

Artikulatorische und akustische Methoden in der Diagnostik und Therapie von Sprechstörungen

Bernd J. Kröger, Alexandra Diem, Miriam Siegert

Klinik für Phoniatrie, Pädaudiologie und Kommunikationsstörungen; Universitätsklinikum der RWTH Aachen
Email: bkroeger@ukaachen.de

Einleitung

Im Bereich der Diagnostik und Therapie von Sprechstörungen (motorische Störungen bei der Sprachentwicklung, Störungen aufgrund von Schädigungen der Sprechorgane, zentrale Sprechstörungen wie Dysarthrien und Sprechapraxien) werden zunehmend akustische und artikulatorische Verfahren eingesetzt. Die Diagnostik von Sprechstörungen ist zumeist perceptiv (Ziegler und Vogel 1998, Ziegler 2002) und damit entweder aufwendig (mehrere Beurteiler) oder nicht einfach objektivierbar.

Es soll in diesem Beitrag am Beispiel der Messung normaler und gestörter Sprache gezeigt werden, inwieweit akustische Methoden im Bereich der Diagnostik von Sprechstörungen eingesetzt werden können. Im zweiten Teil des Beitrages wird der Einsatz eines artikulatorischen Modells zur Visualisierung von Sprechbewegungen in der Therapie von Sprechstörungen vorgestellt.

Akustische Methoden

Zur phonetischen Charakterisierung gestörter Sprache (hier am Beispiel des Störungsbildes der Dysarthrie) definieren wir segmental-akustische Parameter, die jeweils an unterschiedlichen Lauttypen in einem jeweils definierten lautlichen Kontext gemessen werden (Tabelle 1, Spalte 1 und 3). Diese Parameter werden auf der Basis akustischer Signalanalysewerte (Intensität, Formantfrequenzen, Zeitdauer VOT, Jitter, Shimmer, Harmonics-to-Noise-Ratio, siehe Tabelle 1, Spalte 2) an definierten Lautzeitpunkten gewonnen.

Alle angegebenen segmental-akustischen Parameter (Tabelle 1) sind Relativmaße. Es wird entweder die Differenz der Werte eines akustischen Signalparameters zwischen untersuchtem Laut und Nachbarlaut (syntagmatischer Kontrast) oder die Differenz der Werte eines akustischen Signalparameters zwischen unterschiedlichen Lauttypen (z.B. stimmhafter vs. stimmloser Plosiv, hoher vs. tiefer Vokal, vorderer vs. hinterer Vokal) gemessen. Auch die Größe des Vokalraumes ist ein Relativmaß, da diese Größe innerhalb des logarithmierten F1-F2-Feldes gemessen wird. Damit ist eine Normierung in Hinblick auf die unterschiedliche Lage des Vokalraumes bei Männern, Frauen und Kindern aufgrund unterschiedlicher Länge des Ansatzrohres der Sprecher gegeben.

Alle akustischen Analysen wurden mittels der Software „Praat“ (Boersma und Wenink 2003) realisiert. Mittels der dort verfügbaren Skriptsteuerung können die Parameter nach manueller Segmentierung der Testsätze mittels automatischer Prozeduren gewonnen werden.

Die segmental-akustischen Parameter werden anhand einem Korpus von 8 Testsätzen (mit dreimaliger Wiederholung) gewonnen. Sie wurden bisher anhand eines pathologischen Sprechers (schwere Dysarthrie, M. Parkinson) und eines Normalsprechers evaluiert. Alle

Parameter zeigen signifikante Unterschiede zwischen beiden Sprechern auf.

Segmental-akustische Parameter	Akustische Analyse	Lauttyp und Lautkontext
Stärke des Vokal-Verschlusslaut-Kontrastes	Intensität > 0.5 kHz	stimmhafter oder stimmloser Plosiv in VCV-Kontext
Stärke des Plosionsgeräusches	Intensität	stimmlose Plosive in VCV-Kontext
Stärke der Frikativ-Geräuschbildung	Intensität > 5kHz	stimmlose Frikative in VCV-Kontext
Stärke des konsonantischen Stimmkontrastes	Zeitdauer VOT	Plosive: stimmhaft vs. stimmlos in VCV-Kontext
Stärke der konsonantischen Formanttransitionen	Formanttransitionen	Langvokale im CVC-Kontext
Größe des Vokalraums	Formanten F1, F2, F3	Langvokale [i:, e:, o:, u:]
Güte der Stimmqualität	Jitter, Shimmer, HNR	Langvokale

Tabelle 1: Segmental-akustische Parameter zur phonetischen Charakterisierung gestörter Sprache für unterschiedlichen Lauttypen, zugehörige akustische Analyseprozedur und Beschreibung des lautlichen Kontextes für jeden Parameter.

Stärke des Vokal-Verschlusslaut-Kontrastes

Es wird die Intensitätsdifferenz zwischen Maximum der Intensität im vorhergehendem Vokal und Minimum der Intensität in der Verschlussphase des Plosivs gemessen. Vor Durchführung der Messung wird das Signal über 500 Hz hochpassgefiltert, damit im Falle der stimmhaften Plosivlaute der (auch bei Normalsprechern auftretende) Stimmton während der Verschlussbildung nicht mit einfließt.

Stärke des Plosionsgeräusches

Es wird die Intensitätsdifferenz zwischen Maximum der Intensität im Plosionsrauschen (im “noise burst”) des stimmlosen Plosivs und dem Minimum der Intensität in der zeitlich vorhergehenden Verschlussphase dieses Lautes gemessen.

Stärke der Frikativ-Geräuschbildung

Es wird die Intensitätsdifferenz zwischen Maximum der Intensität der Geräuschbildung des stimmlosen Frikatives und Maximum der Intensität im vorhergehenden Vokal gemessen. Es werden nur labiodentale, alveolare und postalveolare Frikative gemessen. Vor Durchführung der Messung wird das Signal über 5 kHz hochpassgefiltert, damit das Rauschsignal des Frikativs gegenüber den vo-

kalischen Signalen (Hauptenergieanteil < 5 kHz) besser hervortritt.

Stärke des konsonantischen Stimmkontrastes

Die Stärke des konsonantischen Stimmkontrastes kann anhand der VOT (voice onset time) gemessen werden. Der Kontrast wird durch die Messung jeweils eines stimmhaften und eines stimmlosen Plosivlautes pro Äußerung ermittelt.

Stärke der konsonantischen Formanttransitionen

Es wird die Differenz der Frequenz des zweiten Formanten zum Zeitpunkt der konsonantischen Verschlusslösung (-bildung) und zum zeitlichen Mittelpunkt des vorhergehenden bzw. nachfolgenden Vokals gemessen.

Größe des Vokalraumes

Es werden die ersten drei Formanten von vier Langvokalen des Deutschen ([i:], [a:], [o:] und [u:]) gemessen. Neben der Berechnung der Größe des Vokalraumes innerhalb der logarithmierten F1-F2-Ebene werden auch direkt die Differenzen der (logarithmierten) Formantfrequenzen des ersten und zweiten Formanten für bestimmte Vokalpaare - z.B. hohe vs. tiefe oder hintere vs. vordere Vokale) als Maß zur Abschätzung der Größe des Vokalraumes verwendet.

Güte der Stimmqualität

Es werden Jitter, Shimmer und Harmonics-to-Noise-Ratio anhand von Phonationsstrecken in Langvokalen berechnet. Um die Stimmqualität bei gesprochener Sprache zu messen, wurden die Stimmqualitätsmessungen nicht anhand von isolierten mit konstanter Grundfrequenz gehaltener Vokale (künstliche Sprechsituation), sondern anhand der im Kontext gegebenen Langvokale durchgeführt.

Ergebnisse

Beim dysarthrischen Sprecher ergeben sich für alle segmental-akustischen Parameter geringere Werte im Vergleich zum Normalsprecher (schwächerer Vokal-Verschlusslaut-Kontrast). Im einzelnen bedeutet dies: Der pathologische Sprecher produziert weniger ausgeprägte Verschlüsse bei Plosiven als der Normalsprecher. Der gestörte Sprecher produziert ein geringeres Rauschen bzw. gar kein Rauschen bei stimmlosen Plosiven und bei Frikativen. Auch die Stärke der Kontrastierung des konsonantischen Artikulationsortes ist beim pathologischen Sprecher geringer, d.h. es treten hier konsonantische Formanttransitionen in geringerem Umfang auf als beim Normalsprecher. Darüber hinaus fällt die Größe des Vokalraumes des gestörten Sprechers im Vergleich zum Normalsprecher geringer aus, d.h. der gestörte Sprecher produziert geringere Vokalkontraste. Ebenfalls liegen beim gestörten Sprecher höhere Jitter- und Shimmerwerte und niedrigere Werte des Harmonics-to-Noise-Ratio vor.

Artikulatorische Methoden

In der Therapie von Dyslalien und Sprechapraxien kommt zunehmend ein Artikulationsmodell (Kröger 1998) mit der Option der Visualisierung von Sprechbewegungen (Kröger 2002 und 2003) zum Einsatz. Dieses Modell verfügt über die Möglichkeit der Darstellung von Sprechbewegungen in der mediosagittalen Schnittebene (Abb. 1). Das Modell erlaubt die Eingabe und Darstellung von

Einzellauten als Lautbilder und von Lautfolgen als visuelle Animation des Sprechbewegungsablaufes.

In der Therapie von Dyslalien wird das Modell zum gezielten Training der velar und palatal gebildeten Konsonanten [k], [g], [ç], [x] und [ŋ] wie auch der alveolar und postalveolar gebildeten Frikative [s], [ʃ], [z], und [ʒ] in unterschiedlich komplexen lautlichen Kontexten genutzt. In der Therapie von Sprechapraxien wird das Modell zum (Wieder-)Erlernen elementarer artikulatorischer und koartikulatorischer Muster auf der Ebene der Silbe (einfache Silbenstrukturen: VC, CV und CVC) genutzt. Der Patient hat die Aufgabe, die entsprechenden Laute bzw. Lautfolgen anhand des visuell dargebotenen (evtl. verlangsamten) Ablaufes der Sprechbewegungen zu erkennen (z.B. Erkennung von CV-Silben bei Beschränkung auf C = [b, d, g] und V = [i:, a:, u:]). Da dem Modell in der gesamten Therapie einer Sprechstörung nur eine begleitende Rolle, ermöglicht die Erkennungsrate einen Rückschluss auf den Lernerfolg insbesondere aufgrund des Einsatzes des Modells.



Abbildung 1: Darstellung des Mediosagittalschnittes des Schwa-Lautes mittels des Artikulationsmodells (Kröger 2002 und 2003)

Literatur

- Boersma P, Wenink D (2003) Praat: doing phonetics by computer. <http://www.praat.org>
- Kröger BJ (1998) Ein phonetisches Modell der Sprachproduktion. Tübingen: Niemeyer
- Kröger BJ (2002) SpeechTrainer – Handbuch. Internetskript PDF unter <http://www.phoniatrie.ukaachen.de> > Lehre > SpeechTrainer
- Kröger BJ (2003) Ein visuelles Modell der Artikulation. Laryngo – Rhino – Otologie (Im Druck)
- Ziegler W (2002) Auditive Methoden in der Neurophonetik. Habilitationsschrift, LMU München
- Ziegler W, Vogel M (1998) Diagnostik dysarthrischer Störungen. In: W Ziegler, M Vogel, B Gröne, H Schröter-Morasch, Dysarthrie. Stuttgart: Thieme, S. 26-52