beendet.

4 Eigene Wörter eingeben

Wenn Ihnen die Auswahl an Wörtern oder kurzen Sätzen in den mitgelieferten Listen nicht ausreicht, so haben Sie die Möglichkeit, mit *SpeechTrainer* eigene Wörter einzugeben und auszuführen. *SpeechTrainer* erzeugt anhand Ihrer Eingabe zur Schreibung oder Lautung eines Wortes die zugehörigen Sprechbewegungen (Animation).

4.1 Texteingabe eines beliebigen Wortes und automatische Transkription

Klicken Sie nach Start des Programms *SpeechTrainer* unten links auf das **Geschrieben-Fenster** („Text“). Klicken Sie auf das Texteingabefeld des Textfensters (weißer Bereich), löschen Sie den dortigen Text (z.B. das Wort „Bitte“) und geben Sie das Wort „hallo“ (oder „Hallo“) ein. Es erscheint gleichzeitig die phonetische Transkription im **Gesprochen-Fenster** („Transcription“, Abb. 2);

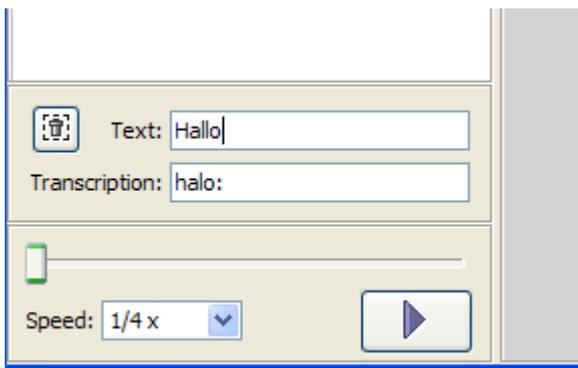


Abbildung 2. Das Gesprochen- und Geschrieben-Fenster von *SpeechTrainer* mit „Play Animation“-Schaltfläche, mit Schieberegler zur Animation und mit Auswahlfenster für die Ablaufgeschwindigkeit.

Anmerkung: Zur einfachen Unterscheidung von Orthographie (Schreibung) und Lautschrift (Lautung, phonetische Transkription) wird die Lautschrift in der phonetischen Literatur meist in eckigen Klammern notiert; z.B. Schreibung: „Hallo“ ergibt die Lautung [halo:]. In diesem Handbuch wird aber aus Gründen der Übersichtlichkeit jede Eingabe des Benutzers und jede Ausgabe des Programms in Anführungszeichen gesetzt („halo:“).

Die von *SpeechTrainer* generierte Lautfolge (Lautung) basiert auf einem sehr einfachen Algorithmus und kann zuweilen fehlerhafte Ergebnisse liefern. Sie können die Lautfolge aber im Texteingabefeld des Lautfensters korrigieren bzw. beliebig ändern.

Beispiel: Klicken Sie das Geschrieben-Fenster an. Tippen Sie im Texteingabefeld das Wort „geben“ ein. Im Texteingabefeld des Gesprochen-Fensters steht nun die von *SpeechTrainer* generierte Lautfolge „gEb@n“.

Klicken Sie auf die Texteingabe im Gesprochen-Fenster und ändern Sie die Lautfolge nach „geb@n“.

Mit dem Anklicken des **Gesprochen-Fensters** wird automatisch auch das **Lautfenster** „View > sounds“ aktiviert (Abb. 3, siehe auch Kap. 5). Damit hat der Benutzer die Möglichkeit, die möglichen phonetischen Eingabesymbole zu überblicken.

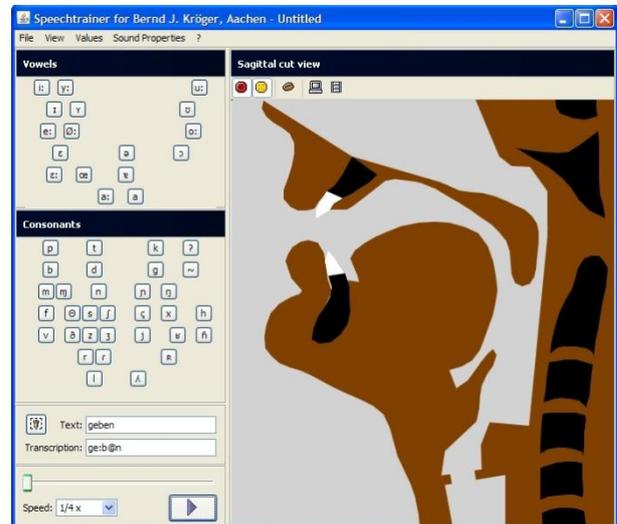


Abbildung 3. Das Lautfenster von *SpeechTrainer*.

4.2 Korrektur der Lauteingabe via Symbolschaltflächen

Falls Sie z.B. die Bedeutung des Klammeraffen in der Lautschrift des Wortes „geben“ nicht verstehen, können Sie dieses Symbol im Gesprochen-Fenster löschen und dann z.B. in der Vokaltabelle des Lautfensters (oben links) auf das Symbol des Schwa-Lautes klicken. Dieser wird dann an der entsprechenden Stelle im Gesprochen-Fenster eingefügt. Sie sehen dann, dass also der Schwa-Laut im Gesprochen-Fenster immer als Klammeraffe ausgedrückt wird.

Auf diese Art können Sie sich die Transkription ganzer Wörter mittels des Anklickens von Lautsymbolen in der Vokal- und Konsonant-Tabelle des Lautfensters (Abb. 3) zusammenstellen.

4.3 Erneute Eingabe eines Textes

Sie können nun damit beginnen, ein neues Wort einzugeben und zu visualisieren. Klicken Sie dazu nun wieder das **Geschrieben-Fenster** an. Löschen Sie den alten Eintrag „geben“ (Schaltfläche des Mülleimers, Abb. 2) und geben Sie das Wort „begabt“ ein. Es erscheint die Lautfolge „bEgapt“. Klicken Sie wieder auf das **Geschrieben-Fenster**. Geben Sie im Texteingabefeld des Textfensters nun „be-gabt“ ein. Es erscheint die weniger fehlerhafte (aber noch nicht vollständig korrekte) Transkription „b@gapt“. Die Korrektur des [a] zum [a:] müssen Sie „per Hand“ durch Änderung im Gesprochen-Fenster generieren.

Anmerkung: Auch fehlerhafte Transkriptionen werden von *SpeechTrainer* ausgeführt. Zum Erlernen der Grundlagen der phonetischen Transkription arbeiten Sie bitte Kapitel 19 durch. Es ist aber auch möglich, die phonetische Transkription anhand der Transkriptionen der Beispielwörter in den mitgelieferten Listen intuitiv zu erlernen.

Anmerkung: Es ist nicht in jedem Fall von Vorteil, Silbengrenzen oder Morphemgrenzen durch „-“ in der Orthographie zu markieren. Probieren Sie die „Text -> Laut“ daher immer zuerst *ohne* Eingabe von Trennungszeichen (siehe auch Kap. 15).

Anmerkung: Eine Trennung von Wörtern kann durch Leerzeichen oder durch „_“ erfolgen. Ein Eintrag, der aus mehreren Wörtern besteht, muss vor der Aufnahme in eine Liste (Kap. 8) aber so aufbereitet werden, dass die Einzelwörter dieses Eintrages in jedem Fall durch „_“ getrennt bzw. verbunden werden. Es dürfen dann nämlich keine Leerzeichen in der Schreibung des Listeneintrages auftreten.

5 Einzellaute darstellen

Lautsymbole Glottisverschlusslaut „ʔ“, Glottalisierung „~“, stimmhafter glottaler Frikativ „h“ und stimmloser glottaler Frikativ „h“ - siehe Kap. 17 bis 19.

Bevor Sie sich intensiver mit Artikulationsbewegungen in Wörtern und kurzen Sätzen befassen, sollten Sie zunächst die *Stellung der Artikulationsorgane* bei Einzellaute (bei Vokalen und Konsonanten) studieren.

5.1 Vokale

Klicken Sie Menü “View > sounds“ an. Es erscheint das **Lautfenster**. Im oberen Bereich des Lautfensters finden Sie die Vokale des Deutschen nach dem Prinzip des phonetischen Vokalraumes angeordnet. Von oben nach unten sind die **hohen bis tiefen Vokale** angeordnet. Von links nach rechts sind die **vorderen bis hinteren Vokale** angeordnet. Die dunkelgrauen Schaltflächen markieren die **gerundeten Vokale**, die hellgrauen Schaltflächen markieren die **ungerundeten Vokale**. (Phonetische Grundlagen werden in Kap. 17 bis 19 erläutert.)

Übung: Klicken Sie nun z.B. im Lautfenster „i:“, „a:“ oder „u:“ an. Das Animationsfenster zeigt dann die zugehörige Einstellung der Artikulationsorgane an.

Anmerkung: Die Kurzvokale und die reduzierten Vokale des Deutschen können nicht in Isolation produziert werden. Sie werden hier im Wortkontext in initialer und medialer Position produziert. Am Ende der Liste werden die reduzierten Vokale Schwa-Laut „@“ und R-gefärbter Schwa-Laut „A“ jeweils in finaler und medialer Position vorgestellt.

5.2 Konsonanten

Klicken Sie Menü “View > sounds“ an. Es erscheint das **Lautfenster**. Im mittleren Bereich des Lautfensters finden Sie die Konsonanten des Deutschen nach phonetischen Merkmalen angeordnet. Die Zeilen geben unterschiedliche **Artikulationsarten** an: **Plosive, Nasale, Frikative, Vibranten, Laterale**. Die Spalten geben unterschiedliche **Artikulationsorte** an: **labial, alveolar, postalveolar, palatal, velar, glottal**. (Phonetische Grundlagen werden in Kap. 17 bis 19 erläutert).

Übung: Klicken Sie nun z.B. im Lautfenster „p“, „n“ oder „j“ an. Das Animationsfenster zeigt dann die zugehörige Einstellung der Artikulationsorgane bei diesen Lauten.

Anmerkung: Zur Bedeutung des gelben Punktes z.B. bei „j“ siehe Kap. 7.

Anmerkung: Viele Konsonanten können nicht in Isolation produziert werden. Sie werden deshalb hier im Kontext von Silben oder von kurzen Wörtern produziert. Es werden für jeden konsonantischen Laut 3 Beispielwörter angegeben, in denen der entsprechende Laut in initialer, medialer bzw. finaler Position auftritt. Der R-Laut wird in 2 Realisierungen (Zungenrücken-R, frikatives R) dargeboten.

Zur genauen Erläuterung der Lautsymbole - insbesondere frikatives R-Realisierung „X“, und der glottalen

6 Die Artikulationsorgane

Die **Artikulationsorgane** (Zungenspitze, Zungenrücken, Lippen, Gaumensegel, Stimmlippen) können in dieser Version 2.0 des *SpeechTrainer* nicht mehr einzeln markiert, wohl aber nach wie vor einzeln bewegt werden (siehe Kap. 7).

Eine Definition der Artikulationsorgane ist somit am einfachsten möglich, indem gleichzeitig die grundlegenden Bewegungsmöglichkeiten, die jeder Artikulator hat, erläutert werden (Kap. 7).

7 Artikulationsorgane bewegen

Wählen Sie Menü „View > Articulators“. Es erscheint das **Artikulatoren-Fenster**. Dieses Fenster zeigt neun Schieberegler („vokalische Gesamtformung“ / „vocalic tract shape“ bis „Glottis“). Mittels der Schieberegler können Sie nun die Lage und die Form einzelner Artikulationsorgane verändern.

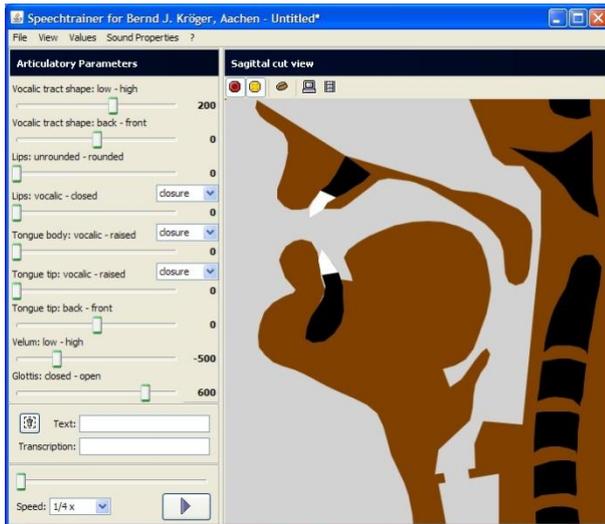


Abbildung 4. Das Artikulatoren-Fenster von *Speech-Trainer*.

In diesem Kapitel werden die mit den Schieberegler verbundenen **artikulatorischen Parameter** erläutert. Am besten können Sie die artikulatorischen Parameter verstehen, wenn Sie mit den Schieberegler spielen und die resultierenden Änderungen im Animationsfenster beobachten.

Übung: Ändern Sie mit der Maus z.B. die Position des Schieberegler „vokalische Gesamtformung: tief - hoch“ („vocalic tract shape: low – high“). Sie können nun im Animationsfenster die Änderung der Zungenlage je nach Einstellung des Schieberegler beobachten.

Anmerkung: Die artikulatorischen Parameter sind **funktionale Parameter** und damit direkt auf das Sprechen ausgerichtet. Es handelt sich nicht um rein physiologische Parameter, die z.B. immer genau einen Artikulator positionieren, sondern um Parameter der *Sprechbewegungen*. Daher sind die artikulatorischen Parameter auch nach den Kategorien „vokalisches“ bzw. „konsonantisches“ organisiert.

Aufgrund der von *SpeechTrainer* vorgegebenen Parameter kann der Schüler bzw. Patient direkt die bei den Einzellaute auftretenden Bewegungen und Einstellungen der Artikulationsorgane erlernen.

Übung: Ändern Sie jeweils nur einen artikulatorischen Parameter und lassen Sie den Schüler bzw. Patienten die zugehörige Artikulationsbewegung bzw. die zugehörige Einstellung der Artikulationsorgane nachvollziehen.

7.1 Vocalic tract shape: low – high / back – front (VTS:Ih) (VTS:bf)

Der Parameter „vokalische Gesamtformung: tief - hoch“ stellt im Sinne des artikulatorischen Vokalraumes (siehe Lautfenster) die Dimension tiefe Vokale (z.B. [a] oder [A]) - hohe Vokale (z.B. [i], [y] oder [u]) dar.

Der Parameter „vokalische Gesamtformung: hinten - vorne“ stellt im Sinne des artikulatorischen Vokalraumes (siehe Lautfenster) die Dimension hintere Vokale (z.B. [u] oder [o]) - vordere Vokale (z.B. [i], [e], [y], [ö]) dar.

Die beiden Parameter der „vokalischen Gesamtformung“ wirken sich in erster Linie auf die Lage und Formung des Zungenrückens aus. Sie beschreiben aber auch die Positionierung des Kiefers sowie die Mitbewegung der Lippen und des Kehlkopfes bei der Produktion von Vokalen. Ebenfalls werden durch diese Parameter leichte Mitbewegungen des Gaumensegels bei der Vokalproduktion beschrieben.

Übung: Stellen Sie den Schieberegler für den Parameter „vokalische Gesamtformung: vorne – hinten“ nach ganz rechts (Wert +1000). Der Parameter „vokalische Gesamtformung: vorne – hinten“ ist damit auf „vorne“ eingestellt. Variieren Sie nun den Parameter „vokalische Gesamtformung: tief – hoch“ langsam von ganz links (Wert -1000) nach ganz rechts (Wert +1000), also von der Parametereinstellung „unten“ bis „oben“. Das Animationsfenster zeigt nun einen Artikulationsverlauf von [a] über [E] und [e] bis [i].

Übung: Klicken Sie Menü „View > sounds“ an. Es erscheint das Lautfenster. Klicken Sie auf „i:“ (oben links). Das Animationsfenster zeigt nun die Artikulation eines [i:] an. Klicken Sie nun Menü „View > articulators“ an. Es erscheint das Artikulatoren-Fenster. Die Schieberegler der beiden Parameter der „vokalischen Gesamtformung“ sind nun nach ganz rechts verschoben (Wert: +1000). Variieren Sie nun den Parameter „vokalische Gesamtformung: tief – hoch“ langsam von ganz rechts (Wert +1000) nach ganz links (Wert -1000), also von der Parametereinstellung „oben“ bis „unten“. Das Animationsfenster zeigt nun einen Artikulationsverlauf von [i:] über [e:] und [E:] bis [a:].

7.2 Lips: unrounded – rounded (Ll:ur)

Der Parameter „Lippen: ungerundet – gerundet“ stellt im Sinne des artikulatorischen Vokalraumes die Dimension ungerundete – gerundete Vokale dar. Ungerundete Vokale sind z.B. [i], [e] oder [a], gerundete Vokale sind z.B. [y], [ö], [o] oder [u]. Der Parameter wirkt sich in erster Linie auf die Formung der Lippen aus, beeinflusst aber zu einem kleinen Anteil auch die Kieferposition.

Übung: Stellen Sie im Artikulatoren-Fenster den Schieberegler „vokalische Gesamtformung: tief – hoch“ auf „hoch“ (nach ganz rechts; Wert +1000) und den Schieberegler „vokalische Gesamtformung: hinten – vorne“ auf „vorne“ (also ebenfalls nach ganz rechts; Wert

+1000). Variieren Sie nun den Schieberegler „Lippen: ungerundet – gerundet“ von „ungerundet“ hin zu „gerundet“ (also von ganz links nach ganz rechts; von Wert 0 zu Wert +1000). Das Animationsfenster zeigt nun den Übergang der Artikulation von [i:] nach [y:].

Übung: Klicken Sie Menü „View > sounds“ an. Es erscheint das Lautfenster. Klicken Sie auf „y:“. Das Animationsfenster zeigt nun die Artikulation des Vokals [y:]. Klicken Sie nun Menü „View > articulators“ an. Es erscheint das Artikulatoren-Fenster. Die Werte der Schieberegler sind nun auf die artikulatorischen Parameter des Vokals [y:] eingestellt. Variieren Sie nun den Schieberegler „Lippen: ungerundet – gerundet“ von „gerundet“ hin zu „ungerundet“ (also von ganz rechts nach ganz links; von Wert +1000 zum Wert 0). Das Animationsfenster zeigt nun den Übergang der Artikulation von [y:] nach [i:].

Anmerkung: Die Zuordnung von *SpeechTrainer*-Lautsymbolen und Phonetikzeichen (nach IPA) wird in Kap. 19 angegeben.

7.3 Lips: vocalic – closed (Ll:vc)

Alle weiteren artikulatorischen Parameter des Zungenrückens und der Zungenspitze sind **konsonantische** artikulatorische Parameter. Der Basiswert 0 repräsentiert bei allen konsonantischen Parametern das Fehlen von konsonantischer Artikulation. Haben alle konsonantischen artikulatorischen Parameter den Wert 0 (oder im Fall der Stimmlippen einen Wert von fast 0), dann richtet sich die Positionierung und Formung aller Artikulationsorgane ausschließlich nach der obigen drei vokalischen artikulatorischen Parametern (**vokalisches Positionierung**, Kap. 7.1 und 7.2).

Bei dem artikulatorischen Parameter „Lippen: vokalisches – verschlossen“ (Lese: „Lippen: vokalisches bis verschlossen“) repräsentiert der Wert 0 des Schiebereglers die momentane vokalische Formung der Lippen. Der Wert +1000 repräsentiert den vollen Verschluss der Lippen (z.B. bei der Artikulation eines [b]).

Übung: Klicken Sie Menü „View > sounds“ an. Es erscheint das Lautfenster. Klicken Sie auf den Vokal „E:“. Im Animationsfenster erscheint die zugehörige Artikulation. Klicken Sie nun Menü „view > articulators“ an. Es erscheint das Artikulatoren-Fenster. Alle Schieberegler repräsentieren nun die artikulatorische Einstellung des Vokals [ε:]. Bewegen Sie nun den Schieberegler „Lippen: vokalisches – geschlossen“ von ganz links („vokalisches“, Wert 0) nach ganz rechts („verschlossen“, Wert +1000). Im Animationsfenster sehen Sie nun eine kontinuierliche Schließbewegung der Lippen.

Übung: Klicken Sie Menü „view > sounds“ an. Es erscheint das Lautfenster. Klicken Sie auf den Vokal „i:“. Im Animationsfenster erscheint die zugehörige Artikulation. Klicken Sie nun Menü „view > articulators“ an. Es erscheint das Artikulatoren-Fenster. Alle Schieberegler repräsentieren nun die artikulatori-

sche Einstellung des Vokals [i:]. Bewegen Sie nun wieder den Schieberegler „Lippen: vokalisches – geschlossen“ von ganz links („vokalisches“, Wert 0) nach ganz rechts („verschlossen“, Wert +1000). Im Animationsfenster sehen Sie nun eine kontinuierliche Schließbewegung der Lippen.

Anmerkung: Die Lippen führen im Fall des [ε:] eine wesentlich größere Schließbewegung aus als im Fall des [i:]. Dies resultiert aus der Tatsache, dass die vokalische Ausgangsstellung der Lippen bei verschiedenen Vokalen unterschiedlich ist. Bei einem [i:] sind die Lippen weniger weit geöffnet als bei einem [ε:].

7.4 Tongue body: vocalic – raised (TB:vr)

Bei dem Parameter „Zungenrücken: vokalisches – gehoben“ repräsentiert der Wert 0 (Schieberegler ganz links) die momentane **vokalisches** Positionierung und Formung des Zungenrückens. Der Wert + 1000 (Schieberegler ganz rechts) repräsentiert den vollen konsonantischen Verschluss des Zungenrückens (z.B. bei der Artikulation eines [k]).

Übung: Verändern Sie den Wert dieses artikulatorischen Parameters durch Betätigung des Schiebereglers.

7.5 Tongue tip: vocalic – raised (TT:vr)

Bei dem Parameter „Zungenspitze: vokalisches – gehoben“ repräsentiert der Wert 0 (Schieberegler ganz links) die momentane **vokalisches** Positionierung und Formung der Zungenspitze. Die Zungenspitze ist in dem Fall nicht aktiv angehoben. Der Wert + 1000 (Schieberegler ganz rechts) repräsentiert den vollen konsonantischen Verschluss durch die Zungenspitze (z.B. bei der Artikulation eines [t]).

Übung: Verändern Sie den Wert dieses artikulatorischen Parameters durch Betätigung des Schiebereglers.

7.6 Tongue tip: back – front (TT:bf)

Der Parameter „Zungenspitze: hinten – vorne“ repräsentiert den genauen Ort der Verschluss- oder Engebildung durch die Zungenspitze. Der Wert -1000 (Schieberegler ganz links) repräsentiert den dentalen Artikulationsort (z.B. beim „th“ im Englischen), der Wert 0 (Schieberegler in der Mitte) repräsentiert den alveolaren Artikulationsort (z.B. [t], [d] oder [s]) und der Wert +1000 (Schieberegler ganz rechts) repräsentiert den postalveolaren Artikulationsort (z.B. [ʃ]).

Übung: Klicken Sie Menü „View > sounds“ an und klicken sie dann im Lautfenster auf „E:“. Klicken sie nun Menü „View > articulators“ an. Es sind nun alle artikulatorischen Parameter auf die Artikulation des [ε:] eingestellt. Schieben Sie nun den Schieberegler „Zungenspitze: vokalisches – gehoben“ nach ganz rechts auf „gehoben“ (Wert +1000). Im Animationsfenster sehen

Sie eine Verschlussbildung durch die Zungenspitze. Verschieben Sie nun den Schieberegler „Zungenspitze: hinten – vorne“ nach ganz links auf „hinten“ (Wert -1000). Sie sehen im Animationsfenster eine postalveolare Verschlussbildung. Verschieben Sie den Schieberegler nun wieder nach rechts zurück in die Mittellage (Wert ca. 0). Sie sehen im Animationsfenster eine alveolare Verschlussbildung. Verschieben Sie den Schieberegler nun nach ganz rechts „vorne“ (Wert +1000). Sie sehen nun im Animationsfenster eine dentale Verschlussbildung.

7.7 Velum: low – high (VE:Ih)

Bei dem Parameter „Gaumensegel: gesenkt – gehoben“ repräsentiert der Wert 0 (Schieberegler in der Mitte) die momentane **vokalische** Positionierung des Gaumensegels. Diese Positionierung kann je nach Vokal unterschiedlich sein. Sie ist bei der Realisierung eines [a] das Gaumensegel meist ein wenig abgesenkt, während es bei der Realisierung eines [i] oder [u] stärker angehoben ist.

Übung: Klicken Sie Menü „View > sounds“ an. Es erscheint das Lautfenster. Klicken Sie auf „a.“. Merken Sie sich die Stellung des Gaumensegels im Animationsfenster. Klicken Sie nun auf „i.“. Sie erkennen eine wesentlich höhere Positionierung des Gaumensegels als beim [a:]. Klicken Sie nun wieder auf „a.“ und abschließend auf „u.“ und betrachten Sie die Variation der Position des Gaumensegels.

Anmerkung: Bei allen Vokalen steht der artikulatorische Parameter „Gaumensegel“ auf dem Wert 0.

Übung: Klicken Sie Menü „View > sounds“ an. Klicken Sie im Lautfenster auf „E.“. Klicken Sie nun Menü „view > articulators“ an. Alle artikulatorischen Parameter sind nun auf [ɛ:] voreingestellt. Schieben Sie den Schieberegler „Gaumensegel: gesenkt – gehoben“ zunächst nach ganz rechts „gehoben“ (Wert +1000). Sie erkennen im Animationsfenster eine leichte Anhebung des Gaumensegels. Es kann nun keine Luft mehr in den Nasenraum entweichen. Diese Einstellung ist wichtig bei der Realisierung von Frikativen und Plosiven. Schieben Sie den Schieberegler „Gaumensegel: gesenkt – gehoben“ nun nach ganz links „gesenkt“ (Wert -1000). Das Gaumensegel wird nun stark abgesenkt. Es kann Luft in den Nasenraum entweichen. Diese Stellung des Gaumensegels kennzeichnet die Nasale ([m], [n], [ŋ]).

7.8 Glottis: closed – open (GL:co)

Bei dem Parameter „Stimmlippen: verschlossen – geöffnet“ repräsentieren Wert um Null (Schieberegler in der Mitte) den lockeren Verschluss der Stimmlippen. Es ist damit **Phonation**, d.h. ein Vibrieren der Stimmlippen und damit Schallentstehung an den Stimmlippen möglich.

Anmerkung: Phonationsschall wird durch einen roten Punkt kurz oberhalb der Stimmritze angedeutet.

Übung: Schieben Sie zunächst den Schieberegler „Stimmlippen: verschlossen – geöffnet“ nach ganz rechts „geöffnet“ (Wert +1000). Sie erkennen im Animationsfenster eine weite Öffnung der Stimmritze. Schieben Sie nun den Schieberegler „Stimmlippen: verschlossen – geöffnet“ nach ganz links „verschlossen“ (Wert -1000). Die Stimmritze ist nun fest verschlossen.

7.9 Die Einstellungen „volle Enge“ (VE), „kritische Enge“ (KE), „laterale Enge“ (LE) und „vibrierende Enge“ (RE)

Bei den drei konsonantischen artikulatorischen Parametern „Lippen: vokalisches – verschlossen“, „Zungenrücken: vokalisches – gehoben“ und „Zungenspitze: vokalisches – gehoben“ können zusätzlich noch unterschiedliche Einstellungen der Artikulatoren bezüglich der *Art* der Engebildung gewählt werden. Die unterschiedlichen Einstellungen können dabei durch Anklicken des Schalters am linken Ende oberhalb des Schiebereglers des jeweiligen artikulatorischen Parameters gewählt werden. Es bedeutet:

Closure (VE: volle Enge) bei Plosiven und Nasalen

Fricative (KE: kritische Enge) bei Frikativen

Lateral (LE: laterale Enge) bei Lateralen

Vibrant (RE: vibrierende Enge) bei Vibranten

Beim artikulatorischen Parameter „Lippen: vokalisches – verschlossen“ (Kap. 7.3) bewirkt die Einstellung des Schalters auf „closure“ labiale Vollverschlussbildung (z.B. bei „b“, „p“ und „m“) und die Einstellung des Schalters auf „fricative“ labiodentale kritische Verschlussbildung (z.B. bei „f“ und „v“).

Übung: Klicken Sie Menü „view > sounds“ an. Klicken Sie im Lautfenster nun auf „f“. Wechseln Sie nun zum Artikulatoren-Fenster (Menü: „view > articulators“). Sie erkennen, dass für den Parameter „Lippen: vokalisches – verschlossen“ ein Wert um 900 und der Schalter auf „fricative“ eingestellt ist. Schieben Sie den Schieberegler dieses Parameters nun nach rechts. Die Engebildung wird aufgehoben. Die Unterlippe wird abgesenkt. Schieben Sie den Schieberegler dieses Parameters nun wieder nach links. Die labiodentale Engebildung entsteht wieder.

Anmerkung: Bei der Ausbildung einer kritischen Enge (Frikativlaute) entstehen vor der artikulatorischen Enge Luftverwirbelungen, die akustisch zu einem Rauschen führen (hörbares Geräusch der Frikativlaute). Diese Luftverwirbelungen werden in der Animation durch einen gelben Punkt vor der Engebildung angedeutet.

Übungsfortsetzung: Klicken Sie nun auf den Schalter des Parameters „Lippen: vokalisches – verschlossen“ und vollziehen einen Wechsel der Artikulationsart von „fricative“ nach „closure“. Gleichzeitig erkennen Sie nun eine bilabiale Verschlussbildung. Schieben Sie den

Schieberegler dieses Parameters nun nach rechts. Die Verschlussbildung wird aufgehoben. Die Unterlippe wird abgesenkt. Schieben Sie den Schieberegler dieses Parameters nun wieder nach links. Die bilabiale Verschlussbildung entsteht wieder.

Beim artikulatorischen **Parameter „Zungenrücken: vokalisches – gehoben“** (Kap. 7.4) bewirkt die Einstellung des Schalters auf „vocalic“ Vollverschlussbildung (z.B. bei „k“, „g“ und „N“), die Einstellung des Schalters auf „fricative“ kritische Verschlussbildung (z.B. bei „c“, „x“, „j“ und „X“) und die Einstellung des Schalters auf „vibrant“ vibrierende Enge (bei „R“).

Übung: Klicken Sie Menü „view > sounds“ an. Klicken Sie im Lautfenster nun auf [ç]. Wechseln Sie nun zum Artikulatoren-Fenster (Menü: „view > articulators“). Sie erkennen, dass für den Parameter „Zungenrücken: vokalisches – gehoben“ der Schieberegler auf den maximalen Wert (+1000) und der Schalter auf „fricative“ eingestellt ist. Sie erkennen die dorsal-palatale Engebildung im Animationsfenster. Schieben Sie den Schieberegler dieses Parameters nun nach rechts. Die Engebildung wird aufgehoben. Der Zungenrücken wird abgesenkt. Schieben Sie den Schieberegler dieses Parameters nun wieder nach links. Die dorsal-palatale Engebildung entsteht wieder.

Übungsfortsetzung: Klicken Sie nun auf den Schalter des Parameters „Zungenrücken: vokalisches – gehoben“ und wechseln Sie von „fricative“ nach „vibrant“. Gleichzeitig erkennen Sie nun eine dorsal-palatale Vollverschlussbildung. Schieben Sie den Schieberegler dieses Parameters nun nach rechts. Die Verschlussbildung wird aufgehoben. Der Zungenrücken wird abgesenkt. Schieben Sie den Schieberegler dieses Parameters nun wieder nach links. Die Vollverschlussbildung entsteht wieder. Klicken Sie ein weiteres Mal auf den Schalter dieses artikulatorischen Parameters. Sie sehen den Wechsel der Artikulationsart von „closure“ nach „fricative“ und erkennen gleichzeitig wieder die kritische Enge im Animationsfenster.

Übungsfortsetzung: Wechseln Sie nun in das Lautfenster (Menü: „view > sounds“). Klicken Sie auf den Laut „R“. Wechseln Sie zurück in das Artikulatoren-Fenster (Menü: „View > articulators“). Sie erkennen, dass für den Parameter „Zungenrücken: vokalisches – gehoben“ der Schieberegler auf den maximalen Wert (+1000) und der Schalter auf „vibrant“ eingestellt ist. Sie erkennen die dorsal-velare Engebildung im Animationsfenster. Die Vibration des Zäpfchens können Sie allerdings nur bei der Filmdarstellung erkennen. Wechseln Sie deshalb zurück in das Lautfenster (Menü: „View > sounds“). Geben Sie im Texteingabefeld des Lautfensters die Lautfolge „RRRR“ ein und klicken Sie auf „Play Animation“. Im Animationsfenster wird nun die Vibration des Zäpfchens angedeutet.

Anmerkung: Um die Vibration des Zäpfchens gut verfolgen zu können, kann es ratsam sein, die Ausführungsgeschwindigkeit der Animation herabzusetzen.

Beim artikulatorischen **Parameter „Zungenspitze: vokalisches – gehoben“** (Kap. 7.5) bewirkt die Einstel-

lung des Schalters auf „closure“ Vollverschlussbildung (z.B. bei „t“, „d“ und „n“), die Einstellung des Schalters auf „fricative“ kritische Verschlussbildung (z.B. bei „s“, „S“, „z“ und „Z“), die Einstellung des Schalters auf „lateral“ laterale Engebildung (bei „l“) und die Einstellung des Schalters auf „vibrant“ vibrierende Engebildung (bei „r“).

Übung: Klicken Sie Menü „View > sounds“ an. Klicken Sie im Lautfenster nun auf „s“. Wechseln Sie nun zum Artikulatoren-Fenster (Menü: „view > articulators“). Sie erkennen, dass für den Parameter „Zungenspitze: vokalisches – gehoben“ der Schieberegler auf den maximalen Wert (+1000) und der Schalter auf „fricative“ eingestellt ist. Sie erkennen die apikal-alveolare Engebildung im Animationsfenster. Schieben Sie den Schieberegler dieses Parameters nun nach rechts. Die Engebildung wird aufgehoben. Die Zungenspitze wird abgesenkt. Schieben Sie den Schieberegler dieses Parameters nun wieder nach links. Die apikal-alveolare Engebildung entsteht wieder.

Übungsfortsetzung: Klicken Sie nun auf den Schalter des Parameters „Zungenspitze: vokalisches – gehoben“ und wechseln Sie die Artikulationsart von „fricative“ nach „vibrant“. Gleichzeitig erkennen Sie nun eine apikal-alveolare Vollverschlussbildung. Schieben Sie den Schieberegler dieses Parameters nun nach rechts. Die Verschlussbildung wird aufgehoben. Die Zungenspitze wird abgesenkt. Schieben Sie den Schieberegler dieses Parameters nun wieder nach links. Die Vollverschlussbildung entsteht wieder. Klicken Sie ein weiteres Mal auf den Schalter dieses artikulatorischen Parameters. Sie sehen den Wechsel der Artikulationsart von „closure“ nach „fricative“ und erkennen gleichzeitig wieder die kritische Enge im Animationsfenster.

Übungsfortsetzung: Wechseln Sie nun in das Lautfenster (Menü: „View > sounds“). Klicken Sie auf den Laut „r“. Wechseln Sie zurück in das Artikulatoren-Fenster (Menü: „View > articulators“). Sie erkennen, dass für den Parameter „Zungenspitze: vokalisches – gehoben“ der Schieberegler auf den maximalen Wert (+1000) und der Schalter auf „vibrant“ eingestellt ist. Sie erkennen die apikal-alveolare Engebildung im Animationsfenster. Die Vibration der Zungenspitze können Sie allerdings nur bei der Filmdarstellung erkennen. Wechseln Sie deshalb zurück in das Lautfenster (Menü: „view > sounds“). Geben Sie im Texteingabefeld des Lautfensters die Lautfolge „rrrr“ ein und klicken Sie auf „Play Animation“. Im Animationsfenster wird nun die Vibration der Zungenspitze angedeutet.

Anmerkung: Um die Vibration der Zungenspitze gut verfolgen zu können, kann es ratsam sein, die Ausführungsgeschwindigkeit der Animation herabzusetzen.

Übungsfortsetzung: Wechseln Sie nun erneut in das Lautfenster (Menü: „View > sounds“). Klicken Sie auf den Laut „l“. Wechseln Sie zurück in das Artikulationsfenster (Menü: „Ansicht > Artikulatoren“). Sie erkennen, dass für den Parameter „Zungenspitze: vokalisches – gehoben“ der Schieberegler auf den maximalen Wert (+1000) und der Schalter auf „lateral“ eingestellt ist.

Die laterale Engebildung (Absenkung des linken und rechten Zungenrandes) wird in unserer 2D-Animationsdarstellung in der Mediosagittalebene nur durch eine zusätzliche Line angedeutet.

Anmerkung: Die Parametrisierung der in diesem Kapitel beschriebenen Einstellungsparameter (Einstellungen) zur Enge erfolgt bei den Target- und Bewegungsdaten (Kap. 16) in der folgenden Weise: Jede mögliche Einstellung wird hier für jeden einzelnen Artikulator getrennt kodiert. So ergeben sich folgende Parameter für die Einstellungen zur Enge:

LI:ke	Lippen: kritische Enge
TB:ke	Zungenrücken: kritische Enge
TB:re	Zungenrücken: vibrierende Enge
TT:ke	Zungenspitze: kritische Enge
TT:re	Zungenspitze: vibrierende Enge
TT:le	Zungenspitze: laterale Enge

Diese Parameter sind in den Tabellen zu den Target-Daten (Menü „Values > Targets“) und im Zeitdarstellungs-Formular (Menü „Values > Movements“) als *farbige Kästchen* zu erkennen.

8 Erstellen einer eigenen Liste

Neben der Benutzung der mitgelieferten Listen kann der Benutzer auch eigene Wörter orthographisch und / oder per Lautschrift (Lautung) eingeben (Geschrieben- und Gesprochen-Fenster, Kap. 4), animieren und in einer eigenen Liste abspeichern.

8.1 Leeren des Listenspeichers von *SpeechTrainer*

Stellen Sie sicher, dass im Hauptfenster von *SpeechTrainer* das **Listenfenster** sichtbar ist. Klicken Sie dazu Menü „View > list“ an. Klicken Sie nun Menü „file > new list“ an. Es sind nun alle Einträge der letzten geladenen Liste gelöscht. Das Listenfenster und auch das Geschrieben-Fenster und das Gesprochen-Fenster sind leer.

Anmerkung: Leeren Sie den Listenspeicher von *SpeechTrainer* durch Anklicken von Menü „File > new list“ auch dann, wenn Sie *SpeechTrainer* gerade erst gestartet haben, um die Einträge „Bitte“, „laden“, „sie“, „eine“, „Liste“ zu löschen.

Anmerkung: Sie können das Leeren des Listenspeichers auch durch Anklicken der sich im Listenfenster befindenden Schaltfläche „create new list“ erreichen.

8.2 Eingabe eines Wortes

Um ein Wort in einer neuen Liste zu erzeugen (oder einer bereits geladenen Liste hinzuzufügen), klicken Sie bitte im Listenfenster die Schaltfläche „add new word to list“ an.

Es öffnet sich ein neues Dialogfenster (Abb. 5). Dieses Dialogfenster hat den Namen „word properties“. In diesem Dialogfenster finden Sie oben links wieder die Eingabemöglichkeit für Text und Lautung (Geschrieben- und Gesprochen-Fenster).

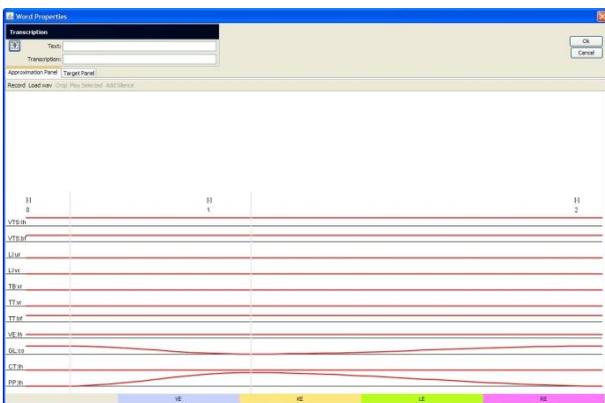


Abbildung 5. Dialogfenster „word properties“.

Klicken Sie das Geschrieben-Fenster an („text“) und geben Sie hier das gewünschte Wort ein (z.B. „Panne“). Es erscheint im Gesprochen-Fenster gleichzeitig die

gewünschte Transkription. Wenn Sie mit der Transkription zufrieden sind, schließen Sie das Dialogfenster durch Anklicken des „ok“-Button. Das neue Wort befindet sich dann im Listenfenster. Sie können nun auf die gleiche Art und Weise weitere Wörter der Liste hinzufügen.

Anmerkung: Wenn die im Dialogfenster „word properties“ entstehende Transkription (Gesprochen-Fenster) nicht zufriedenstellend sein sollte, können Sie den Vorgang abbrechen, indem Sie das Dialogfenster mit dem „cancel“-Button schließen. Sie können alternativ nun das Wort auch wie in Kap. 4 beschrieben im Geschrieben-Fenster des Hauptfensters eingeben und unter Zuhilfenahme von „view > sounds“ im Gesprochen-Fenster korrigieren. Gehen Sie dann durch „view > list“ wieder zurück ins Listenfenster und klicken Sie dann dort die Schaltfläche „add new word to list“ an und bestätigen Sie nun im dialogfenster „word properties“ die Eingabe durch Anklicken des „ok“-Buttons.

8.3 Einfügen eines weiteren Eintrages (Menü „Bearbeiten > Einfügen“)

Sie können aktuell editierte Wörter aber auch an beliebige Positionen der aktuell geladenen Liste einfügen. Dazu muss zunächst der Eintrag im Listenfenster markiert werden, *über* dem das neue Wort eingefügt werden soll. Gehen Sie dann wie gewohnt vor, indem Sie die Schaltfläche „add new word to list“ anklicken. Das Wort erscheint anschließend an der gewünschten Position in der Liste.

8.4 Abspeichern der aktuellen Liste

Es ist sinnvoll, eine aktuell in *SpeechTrainer* generierte oder bearbeitete Liste zwischendurch zu speichern, damit die gerade durchgeführten Arbeiten nicht verloren gehen. Klicken Sie dazu im Listenfenster auf die Schaltfläche „save list to new file“ oder klicken Sie in der Menüleiste „file > save list“ an.

Anmerkung: Der Name der von *SpeechTrainer* (aktuell) geladenen Liste erscheint auch immer in der Titelseile des Hauptfensters von *SpeechTrainer*.

Anmerkung: Eine Liste, die sich in Bearbeitung befindet, und noch nicht abgespeichert ist (d.h. deren Arbeit noch nicht gesichert ist), ist in der Titelseile des Programmfensters von *SpeechTrainer* durch einen **Stern** hinter dem Listennamen gekennzeichnet.

8.5 Bearbeiten von Listeneinträgen

Neben dem Hinzufügen und Einfügen von Listeneinträgen ist auch das nachträgliche Bearbeiten einer Liste bzw. eines Listeneintrags möglich.

Klicken Sie dazu im Listenfenster auf die Schaltfläche „edit / record selected word“. Es öffnet sich das bereits bekannte Dialogfenster „word properties“. Hier kann wie gewohnt die Schreibung oder Lautung des aktuellen Wortes geändert werden. Nach Schließen des Fensters

via „ok“-Button erscheint die Änderung im Listenfenster.

8.6 Löschen von Listeneinträgen

Markieren Sie die zu löschende Zeile durch Anklicken des Wortes im Listenfenster. Klicken Sie nun im Listenfenster die Schaltfläche „delete selected word“ und das markierte Wort wird gelöscht.

den grauen Bereich und halten Sie die Maustaste gedrückt. Sie können nun das graue Fenster über das Signal verschieben. Wenn Sie es über das Signal geschoben haben, dass Signal aber größer als der blaue Bereich ist, wandern Sie mit dem Mauszeiger auf den Beginn bzw. auf das Ende des blauen Bereichs (rote Begrenzungslinie). Sie können nun den Anfang bzw. das Ende des blauen Bereichs verschieben und damit die Länge des blauen Fensters variieren, bis das Signal zu beiden Seiten gut im blauen Fenster liegen (siehe Abb. 8).

Durch Anklicken der Schaltfläche „Play selected“ können Sie nun noch einmal auditiv kontrollieren, ob das gesamte Wort hörbar ist.

Wenn das der Fall ist, können Sie dieses Dialogfenster nun mittels Anklicken des „ok“-Button verlassen. Im **SpeechTrainer**-Listenfenster erscheint nun vor dem Wort ein Laugsprechersymbol. Dieses Symbol zeigt an, dass der Worteintrag „vertont“ ist.

Stellen Sie nun im Listenfenster (unten) die Wiedergabegeschwindigkeit „speed“ auf „1x“ ein und spielen Sie dann die Animation ab (Schaltfläche „Play animation“). Sie sollten nun neben dem visuellen Ablauf der Artikulation auch das eingesprochene Wort hören können. Das hörbare Wort und die Animation der Artikulationsbewegungen müssen jetzt aber noch nicht unbedingt synchron ablaufen. Die Herstellung von Synchronität wird im nächsten Kapitel erläutert.

Doppelklicken Sie nun den nächsten Listeneintrag und führen Sie die Vertonung dieses und aller weiteren Listeneinträge nun auf die gleiche Art durch, wie oben für das Wort „Mama“ ausgeführt.

Anmerkung: Sie müssen akustische Sprachsignale nicht selbst per Mikrofon einsprechen, sondern via Anklicken der Schaltfläche „load wav“ können auch **externe Sounddateien** einladen. Der Typ dieser Sounddatei sollte sein: WAV-Format, 44100 Hz, 16 bitt, mono.

9.3 Speichern der Liste nicht vergessen

Vergessen Sie nicht, zwischendurch und am Ende der Sitzung nach Vertonen des letzten Wortes die Liste erneut zu speichern („file > save list“).

Anmerkung: Nicht gespeicherte Arbeit erkennt man durch den Stern hinter dem Namen der Liste in der Titelzeile des Programmfensters.

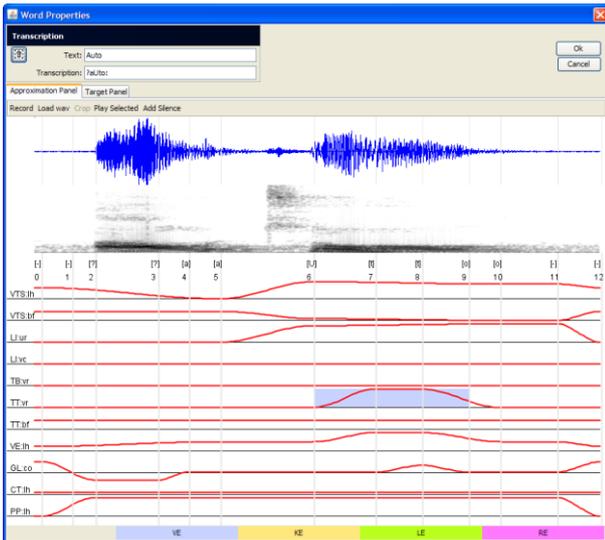


Abbildung 11. Das Dialogfenster „word properties“ für den Listeneintrag „Auto“ direkt nach Verkürzung des stummen Vor- und Nachlaufes des Signals durch „crop“.

Um nun zunächst die Bedeutung jedes Labels verstehen, ist am oberen Ende jedes Label die laufende Nummer des Label und das Lautsymbol des Label angegeben (z.B. [-] 1, [-] 2, [?] 3 usw.).

Führen Sie nun den Mauszeiger auf die graue vertikale Linie eines Label ohne eine Maustaste zu drücken. Das Maussymbol zeigt an, das Sie dieses Label nun bei gehaltener Maustaste verschieben können.

Anmerkung: Die Bedeutung aller unterschiedlichen Label wird detailliert in Kap. 16 erläutert.

Durch die Synchronisation soll die Position eines Labels immer mit dem Anfang, dem zeitlichen Mittel oder dem Ende eines Lautes zusammenfallen. Diese Zeitpunkte können oft direkt anhand des Oszillogramms und Spektrogramms erkannt werden.

10.4 Versetzen eines Labels

Zum Versetzen eines Labels müssen Sie also den Mauszeiger über das Label führen, bis der Mauszeiger als horizontaler Doppelpfeil angezeigt wird. Führen Sie nun das Label bei gedrückter Maustaste an die neue gewünschte Position (dies können Sie z.B. durch verschieben des Label 3 ausprobieren).

10.5 Justieren aller Label auf ein Mal

Zum ersten Justieren der Label können sie zunächst das **erste Label** (Label 0) „anfassen“ (erscheinen des Doppelpfeils). Bei verschieben dieses Labels werden alle nachfolgenden Label parallel verschoben. Sie können hierdurch also den Anfangszeitpunkt für die **gesamte Labelstrecke verschieben**.

Wenn Sie den Anfang eingestellt haben können sie nun – bevor sie einzelne Label einstellen – noch das letzte Label „anfassen“ (Erscheinen des Doppelpfeils). Durch verschieben dieses letzten Labels wird nun die **Länge der Labelstrecke variiert**.

So können Sie also durch eine grobe Justierung des ersten und letzten Labels die gesamte Labelstrecke grob mit dem akustischen Signal synchronisieren. Erst danach sollten Sie mit der Feinjustierung weiterer einzelner Label beginnen.

Anmerkung: Ein Verschieben des ersten und letzten Labels beeinflusst immer die Lage aller Label. Wenn dies nicht gewünscht ist, darf das Label per Maus nur dann „angefasst“ werden, wenn gleichzeitig die Taste „Strg“ gedrückt gehalten wird.

10.6 Hören eines Signalausschnittes

In der gleichen Weise wie beim Ausschneiden von Signalen (Kap. 9) kann auch hier bei der Synchronisation nicht nur das ganze Audiosignal gehört werden, sondern auch jeder beliebige Signalausschnitt.

Klicken Sie dazu einfach auf den Anfangszeitpunkt des Signalausschnittes im Bereich des Oszillogramms und halten Sie die Maustaste gedrückt. Ziehen Sie die Maus bei gedrückter Maustaste nun bis zum Ende des Signalausschnittes und lassen Sie die Maustaste dann los. Der entsprechende Signalausschnitt ist nun als blauer Bereich markiert (Abb. 12).

Klicken Sie nun den Button „Play selected“ an und Sie hören genau den markierten Ausschnitt.

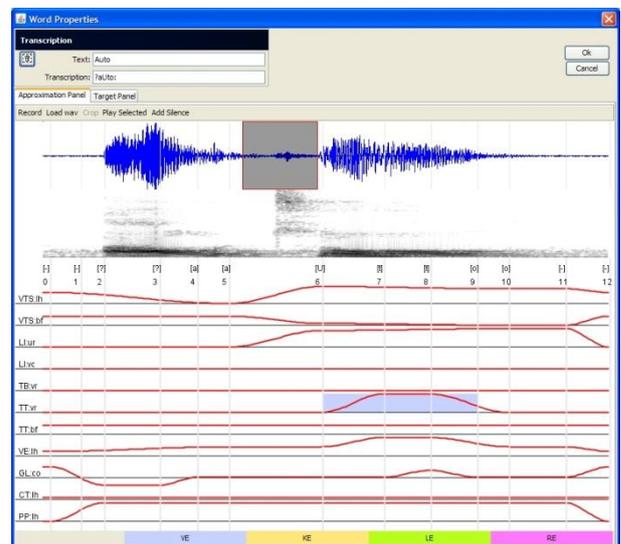


Abbildung 12. Das Dialogfenster „word properties“ für den Listeneintrag „Auto“ nach Auswahl eines Signalausschnittes zum Hören und zur Orientierung im akustischen Signal. Hier wurde der Signalausschnitt des Lautes [t] aus Auto markiert.

Anmerkung: Um die Markierung des Signalausschnittes wieder aufzuheben, klicken Sie im Bereich des Oszillogramms den Anfang des Oszillogramms an.

Anmerkung: Bevor Sie dieses Dialogfenster mit „ok“ verlassen, müssen Sie unbedingt sicherstellen, dass kein Signalausschnitt mehr markiert ist. Ansonsten wird nur der Signalausschnitt gespeichert.

10.7 Zu kurzer Vorlauf am Signalanfang

Im aktuellen Beispiel „Auto“ ist es erforderlich, dass die Label 0 bis 4 vor dem Beginn des Vokals gesetzt werden (siehe Kap. 16). Daher ist wünschenswert, dass das der Vorlauf des Signals wieder verlängert wird. Dies kann durch Anklicken der Schaltfläche „add silence“ erreicht werden (Ergebnis siehe Abb. 13).

Anmerkung: Danach kann das Fenster durch Klicken des „ok“-Button geschlossen werden, erneut geöffnet werden und es kann noch einmal „add silence“ durchgeführt werden, um den stummen Vorlauf evtl. noch weiter zu verlängern.

Nach zweimaligem Durchlauf von „add silence“ ist der stumme Vorlauf des akustischen Signals nun wieder ausreichend lang, um die Label 0 bis 4 vor Beginn des Signals (des Vokals „a“) zu positionieren.

Abb. 14 zeigt die erfolgreich durchgeführte Synchronisation aller Label zum Wort „Auto“ (siehe dazu Kapitel 16).

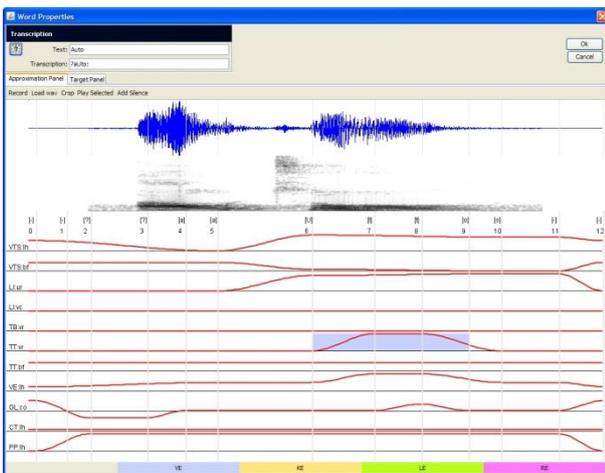


Abbildung 13. Das Dialogfenster „word properties“ für den Listeneintrag „Auto“ nach zweimaligem Durchlauf von „add silence“.

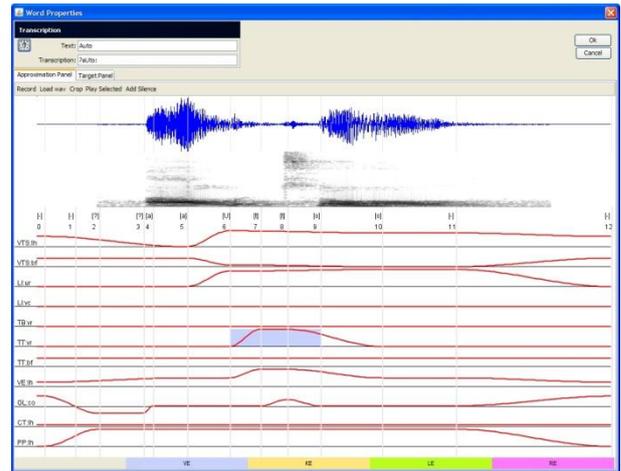


Abbildung 14. Das Dialogfenster „word properties“ für den Listeneintrag „Auto“ nach erfolgreicher Synchronisation durch Versetzen aller Label.

Schließen Sie danach das Dialogfenster durch Klicken des „ok“-Button und lassen Sie dann die Animation des gerade synchronisierten Listeneintrages ablaufen. Die audio-visuelle Synchronisation sollte nun zufriedenstellend sein.

Ansonsten müssen Sie das Dialogfenster „word properties“ für diesen Listeneintrag erneut öffnen und die Label an andere Positionen setzen.

Anmerkung: Eine fehlerhafte Synchronisation kann aber auch in einer mangelhaften zeitlichen Synchronisation der Grafikkarte und der Audiokarte begründet liegen. Dies tritt aber in der Regel nur bei älteren oder leistungsschwachen Rechnern (evtl. bei Netbooks) auf.

10.8 Weitere Probleme nach dem Synchronisieren

Trotz richtigem Setzen der Label kann es nach Durchführung der Synchronisation in einigen Fällen zu noch wenig befriedigenden Animationen kommen. (Falsche Koartikulation, unbefriedigende artikulatorische Positionen für bestimmte Laute ...). In diesen Fällen ist eine Korrektur durch weitergehende Nachbearbeitung der Artikulatorbewegungen der Animation auf **Target-Ebene** möglich (siehe Kap. 16).

11 Zusammenstellen einer eigenen Liste

Sie können aus mehreren selbsterstellten (vertonten) Listen auch weitere neue (vertonte) Listen erstellen, ohne dass Sie dazu die einzelnen Listeneinträge (Wörter) wieder neu einsprechen müssen. Es entfällt damit die Arbeit der Erstellung von Schreibung und Lautung, der Vertonung und der Synchronisation.

11.1 Anlegen einer eigenen Liste

Kopieren Sie zunächst eine erste Liste als Basis für Ihre eigene Liste. Öffnen Sie dazu zunächst die erste Liste (Menü: „File > open list“; z.B. „liste_02.st“) und speichern Sie diese Liste nun unter einem von Ihnen gewählten Namen ab (Menü: „file > save list“; z.B. als „liste_11.st“).

Die nun aktuell in *SpeechTrainer* geladene Liste ist „liste_11.st“ (siehe Titelzeile von *SpeechTrainer*).

11.2 Löschen von Listeneinträgen

Löschen Sie nun aus Ihrer Liste alle nicht benötigten Einträge (siehe auch Kap. 8). Klicken Sie dazu in der Liste auf einen Eintrag (z.B. auf das Wort „Tomate“) und klicken Sie dann auf die Schaltfläche „delete selected word“.

Sichern Sie nun diese Arbeit durch Speichern der Listendatei (Menü: „file > save list“).

11.3 Einfügen von weiteren Listeneinträgen aus einer anderen Liste

Um nun weitere Einträge aus einer anderen Liste in „liste_11.st“ zu integrieren, gehen Sie wie folgt vor: (1) Hängen Sie eine weitere Liste an die aktuell bearbeitete Liste an. (2) Löschen Sie dann in der neu entstandenen Gesamtliste die Einträge aus der angehängten Liste, die Sie nicht in der aktuellen Liste stehen haben wollen.

Klicken Sie auf Menü („file > append list“) und hängen Sie eine weitere vertonte Liste (z.B. „Liste_04.st“) an die aktuelle Datei an.

Der aktuelle Dateiname bleibt (z.B. „liste_11.st“) bleibt bestehen; allerdings sollten Sie die Arbeit speichern (Menü: „file > save list“).

Löschen Sie nun aus der angehängten Datei die Wörter, die Sie nicht in der aktuellen Liste stehen haben möchten und speichern Sie die Liste dann erneut.

11.4 Umsortieren von Einträgen in einer aktuell geladenen Liste

Nachdem Sie Listen aneinander angehängt haben, können Sie Listeneinträge auch beliebig umsordieren, indem Sie im Listenfenster mit der Maus einen Listeneintrag markieren und das Wort mit gedrückter Maustaste an eine andere Position in der Liste ziehen.

11.5 Löschen einer ganzen Liste

Nachdem Sie nun Ihre eigene Liste fertiggestellt haben, können Sie ältere Listen von Ihrem Speichermedium (z.B. Festplatte Ihres Rechners) löschen.

Dieses Löschen können Sie nicht innerhalb des Programms *SpeechTrainer* durchführen, sondern nur außerhalb des Programms durch die entsprechende Befehle Ihres Betriebssystems.

Anmerkung: Das Löschen einer Liste innerhalb des *SpeechTrainers* durch Klicken der Schaltfläche „create new list“ betrifft nur den internen Speicher des Programms *SpeechTrainer*. Allerdings geht die zu dem Zeitpunkt dann die **aktuell im *SpeechTrainer* existierende Liste** verloren, wenn diese nicht zuvor gespeichert wurde („file > save list“).

12 Nachträgliches Ändern eines Listeneintrages

Sie können nachträglich jeden Listeneintrag in einer nicht vertonten oder auch in einer vertonten und synchronisierten Liste ändern.

Änderungen können auf allen Ebenen, beginnend mit der Schreibung über die Lautschrift bis hin zu Details der Animation (artikulatorische Bewegungen, artikulatorische Targets) durchgeführt werden.

Allerdings zieht eine Änderung auf einer bestimmten Ebene immer eine Neugenerierung des Listeneintrages auf allen darunter liegenden Ebenen nach sich. Die Reihenfolge der Ebenen ist: Schreibung, Lautung, Zeitpunkte von Labels, Positionswerte der Artikulatoren an einem Label (Label werden nun auch **Targets** genannt; siehe auch Kap. 14). Je nach Ebene wird dann evtl. auch eine erneute Vertonung und Synchronisation nötig.

12.1 Ändern der Schreibung oder Lautung eines Listeneintrages

Doppelklicken Sie einen Eintrag in der aktuellen Liste. Es erscheint das bekannte Dialogfenster „word properties“.

Hier können Sie nun im Feld „Text“ und im Feld „transcription“ die Schreibung und Lautung des Listeneintrages ändern. Verlassen Sie dann das Dialogfenster mittels Anklicken des „ok“-Button.

Anmerkung: wenn Sie aber bereits eine Synchronisation durchgeführt haben, machen Sie diese Arbeit dadurch zunichte. Die Labelpositionen werden wieder wie vor dem Synchronisieren gesetzt, auch wenn das im Dialogfenster erst beim erneuten Öffnen des Dialogfensters angezeigt wird.

12.2 Ändern der Animation

Sehr effektiv und relativ einfach kann die Ausführung der Animation (können die Artikulatorbewegungen) durch **Ändern von Targets (Zeitpositionen)** eines Labels; **Targetwerte der Artikulatoren** an einem Label geändert werden. Targets sind artikulatorische Parameterwerte, die zum Zeitpunkt eines Label gesetzt sind (siehe Kap. 16). Durch Ändern eines Targets (Targetwertes) ändern sich dann auch die artikulatorischen Bewegungsabläufe und damit die Animation.

Letztlich ist die Synchronisation, also das Versetzen bzw. das Definieren der *Labelzeitpunkte* als Teil der Änderung der Animation zu verstehen. Schließlich wird durch die Synchronisation die Standardanimation für eine Lautfolge in ihrem zeitlichen Ablauf in Hinblick auf das Timing des vom Benutzer gesprochenen Audiosignals geändert. Durch die Synchronisation werden somit die *Zeitpunkte der Targets* (die *Labelzeitpunkte*) geändert.

Wenn trotz sorgfältig ausgeführter Synchronisation die Animation unbefriedigend ist, so können die Targetwerte selbst und deren **Pertinenz** geändert werden. Dies führt sehr einfach und sehr effektiv zu besseren Artikulatorbewegungen (siehe Kap. 16).

Zunächst muss allerdings ergründet werden, welcher Artikulator in welchen Zeitbereich (bei welchen Lauten) eine falsche oder unbefriedigende Artikulationsbewegung ausführt.

Dazu muss für den aktuellen Listeneintrag durch Doppelklicken des Listeneintrages das bekannte Dialogfenster „word properties“ geöffnet werden. Hier können nun im Bereich unter dem Oszillogramm und Spektrogramm die artikulatorischen Bewegungsverläufe anhand der roten **Bewegungstrajektorien** (Bewegungsverläufe) nachvollzogen werden. Die Abkürzungen zu den Bewegungstrajektorien an der rechten Seite bezeichnen die **artikulatorischen Parameter** (siehe Kap. 7).

Kontrollieren Sie die Bewegungsverläufe und machen Sie den Zeitbereich ausfindig, in dem es zur verbesserungswürdigen Artikulation kommt. Merken Sie sich die Namen der Label, die diesen Zeitbereich markieren. Merken Sie sich auch den artikulatorischen Parameter (z.B. VG: hv), der zu der fehlerhaften Artikulation führt.

Klicken Sie nun in diesem Dialogfenster auf die Schaltfläche „**target panel**“ und Sie erhalten eine Liste der Label, Ihrer Zeitwerte und der zugehörigen **Targetwerte**. Wenn ein Zahlenwert rot markiert ist, ist dieser Zahlenwert ein Targetwert, der von der Bewegungstrajektorie zum Zeitpunkt dieses Labels angesteuert wird. Alle anderen (schwarz geschrieben) Targetwerte sind „**transparent**“, d.h. haben für die Ausbildung (Berechnung) der artikulatorischen Bewegungsabläufe keine Bedeutung.

Durch linken Mausklick kann ein **Targetwert** im „target panel“ editiert werden. Durch rechten Mausklick kann die **Pertinenz** geändert werden, d.h. kann ein Targetwert von „transparent“ auf „nicht-transparent“ gesetzt werden und umgekehrt.

Dieses System der transparenten und nicht-transparenten Targetwerte trägt der Tatsache Rechnung, dass nicht jedes Label *alle* artikulatorischen Targetpunkte definieren muss. So wird beispielsweise im Wort „Banane“ („liste_02.st“) der artikulatorische Parameter VST:lh (der erste vokalische Parameter, siehe Kap. 7) nicht von den Labeln 1, 2, und 3, wohl aber von den Labeln 0 und 4 beeinflusst.

Wenn Sie Änderungen im Dialogfeld „word properties“ im Bereich „target panel“ vornehmen, so können Sie die Auswirkungen dieser Änderungen auf den Verlauf der Artikulationsbewegungen (auf die Bewegungstrajektorien) sehen, indem Sie nach erfolgreicher Änderung eines Wertes oder einer Pertinenz im Dialogfeld „word properties“ auf „**trajectory panel**“ klicken.

Durch laufendes Umschalten zwischen „trajectory panel“ und „target panel“ können Sie so also die Verände-

zung des Verlaufes der Bewegungstrajektorien verfolgen.

Nachdem alle Änderungen an den Verläufen der Bewegungstrajektorien durchgeführt wurden, verlassen Sie das Dialogfenster „word properties“ durch Anklicken des „ok“-Buttons und können nun die Änderung der artikulatorischen Bewegungsabläufe auch in der Animation („play animation“) erleben.

Anmerkung: Im „target panel“ wie auch im „trajectory panel“ innerhalb des Dialogfensters „word properties“ werden die **Zusatzeinstellungen der konsonantischen Parameter** („closure, fricative, lateral, vibrant“) durch unterschiedliche Hintergrundfarben angegeben. Auch diese Zusatzeinstellung können im „target panel“ durch rechten Mausklick geändert werden.

13 Überblick: Menüleiste und Programmfenster

In diesem Kapitel werden die Menübefehle und die durch die Menübefehle geöffneten Formulare der Reihenfolge nach vorgestellt. Eine detaillierte Einführung in die themenbezogene Arbeitsweise mit *SpeechTrainer* wurde in den vorangehenden Kapiteln gegeben.

13.1 Das Menü „file“

„file > new list“

Durch Menü „file > new list“ wird die aktuell in *SpeechTrainer* geladene Liste aus dem Programmspeicher gelöscht. Aktuelle Änderungen gehen dadurch verloren, wenn Sie nicht zuvor abgespeichert wurden.

„file > open list“

Durch Menü „file > open list“ können Sie eine Liste auswählen, und diese dann in *SpeechTrainer* laden. Dabei wird automatisch die vorher in *SpeechTrainer* geladene Liste aus dem Programmspeicher entladen.

„file > save list“

Durch Menü „file > save list“ kann die aktuell in *SpeechTrainer* geladene und evtl. bearbeitete Liste gespeichert werden. Auf diese Weise können aktuell mit *SpeechTrainer* durchgeführte Änderungen an der Liste gesichert werden.

„file > append list“

Durch Menü „file > append list“ kann eine weitere Liste auszuwählende Liste an die bereits in *SpeechTrainer* geladene Liste angehängt. Der Name der in *SpeechTrainer* geladenen Liste ändert sich hierdurch nicht.

„file > quit“

Durch Menü „file > quit“ wird das Programm *SpeechTrainer* beendet. Stellen Sie sicher, dass Sie zuvor die aktuell geöffnete Liste speichern um evtl. durchgeführte Änderungsarbeiten zu sichern.

13.2 Das Menü „view“

Durch das Menü „view“ können die unterschiedlichen **Unterfenster** (linke Seite im Programmfenster) von *SpeechTrainer* aktiviert werden.

„view > list“

Durch Menü „view > list“ wird das **Listfenster** aktiviert. Das Listfenster gibt die in *SpeechTrainer* aktuell geladene (bzw. geöffnete) Liste wieder.

„view > sounds“

Durch Menü „view > sounds“ wird das **Lautfenster** aktiviert. Das Lautfenster ermöglicht das Anzeigen von Einzellauten und kann als Hilfe bei der Transkription von Listeneinträgen herangezogen werden.

Anmerkung: Alternativ kann das Lautfenster auch durch Anklicken des **Gesprochen-Fensters** („transcription“ links unten) aktiviert werden.

„view > articulators“

Durch Menü „view > articulators“ wird das **Artikulatoren-Fenster** aktiviert. Das Artikulatoren-Fenster zeigt die artikulatorischen Parameter des Artikulationsmodells des *SpeechTrainer* und ermöglicht das interaktive Studium der Auswirkung einer Änderung des Wertes eines artikulatorischen Parameters auf die Animationsdarstellung.

„Ansicht > Übersicht“

Durch Menü „Ansicht > Übersicht“ wird das Übersichtsfenster aktiviert. Im Übersichtsfenster wird dargestellt, bis zu welcher Ebene der aktuelle Listeneintrag gerade editiert wurde. Durch Anklicken der Schaltflächen kann der Editiervorgang auf einer bestimmten Ebene gestartet werden (Kap. 14.6).

13.3 Das Menü „values“

Im Menü „values“ können die **Parameter-Einstellungen des Artikulationsmodells für einen aktuellen Listeneintrag** eingesehen werden.

„values > sounds > table“

Durch Menü „values > sounds > table“ wird die **Lautfolge-Tabelle** (Dialogfenster „sound table“) geöffnet. In dieser Tabelle ist die Lautung des aktuellen Listeneintrages angegeben. Zusätzlich werden noch weitere Daten der Lautebene dargestellt. Diese Daten werden auf der Lautebene nach Eingabe der Lautfolge (der Lautung) generiert und dienen als Vorbereitung für die Generierung der Label (Abb. 15).

No.	Phoneme	Vowel	Voiced	Manner	Place	Length	Beg-Time	Mid-Time	End-Time	Beg-Tar...	End-Tar...
0	-	Yes	No	-	-	short	0	25	30	0	0
1	ʔ	No	No	plosive	glottal	short	50	100	150	0	0
2	ɐ	Yes	Yes	-	-	short	150	200	250	0	0
3	f	No	No	fricative	labiodental	short	250	300	350	0	0
4	n	No	Yes	nasal	alveolar	short	350	400	450	0	0
5	@	Yes	Yes	-	-	short	450	500	550	0	0
6	n	No	Yes	nasal	alveolar	short	550	600	650	0	0
7	-	Yes	No	-	-	short	650	685	720	0	0

Abbildung 15: Dialogfenster „sound table“ für das Wort „öffnen“ (aus der Default-Liste).

Die Zeilen der Lautfolge-Tabelle (Abb. 15) entsprechen jeweils einem Lautsymbol des aktuellen Listeneintrages, allerdings bereits um Prä- und Postphonation (Kap. 16) erweitert. Die erste Spalte („No.“) gibt die fortlaufende Lautnummer innerhalb des aktuellen Listeneintrages an. Die zweite Spalte („phoneme“) gibt das Lautsymbol an. Erste und letzte Zeile stehen hier für Prä- bzw. Postphonation. Die dritte und vierte Spalte gibt an, ob der aktuelle Laut ein Vokal ist oder nicht („vowel“) und ob der Laut stimmhaft ist oder nicht („voiced“). Die fünfte und sechste Spalte gibt die Artikulationsart („manner“) und den Artikulationsort („place“) an. Die siebte Spalte gibt an, ob der Laut kurz oder lang realisiert wird („length“). Die achte, neunte und zehnte Spalte gibt die berechneten Zeitpunkte für Lautbeginn („Beg-Time“), Lautmitte („Mid-Time“) und Lautende („End-Time“) an. Die elfte und zwölfte Spalte geben an, wie viele Label pro Laut gesetzt werden.

Anmerkung: Die Daten in diesem Dialogfenster („sound table“) können nicht editiert werden.

„values > targets > table“

Durch Menü „values > targets > table“ wird das Dialogformular („target table“) für die **Targetwerte** (vgl. Kap. 12) geöffnet. Diese Tabelle stellt die Target-Daten für alle Label des aktuellen Listeneintrages dar (vgl. Beschreibung des „target panel“ im Dialogfenster „word properties“, Kap. 12).

Anmerkung: Die Target-Daten können in diesem Dialogfenster („target table“) nicht editiert werden, wohl aber im „target panel“ des Dialogfensters „word properties“, vgl. Kap. 12).

„values > targets > graph“

Durch Menü „values > targets > graph“ wird das Dialogfenster („target graph“) geöffnet. Hier sind die **artikulatorischen Bewegungsverläufe** (artikulatorische Trajektorien) wie im „trajectory panel“ des Dialogfensters „word properties“ abgebildet.

„values > movements > table“

Durch Menü „values > movements > table“ wird das Dialogfenster („parameter table“) geöffnet. In diesem Fenster sind die Parameterwerte aller artikulatorischen Parameter im Zeitabstand von ca. 3 msec angegeben. Auf der Basis dieser Zahlenwerte werden die artikulatorischen Bewegungstrajektorien gezeichnet.

Anmerkung: Die Daten in diesem Dialogfenster („parameter table“) sind nicht editierbar.

„values > movements > graph“

Durch Menü „values > movements > graph“ wird das Dialogfenster („parameter graph“) geöffnet. In diesem Fenster sind die artikulatorischen Bewegungstrajektorien eingezeichnet, im Unterschied zu anderen Darstellungen nun aber ohne gleichzeitige Einzeichnung der Lautlabel.

13.4 Das Menü „sound properties“

„sound properties > mic calibration“

Durch Menü „sound properties > mic calibration“ wird das Dialogfenster („microphone calibration“) geöffnet. Dieses Fenster dient der Überprüfung der Funktion von Mikrofon und Soundkarte (vgl. Kap. 9).

13.5 Das Menü „?“

„? > about“

Durch Menü „? > about“ wird das Dialogfenster „Info zum *SpeechTrainer*“ geöffnet. Dieses Formular zeigt Ihnen die aktuelle Versionsnummer des Programms an, mit dem Sie arbeiten.

13.6 Schaltflächen im Listenfenster

Über Menü „view > list“ erscheint links im Programmfenster das **Listenfenster**. Im Folgenden werden die Schaltflächen des Listenfensters erläutert.

„create new list“

Durch Anklicken der Schaltfläche „create new list“ wird die aktuelle Liste gelöscht. Wenn in dieser Liste Änderungen vorgenommen worden sind, sollte die Liste also vorher gespeichert werden. Nach dem Löschen der aktuellen Liste kann eine neue Liste eingegeben werden.

„open list from file“

Durch Anklicken der Schaltfläche „open list from file“ kann eine vorbereitete Liste z.B. von der Festplatte in den *SpeechTrainer* eingeladen werden. Dabei wird automatisch die vorher in *SpeechTrainer* geladene Liste aus dem Programmspeicher entladen.

„save list to file“

Durch Anklicken der Schaltfläche „save list to file“ kann die aktuell in *SpeechTrainer* geladene Liste und evtl. bearbeitete gespeichert werden. Auf diese Weise können aktuell mit *SpeechTrainer* durchgeführte Änderungen an der Liste gesichert werden.

„append list to existing list“

Durch Anklicken der Schaltfläche „append list to existing list“ kann eine weitere Liste auszuwählende Liste an die bereits in *SpeechTrainer* geladene Liste angehängt. Der Name der in *SpeechTrainer* geladenen Liste ändert sich hierdurch nicht.

„add new word to list“

Durch Anklicken der Schaltfläche „add new word to list“ wird das aktuell im Geschrieben- und Gesprochen-Fenster eingetragene Wort als Kandidat für ein an die aktuelle Liste zu hängendes weiteres Wort hergenommen. Es öffnet sich dann das Dialogfenster „word properties“. Hier kann das Wort nochmals bearbeitet werden. Wird das Dialogfenster durch Anklicken des „ok“-Button geschlossen, erscheint das in diesem Dialogfenster bearbeitete Wort als neuer Eintrag in der aktuellen Wortliste.

„edit/record selected word“

Durch Anklicken der Schaltfläche „edit/record selected word“ wird das in der Wortliste aktuell markierte Wort im Dialogfenster „word properties“ angezeigt. Das Wort kann hier nun bearbeitet und/oder vertont und synchronisiert werden. Verlässt man das Dialogfenster „word properties“ durch Anklicken des „ok“-Button, so werden alle im Dialogfenster durchgeführten Änderungen an diesem Listeneintrag (Wort) übernommen.

„delete selected word“

Durch Anklicken der Schaltfläche „delete selected word“ wird der aktuell in der Wortliste markierte Eintrag (das aktuell markierte Wort) gelöscht.

13.7 Die Schaltflächen im Bereich des Programmfensters links unten

Die Schaltflächen in diesem Bereich (d.h. unter den wechselnden Funktionsfenstern) dienen der Steuerung der Animation („sagittal cut view“, rechte Seite im Programmfenster).

„Text“

Durch Anklicken des **Textfensters (Geschrieben-Fenster)** kann ein Wort oder eine kurze Äußerung eingetippt werden. Es entweht gleichzeitig ein Vorschlag für die Transkription im **Transkriptionsfenster (Gesprochen-Fenster)**.

„Transcription“

Durch Anklicken des Transkriptionsfensters (Gesprochen-Fenster) kann die Transkription überarbeitet werden. Als Hilfe wird gleichzeitig das Lautfenster geöffnet.

„clear text and transcription fields“

Durch Anklicken der Schaltfläche „clear text and transcription fields“ wird die aktuell im Text- und Transkriptionsfeld durchgeführte Editierarbeit gelöscht.

Schieberegler

Durch Anklicken und Verschieben des Schiebereglers wird die aktuell im Gesprochen-Fenster ausgegebene Transkription als Animation umgesetzt und realisiert. Dabei kann die Artikulation vorwärts und rückwärts mit beliebigem Tempo realisiert werden und es kann an jedem beliebigen Punkt (Zeitpunkt) gestoppt werden.

„play animation“

Durch Anklicken der Schaltfläche „play animation“ wird die aktuell im Gesprochen-Fenster ausgegebene Transkription als Animation mit der unter „speed“ angegebenen Realisierungsgeschwindigkeit artikuliert.

„speed“

Durch Anklicken der Schaltfläche „speed“ kann eine Ausführungsgeschwindigkeit von normal („1x“) bis sehr langsam („1/32x“) gewählt werden.

Zum langsamen Mitsprechen und zur einfachen visuellen Wahrnehmung der Artikulationsbewegungen ist die Geschwindigkeitsstufe („1/4“) voreingestellt.

13.8 Schaltflächen im Lautfenster

Über Menü „view > sounds“ erscheint links im Programmfenster das **Lautfenster**. Im Folgenden werden die Schaltflächen des Lautfensters erläutert.

Vokale / Konsonanten

Jede Schaltfläche kann angeklickt werden. Der zugehörige Laut erscheint auch im Gesprochen-Fenster und die zugehörige (statische) Artikulation wird auf der rechten Seite (unter „sagittal cut view“) dargestellt.

13.9 Schaltflächen im Artikulatoren-Fenster

Über Menü „view > articulators“ erscheint links im Programmfenster das **Artikulatoren-Fenster**. Im Folgenden werden die Schaltflächen des Artikulatoren-Fensters erläutert.

Schieberegler

Durch Betätigung der Schieberegler können die Werte einzelner artikulatorischer Parameter geändert werden. Dementsprechend ändert sich auch die statische Darstellung der Artikulationsorgane im Animationsfenster. Die artikulatorische Parameter wurden in Kap. 7 erläutert.

Schaltflächen

Durch Anklicken der Schaltflächen kann der Typ der Engebildung für konsonantische Parameter verändert werden (siehe hierzu Kap. 7)

13.10 Schaltflächen im Animationsfenster

Das **Animationsfenster** („sagittal cut view“) nimmt die rechten Seite des Programmfensters ein.

„show phonation / aspiration points“

Durch Anklicken dieser Schaltflächen wird die Darstellung der Phonation (roter Punkt) bzw. die Darstellung des Reibegeräuschs bei Frikativen und Plosiven (gelber Punkt) aus- bzw. wieder eingeschaltet.

„switch color scheme“

Durch Anklicken dieser Schaltfläche wird die farbliche Darstellung im Animationsfenster geändert.

14 Die Ebenen des Artikulationsmodells

Die Berechnung der Animation im Artikulationsmodell beginnt mit der Textebene und führt über die Lautebene und die Ebene der artikulatorischen Targets bis hin zu den Artikulationsbewegungen.

Textebene:

Die Schreibung eines Wortes oder einer kurzen Äußerung ist im Textfeld (Geschrieben-Fenster) sichtbar.

Lautebene:

Die zugehörige Lautung (phonetische Transkription) ist im Lautfeld (Gesprochen-Fenster) sichtbar. Weitere Parameter, die auf dieser Ebene voreingestellt werden, können unter Menü „values > sounds > table“ eingesehen werden.

Ebene der artikulatorischen Targets:

Aufgrund der phonetischen Transkription (Lautung) werden die **Label** gesetzt und zugehörige artikulatorische Targetwerte spezifiziert. Diese Werte können unter Menü „values > targets > table“ eingesehen werden.

Jede Zeile enthält die Targetwerte, ihre Pertinenz und den zugehörigen Zeitpunkt für ein Label. Jede Zeile stellt somit die Information eines Labelpunktes dar. Die artikulatorischen Targets stellen somit die Ankerpunkte für die artikulatorischen Bewegungen und damit für den Ablauf der Animation dar. Das Konzept der artikulatorischen Targets und Bewegungen wird detailliert in Kap. 16 erläutert.

Ebene der artikulatorischen Bewegungen:

Aufgrund der artikulatorischen Targetwerte werden die artikulatorischen Bewegungen berechnet. Diese Bewegungen können unter Menü „values > movements > table“ eingesehen werden.

Jede Zeile enthält Zeitpunkt und artikulatorische Parameterwerte für eine Stützstelle. Diese Tabelle enthält somit die Daten aller **artikulatorischen Positionen** aller Artikulatoren für alle Zeitpunkte der Äußerung. Dies sind genau die Daten der **artikulatorischen Bewegungen** und damit die Daten der **Animation**.

15 Die Text-nach-Laut-Konversion

Die Text-nach-Laut-Konversion ist vorläufig. Insbesondere fehlt unserem Konversionsalgorithmus noch ein umfangreiches Ausnahmelexikon. Daher bedarf es etwas an Übung, um mit der hier vorhandenen Text-nach-Laut-Konversion schnell zu brauchbaren Transkriptionen zu gelangen.

Hier sollen deshalb nochmals kurz Regeln zum Umgang mit der Text-nach-Laut-Konversion zusammengestellt werden.

15.1 Abkürzungen und Zahlen

Abkürzungen oder Zahlen können nicht konvertiert werden.

Geben Sie beispielsweise „bzw.“ als „beziehungsweise“ ein.

Geben Sie „12“ als „zwölf“ ein.

15.2 Fremdwörter

Fremdwörter können in dieser Version von *SpeechTrainer* noch nicht konvertiert werden.

Beispielsweise „Computer“ als nicht aus dem Deutschen stammendes Wort kann nicht richtig konvertiert werden.

15.3 Wortgrenzen

Wortgrenzen können durch Leerzeichen oder durch das Underline-Zeichen markiert werden.

Beispielsweise: „Er malt“ oder „Er_malt“

Anmerkung: Wenn Einträge in einer Liste abgespeichert werden sollen, so dürfen in der Schreibung eines Listeneintrages keine Leerzeichen auftreten.

15.4 Silben- und Morphemgrenzen

Silben- oder Morphemgrenzen können durch Bindestrich markiert werden.

In vielen Fällen ist es sinnvoll, dies zu tun. Beispiele folgen:

Zwei Konsonanten nach Vokal führen häufig zu Kurzvokal:

Beispiel: „kamt“; richtige Transkription erfolgt bei Eingabe von: „kam-t“:

<kamt> -> [kamt] (falsch), <kam-t> -> [ka:mt] (richtig).

Ebenso: „fragt“ -> „frag-t“, „fackeln“ -> „fackel-n“.

Die Präfixe „be-“ oder „ge-“ sollten immer abgetrennt werden:

Beispiel: „begibt“; richtige Transkription erfolgt bei Eingabe von „be-gibt“

„begibt“ -> [bEgIpt] (falsch); „be-gibt“ -> [b@gipt] (richtig)

Beispiel: „gerecht“; richtige Transkription erfolgt bei „ge-recht“

„gerecht“ -> [gEAect] (falsch); „ge-recht“ -> [g@rEct] (richtig)

„...e-“ wird als Langvokal [e:] erkannt, „...e...“ wird als Kurzvokal [E] erkannt:

Beispielsweise: „le-bend“, aber „Le-ge Hühner“.

Oder: „e-ben“.

Dagegen sind „be-“ und „ge-“ im internen (kleinen) Ausnahmelexikon als [b@] bzw. [g@] gespeichert.

Deshalb nicht „eigne-t“ sondern „eigne t“

Nicht „abgeschnitten“ sondern „ab ge schnitten“ oder „ab ge-schnitten“, aber auch nicht „ab-ge-schnitten“, weil „-ge-“ nicht als Prefix behandelt wird.

„Ebene“ als „e-be ne“. Anders geht es nicht, weil das finale „e“ immer als [@] konvertiert wird.

Manchmal sollte eher morphemisch als silbisch getrennt werden:

Richtig: „ver-schieden-e“ oder „ver-schieden e“, nicht aber „ver-schiede-ne“. Es geht auch: „ver-schiede ne“ oder „ver-schiede nes“.

Es geht nur: „an ge-bunde nem“.

„e-be nem“ liefert Schwalaut [@]; dagegen: „ge-be-ten“ liefert langes [e:].

Es geht: „ein ge-schlagen“, „be-haupte t“, „le-ben-den“, „re-gne t“, „hatte st“.

Niemals Doppelkonsonanten durch Silbengrenze markieren:

Nicht „Ebbe“ oder „Eb-be“, sondern „Ebe“.

Doppelkonsonanten nur trennen, wenn dazwischen eine Morphemgrenze liegt:

„abbekommen“ oder „ab-bekommen“.

Wichtigste Regel: Versuchen Sie es erst mal ohne jegliche Trennung im Wort!

Denn: So schlecht ist unsere Text-Nach-Laut-Komponente nun auch wiederum nicht.

Bevor Sie sich zu viele Trennungsregeln merken: Viele Wörter werden auch direkt von unserer Text-Nach-Laut-Komponente richtig konvertiert, da diese die am häufigsten auftretenden Lautkombinationen richtig behandelt.

16 Artikulatorische Targets und Bewegungen

Die artikulatorischen Targets liefern das Gerüst für den Ablauf der artikulatorischen Bewegungen und damit auch die Basis für den zeitlichen Ablauf der Animation. Zunächst soll in diesem Kapitel das Konzept der Targets und der Erzeugung der artikulatorischen Bewegungen erläutert werden.

Zunächst ist es sehr wichtig zu verstehen, dass die Produktion der meisten Sprachlaute nicht allein auf eine statische Einstellung der Artikulationsorgane zurückgeführt werden kann. Vielmehr ist eine koordinierte *Bewegung* mehrerer Artikulationsorgane für das Gelingen der Produktion eines Lautes entscheidend. Deshalb beinhaltet *SpeechTrainer* auch ein Modell (segmentales Modell, segmentales Konzept) zur Erzeugung artikulatorischer Bewegungen. In diesem Konzept wird davon ausgegangen, dass maximal drei Zeitpunkte pro Laut nötig sind, um die wichtigsten artikulatorischen Ereignisse (z.B. Beginn oder Ende einer konsonantischen Verschlussbildung, Beginn oder Ende der Phonation etc.) zu beschreiben. Diese Zeitpunkte werden im segmentalen Konzept von *SpeechTrainer* durch die **Label** definiert.

Übung: Starten Sie das Programm *SpeechTrainer* und öffnen Sie die Liste „Listen\ Start\ Liste_05.st“ (Menü „file > open list“). Klicken Sie im Listenfenster den Eintrag „Igel“ an. Klicken Sie nun „Ausführen“ an. Sie sehen und hören die Animation des Wortes „Igel“ (Setzen Sie zum Studium der Animation evtl. die Ausführungsgeschwindigkeit herab; Menü „Einstellungen > Geschwindigkeit“). Öffnen sie nun das Dialogfenster „word properties“ durch Doppelklicken auf den Listeneintrag. Sie sehen in diesem dann das Oszillogramm und die Label des Wortes (des Listeneintrages) „Igel“. Die Funktion von Label 4, 6, 7 und 10 können anhand des Oszillogramms leicht erkannt werden. Label 4 („V Start [i]“) markiert den Zeitpunkt des Einsetzens der Phonation. Label 6 („B Start [g]“) markiert den Beginn der dorsalen Verschlussbildung.

Anmerkung: Die vollständige Bezeichnung eines Label wird in dieser Version 2.0 leider nicht in Dialogfenster „word properties“ dargestellt. Um mehr Information über ein Label zu bekommen, müssen Sie Menü „values > targets > table“ aufrufen (Dialogfenster „target table“).

Label 7 („B Ende [g]“) markiert das Ende der Verschlussbildung. Zu diesem Zeitpunkt setzt dann auch das mit der Verschlusslösung einhergehende Frikationsrauschen ein. Label 10 („L Ende [l]“) markiert den Endzeitpunkt der Phonation. Betrachten wir nun auch noch Label 5. Dieses Label wird mit „V Mitte [i]“ bezeichnet. Es befindet sich ungefähr im zeitlichen Mittelpunkt der Vokalproduktion des [i:].

Wir können also erkennen, dass für jeden Laut meist nur zwei, maximal aber drei Label festgelegt werden. Die Label pro Laut werden intern mit „Start“, „Mitte“ oder

„Ende“ bezeichnet. Es kann jedoch sein, dass eben eines (oder manchmal auch zwei) der Label für einen Laut gar nicht festgelegt werden (gar nicht erscheinen; siehe unten).

Zusätzlich werden die Label intern nach der Lautgruppe des Lautes, den sie repräsentieren gekennzeichnet. Beispielsweise steht „V“ für Vokale. Nachfolgend wird eine Übersicht über die möglichen Lautgruppen gegeben:

V	Vokal
P	stimmlose Plosive
B	stimmhafte Plosive
F	stimmlose Frikative
W	stimmhafte Frikative
N	Nasale
R	Vibranten
L	Laterale
G	Glottallaute
PR	Präphonation
PO	Postphonation

In der Kennzeichnung eines Labels in der Zeitdarstellung (z.B. Dialogfenster „word properties“) wird das konkrete Lautsymbol, z.B. „[i]“ in rechteckigen Klammern und darunter stehend die fortlaufende Nummer des Label angegeben.

Anhand der Bewegungsverläufe z.B. im dialogfenster „word properties“ können wir auch erkennen, dass nur zu den Zeitpunkten der Label artikulatorische Parameter (Positionen der Artikulatoren) festgelegt wurden. In den Zeitbereichen zwischen zwei Labels hingegen führen die Artikulatoren kontinuierliche Bewegungen hin zum nächsten Target durch. Beispielsweise erkennen wir eine konsonantische Hebung des Zungenrückens zwischen Label 5 und Label 6 (Parameter „TB:vr“).

Wir können auch erkennen, dass nicht jedes Label für jeden artikulatorischen Parameter einen Targetwert trägt. So sind beispielsweise die Label 1, 2, 3 und 4 für die Parameter der vokalischen Gesamtformung („VTS:lh“ und „VTS:bf“) transparent. Die vokalische Gesamtformung wird nur bei Label 0, bei Label 5 und bei Label 8 in Form von Targetwerten festgelegt.

Im Folgenden werden alle Lautgruppen und die zugehörige Generierung von Labels und von Targets erläutert.

Übungsfortsetzung: Klicken Sie nun im Dialogfenster „word properties“ auf „target panel“. Sie sehen nun die Tabellendarstellung der Targets. Jede Zeile steht hier für ein Label. Wir erkennen, dass bei jedem Label für jeden artikulatorischen Parameter ein Wert (Targetwert) gesetzt ist. Entweder ist dieser rot gekennzeichnet (Pertinence = nicht transparent) oder schwarz gekennzeichnet (pertinence = transparent). „Nicht transparent“ (rot) bedeutet, dass der Targetwert für diesen Parameter und dieses Label gilt. „Transparent“ (schwarz) bedeutet,

dass dieser Targetwert für diesen Parameter und dieses Label nicht gilt. Das entsprechende Label ist für diesen artikulatorischen Parameter also „transparent“.

Übungsfortsetzung: Sie erkennen beispielsweise, dass die Label 1, 2, 3 und 4 bezüglich der Parameter der vokalischen Gesamtformung („VST:lh“ und „VST:bf“) „transparent“ sind. Hingegen sind die Label 0 und 5 für diesen Parameter „nicht transparent“. Label 0 und Label 5 definieren also die vokalische Gesamtformung.

16.1 Vokale („V“)

Vokale werden je nach lautlichem Kontext durch ein bis drei Label beschrieben. Es tritt immer das Label „V Mitte“ auf. Dieses Label definiert die vokalische Gesamtformung (Parameter „VST:lh“, „VST:bf“ und „LI:vc“). Zu dem durch dieses Label festgelegten Zeitpunkt wird die Zungenrücken- und Lippenposition des zugehörigen Vokals eingenommen (so wie Sie durch Anklicken der Lautsymbole im Lautfenster gezeigt wird). Zusätzlich definiert das Label „V Mitte“ die Neutralstellung des Gaumensegels. Falls vor dem Vokal ein stimmloser Laut auftritt, wird zusätzlich ein Label „V Start“ hinzugefügt. Dieses Label hat nur die Funktion der Festlegung des Beginns der Phonation (Parameter „GL:co“). Falls nach dem Vokal ein stimmloser Frikativ auftritt, wird zusätzlich ein Label „V Ende“ hinzugefügt. Dieses Label hat nur die Funktion der Festlegung des Endes der Phonation (Parameter „GL:co“).

Übung: Laden Sie die „Listen\ Start\ Liste_05.st“ (Menü „file > open list“). Klicken Sie im Listenfenster den Eintrag „Efeu“ doppelt an. Es öffnet sich das Dialogfenster „word properties“ für dieses Wort. Sie erkennen, dass die Label 4, 5 und 6 für den Vokal [e:] die oben beschriebenen Funktionen übernehmen.

Übungsfortsetzung: Schließen Sie das Dialogfenster „word properties“ und Doppelklicken Sie im Listenfenster den Eintrag „Igel“. Sie erkennen nun im Dialogfenster „word properties“, dass für den Vokal [ə] nur ein Label (Label 8) generiert wurde. Dies liegt daran, dass der Schwa-Laut hier von stimmhaften Lauten umgeben ist.

16.2 Präphonation und Postphonation

Mit Präphonation („PR“) wird der Beginn und mit Postphonation („PO“) das Ende der Äußerung (des Wortes bzw. des kurzen Satzes) festgelegt.

Am Beginn jeder Äußerung / jedes Wortes, also im Bereich der Präphonation (PR) sind alle Artikulatoren in Neutralstellung und weisen damit ungefähr die artikulatorische Einstellung des Schwa-Lautes auf. Lediglich die Stimmritze ist geöffnet um die Einatmung darzustellen (GL:co = 600). All diese Einstellungen werden beim Label „PR Start [-]“ vorgenommen. Zusätzlich wird noch ein Label „PR Ende [-]“ generiert. Dieses Label ist dazu da, den Lungendruck der Äußerung so einzustellen,

dass genügend Luftdruck und Luftstrom während der Produktion der Äußerung zur Verfügung gestellt werden kann.

Übungsfortsetzung: Betrachten Sie das Ende der Zeitdarstellung des Eintrages „Igel“ im dialogfenster „word properties“.

Am Ende der Äußerung (Postphonation PO, Label 11) kehren alle Artikulatoren in die Neutralstellung zurück und die Stimmritze wird zum erneuten Einatmen wieder geöffnet. Diese Einstellungen werden mit dem Label „PO Ende [-]“ realisiert. Falls die Äußerung vokalisches endet, wird noch ein weiteres postphonatorisches Label generiert. Dieses Label „PO Start [-]“ definiert dann die Artikulatorstellung des letzten Vokals und definiert das Ende der Phonation.

Übungsfortsetzung: Betrachten Sie dazu den Endbereich der Zeitdarstellung des Wortes „Igel“ (Label 11) und des Wortes „Vieh“ (Label 6 und 7) in diesem Dialogfenster.

16.3 Stimmhafte Plosive („B“)

Stimmhafte Plosive werden immer durch zwei Label beschrieben: „B Start“ und „B Ende“. Das Start-Label definiert den Beginn der oralen Verschlussbildung und den Beginn (bzw. je nach vorangehendem Laut die Fortführung) der Phonation. Das Ende-Label definiert das Ende der oralen Verschlussbildung und das Ende (bzw. je nach nachfolgendem Laut die Fortführung) der Phonation. Zusätzlich definieren beide Label auch die maximal mögliche Anhebung des Gaumensegels (VE:lh).

Übungsfortsetzung: Betrachten Sie Label 6 und Label 7 des Wortes „Übel“ und Label 2 und Label 3 des Eintrages „beten“ in der Zeitdarstellung im selben Dialogfenster.

16.4 Stimmlose Plosive („P“)

Stimmlose Plosive werden immer durch zwei Label beschrieben: „P Start“ und „P Ende“. Beide Label definieren die orale Verschlussbildung und die maximale Anhebung des Gaumensegels. Das Ende-Label definiert zusätzlich eine maximale Öffnung der Stimmritze (GL:co = 400) zur Gewährleistung eines gut hörbaren Friktionsrauschens zum Zeitpunkt der oralen Verschlusslösung. Im Falle vorangehender stimmhafter Laute definiert das Start-Label zusätzlich das Ende der Phonation.

Übungsfortsetzung: Betrachten Sie Label 5 und Label 6 in der Zeitdarstellung des Wortes „bieten“.

16.5 Stimmhafte Frikative („W“)

Stimmhafte Frikative werden immer durch zwei Label beschrieben: „W Start“ und „W Ende“. Beide Label definieren die orale Engebildung, die maximale Anhe-

bung des Gaumensegels und die Sicherstellung der Phonation.

Übungsfortsetzung: Betrachten Sie Label 2 und Label 3 in der Zeitdarstellung des Wortes „sah“.

16.6 Stimmlose Frikative („F“)

Stimmlose Frikative werden wie stimmhafte Frikative gebildet. Beide Label definieren ebenfalls orale Engebildung und maximale Anhebung des Gaumensegels, aber bezüglich der Stimmritze die Sicherstellung einer Stimmritzenöffnung (GL:co = 400).

Übungsfortsetzung: Betrachten Sie Label 7 und Label 8 in der Zeitdarstellung des Wortes „Höfe“.

16.7 Nasale („N“)

Nasale werden immer durch drei Label beschrieben: „N Start“, „N Mitte“ und „N Ende“. Das Mitte-Label definiert die maximale Senkung des Gaumensegels. Die Label „N Start“ und „N Ende“ definieren (genau wie bei stimmhaften Plosiven) die orale Verschlussbildung und die Sicherstellung der Phonation.

Übungsfortsetzung: Betrachten Sie Label 7, 8 und 9 in der Zeitdarstellung des Wortes „Eimer“.

16.8 Laterale („L“)

Laterale werden durch zwei Label beschrieben: „L Start“ und „L Ende“. Die Label definieren jeweils Beginn und Ende der oralen Konstriktionsbildung. Zusätzlich stellen die Label die Phonation sicher.

Anmerkung: Zusätzlich wird mit der Konstriktion für den konstriktionsbildenden Artikulator (Lippen, Zungenrücken oder Zungenspitze) auch immer die Art der Engebildung („closure“ bei Plosiven und Nasalen, „fricative“ bei Frikativen, „lateral“ bei Lateralen und „vibrant“ bei Vibranten) definiert (Kap. 7). Die Verläufe dieser Parameter werden auch in der Zeitdarstellung als farbiger Hintergrund markiert. Die Art der Engebildung wird bei Vokalen immer für alle konstriktionsbildenden Artikulatoren wieder auf „closure“ zurückgestellt. Im Falle benachbarter Konsonanten mit gleichem konstriktionsbildendem Artikulator wird die Art der Engebildung für diesen Artikulator natürlich durch diesen Nachbarlaut geändert.

Übungsfortsetzung: Betrachten Sie Label 6 und Label 7 in „Öl“. Die Artikulationsart „lateral“ wird durch die hellgrüne Hintergrundfarbe für den Parameter (TT:vr) angedeutet.

16.9 Vibranten („R“)

Vibranten werden wie Laterale durch zwei Label beschrieben: „B Start“ und „B Ende“. Die Label definieren jeweils Beginn und Ende der oralen Konstriktionsbildung. Zusätzlich stellen die Label die Phonation sicher. Die Artikulationsart „Vibrant“ wird durch die

rosa Hintergrundfarbe für den entsprechenden Artikulator definiert.

Übungsfortsetzung: Betrachten Sie Label 4 und Label 5 in „froh“.

16.10 Glottallaute

Als Glottallaute werden der Glottisverschlusslaut [ʔ], die Glottalisierung [~], der stimmhafte glottale Frikativ [H] und der stimmlose glottale Frikativ [h] bezeichnet. Glottallaute werden durch zwei Label beschrieben: „G Start“ und „G Ende“. Die Label definieren jeweils Beginn und Ende der glottalen Konstriktionsbildung. Die Label dieser Laute bleiben für alle anderen Artikulatoren transparent.

Übungsfortsetzung: Betrachten Sie Label 2 und Label 3 in „Höfe“. Betrachten Sie Label 2 und Label 3 in „Efeu“.

17 Grundlagen der artikulatorischen und akustischen Phonetik

Die zum Umgang mit *SpeechTrainer* nötigen phonetischen Grundlagen sollen in diesem und in den folgenden Kapiteln vermittelt werden. In diesem Kapitel wollen wir uns mit den Grundlagen der Artikulation und dem daraus resultierenden akustischen Signal beschäftigen.

Der Sprechapparat des Menschen kann funktional wie anatomisch in 3 Bereiche unterteilt werden. Die **Lunge** stellt aerodynamische Energie (Luftdruck und potentielle Möglichkeit des Luftstroms) zur Verfügung (Initiation, siehe Catford 1988). Im **Kehlkopf** wird mittels dieser Energie ein akustisches Signal produziert (Phonation) und im **Ansatzrohr** (Rachen-, Mund- und Nasenraum) wird dieses Signal moduliert (Artikulation). Somit kann der Sprechapparat in den **sublaryngalen Bereich** (Lunge, Bronchien und Luftröhre), den **Larynx** (Kehlkopf) und den **supralaryngalen Bereich** (Rachen-, Mund- und Nasenraum) unterteilt werden (Abb. 16).

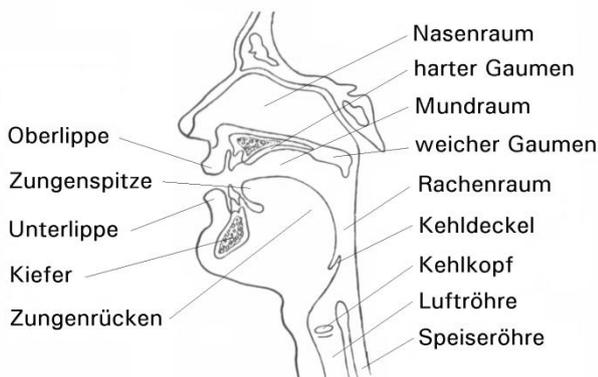


Abbildung 16. Der Sprechapparat des Menschen

17.1 Initiation

Die Lungenaktivität während des Sprechens hat – neben den primär lebenserhaltenden Funktionen (Gasaustausch mit dem Blutkreislauf) - aus der Sicht der Sprachproduktion zwei wichtige Funktionen: zum einen die Herstellung und Aufrechterhaltung eines zeitlich konstanten **Luftüberdrucks** in der Lunge, zum anderen die Bereitstellung einer bestimmten Luftmenge als Basis für den zum Sprechen nötigen **Luftstrom**. Die Aktivität der Gesamtheit der Atmungsmuskulatur ist während des Sprechens darauf ausgerichtet, den Lungenluftdruck annähernd konstant zu halten (siehe Ohala 1990).

Anatomisch wird die Lunge vom Brustkorb (Brustbein und Rippen) und nach unten vom Zwerchfell gegenüber dem Bauchraum abgegrenzt. Entsprechend unterscheiden wir **Brustatmung** (Hebung und Senkung des Brustkorbes) und **Bauchatmung** (Hebung und Senkung des Zwerchfells). Wichtige Muskeln der Einatmung

sind der Zwerchfellmuskel (Diaphragma), dessen Kontraktion die Absenkung des Zwerchfells bewirkt, und die äußeren Zwischenrippenmuskeln (Intercostalsi externi), dessen Kontraktion die Hebung des Brustkorbes bewirkt. Darüber hinaus tragen Teile der Hals-, Schulter- und Rückenmuskulatur zur Einatmung bei. Wichtige Muskeln der Ausatmung sind die inneren Zwischenrippenmuskeln (intercostalis interni) aber auch Teile der Bauch- und Rückenmuskulatur. (siehe Abb. 17). Die Muskulatur dient somit der Änderung des Lungenvolumens, steuert aber damit den Luftdruck in der Lunge.

Die Aktivität der Atmungsmuskulatur beim Sprechen dient der Aufrechterhaltung eines **zeitlich konstanten Überdrucks** in der Lunge. Nach dem Einatmen ist der Brustkorb wie ein Luftballon gespannt, so dass auch ohne Beteiligung von Muskulatur ein Überdruck existiert (Ballondruck). Dieser Ballondruck sinkt mit Abnahme des Lungenvolumens (Abb. 17, gestrichelte Linie). Zur Realisierung eines zeitlich konstanten Drucks beim Sprechen wirkt zunächst die Muskulatur der Einatmung dem Ballondruck entgegen. Sinkt der Ballondruck unter den erwünschte Sprechluftdruck (Zeitpunkte rechts vom vertikalen Doppelpfeil in Abb. 17), so wird die Ausatmungsmuskulatur aktiv.

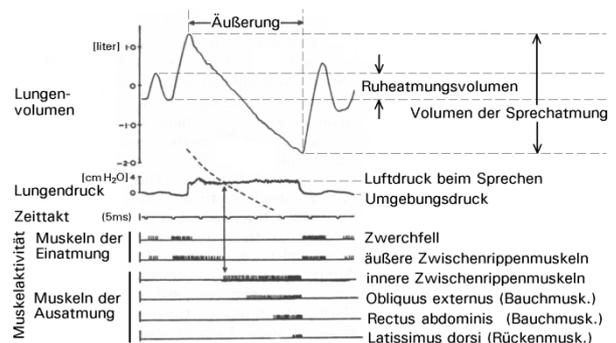


Abbildung 17. Lungenvolumen, Lungenluftdruck, Zeittakt eines Zeitgebers und die Aktivität der Atmungsmuskulatur während der Produktion einer Äußerung (nach Draper, Ladefoged et al. 1959). Die Zeitachse verläuft horizontal von links nach rechts.

Der Lungenluftdruck wird über Änderungen des Lungenvolumens gesteuert. Das maximale Volumen der Lunge eines Menschen liegt bei rund 7 Litern (Maximalvolumen). Nach Ausatmen mit maximaler Muskelkraft liegt ein Lungenvolumen von rund 2 Litern vor (Restvolumen). Die Differenz von 5 Liter wird als **Vitalkapazität** bezeichnet. Die Kapazität der **Ruheatmung** liegt bei rund 1 bis 2 Litern, die der **Sprechatmung** liegt je nach Lautstärke um 3 Liter (siehe auch Abb. 17).

Die Stärke des während des Sprechens durch Kehlkopf und Ansatzrohr ausströmenden Luftstroms hängt aber nicht nur von der Aktivität der Atmungsmuskulatur sondern auch wesentlich von dem **Grad der Engebildung** (Konstriktionsbildung) im Bereich des Kehlkopfes und des Ansatzrohres ab. Bei starker Engebildung kann Luft eben nur sehr langsam ausströmen.

17.2 Phonation

Im Kehlkopf wird ein Teil der aerodynamische Energie (Überdruck unterhalb des Kehlkopfes und Luftstrom durch den Kehlkopf) in **Schwingungen der Stimmlippen** und in ein **akustisches Signal** (Schwingungen der Luftsäule im Ansatzrohr) gewandelt. Die Entstehung von Stimmlippenschwingung und damit die Entstehung eines akustischen Signals im Kehlkopf wird als **Phonation** bezeichnet. Der Kehlkopf befindet sich zwischen Luftröhre und Rachenraum (Abb. 16). Er besteht aus Knorpeln, Gewebe und Muskulatur (Abb. 18). Der Ringknorpel (Cricoid) bildet den oberen Abschluß der Luftröhre. Mit dem Ringknorpel verbunden sind der nach vorne kippbare Schildknorpel (Thyroid) und zwei Stellknorpel (Arytenoid). Zwischen Schild- und Stellknorpel verlaufen die Stimmlippen. Die Stimmlippen bestehen aus Muskelfasern (Vocalis-Muskel oder Thyroarytenoideus (TA)) und Gewebehaut. Durch Änderung des Abstandes zwischen den Stellknorpeln ändert sich der Abstand der Stimmlippen zueinander (Abb. 18). Die Öffnung zwischen den Stimmlippen wird auch als **Stimmritze** (Glottis) bezeichnet. Aus funktional phonetischer Sicht sind zwei Stimmlippenparameter wichtig: die **Längsspannung der Stimmlippen** und der **Abstand der Stimmlippen** zueinander. Durch Kippen des Schildknorpels kann die Längsspannung der Stimmlippen geändert werden. Die Positionierung von Schild- und Stellknorpel geschieht durch die innere Kehlkopfmuskulatur. Wichtige Kehlkopfmuskeln zur Positionierung der Stellknorpel (und damit zur Einstellung des Abstandes der Stimmlippen zueinander) sind der Stimmlippenöffner (Cricothyroideus posterior PCA) und die Stimmlippenschließer (Cricothyroideus lateralis LCA, Interarytenoideus IA). Die Positionierung des Schildknorpels wird in erster Linie vom Stimmlippenspanner (Cricothyroideus, CT) geleistet.

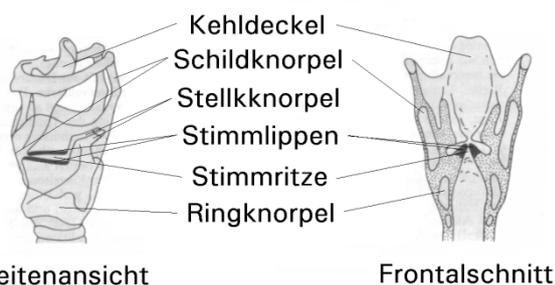


Abbildung 18. Kehlkopf: Seitenansicht und hinterer Frontalschnitt

Für die Erzeugung von regelmäßigen Schwingungen der Stimmlippen und damit für die Erzeugung von Schall, **Phonation** genannt, ist neben der Bereitstellung von aerodynamischer Energie auch eine geeignete Positionierung der Stimmlippen, nämlich ihre lockere Verschiebung und eine Spannung der Stimmlippen längs der Stimmritze nötig (Abb. 19b). Je höher die Längsspannung der Stimmlippen, umso höher die Tonhöhe

des erzeugten Schallsignals. Man unterscheidet aktive Spannung der Stimmlippen (Aktivität des TA) und passive Spannung der Stimmlippen (durch Kippen des Schildknorpels, Aktivität des CT). Der richtige Abstand der Stimmlippen zueinander (Stimmlippenabstand) entscheidet über das Auftreten von Phonation. Bei zu festem Verschluss der Stimmlippen, d.h. beim Zusammenpressen der Stimmlippen (Glottisverschluss) ist das Schwingen der Stimmlippen nicht möglich. Aber auch bei zu weitem Öffnen der Stimmritze endet die Schwingung der Stimmlippen.

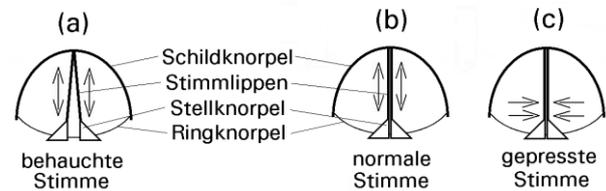


Abbildung 19. Aufsicht auf die Stimmritze.

Bei der **Phonation** werden die Stimmlippen durch die Kehlkopfmuskulatur wie oben beschrieben eingestellt und gespannt. Die Schwingung der Stimmlippen führt zu einer Abfolge von sich öffnender und wieder verschließender Stimmritze (Abb. 20). Die Stimmlippenschwingung entsteht aber nicht direkt aufgrund der Aktivität der Kehlkopfmuskulatur (schnelle Abfolge von Muskelkontraktionen), sondern aufgrund der Bereitstellung aerodynamischer Energie zusammen mit der richtigen laryngalen Voreinstellung (lockere (nicht zu feste) Verschiebung der Stimmlippen und leichte Längsspannung). Es sind insbesondere zwei Mechanismen der Energieübertragung zu nennen, die in geeigneter Weise zeitlich geschickt (phasenrichtig) zusammenarbeiten, um die Schwingung einzuleiten und aufrecht zu erhalten. (1) Der Luftdruck unterhalb der Stimmlippen drückt die geschlossenen Stimmlippen auseinander (Öffnungskraft innerhalb der Verschlussphase). (2) Die hohe Strömungsgeschwindigkeit der Luft zwischen den (dann) geöffneten Stimmlippen sorgt für die Ausbildung eines Unterdrucks zwischen den Stimmlippen (Bernoulli-Effekt) und damit für eine Anziehung der Stimmlippen (Schließkraft innerhalb der Offenphase). Der Vorgang beginnt dann erneut.

Anmerkung: Dieser Mechanismus zur Erzeugung und Aufrechterhaltung der Stimmlippenschwingung beim Sprechen (auch beim Singen) existiert in vergleichbarer Weise auch bei der Tonerzeugung des Blechbläfers mit seinem Blechblasinstrument. Hier wird nun die Funktion der Stimmlippen von den Lippen des Bläfers übernommen.

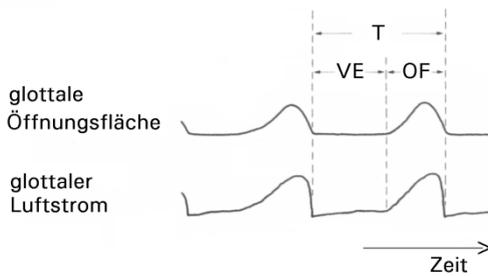


Abbildung 20. Öffnungsfläche der Stimmritze und zugehöriger Luftstrom als Funktion der Zeit. T ist die Schwingungsdauer einer Schwingungsperiode, OF kennzeichnet die Offenphase der Stimmlippen und VE die Verschlussphase der Stimmlippen innerhalb einer Schwingungsperiode.

Die **Tonhöhe** der Stimme wird in erster Linie durch die Stärke der Stimmlippenspannung (Längsspannung der Stimmlippen) determiniert, während die **Lautstärke** hauptsächlich durch die Höhe des Lungenluftdrucks und damit des Luftdrucks unterhalb der Stimmlippen determiniert wird. Je höher der Lungenluftdruck ist, um so lauter erklingt die Stimme.

Der Abstand der Stimmlippen hingegen determiniert die **Stimmqualität**. So wird die Stimme als „knarrend“ oder „gepresst“ bezeichnet, wenn starke Verschlusskräfte senkrecht zur Stimmritze wirken (Abb. 19c). Dies tritt auch bei der **Glottalisierung** auf. Sind diese Verschlusskräfte sehr stark, kann keine Schwingung der Stimmlippen mehr auftreten. Die Stimmlippen sind nun fest verschlossen. Dieser feste Verschluss der Stimmritze ist das Merkmal des **Glottisverschlusslautes** [ʔ] (IPA 1993) und tritt z.B. im Wörtern mit anlautendem Vokal auf (z.B. das Gegensatzpaar „Heiner“ und „Einer“ am Wortanfang von „Einer“). Sind die Stimmlippen hingegen leicht geöffnet (und zwar bereits vor Einsetzen der Schwingung der Stimmlippen), so ergibt sich die Stimmqualität „behaucht“ (Abb. 19a). Hier dürfen die Stimmlippen nicht zu weit geöffnet sein, da sonst keine Stimmlippenschwingung mehr auftritt. Das Vorhandensein bzw. Das Nichtvorhandensein von Stimmlippenschwingung unterscheidet genau die Lautkategorien **stimmhafte und stimmlose Laute** (stimmhaft gegen stimmlos: z.B. [t] gegen [d] in „Tank“ kontra „Dank“). So ist im Fall der **stimmlosen Laute** (mit Ausnahme des [ʔ]) eine weite Öffnung der Stimmritze gegeben. Normale Stimmqualität wird, wie oben beschrieben, bei locker verschlossenen Stimmlippen produziert (Abb. 19b).

Anmerkung: Auch im Fall des Parameters der Stimmlippenspannung und der Öffnungsweite der Stimmritze gibt es wieder Analogien zur Tonerzeugung bei Blechblasinstrumenten. Die Lippenspannung ist (neben dem Drücken der Ventile) zur Einstellung der Tonhöhe wichtig und der Grad der Lippenöffnung im Mundstück entscheidet über die Tonqualität.

Neben der Erzeugung von Sprechschall aufgrund der Phonation (z.B. bei der Produktion von Vokalen) kann Sprechschall auch durch **Friktion** erzeugt werden (z.B.

bei der Produktion von stimmlosen **Frikativen** wie [s] oder [f]). Diese Form des Sprachschalls entsteht im Ansatzrohr aufgrund einer starken Engebildung (z.B. bei [f]: Enge zwischen Oberlippe und unterer Zahnreihe). Aufgrund einer solchen Enge wird die durch das Ansatzrohr fließende Luft stark beschleunigt und nach Ausströmen aus der Enge entstehen (aufgrund der Luftreibung: Friktion) Luftwirbel, die wiederum (hörbare) Luftdruckschwankungen verursachen. Ein wesentlicher Unterschied zwischen Phonations- und Friktionsschall ist die Periodizität. Phonationsschall ist periodisch (Abb. 20). Es handelt sich um einen Klang mit definierter Tonhöhe. Friktionsschall ist hingegen ein Rauschen, also ein nichtperiodisches Signal.

Merke: **Phonationsschall** tritt beispielsweise bei **Vokalen, Nasalen, Lateralen, Vibranten, stimmhaften Frikativen und stimmhaften Plosiven** auf. **Friktionsschall** tritt bei allen **Frikativen und Plosiven** auf.

Die Parameter der Initiation und Phonation wie z.B. Tonhöhe, Lautstärke und Stimmqualität (z.B. gepresste, normale, behauchte Stimme) kodieren **prosodische** Information (z.B. Frage kontra Aussage, emotionaler Zustand des Sprechers etc.). Daneben existiert die (oft als primär betrachtete) **segmentale** Information. diese wird in erster Linie durch die Parameter der Artikulation kodiert. Durch unterschiedliche Einstellungen und Bewegungen der supralaryngalen Artikulatoren (Lippen, Zunge, etc.) werden Laute realisiert und damit Bedeutungsunterschiede auf Wortebene signalisiert.

17.3 Artikulation

Das an der Stimmritze erzeugte akustische Signal – also die akustische Konsequenz der Stimmlippenschwingung – wird im Ansatzrohr moduliert. Die Stimmritze stellt eine **akustische Quelle** und das Ansatzrohr (Rachen-, Mund- und Nasenraum, Abb. 16) ein **akustisches Filter** dar. Die Filterwirkung des Sprechtraktes ist frequenzabhängig. Die spektrale Analyse des Quellsignals ergibt ein Linienspektrum mit kontinuierlich abnehmender Amplitude der Spektrallinien (Abb. 21a). Die **Tonhöhe** (f_0) des hier dargestellten Klanges kann an der Frequenz des ersten Teiltones (erste Spektrallinie) oder auch am Abstand der Spektrallinien abgelesen werden. Die **Klangfarbe** hingegen wird durch die Gesamtform des Spektrums, also durch den Gesamtverlauf der Amplituden der Spektrallinien determiniert (siehe gestrichelte Linie in Abb. 21a). Die vom Ansatzrohr bewirkte akustische Modulation des Quellsignals ist ein Resultat der Ausbildung von stehenden Schallwellen im Ansatzrohr. Diese stehenden Wellen werden auch Resonanzen des Ansatzrohres oder **Formanten** genannt. Die Frequenz jedes Formanten hängt von der Geometrie des Ansatzrohres, d.h. von der Länge des Ansatzrohres von Stimmritze bis Lippen und von der Formung des Ansatzrohres aufgrund von Lippen und Zungenstellung ab. Eine Resonanz des Ansatzrohres (Formant) bedingt ein akustisches Amplitudenmaximum. Die Amplitudenmaxima der ersten drei Formanten eines Ansatzrohres können anhand der Filtercharakteristik des Ansatzrohres (Abb.

21b) erkannt werden. Die zugehörigen Formantfrequenzen (F1, F2, F3) können ebenfalls anhand dieser Maxima ermittelt werden. Im Unterschied zu der kontinuierlich abnehmenden Amplitude der Spektrallinien des Quellsignals bewirkt die akustische Filterung durch das Ansatzrohr, dass die Amplitudenmaxima und damit die Formanten auch anhand des Spektrums des vom Munde abgestrahlten Schallsignals abgelesen werden können (Abb. 21c). Die Modulation (Filterung) des Quellsignals durch das Ansatzrohr ändert somit den Gesamtverlauf der Amplituden der Spektrallinien und damit also insbesondere den Klang des akustischen Signals. Die Tonhöhe und damit die Lage der Spektrallinien auf der Frequenzachse bleiben hingegen unverändert.

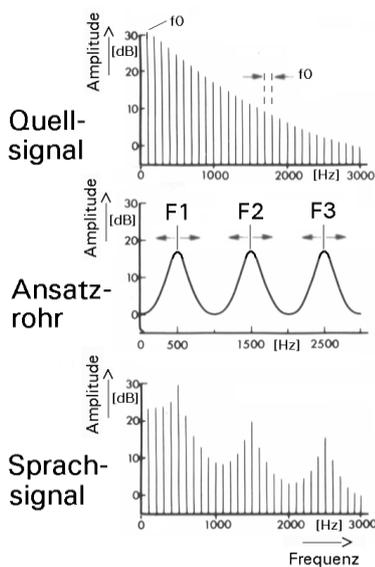


Abbildung 21. Darstellung des Spektrums (a) des Quellsignals, (b) der Filtercharakteristik des Ansatzrohres und (c) des Sprachsignals.

Anmerkung: Ein systematischer Unterschied in der Klangerzeugung beim Sprechen und Singen gegenüber der Klangerzeugung bei Blasinstrumenten ist aber die Unabhängigkeit bzw. Abhängigkeit der Tonhöhe von der Länge des Ansatzrohres. Während die Länge des (Ansatz-)Rohres bei Blasinstrumenten die Tonhöhe determiniert, ist dies beim Sprechen und Singen nicht der Fall. Hier wird die Tonhöhe einzig und allein vom pulmonalen und laryngalen System determiniert. Grund hierfür ist, dass beim Sprechen und Singen zumeist keine starke Kopplung zwischen Stimmlippen und Ansatzrohr besteht, während bei der instrumentalen Tonerzeugung zwischen den Lippen und dem Rohr des Instrumentes eines Blechbläusers oder zwischen den Blättern und dem Rohr des Instrumentes eines Holzbläusers wohl starke Kopplungen auftreten. Die Grundtonhöhe f_0 ist beim Sprechen somit völlig unabhängig von den Formantfrequenzen F1, F2, F3 usw. Wohl determiniert f_0 den Abstand aller Teiltöne (Abb. 21a, 21c). Insbesondere muß einer Formantfrequenz auch nicht ein ganzzahliges Vielfaches von f_0 sein und die Formantfrequenzen müssen auch untereinander nicht in einem bestimmten Verhältnis zueinander stehen. Vielmehr können sie aufgrund der

Variabilität der Ansatzrohrform um die hier dargestellten Mittelwerte variieren (Abb. 21b).

Die Formänderung des Ansatzrohres wird durch Lageänderung der aktiv beweglichen **Artikulatoren** – vornehmlich der **Zunge**, der **Lippen** und des **Gaumensegels** (Abb. 16) – bewirkt. Zunge und Lippen sind komplex aufgebaute Muskelsysteme.

Die **Zunge** unterteilt sich in intrinsische und extrinsische Zungenmuskulatur. Die intrinsische Zungenmuskulatur baut die Zunge auf und bewirkt hauptsächlich eine Änderung der Zungenform. Insbesondere kann eine Hebung der **Zungenspitze** zur **konsonantischen Artikulation** realisiert werden (Musculus longitudinalis). Die extrinsische Zungenmuskulatur verbindet die Zunge mit Kiefer, Schädelknochen und über das Zungenbein auch mit dem Kehlkopf. Sie bewirkt vor allem die Änderung der Zungenlage. Wichtige extrinsische Zungenmuskeln sind z.B. der die Zunge mit dem Kieferknochen verbindende Genioglossus, und der die Zunge mit dem Schädel verbindende Styloglossus.

Die **Lippen** werden hauptsächlich vom Lippenringmuskel (orbicularis oris) aufgebaut, der die Lippenverschlussbildung sowie die Lippenrundung realisiert. Weitere Lippenmuskeln realisieren die Oberlippenhebung, Unterlippensenkung, Lippenstülpung und das Zurückziehen und anhaben der Mundwinkel.

Das **Gaumensegel** (Velum) umfasst den Bereich des **weichen Gaumens** (Abb. 16). Am Ende des Gaumensegels befindet sich das **Zäpfchen** (Uvula). Das Gaumensegel wird durch den levator palatini und tensor palatini angehoben und gestrafft. Seine Funktion beim Sprechen ist vor allem die Verbindung bzw. Trennung des Nasenraumes vom Mund- und Rachenraum. Ist das Velum gesenkt (Abb. 16), so ist eine akustische Ankopplung des Nasenraumes gegeben. Diese ist z.B. bei der Realisierung von **Nasalen** (z.B. [m], [n]) wichtig.

Unterschiedliche Laute werden durch unterschiedliche Bewegungen der supralaryngalen Artikulatoren realisiert. **Konsonanten** werden nach Artikulationsort und Artikulationsart unterschieden.

Als **Artikulationsorte** werden u.a. **bilabial** (an den Lippen gebildet, z.B. [p, b, m]), **labiodental** (an den Zähnen gebildet, z.B. [f, v]), **alveolar** (am Zahndamm gebildet, z.B. [t, d, n]), **palatal** (am harten Gaumen gebildet, z.B. [c, j]), **velar** (am weichen Gaumen gebildet, z.B. [k, x, ŋ]), **uvular** (am Zäpfchen gebildet, z.B. [ʀ]) und **glottal** (an der Stimmritze gebildet, z.B. [ʔ, h]) unterschieden.

Bezüglich der **Artikulationsart** werden u.a. **Plosivlaute (Plosive)** (vollständiger Verschluss, z.B. [p, b, t, d]), **Frikativlaute (Frikative)** (Engebildung mit Luftverwirbelung, z.B. [f, v, s, z]), **Nasallaute (Nasale)** (gesenktes Velum, z.B. [m, n]) unterschieden. Unterschiedliche **Artikulationsart** kennzeichnet zum Teil den Grad der Engebildung (z.B. Plosive kontra Frikative) aber auch die Dauer einer Engebildung und damit die Art der

Artikulatorbewegung (z.B. sehr kurze Engebildung bei Approximanten). Darüber hinaus differenziert die Beteiligung der Stimme das System der Konsonanten in **stimmhafte und stimmlose Laute** (z.B. [b] vs. [p], [s] vs. [z]).

Weitere Artikulationsarten sind **Laterallaute (Laterale)** (Verschlussbildung in der Medianebene aber gleichzeitig seitliche Öffnung, z.B. [l]) und Laute mit vibrierendem Artikulator (**Vibranten**), z.B. [r]. Hier ergeben sich in der Tat sehr unterschiedliche Arten der Engebildung.

Während Konsonanten artikulatorisch durch lokale Engebildung im Ansatzrohr charakterisiert werden, können **Vokale** nur durch die Formung des gesamten Ansatzrohrform wird vor allen durch die **Zungenlage (Höhe und Grad der Vor- bzw. Rückverlagerung)** und die **Lippenformung (Öffnungsweite und Grad der Rundung bzw. Spreizung)** determiniert. Eine ausführliche und vollständige Darstellung aller Lautklassen gibt z.B. Pompino-Marschall (1995).

Die **Formanten** – insbesondere die Formantfrequenzen – werden durch die **Artikulation** geändert. Ein wichtiger Faktor ist die Ansatzrohrlänge. Haben wir ein Ansatzrohr konstanten Querschnitts (dies ist am ehesten beim „Murmelvokal“ [↔] erfüllt), so ergeben sich die Formantfrequenzen bei einer typischen Ansatzrohrlänge von 17,5 cm in regelmäßigen Abständen von 1000 Hz beginnend mit dem ersten Formanten bei 500 Hz (Abb. 21b). Durch Änderungen der Lippen- und Zungenstellung können die Formanten nun auf der Frequenzachse verschoben werden (Pfeile in Abb. 21b). Die Abhängigkeit der Formantfrequenzen von der Geometrie des Ansatzrohres ist sehr komplex. Es können aber zwei Faustregeln angegeben werden. Zum einen bewirkt die zunehmende Mundöffnung (Kiefersenkung) eine Absenkung der Frequenz des ersten Formanten. So zeichnet den Vokal [a:] eine tiefe, die Vokale [i:] und [u:] hingegen eine hohe Frequenz des ersten Formanten aus (Abb. 22, vertikale Pfeile). Zum anderen bewirkt die zunehmende Vorverlagerung des Zungenrückens in den palatalen Bereich eine Anhebung und die Rückverlagerung in den velaren Bereich eine Absenkung der Frequenz des zweiten Formanten. So zeichnet sich der Vokal [i:] durch eine hohe, der Vokale [u:] hingegen durch eine niedrige Frequenz des zweiten Formanten aus (Abb. 22, horizontale Pfeile).

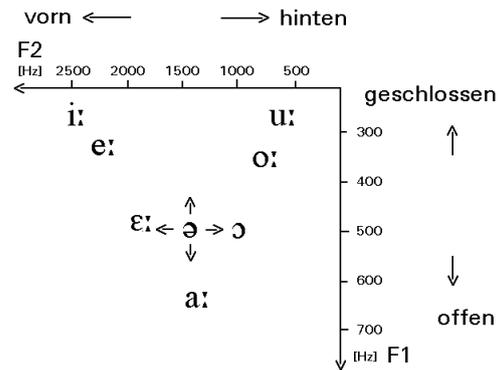


Abbildung 22. Zum Zusammenhang zwischen akustischem und artikulatorischem Vokalraum. Durchschnittliche Frequenzen des ersten und zweiten Formanten von acht Vokalen des Deutschen und zugehörige artikulatorische Kategorien sind gekennzeichnet.

Anmerkung: Anstelle von geschlossen und offen werden von *SpeechTrainer* die Begriffe hoch und tief benutzt: hohe bzw. tiefe Vokale.

Ähnliche Gesetzmäßigkeiten können auch für die **lokale Engebildungen** und damit für die Produktion von Konsonanten aufgestellt werden. Durch eine lokale Engebildung kommt es ebenfalls zu Änderungen (Verschiebungen) von Formantfrequenzen. Diese Formantfrequenzänderungen sind umso stärker, je stärker die Engebildung wird. Da im zeitlichen Ablauf der Artikulation in allen Übergangsbereichen von Vokalen nach Konsonanten und von Konsonanten nach Vokalen eine Zunahme bzw. Abnahme des Grades der Engebildung vorliegt, entsteht hier eine kontinuierliche Änderungen der Formantfrequenzen; Es treten **Formanttransitionen** auf (Abb. 23b). Die Formanttransitionen in Abb. 23b stellen das akustische Abbild der **Artikulatorbewegung** zur Lösung einer artikulatorischen Enge von Konsonant zum Vokal dar, nämlich der Öffnungsbewegung der Lippen bei [b], der Senkung der Zungenspitze bei [d] und des Zungenrückens bei [g] (Abb. 23a). Es zeigt sich, dass je nach Lage der artikulatorischen Enge (Artikulationsort labial, alveolar oder palatal/velar) Änderungen der Formantfrequenzen in unterschiedliche Richtung erfolgen. So ändert sich der erste Formant zwar immer gleichförmig (Absenkung mit zunehmender Engebildung) der zweite Formant zeigt aber beispielsweise bei der Ausbildung einer labialen Enge eine Absenkung, ansonsten aber eine Anhebung. Der dritte Formant zeigt nur bei einer alvelaren Enge eine Anhebung und eine Absenkung sonst (Abb. 23b). Somit können insbesondere diese drei konsonantischen **Artikulationsorte** allein über die **Richtung der Formanttransitionen** der ersten drei Formanten unterschieden (kodiert) werden.

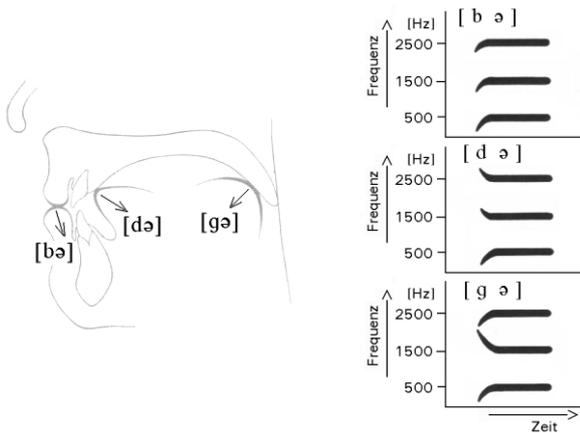


Abbildung 23. Die Formantentransitionen beim Übergang vom labialen, apikalen bzw. velaren Verschluss zum Murrelvokal. Die Pfeile im linken Bildteil deuten die artikulatorische Öffnungsbewegung beim Übergang vom Plosiv in den Vokal an. Die zugehörigen Formantentransitionen sind im rechten Bildteil angegeben.

17.4 Zur phonetischen Realisierung ganzer Äußerungen

Die Artikulationsorgane sind während der Lautproduktion in ständiger **Bewegung**. Statische Artikulatorpositionen treten während der Produktion von fließender Rede (Wörter oder Sätze) auch in vokalischen Bereichen einer Äußerung nicht auf. Betrachten wir die Frequenzverläufe der Formanten einer Silbe (Abb. 24), so sehen wir, dass sich die Formantfrequenzen als Folge von kontinuierlichen Artikulatorbewegungen und damit als Folge einer ständigen Wandlung der Ansatzrohrform kontinuierlich ändern.

Darüber hinaus ist die artikulatorische, aber auch die akustische Realisierung von Lauten stark **kontextabhängig**. So kann insbesondere die artikulatorische und akustische Gestalt von Konsonanten stark von den umliegenden Vokalen abhängen. Aber auch die prosodischen Eigenschaften wie Sprechtempo oder Betonungsgrad beeinflussen Lautrealisierungen.

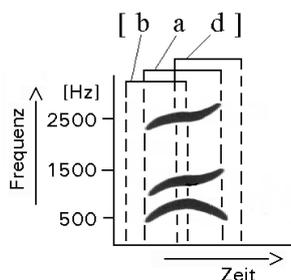


Abbildung 24. Verläufe des ersten, zweiten und dritten Formanten der ersten drei Laute des Logatoms „Badle“. Oben eingezeichnet: Zeitbereiche der akustischen Merkmale der Laute.

Somit ist ein (durch die Lautschrift evtl. naheliegendes) Modell einer nicht überlappenden Aneinanderreihung diskreter und invarianter Einheiten in der akustischen wie auch in der artikulatorischen Domäne nicht richtig. Neben der gegenseitigen Beeinflussung haben wir auf der artikulatorischen aber auch auf der akustischen Ebene ein **zeitliches Überlappen von Lautmerkmalen**. So kodiert beispielsweise eine Formantentransition den Artikulationsort eines Plosivs. Andererseits gehört die Gesamtheit aller Formantentransitionen während der Produktion eines Vokals (z.B. Strecke der Phonation zwischen zwei konsonantischen Verschlussphasen, Abb. 24) auch zum Vokal.

Das in *SpeechTrainer* realisierte segmentale Artikulationsmodell trägt diesen Gegebenheiten Rechnung. Targetwerte werden nur zu den durch die Label definierten Zeitpunkten erreicht. Somit bewegen sich die Artikulatoren ständig von Target zu Target. Die somit resultierenden Artikulatorbewegungen liefern genau die oben beschriebenen Formantentransitionen. Somit ist die Lautartikulation auch in dem von uns benutzten segmentalen Artikulationsmodell immer kontextabhängig und zeitlich überlappend, da die Artikulatorbewegungen in diesem Modell nicht durch ein isoliertes (einen Laut repräsentierendes) Target, sondern erst durch den Übergang von Target zu Target festgelegt werden.

Das segmentale Artikulationsmodell ist zwar aus Sicht der phonetischen Theoriebildung recht simpel, aber genau darin liegt auch seine Stärke: Die Vorzüge des segmentalen Ansatzes liegen vor allem in seiner einfachen und robusten Handhabbarkeit. Die Synchronisation (Kap. 10) kann durch alleiniges Versetzen von Labels realisiert werden. Neben dem segmentalen Modell wurde aber mittlerweile auch ein vollständiges **gestisches Artikulationsmodell** realisiert (Kröger 1998). Dieser Ansatz basiert auf der Theorie der artikulatorischen **Geste** (auch Artikulatorische Phonologie nach Browman und Goldstein 1989, 1990 und 1992) und gibt insbesondere auch Antworten auf die Frage nach invarianten Einheiten der Sprachproduktion.

Die invarianten Einheiten gesprochener Sprache sind aus der Sicht der Sprachproduktion nicht statische Artikulatorpositionierungen sondern charakteristische zielgerichtete **Artikulatorbewegungen**. Betrachten wir die Realisierung des Wortes „Panne“, so geschieht dies durch die zeitliche Koordinierung mehrerer Gesten (siehe Abb. 25). Die Realisierung des [p] geschieht durch eine zeitliche Überlappung einer Geste zur Verschlussbildung durch die Lippen (vbli, siehe auch Tab. 1) mit einer Geste zur Öffnung der Glottis (oegl). Die Realisierung des Vokals geschieht durch die Überlappung einer Geste zur Zungenrückensenkung (sezr) usw. Zusätzlich ist festgelegt, dass im Falle des Fehlens von gestischer Aktivität die Zunge die Stellung des Murrelvokals [ə] realisiert, also das Gaumensegel gehoben ist und die Stimmlippen in Phonationsstellung sind. So wird im Falle von Vokalen auf die Extra-Bezeichnung einer glottalen Geste und im Falle von nicht nasalisierten Lauten auf die Extrabezeichnung einer Geste des Gaumensegels verzichtet. Zur Realisierung des gesamten

Wortes ist nun wiederum die geschickte zeitliche Koordinierung all dieser Gesten erforderlich. Dabei kommt es immer auch zur zeitlichen Überlappung von Gesten unterschiedlicher Laute.

Anmerkung: Der Schwa-Laut zeichnet sich durch das Fehlen von zur bezeichnenden Gesten aus. Hier wird lediglich zur Wahrung der richtigen zeitlichen Struktur der Äußerung auf der phonologischen Ebene ein Platzhalter (*) gesetzt.

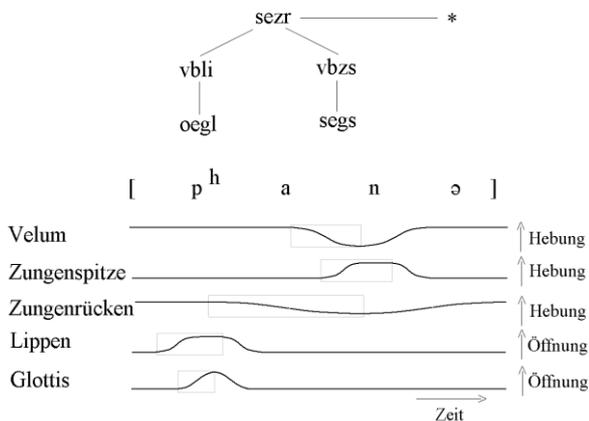


Abbildung 25. Die zur Realisierung des Wortes „Panne“ benötigten Gesten. (a) Phonologische Ebene, (b) Ebene der artikulatorischen Realisierung mit Transkription. Die Rechtecke kennzeichnen die zeitliche Ausdehnung (Dauer) einer Geste. Die Linien kennzeichnen die Bewegung einzelner Artikulatoren.

Abb. 25 zeigt zwei Arten der gestischen Spezifikation eines Wortes. Auf der **phonologischen Ebene** können Gesten als abstrakte Einheiten, die in bestimmter Form aufeinander bezogen sind, dargestellt werden. Gesten werden ähnlich wie in autosegmental-phonologischen Konzepten auf drei Ebenen angeordnet. Die **Assoziationslinien** verbinden Gesten, die miteinander zeitlich koordiniert werden. Auf der **Ebene der artikulatorischen Realisierung** sind die Gesten dann konkrete Ereignisse, eben definierte Artikulatorbewegungen z.B. mit definierter Dauer versehen und auf definierte Artikulatoren wirkend.

Die Geste tritt somit in diesem Konzept sowohl als **abstrakte linguistisch-phonologische Größe** als auch als **konkrete Größe der phonetisch-artikulatorischen Realisierung** auf. Auf der phonologischen Ebene kann ein System diskreter vierbuchstabiger Gestensymbole realisiert werden (Tab. 1). Die symbolische Darstellung grenzt unterschiedliche Gesten in diskreter Weise voneinander ab und führt in ähnlicher Weise wie die bekannten segmentalphonologischen Systeme zu einem bedeutungsunterscheidenden (distinktiven) System von Gesten.

Geste zur....

z.B. zur Realisierung von ...

hhzr	hohen Hebung des Zungenrückens /i:/, /u:/
hezr	Hebung des Zungenrückens /e:/, /o:/
seZR	Senkung des Zungenrückens /a:/
rvzr	Rückverlagerung des Zungenrückens /u:/, /o:/
vvzr	Vorverlagerung des Zungenrückens /i:/, /e:/
ruli	Rundung der Lippen /u:/, /y:/
vbli	Verschlussbildung durch die Lippen /b/, /p/, /m/
vbzs	Verschlussbildung durch die Zungenspitze /d/, /t/, /n/
vbzr	Verschlussbildung durch den Zungenrücken /g/, /k/
eazs	alveolaren Engebildung durch die Zungenspitze /s/, /z/
epzs	postalveolaren Engebildung durch die Zungenspitze /ʃ/, /ʒ/
oegl	Geste zur Öffnung der Glottis stimmlosen Lauten
segs	Geste zur Senkung des Gaumensegels Nasalen

Tabelle 1. Ausschnitt aus dem phonologischen System der Gesten des Deutschen

Gleichzeitig können dieser symbolischen Darstellung einer Geste aber auch konkrete Werte z.B. über Dauer und räumlich-zeitlichen Verlauf der durch die Geste realisierten Artikulatorbewegungen zugeordnet werden. Somit löst dieser Ansatz auch die **Problematik der phonetisch-phonologischen Schnittstelle** in eleganter Weise: Die Geste ist sowohl diskrete linguistische Einheit als auch konkrete quantifizierbare artikulatorische Größe. Aufgrund der Quantifizierbarkeit konnte dieses Konzept auch bereits erfolgreich zur Steuerung von Sprachsynthesatoren eingesetzt werden (z.B. Kröger 1998)

Auch löst dieser Ansatz das oben angedeutete **Problem der fehlenden Invarianz von akustischen Lautmerkmalen**: Perzeptive Untersuchungen ergaben, dass z.B. im Fall stimmhafter Plosive vor allem die Formanttransitionen das akustische Merkmal des Artikulationsortes sind (siehe Kap. 3). Andererseits ist bekannt, daß z.B. die Transition des zweiten Formanten für den Artikulationsort „alveolar“ je nach Vokalkontext sehr unterschiedlich sein kann (Öhmann 1966). Es ist also eine fehlende Invarianz zwischen der phonetisch-phonologischen Kategorie und dem akustischen Merkmal zu beklagen. Erst in der artikulatorischen Domäne tritt das invariante Lautmerkmal zu Tage, nämlich eben der artikulatorische Zielpunkt der Geste zur Verschlussbildung durch die Zungenspitze (vbzs). Diese Resultate führten zu der Annahme einer engen Kopplung zwischen Artikulation und Wahrnehmung: **Artikulatorische Gesten sind die Einheiten auf die unser Wahr-**

nehmungsmechanismus trainiert ist. (Lieberman und Mattingly 1985, Mattingly 1990).

Ein weiteres Problem segmental-phonologischer Konzepte ist die Beschreibung von **reduzierten Formen**: Es ist bekannt, dass bei schneller Sprechweise eine Vielzahl von Assimilations- und Elisionseffekten auftreten können (Kohler 1990). Beispielsweise kann die Äußerung „Er hat mir geholfen“ bei langsamer Sprechweise als [ʔɛvhat^hmɪrəgøhølfən] und bei schneller Sprechweise phonetisch als [ʔɛabmɪrəgøhølfən] realisiert werden. Wir sehen hier beispielsweise einen **Assimilationsprozess** des Artikulationsortes wie auch der Stimmhaftigkeit den Übergang von [t] in „hat“ nach [b] vor nachfolgendem [m]. Darüber hinaus können auch **Elisionen** (Wegfall von Segmenten) erkannt werden, da sich z.B. das Wort „hat“ auf zwei Segmente und das Wort „mir“ auch auf zwei Segmente reduziert.

Somit ist also gerade das **Segment** - als Realisierung der zugrundeliegenden Einheit Phonem - extremen Änderungen aufgrund des (nicht-phonologischen) Faktors Sprechtempo ausgesetzt. Die segmentale Darstellung ist somit nicht invariant gegenüber Änderungen des Sprechtempos. Im gestischen Ansatz kann hingegen gezeigt werden, dass bei Sprechtempoänderungen die zugrundeliegende **gestische Struktur** erhalten bleibt (Browman und Goldstein 1990, Kröger 1998). **Hier werden die Reduktionen allein durch Änderung der zeitlichen Koordinierung von Gesten erreicht.**

Neben der Organisation der zeitlichen Abläufe zur Lautbildung tritt die Produktion von **prosodischen Merkmalen** wie **Akzentuierung (Betonung)** und der **Intonation**. Aus phonetischer Sicht soll unter Akzent (Betonung) die lokale Hervorhebung einer Silbe gegenüber benachbarten Silben innerhalb einer Äußerung verstanden werden. Diese Hervorhebung erfolgt in den meisten Sprachen in erster Linie über markante Änderungen der Grundfrequenz und Erhöhung der Silbendauer und damit Änderung der Lautdauern der in der Silbe auftretenden Laute. Die Intonation einer Äußerung manifestiert sich phonetisch in erster Linie durch den Gesamtverlauf der Grundfrequenz (f0). In vielen Sprachen ist beispielsweise der Anstieg der Grundfrequenz am Äußerungsende ein wichtiges Merkmal der Frageintonation, ein Abfallen der Grundfrequenz hingegen signalisiert das Ende einer Aussage. Hier sind auch bereits die organisatorischen Einheiten der Sprachproduktion oberhalb des Lautes, nämlich **Silbe** und **ganze Äußerung** genannt. Man kann zusätzlich noch die Einheit der **rhythmischen Gruppe** annehmen, die jeweils von einer betonten Silbe bis zur nächsten reicht.

Die Domäne der Intonation ist die Äußerung. Eine Äußerung kann aus einer oder mehreren rhythmischen Gruppen aufgebaut sein. Eine rhythmische Gruppe wiederum besteht aus einer oder mehreren Silben. Die Domäne der Akzentuierung ist die Silbe. Eine Silbe wiederum ist aus mehreren Lauten aufgebaut. Es ist anzumerken, dass sich Intonation und Akzentuierung gegenseitig beeinflussen und bedingen. Darüber hinaus gibt es auch Wechselwirkungen mit der Lautbildung. So konnte

beispielsweise gezeigt werden, dass die Artikulation von Lauten in betonten Silben mit höherer Präzision und mit höherem artikulatorischem Aufwand verbunden ist. Die Artikulatoren nähern sich in betonten Silben den (fiktiven) artikulatorischen Zielpunkten der Laute stärker als in unbetonten Silben. Somit sollte eine Theorie der Lautproduktion die prosodischen Aspekte nicht unbeachtet lassen (Kröger 1998, S. 223ff).

17.5 Literatur zur Phonetik

- Browman, C.P., Goldstein, L. (1989): "Articulatory gestures as phonological units", *Phonology* **6**, 201-251.
- Browman, C.P., Goldstein, L. (1990): "Tiers in articulatory phonology, with some implications for casual speech", in: J. Kingston, M.E. Beckman (Hrsg.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (Cambridge University Press, Cambridge), S. 341-376. Auch in: *Haskins Laboratories Status Report on Speech Research SR-92* (1987), 1-30.
- Browman, C.P., Goldstein, L. (1992): "Articulatory phonology: An overview", *Phonetica* **49**, 155-180.
- Catford, J.C. (1988): *A Practical Introduction to Phonetics* (Oxford University Press, New York).
- Draper, M.H., Ladefoged, P., Whitteridge, D. (1959): "Respiratory muscles in speech", *Journal of Speech and Hearing Research* **2**, 16-27.
- IPA (1993): "Revision of the IPA", *Journal of the International Phonetic Association* **23**, 32-34.
- Kohler, K.J. (1990): "Segmental reduction in connected speech in German: Phonological facts and phonetic explanations", in: W.J. Hardcastle, A. Marchal (Hrsg.), *Speech Production and Speech Modelling* (Kluwer Academic Press, Dordrecht), S. 69-92.
- Kröger, B.J. (1998): *Ein phonetisches Modell der Sprachproduktion* (Niemeyer, Tübingen).
- Lieberman, A.M., Mattingly, I.G. (1985): "The motor theory of speech perception revised", *Cognition* **21**, 1-36.
- Mattingly, I.G. (1990): "The global character of phonetic gestures", *Journal of Phonetics* **18**, 445-452.
- Öhman, S.E.G. (1966): "Coarticulation in VCV utterances", *Journal of the Acoustical Society of America* **39**, 151-168.
- Ohala, J.J. (1990): "Respiratory activity in speech", in: W.J. Hardcastle, A. Marchal (Hrsg.), *Speech Production and Speech Modelling* (Kluwer Academic Press, Dordrecht), S. 23-54.
- Pompino-Marschall, B. (1995): *Einführung in die Phonetik* (DeGruyter, Berlin).

18 Artikulatorische und akustische Lautmerkmale

Grundsätzlich basiert die Produktion eines Lautes bzw. einer Lautfolge auf einer Folge von zeitlich gut koordinierten Bewegungen aller Artikulationsorgane (aller Artikulatoren). Dabei können die Artikulationsorgane zu drei funktionalen Systemen zusammengefasst werden: sublaryngales (oder auch pulmonales) System, laryngales System und supralaryngales System, wobei das supralaryngale System noch in orales und nasales System unterteilt werden kann. Die Funktion des **sublaryngalen Systems** ist die Bereitstellung eines ausreichenden (subglottalen) Drucks bzw. Luftstroms. Dieser Druck bzw. Luftstrom wird als Energielieferant für die primäre Schallquelle (Phonation) bzw. für die sekundären Schallquellen (Friktionsrauschen an Engstellen) benötigt. Die Funktion des **laryngalen Systems** ist die Generierung von Phonationsschall bei locker geschlossener Stimmritze aber auch von Friktionsschall bei etwas stärker geöffneter Stimmritze. Insbesondere ist es dem laryngalen System möglich sehr schnell zwischen diesen Zuständen zu wechseln oder auch gar keinen Schall zu produzieren. Hauptfunktion des **oralen supralaryngalen Systems** ist die Ausbildung von für einzelne Laute charakteristischer Formung des Mund- und Rachenraumes. Dies führt dann auf der akustischen Seite zu lautcharakteristischen Hohlraumresonanzen (Formanten, Kap. 17) und zur geeigneten Überformung des an den Stimmklappen erzeugten Primärschalls. Diese Funktion führt insbesondere zur Differenzierung zwischen einzelnen Vokalen oder zur Differenzierung von Konsonanten der gleichen Artikulationsart. Die Artikulatoren des oralen supralaryngalen Systems sind: Lippen, Zungenrücken und Zungenspitze. Aber auch das Gaumensegel mit Zäpfchen trägt zur Formung von Mund- und Rachenraum bei. Die Funktion des **nasalen supralaryngalen Systems** ist die akustische An- bzw. Abkopplung des Nasenraumes vom Mund- und Rachenraum. Der Artikulator dieses Systems ist das Gaumensegel. Bei Ankoppelung des Nasenraumes werden Nasallaute, bei Abkopplung Orallaute gebildet. Eine strenge Abkopplung des Nasenraumes ist bei Obstruenten (Frikativen und Plosiven) nötig.

Das sublaryngale System wird in *SpeechTrainer* zur Zeit noch nicht modelliert, da die Aerodynamik der Lautproduktion ebenfalls noch nicht explizit modelliert wird. Wir können davon ausgehen, dass während der gesamten Äußerung ein annähernd konstanter Lungen- druck vorliegen sollte. Dieser Druck ist die Basis für den Aufbau eines Luftstroms. Die Stärke des Luftstroms richtet sich insbesondere nach Engebildung bzw. Öffnung von laryngalem bzw. supralaryngalem System. Im folgenden soll aber die zeitliche Koordinierung von laryngalen, oral supralaryngalen und nasal supralaryngalen Artikulatoren für die einzelnen Lautgruppen detailliert besprochen werden.

18.1 Vokale („V“)

Die Artikulation und Akustik der Vokale wurde im Wesentlichen bereits in Kap. 17 erörtert. Während der Produktion eines Vokals ist die Stimmritze locker geschlossen sodass bei ausreichendem subglottalen Druck eine Phonationsschwingung der Stimmklappen entsteht. Diese Schwingung sorgt für den Phonationsschall (Primärschall). Die oralen supralaryngalen Artikulationsorgane werden für jeden Vokal so eingestellt, dass sich akustisch die entsprechende Vokalqualität (ein für den Vokal charakteristisches Formantmuster) ergibt (z.B. hohe Zungenlage bei [i] oder [u], tiefe Zungenlage bei [a], siehe Kap. 7.1 und 7.2).

Im akustischen Signal zeichnet sich ein Vokal vor allem durch regelmäßige Stimmtonimpulse und durch die Ausbildung von Formanten aus. Die regelmäßigen - also periodisch wiederkehrenden - Stimmtonimpulse können in der Zeitdarstellung, die Formanten am besten in der Frequenzdarstellung (Sonagramm) erkannt werden.

Übung: Starten Sie *SpeechTrainer*. Öffnen Sie die Liste „Listen\ Laute_Kontext\ Langvokale.txt“. Klicken Sie im Listenfenster auf den Eintrag „baten“. Öffnen Sie nun das Zeitdarstellungs-Formular durch Anklicken von Menü „Audio > Synchronisieren“. Sie sehen in diesem Formular das Oszillogramm des akustischen Sprachsignals von „baten“. Der Vokal [a:] erstreckt sich hier von Label 4 bis Label 6. Sie erkennen gut die einzelnen Stimmtonimpulse (lange vertikale Impulslinien im Oszillogramm; ca. 35 Impulse).

18.2 Präphonation und Postphonation

Präphonation („PR“) tritt am Beginn und Postphonation („PO“) das Ende einer Äußerung (des Wortes bzw. des kurzen Satzes) auf. Sowohl am Beginn wie auch am Ende einer Äußerung sind die Stimmklappen weit geöffnet um das Atmen zu ermöglichen. Jetzt kann die Lunge mit Luft gefüllt werden. Dies ist die Basis für den Aufbau eines Lungendrucks und ermöglicht dann auch den beim Sprechen auftretenden ausströmenden Luftstrom. Zu Beginn wie auch am Ende eine Äußerung (eines Sprechaktes) sind die oralen supralaryngalen Artikulatoren in Neutralstellung und der Nasenraum ist nicht angekoppelt (das Gaumensegel ist also gehoben). Es wird somit die artikulatorische Stellung eines Schwa-Lautes realisiert.

Akustisch zeichnet sich Prä- wie Postphonation durch das Fehlen von Signal aus. Präphonation ist der Zeitbereich kurz vor der Produktion des ersten Lautes und Postphonation der Zeitbereich kurz nach der Produktion des letzten Lautes einer Äußerung.

Übungsfortsetzung: Betrachten Sie nun im Zeitdarstellungs-Formular den Bereich von Label 1 bis Label 2 (und bis Label 3) für das Wort „baten“. Dies ist der Bereich der Präphonation. Sie erkennen, dass kein akustisches Signal vorhanden ist.

Übungsfortsetzung: Betrachten Sie nun im selben Formular den Bereich um Label 13 für das Wort „baten“. Dies ist der Bereich der Postphonation. Hier ist ebenfalls (evtl. bis auf ein paar Neben- oder Störgeräusche) kein akustisches Signal vorhanden.

Anmerkung: Die **artikulatorische Neutraleinstellung** kurz vor bzw. kurz nach einem Sprechakt ist von der artikulatorischen Ruhestellung zu unterscheiden. Die **Ruhestellung** wird eingenommen, wenn längere Zeit nicht gesprochen wird. Sie zeichnet sich bei den meisten Menschen dadurch aus, dass der Zungenrücken angehoben ist und am Gaumen anliegt. Zusätzlich ist das Gaumensegel abgesenkt um Nasenatmung zu ermöglichen. Diese Ruhestellung der Artikulationsorgane wird in *SpeechTrainer* nicht dargestellt.

18.3 Stimmhafte Plosive („B“)

Plosivlaute zeichnen sich im Wesentlichen durch eine orale Vollverschlussbildung aus. Diese Vollverschlussbildung kann durch die Lippen ([b]), die Zungenspitze ([d]) oder durch den Zungenrücken ([g]) realisiert werden. Bei stimmhaften Plosiven sind zusätzlich die Stimmlippen locker geschlossen, sodass die Stimmlippen schwingen können. Bei länger anhaltender Verschlussbildung wird die Stimmlippenschwingung aber durch den Aufbau eines Druckes im Rachen- und Mundraum (also oberhalb der Stimmritze) zunehmend erschwert. Zusätzlich ist noch das Gaumensegel während der oralen Verschlussbildung stark angehoben, um ein Entweichen von Luft durch den Nasenraum trotz Aufbau des Drucks in Rachen- und Mundraum zu verhindern.

Akustisch zeichnet sich ein stimmhafter Plosiv wegen der oralen Verschlussbildung weitgehend durch das Fehlen von akustischem Signal aus. Gelegentlich kann insbesondere im Anfangsbereich der Verschlussbildung ein kleines periodisches Signal auftreten. Das Pendant im Sonagramm wird als Stimmbalken oder „voice bar“ bezeichnet. Dieses Signal resultiert aus der Stimmlippenschwingung. Das Phonationsgeräusch wird hier durch die zu einem kleinen Teil durch die Backen und den geschlossenen Mund etc. abgestrahlt.

Übungsfortsetzung: Betrachten Sie nun im Zeitdarstellungs-Formular den Bereich zwischen Label 3 und Label 4 des Wortes „baten“. Sie erkennen hier einen stimmhaften Plosiv [b]. Wegen der oralen Verschlussbildung fehlt ein akustisches Signal weitgehend. Im Anfangsbereich des Plosivlautes (kurz hinter Label 3) kann ein kleines periodisches Signal festgestellt werden.

Übungsfortsetzung: Schließen Sie nun das Zeitdarstellungs-Formular (Formular-Menü „Datei > Beenden“) und klicken Sie im Listenfenster auf den Eintrag „aber“. Öffnen Sie nun wieder das Zeitdarstellungs-Formular (Menü „Audio > Synchronisieren“). Sie erkennen einen intervokalischen stimmhaften Plosiv [b] zwischen Label 7 und Label 8.

18.4 Stimmlose Plosive („P“)

Wie bei den stimmhaften Plosiven zeichnen sich auch die stimmlosen Plosiven im Wesentlichen durch eine orale Vollverschlussbildung aus. Diese Vollverschlussbildung kann von den Lippen ([p]), von der Zungenspitze ([t]) oder aber vom Zungenrücken ([k]) realisiert werden. Wie bei den stimmhaften Plosiven ist bei den stimmlosen Plosiven ebenfalls das Gaumensegel stark angehoben, um während der Verschlussbildung ein Entweichen von Luft aus dem Rachen- und Mundraum in den Nasenraum zu verhindern. Anders als bei den stimmhaften Plosiven öffnet sich die Stimmritze aber während der Verschlussbildung (oder ist bei vorangehenden stimmlosen Lauten bereits geöffnet). Die Öffnung der Stimmritze ist zum Zeitpunkt der Verschlusslösung maximal. Dies garantiert zum Zeitpunkt der Verschlusslösung und kurz danach die Bildung von Friktionsrauschen und evtl. Aspirationsrauschen.

Akustisch zeichnet sich der stimmlose Plosiv durch mehrere zeitlich aufeinander folgende Phasen aus. Deutlich erkennbar ist die Phase des Vollverschlusses. Während dieser Zeit fehlt jegliches akustisches Signal. Geht ein stimmhafter Laut voraus, so ebbt die Phonation zu Beginn der Verschlussbildung ab. Es schließt hieran die Phase der Verschlusslösung an. Da sich durch die geöffnete Stimmritze während der Verschlusszeit ein Druck in Mund- und Rachenraum aufgebaut hat, führt die Verschlusslösung anfangs zu einer Ausbildung eines starken Luftstroms und damit zur Ausbildung von Friktionsrauschen kurz vor der oralen Engstelle. Es entsteht ein kurzzeitiges Friktionsrauschen und nachfolgend evtl. noch (glottales) Aspirationsrauschen.

Übungsfortsetzung: Schließen Sie erneut das Zeitdarstellungs-Formular, markieren Sie im Listenfenster das Wort „bieten“ und öffnen Sie das Zeitdarstellungs-Formular erneut. Sie erkennen zwischen Label 7 und Label 9 die zwei Phasen des stimmlosen Plosivlautes [t]. Zwischen Label 7 und Label 8 befindet sich die Phase der Vollverschlussbildung. Es fehlt jegliches akustisches Signal. Die Stimmritze öffnet sich während dieser Zeit (siehe Bewegungsverläufe: Menü „Ansicht > Bewegungsverläufe“) sodass in dieser Phase keine Phonation stattfindet. Hinter Label 8 beginnt die Phase der Verschlusslösung. Sie erkennen gut das Friktionsrauschen. Im Unterschied zu Stimmtonschall (z.B. bei Vokalen) fehlen hier die regelmäßig auftretenden Impulse (es fehlt die Periodizität). Das sich der Verschluss relativ schnell löst, klingt auch das Friktionsrauschen relativ schnell ab.

18.5 Stimmhafte Frikative („W“)

Frikative zeichnen sich durch die Ausbildung einer oralen kritischen Enge aus. Kritisch bedeutet, dass die orale Enge genau so groß ist, dass ein stark gebündelter und schneller Luftstrom entsteht, der nach Verlassen der Enge verwirbelt und Friktionsrauschen (Sekundärschall) bildet. Zusätzlich ist bei Frikativen (wie bei Plosiven) während der oralen Engebildung das Gaumensegel stark

angehoben, um ein Entweichen von Luft durch den Nasenraum zu verhindern. Die Stimmritze ist während der oralen Engebildung locker verschlossen um Phonation zu ermöglichen.

Akustisch zeichnen sich stimmhafte Frikative durch eine Überlagerung von glottalem Phonationsschall und oralem Friktionsrauschen aus. Der Phonationsschall weist periodisch wiederkehrende (regelmäßig auftretende) Impulse auf, die aber im akustischen Signal wegen der oralen Enge nicht so stark ausgebildet sind wie beispielsweise bei Vokalen. Der Friktionsschall wird als nichtperiodisches Rauschsignal erkennbar.

Übungsfortsetzung: Betrachten Sie das Oszillogramm zwischen Label 3 und Label 4 im Wort „sah“. Dies ist der Bereich der Engebildung für den stimmhaften Frikativ [z]. Sie erkennen, dass sich das akustische Signal aus periodisch wiederkehrenden Stimmtonimpulsen und aus nichtperiodischem Rauschen zusammensetzt.

18.6 Stimmlose Frikative („F“)

Stimmlose Frikative werden wie stimmhafte Frikative gebildet: Ausbildung einer oralen kritischen Enge und starke Anhebung des Gaumensegels in diesem Zeitintervall. Jedoch ist die Stimmritze während der Engebildung geöffnet. Dies ermöglicht einen höheren Luftstrom als im Fall der stimmhaften Frikative und damit ein stärkeres Friktionsrauschen vor der oralen Engstelle. Phonation tritt wegen der Öffnung der Stimmritze nicht auf.

Akustisch zeichnen sich stimmlose Frikative durch das Auftreten von Friktionsrauschen über den gesamten Zeitbereich der Engebildung aus.

Übungsfortsetzung: Betrachten Sie Label 8 und Label 9 im Wort „Höfe“. Sie sehen in diesem Bereich das Oszillogramm des Friktionsrauschens des stimmlosen Frikativs [f].

Übungsfortsetzung: Betrachten Sie nun Label 3 und Label 4 im Wort „sputen“. Sie erkennen, dass das Friktionsrauschen des Lautes [S] wesentlich kräftiger ist, als das des Lautes [f]. Suchen und betrachten Sie nun weitere stimmlose Frikative, z.B. das häufig auftretende [s].

18.7 Nasale („N“)

Nasale zeichnen sich artikulatorisch wie stimmhafte Plosive durch die Ausbildung eines oralen Vollverschlusses und durch die lockere Verschließung der Stimmritze aus. Im Unterschied zu anderen Lauten ist bei den Nasalen allerdings das Gaumensegel stark abgesenkt. Der Luftstrom wird somit durch Rachen- und Nasenraum geführt. Wegen des lockeren Verschlusses der Stimmlippen und des Fehlens einer Hemmung des Luftstroms oberhalb der Glottis tritt bei den Nasalen normale Phonation auf.

Akustisch ist ein Nasal ähnlich wie ein Vokal durch die periodischen Impulse des Phonationsschalls gekennzeichnet. Tritt ein Nasallaut in Anschluss an einen Vokal auf, so kann im Oszillogramm der Übergang von Vokal zu Nasal meist durch die abrupte Änderung der Amplitude der Phonationsimpulse erkannt werden.

Übungsfortsetzung: Betrachten Sie Label 8, 9 und 10 in „Einer“. Dieses Zeitintervall charakterisiert den Nasal [n]. Sie erkennen die Periodizität des Nasalsignals. Es unterscheidet sich nur geringfügig von den benachbarten Vokalsignalen. Die abrupte Verringerung der Amplitude der Phonationsimpulse kann beim Übergang von [i] nach [n] erkannt werden.

18.8 Laterale („L“)

Laterale zeichnen sich artikulatorisch durch die Ausbildung einer lateralen Enge bei lockerer Verschließung der Stimmritze aus. Wie bei den Nasalen tritt auch hier normale Phonation auf. Laterale Enge bedeutet, dass in der Medianebene (wie bei Plosiven und Nasalen) ein Verschluss gebildet wird, an beiden Seiten die Zungenränder aber abgesenkt sind, sodass an beiden Seiten Luft strömen kann.

Im Oszillogramm weist ein Lateral ähnliche Eigenschaften wie ein Nasal auf. Der Übergang von Vokal nach Lateral ist aber oftmals schwerer zu erkennen als beispielsweise der Übergang von Vokal nach Nasal. Akustisch unterscheiden sich Laterale und Nasal auch durch ihre spektrale Struktur (im Sonagramm erkennbar).

Übungsfortsetzung: Betrachten Sie das Oszillogramm des Wortes „Öl“ im Bereich von Label 7 und Label 8.

Anmerkung: Es ist geplant, *SpeechTrainer* auch um spektrale Darstellungen des akustischen Sprachsignals (z.B. Sonagramm-Darstellung) zu erweitern. Dies wird die Erkennung von Lautübergängen speziell im Fall Vokal, Nasal, Lateral wesentlich vereinfachen. Bisher muss zur Erkennung von Lautgrenzen in schwierigen Fällen auf das Anhören von Signalauschnitten zurückgegriffen werden.

18.9 Vibranten („R“)

Vibranten zeichnen sich akustisch durch die Ausbildung einer oralen Enge mit vibrierendem Artikulator (Zungenspitze oder Zäpfchen) innerhalb dieser Enge aus. Gleichzeitig ist die Stimmritze locker geschlossen und es tritt Phonation auf. Das Vibrieren des oralen Artikulators wird – ähnlich wie das Vibrieren der Stimmlippen – durch den Luftstrom und nicht durch muskuläre Impulse initiiert. Dabei ist die Stimmlippen-schwingung aber ca. 3 – 15 Mal schneller als die Vibration von Zungenspitze oder Zäpfchen bei Vibranten.

Akustisch weisen Vibranten wegen Ihrer Stimmhaftigkeit wiederum periodisch (regelmäßig) wiederkehrende Glottisimpulse auf. Die Amplitude der glottalen Impulse wird aber mit der (langsameren) Rate der oralen Vibration moduliert.

Übungsfortsetzung: Betrachten Sie dazu den Bereich zwischen Label 5 und Label 6 in „froh“. Sie erkennen hier sowohl die glottalen Impulse aufgrund der Phonation (ca. 15 bis 20 Impulse für das [R]). Darüber hinaus erkennen Sie auch die Modulation der Amplitude der glottalen Impulse aufgrund der oralen Vibration (2 Maxima und 2 Minima für dieses [R]).

18.10 Glottallaute („G“)

Als Glottallaute werden der Glottisverschlusslaut [ʔ], die Glottalisierung [~] (in älteren Dateien auch als [β] notiert), der stimmhafte glottale Frikativ [H] und der stimmlose glottale Frikativ [h] bezeichnet.

Glottisverschlusslaut und Glottalisierung zeichnen sich artikulatorisch durch Ausbildung eines starken bzw. etwas weniger starken glottalen Verschlusses aus. Bei dem Glottisverschlusslaut kann keine glottale Schwingung auftreten. Bei der Glottalisierung sind glottale Impulse in unregelmäßigen Zeitabständen möglich. Diese Zeitabstände sind meist größer als im Falle normaler Phonation. Stimmhafter und stimmloser glottaler Frikativ zeichnen sich durch glottale Engebildung aus, wobei die Engebildung beim stimmhaften glottalen Frikativ stärker ist als im Fall des stimmlosen glottalen Frikativs.

Akustisch erscheinen die vier glottalen Frikative recht unterschiedlich. Im Fall des Glottisverschlusslautes tritt keinerlei akustisches Signal auf. Die Stimm Lippen sind fest verschlossen. Es tritt keine glottale Schwingung auf und es tritt auch kein Luftstrom durch die Stimmritze.

Bei der Glottalisierung hingegen treten glottale Impulse auf. Im Unterschied zu Vokalen (mit normaler Phonation) sind die Zeitabstände zwischen den glottalen Impulsen bei der Glottalisierung aber recht groß und die glottalen Impulse treten unregelmäßig auf.

Stimmhafte glottale Frikative erscheinen akustisch ähnlich wie andere stimmhafte Frikative. Es treten glottale Impulse auf, allerdings mit geringer Amplitude. Zusätzlich ist das akustische Signal mit nichtperiodischem Rauschen überlagert.

Stimmlose glottale Frikative erscheinen akustisch ähnlich wie andere stimmlose Frikative: Es treten keine glottalen Impulse auf. Wir finden aber Friktionsrauschen, dass im Fall der Entstehung an der Glottis auch als Aspiration bezeichnet wird.

Übungsfortsetzung: Betrachten Sie Label 3 und Label 4 in „Höfe“. Sie erkennen das Friktionsrauschen (die Aspiration) eines stimmlosen glottalen Frikativs [h].

Übungsfortsetzung: Betrachten Sie nun Label 3 und Label 4 in „Efeu“. Sie erkennen das Fehlen von Signal im Fall des Glottisverschlusslautes [ʔ].

Übungsfortsetzung: Betrachten Sie Label 7 und Label 8 in „behagen“ (Liste „Listen\ Laute_Kontext\ Orte-Fri.txt“). Sie erkennen abgeschwächtes Phonationssignal und Friktionsrauschen des stimmhaften glottalen Frikativs [H].

Übungsfortsetzung: Betrachten Sie Label 10 und Label 11 in „Essecke“ (Liste „Listen\ Listen_a-m\ Liste.txt“). Sie erkennen eine (leider sehr kurze) Glottalisierung [~].

Übungsfortsetzung: Betrachten Sie Label 22 und Label 23 in „Ich_brauche_etwas“ (Liste „Listen\ Alltagswörter\ Ich\ Ich_brauche_etwas.txt“). Sie erkennen eine Glottalisierung (hier noch in veralteter Schreibweise als [β] notiert).

Glottal: Enge zwischen den Stimmlippen, z.B. [ʔ] in „_Apfel“, [h] in „_Hallo“, [ɦ] in „_aha“.

Artikulationsart:

Die Artikulationsart (auch Artikulationsmodus) beschreibt die *Art* der Enge (der Konstriktion), die gebildet wird. Hieraus ergeben sich die folgenden Lautklassen:

Verschlusslaut, Plosiv (stop): vollständiger Verschluss des Ansatzrohres; Unterbrechung des Luftstroms; danach: Verschlusslösung (evtl. mit Friktions- und Aspirationsgeräusch).

Nasallaut, Nasal (nasal): vollständiger Verschluss des Mundraumes; Luft kann aber durch Nasenraum entweichen; Also: keine Hemmung des Luftstroms.

Reibelaut, Frikativ (fricative): schmale Passage für den Luftstrom (aufgrund einer Enge), Entstehung von Rauschen am Ausgang der Enge aufgrund von Luftreibung (Friktion) und daraus resultierender Bildung von Luftturbulenz (Friktionsrauschen, Geräuschbildung).

lateral Engelaut, lateraler Frikativ (lateral fricative): schmale Passage für den Luftstrom aufgrund einer lateralen Enge (mit Verschluss in der Medianebene), Entstehung von Rauschen am Ausgang der Enge.

Vibrant, mehrschlägiger Verschlusslaut, (trill): Luftstrom generiert mehrere schnelle Verschließ- und Öffnungsbewegungen eines Artikulators in der Enge. Der entsprechende Artikulator (Zungenspitze oder Zäpfchen) vibriert dann durch den Luftstrom.

einschlägiger Vibrant, (tap, flap): Luftstrom generiert genau *eine* schnelle Verschließbewegung des Artikulators in der Enge.

Approximant, (approximant): Ausbildung einer (kurzzeitigen) Enge, der Luftstrom an der Enge wird gerade nicht so stark behindert, dass Friktionsrauschen auftritt.

lateral Approximant, (lateral approximant): Ausbildung einer (kurzzeitigen) lateralen Enge (bei Verschluss in der Medianebene), der Luftstrom an der Enge wird gerade nicht so stark behindert, dass Friktionsrauschen auftritt.

Stimmbeteiligung:

Dieser Parameter differenziert die Laute je nachdem, ob die Stimmlippen während der Lautproduktion schwingen (Phonation) oder nicht; Resultat: ob an den Stimmlippen Phonationsschall gebildet wird oder nicht.

stimmhaft: Stimmlippenschwingung vorhanden; Luft wandert stoßweise in das Ansatzrohr: Phonation, Phonationsschallbildung.

stimmlos: keine Stimmlippenschwingung; keine Phonation; 2 unterschiedliche artikulatorische Möglichkeiten bestehen:

Fall i) (weit) geöffnete Stimmlippen: Luft gelang ungehindert durch das Ansatzrohr, Geräuschbildung direkt oberhalb der Stimmritze möglich: Aspiration, z.B. bei [h] in „_hallo“

Fall ii) fest verschlossene Stimmlippen: glottaler Plosiv: Glottisverschlusslaut [ʔ], z.B. in „_aber“

19.3 Die Vokale

Die IPA-Tabelle stellt die artikulatorischen Ordnungsprinzipien der Vokale aller Sprachen dar: Zungenposition (offen - geschlossen, vorne - hinten) und Lippenrundung (ungerundet - gerundet). Das sprachübergreifende Vokalsystem wird auch als System der Kardinalvokale bezeichnet.

Anmerkung: Die Kategorie „offen - geschlossen“ wird in *SpeechTrainer* durch die Begriffe „hoch - tief“ beschrieben, da „geschlossen - offen“ im Vokalsystem des Standarddeutschen für den Kontrast von Lang- und Kurzvokal reserviert ist.

offen - geschlossen (hoch - tief): gehobene(r) bzw. gesenkte(r) Zungenrücken, Kiefer und Unterlippe. Die Abfolge von hohen zu tiefen Vokalen geschieht in vier Hauptstufen: [i, e, ε, a], z.B. in „_Miete, Leben, Ähre, aber“. Diese vier Hauptstufen der Höhe werden als **hoch, halbhoch, halbtief** und **tief** bezeichnet.

vorne - hinten: Wir unterscheiden: **Vorderzungenvokale** (palatale Enge), z.B. [i, e, y] in „_Igel, edel, übel“ und **Hinterzungenvokale** (velare Enge), z.B. [u, o] in „_Mut, Not“. Dazwischen gibt es die **Mittelzungenvokale**, z.B. [ə, ɐ] in „_meine, Vater“.

gerundet - ungerundet (gespreizt): Wir unterscheiden gerundete Vokale, z.B. [y], und ungerundete Vokale, z.B. in [i, a]. Bei hohen ungerundete Vokalen tritt auch Lippen spreizung auf.

19.4 Artikulation und Transkription der Konsonanten des Standarddeutschen

Plosive:

Im Standarddeutschen treten stimmlose und stimmhafte Plosive mit dem Artikulationsort bilabial, alveolar, velar auf; stimmhaft: [b, d, g] z.B. in „_Rebe, Rede, rege“; stimmlos: [p^h, t^h, k^h] z.B. in „_Mappe, Matte, Macke“.

Die stimmlosen Plosive werden meist aspiriert realisiert. Dies wird durch ein diakritisches Zeichen (hochgestelltes „h“) notiert: [p^h].

Die stimmhaften Plosive sind initial oft (teil-)entstimmte z.B. [b̥] in „_Bad“ (siehe diakritische Zeichen).

Im Deutschen unterscheiden wir im Fall der Plosive weniger nach Stimmbeteiligung sondern eher nach artikulatorischem Aufwand: Stimmlose Plosive sind „Fortis-Laute“. Sie werden mit höherem artikulatorischem Aufwand als stimmhafte Plosive („Lenis“-Laute). Dem-

entsprechend ist das Verschlussintervall bei stimmlosen Plosiven meist länger als bei stimmhaften.

Der glottale Plosiv [ʔ] (Glottisverschlusslaut) wird im Standarddeutschen aus phonologischen Gründen oft auch isoliert behandelt (Frage nach Phonemstatus dieses Lautes je nach Schule unterschiedlich: [ʔ] kann bei vorgegebenen Morphemgrenzen aus Regeln abgeleitet werden). Wichtig ist, dass der glottale Plosiv im Deutschen oft vor Vokalen am Wort- oder Silbenanfang auftritt; z.B. in „aber“, „ich“, „alle“, Silbenanfang: „ver-eisen“, nicht aber „verreisen“.

Der Plosivlaut ist artikulatorisch-akustisch recht komplex. Bei stimmlosen Plosiven in Vokalumgebung sind vier artikulatorisch-akustischen Phasen zu unterscheiden:

Phase 1: Bewegung des Artikulators in den Verschluss hinein; Formanttransitionen des Plosives noch während der vokalischen Phonation.

Phase 2: Halten des Verschlusses (kein Signal); dabei: langsame Öffnung der glottalen Enge (von stimmhaft nach stimmlos) und gleichzeitig: Druckaufbau im Mundraum; Dabei auch fester Verschluss des Nasenraumes: Anhebung des Gaumensegels. Grund: Es darf keine Luft durch den Nasenraum entweichen (sonst: Ausbildung des homorganen Nasals)).

Phase 3: Lösung des Verschlusses; Lösegeräusch: kurzzeitiges Friktionsgeräusch, genannt: „noise burst“.

Anmerkung: Der Verschluss wird aber durch aktive Artikulatorbewegung gelöst, nicht durch aerodynamische „Sprengung“; evtl. leichte „(Ex-)Plosion“ aufgrund des vorhergehenden Druckaufbaus im Mundraum.

Phase 4: Weitere (Öffnungs-)Bewegung des Artikulators zur vokalischen Artikulatorposition (Öffnung); Formanttransitionen aufgrund der Lösung der Enge in den Vokal hinein; Während dieser Zeit tritt auch Aspiration auf: die Glottis ist noch weit geöffnet.

Anschließend: Einsetzen der Phonation definiert den akustischen Anfang des nachfolgenden Vokals; Voice Onset Time (VOT): Zeitbereich vom Beginn des Lösegeräusches (Beginn der Verschlusslösung) bis zum Einsetzen der Phonation beim nachfolgenden stimmhaften Laut.

Nasale:

Im Deutschen treten bilabiale, alveolare, und velare Nasale auf: [m, n, ŋ] z.B. in „Maat“, „Naht“, Menge“)

Die Verschlussbildung ist artikulatorisch ähnlich wie im Fall der Plosive. Aber durch die gleichzeitige Senkung des Gaumensegels tritt eine Ankoppelung des Nasenraumes an den Rachen- und Mundraum auf. Somit kann die Luft – im Unterschied zu Plosiven – bei Nasalen (ähnlich wie bei Vokalen) ungehindert strömen. Im Unterschied zu den Vokalen nun aber nicht durch Rachen- und Mundraum sondern durch Rachen- und Nasenraum.

Als Besonderheit kann es im Standarddeutschen bei Wegfall des Kurzvokals zwischen Plosiv und Nasal (z.B. in „hatten“) zur „nasalen Plosion“ kommen. Transkription: mit Kurzvokal: [hat^hən]; bei Wegfall des Vokals [ə]: [hatⁿŋ] (mit silbischem „n“).

Frikative:

Die Luft wird in der Engstelle gebündelt und beschleunigt. Die hinter der Enge ausströmende Luft verwirbelt und führt zur Geräuschbildung. Die Reihe der Frikative ist in der IPA-Tabelle die vollständigste Reihe. Stimmhafte und stimmlose Frikative können an allen Artikulationsorten gebildet werden. Auch im Lautsystem des Deutschen bilden die Frikative die umfangreichste Lautreihe.

Artikulationsorte: labiodental, alveolar, postalveolar, palatal, velar, (uvular als r-Realisation), glottal. Laute: [f v, s z, ʃ ʒ, ç ʝ, x, ʁ, h, ɦ] z.B. in „falten, walten, rei-ßen, reißen, Masche, Gage, brechen, jagen, tauchen, Ehre, heben, beheben“

Besonderheiten:

Die postalveolaren Frikative [ʃ, ʒ] treten im Deutschen immer mit zusätzlicher Lippenrundung auf [ʃ^w] (labialisiert). (Diese wird allerdings in der Transkription oft gar nicht extra vermerkt.)

Der stimmhafte postalveolare Frikativ [ʒ] tritt nur in Lehnwörtern auf (z.B. „Garage, Gage, Blamage, ...).

Der stimmlose palatale und velare Frikativ [ç, x] („ich-Laut“ und „ach-Laut“) sind allophonische Varianten des Phonems (/x/). Die Realisierungsnorm lautet wie folgt: Der „ich-Laut“ tritt nach nicht-tiefen Vorderzungenvokalen (z.B. [i, e, ε, y, ø]), der „ach-Laut“ nach allen anderen Vokalen auf.

Der stimmhafte palatale Frikativ [j] (z.B. in „jeder“) alterniert situations- und sprecherabhängig mit dem palatalen Approximanten [j]. dies gilt auch für den labiodentalen stimmhaften Frikativ [v]. Dieser alterniert mit dem labiodentalen Approximanten [ʋ] (z.B. in „Vase“).

Der stimmhafte uvulare Frikative [ʁ] kann als /r/-Realisierung auftreten, z.B. in „sie wren“. Dies ist ebenfalls sprecher- bzw. situationsabhängig. Das /r/-Phonem kann aber auch als Vibrant realisiert werden (Bühnensprache).

Das Phonem /h/ wird zumeist stimmlos realisiert [h] (z.B. am Wortanfang). In intervokalischer Stellung kann es aber auch stimmhaft ([ɦ]) realisiert werden (z.B. [h] in „heiß“ aber [ɦ] in „verheißen“).

Vibranten:

Die Artikulation der Vibranten ist relativ aufwendig. Es liegen ein gespannter Artikulator zur Ausbildung der Enge und ein ungespannter, schwingungsfähigen

Artikulator in der Enge vor. Der ungespannte Artikulator wird passiv vom Luftstrom zur Vibration angeregt.

Im Standarddeutschen tritt das „Zungenspitzen-r“ ([r]: Zungenspitze in Medianebene entspannt, dagegen gespannter Zungenrand) und das „Zäpfchen-r“ ([ʀ]: Zäpfchen entspannt, dagegen gespanntes Gaumensegel und gespannten Zungenrücken) auf.

Das /r/-Phonem wird im Deutschen aber oft vokalisch oder als Frikativ realisiert. Die vokalische Realisierung des /r/ tritt in postvokalischer Position als [ɐ] z.B. in „Vater, verlieren“. Die frikativische Realisierung kann als sprecherspezifische Realisierung des Vibranten (artikulatorische Erleichterung) in allen anderen Positionen auftreten ([ʁ] z.B. in „rollen“). Nach stimmlosem Konsonant kann sogar ein stimmloser Frikativ als /r/-Realisierung auftreten (z.B. [χ] in „Frucht“).

Laterale:

Beim Lateral (lateraler Approximant) wird eine laterale Enge bei gleichzeitigem Verschluss der Medianebene gebildet. Der Luftstrom wird durch die Enge nicht so stark behindert, dass Turbulenzen auftreten; Er kann schließlich lateral fließen.

Im Deutschen tritt nur der alveolare Laterale [l] (z.B. in „Laternen“) auf.

Approximanten:

Approximanten bilden eine starke Enge oder sogar einen Verschluss aus. Die Enge bzw. der Verschluss werden aber derartig kurzzeitig gebildet, dass hierdurch der Luftstrom nicht sehr stark behindert wird. Die Luft kann also wegen der Kurze der Konstriktionsbildung annähernd ungehindert strömen. Es passiert insbesondere kein Druckaufbau im Mundraum.

Im Standarddeutschen können Approximanten als sprecherspezifische Variante zu stimmhaften Frikativen auftreten. Es existiert der labiodentale Approximant [ʋ] (z.B. in „sie waren“) und der palatale Approximant [j] (z.B. in „diejenigen“). Diese Approximanten treten dabei vor allem bei höherem Sprechtempo auf. Sie können auch als verschliffene Realisierung des entsprechenden Phonems /v/ bzw. /j/ angesehen werden.

19.5 Die Artikulation und Transkription der Vokale und Diphthonge des Standarddeutschen

Vokale zeichnen sich durch das Fehlen einer ausgeprägten Engstelle im Ansatzrohr aus. Der Luftstrom wird nicht behindert. Die gesamte Formbildung des Ansatzrohres determiniert die Vokalqualität (d.h. die Lage der Formantfrequenzen).

Das Vokalsystem des Standarddeutschen kennt Langvokale und Kurzvokale (auch als gespannte bzw. ungespannte oder geschlossene bzw. offene Vokale bezeichnet). Beispielsweise: gespanntes [ɛ:] („er nähte“) vs. ungespanntes [ɛ] „der Nette“. Darüber hinaus treten noch reduzierte Vokale (z.B. [ɔ] in „hätte“) und Diphthonge (d.h. Vokale mit hörbarer Artikulationsbewegung; z.B. [au] in „Haus“) auf.

Die Lang- und Kurzvokale:

Im Standarddeutschen existieren acht Langvokale: [i: e: ε: a: o: u: y: ø:] (z.B. in „Liebe, Leben, Lähmung, Laden, Lob, Bube, Mühe, Morgenröte“) und sieben Kurzvokale: [ɪ ɛ a ɔ ʊ ʏ œ] (z.B. in „Lippen, lecker, Land, locker, Lust, Kümmel, Köpfe“).

Die Qualität der hohen Kurzvokale ist „offener“ und „zentralisierter“ als die der entsprechenden Langvokale: (siehe [i:] vs. [ɪ] in „Miete“ vs. „Mitte“, [y:] vs. [ʏ] in „Fühler“ vs. „Füller“). Die zu den halbhohen Langvokalen gehörenden Kurzvokale sind sogar wesentlich offener. Sie gehören bereits zur Kategorie halbtief: (siehe [e:] vs. [ɛ] in „Beet“ vs. „Bett“, [o:] vs. [ɔ] in „Robe“ vs. „Röbbe“, [ø:] vs. [œ] in „Röslein“ vs. „Rösslein“).

Die (hohen) Kurzvokale werden im Deutschen deshalb auch als offene Vokale bezeichnet, die (hohen) Langvokale als geschlossene Vokale. Da diese Bezeichnung aber mit der IPA-Konvention („closed“ = „hoch“ und „open“ = „tief“) kollidiert, wollen wir uns der Konvention anschließen, die Langvokale auch als **gespannte Vokale** und die Kurzvokale auch als **ungespannte Vokale** zu bezeichnen. Dahinter steht die Idee, dass gespannte Laute mit höherem artikulatorischem Aufwand gebildet werden als ungespannte Laute. Somit kann das Begriffspaar „gespannt – ungespannt“ dann auch erklären, warum ungespannte Vokale mit zentralerer Vokalqualität produziert werden.

Die Qualität der tiefen Langvokale bleibt allerdings beim Übergang zum entsprechenden Kurzvokal annähernd erhalten, z.B. [ɛ:] vs. [ɛ] in „schälen“ vs. „schellen“, [a:] vs. [a] in „Schal“ vs. „Schall“. Bei einigen Sprechern wird allerdings das Phonem /a:/ nicht vorne (als [a:]) sondern etwas weiter hinten als [ɑ:] realisiert.

Die vollständige Bezeichnung der Lang- und Kurzvokale ist somit wie folgt:

[i:] gespannter hoher ungerundeter Vorderzungenvokal
 [ɪ] ungesp. hoher ungerundeter Vorderzungenvokal
 [y:] gespannter hoher gerundeter Vorderzungenvokal
 [ʏ] ungesp. hoher gerundeter Vorderzungenvokal
 [u:] gespannter hoher gerundeter Hinterzungenvokal
 [ʊ] ungesp. hoher gerundeter Hinterzungenvokal
 [e:] gespannter halbhoher unger. Vorderzungenvok.
 [ɛ:] gespannter halbhoher gerundeter Vorderzungenvok.
 [o:] gespannter halbhoher gerundeter Hinterzungenvok.
 [ɔ] ungesp. halbtiefer gerundeter Hinterzungenvokal
 [œ] ungesp. halbtiefer gerundeter Vorderzungenvokal
 [ɐ:] gespannter halbtiefer unger. Vorderzungenvokal

[ɛ] ungesp. halbtiefer ungerundeter Vorderzungenvokal
 [a:] gespannter tiefer ungerundeter Vorderzungenvokal
 [a] ungespannter tiefer unger. Vorderzungenvokal

Des Weiteren gilt: Kurzvokale können nie am Silbenende auftreten. Ihnen folgt in der Silbe immer ein Konsonant. Dies führte zu der Idee, Kurzvokale auch als durch einen nachfolgenden Konsonanten „abgeschnittene“ Vokale aufzufassen.

Darüber hinaus gibt es im Standarddeutschen auch ein paar „kurze gespannte Vokale“, allerdings vorwiegend in Lehnwörtern, z.B. [o] in „Musikologie“ oder [e] in „Systemematik“.

Die reduzierten Vokale:

Im Standarddeutschen treten zwei reduzierte Vokale auf:

[ə]: Zentralvokal, Schwa-Laut, Murmelvokal

[ɐ]: r-gefärbter Schwa-Laut, vokalisierter r-Laut

Der Schwa-Laut nimmt artikulatorisch eine „Neutrallage“ oder „Zentrallage“ im Vokalraum ein. Deshalb wird er auch als Neutral- oder Zentralvokal bezeichnet. Dieser Laut kann oft auch ganz entfallen, z.B. „Mantel“ [mantəl] geht über in: [mantl] (genauer: [mant^l]): silbisches „l“ und laterale Lösung des Plosivs).

Daneben tritt noch der vokalisierte r-Laut [ɐ] (z.B. in „bitter“ vs. „bitte“, „Vater“ vs. „Pate“) auf. Dieser Laut kann nicht entfallen, da er meist stärker bedeutungstragend als der Schwa-Laut ist.

Reduzierte Vokale sind die einzigen Mittelzungenvokale des Standarddeutschen. Sie treten nur in unbetonten Silben auf.

Die Diphthonge:

Diphthonge können als „Vokale mit hörbarer Artikulationsbewegung“ bezeichnet werden. Im Standarddeutschen ist dies immer eine Schließbewegung (Hebung des Kiefers und des Zungenrückens).

Im Standarddeutschen treten drei 3 Diphthonge auf: [aɔ] [aɛ] [ɔø] „Haus“, „heiß“, „Heuß“.

Oft treten Sie auch in etwas geschlossenerer Realisierung (weniger reduzierter Realisierung) als [aʊ] [aɪ] [ɔʏ] auf.

Anmerkung: Beim Studium der Artikulationsbewegungen in *SpeechTrainer* können wir erkennen, dass die meist recht starken Artikulationsbewegungen zur Realisierung eines Vokals immer hinter einer konsonantischen Konstriktion versteckt sind. Dies führt zu einer als konstant wahrgenommener Vokalqualität. (z.B. „Mallus“: [malus]: Der Übergang von [a] nach [ʊ] geschieht während der Produktion des [l].) Fehlt nun der Konsonant zwischen zwei Vokalen, gelangen wir zu einem Diphthong [aʊ] (z.B. in „Maus“ [maʊ]).

Diphthonge entsprechen in ihrer Länge den Langvokalen. Sie können deshalb phonotaktisch wie Langvokale angesehen werden.

19.6 Tastatursymbole

Die hier und in Kap. 17 (phonetische Grundlagen) benutzen Transkriptionssymbole entsprechen der IPA-Norm. In *SpeechTrainer* werden einige dieser Symbole allerdings durch möglichst ähnliche Symbole des Standardzeichensatzes (Tastatursymbole) ersetzt. Eine Tabelle der Ersetzungen wird hier gegeben.

Es werden hier nicht alle Phonetikzeichen repräsentiert. Feinheiten der Artikulation müssen in *SpeechTrainer* dann auf der Ebene der artikulatorischen Targets eingestellt werden.

Tastatur	IPA-Zeichen
Vokale:	
i:	i:
I	ɪ
e:	e:
E:	ɛ:
E	ɛ
a	ɑ
a:	ɑ:
A, 6	æ
u:	u:
U	ʊ
o:	o:
O	ɔ
y:	y:
Y	ʏ
ö:, 2:	ø:
Ö, 9	œ
@	ə
Konsonanten	
b	b
d	d
g	g
~ (alt: ß)	Glottalisierung
p	p
t	t
k	k
?	ʔ
m	m
n	n
N	ŋ
f	f
v	v
s	s
z	z
S	ʃ
Z	ʒ
c	ç
x	x
X	ɣ
h	h
H	ɦ
r	r
R	ʀ
l	l
j	j, ɰ

Tabelle 3. Tastatursymbole und Phonetikzeichen (nach IPA).