

OBJEKTIVE BEWERTUNG VON /S/-ALLOPHONEN

EBERHARD STOCK

UWE HOLLMACH

Bereich Sprechwissenschaft
Martin-Luther-Universität
Halle, DDR 4020

ZUSAMMENFASSUNG

Es wird über die Entwicklung eines in der Praxis verwendbaren, möglichst billigen und möglichst einfachen computergestützten Verfahrens zur Rationalisierung sprechwissenschaftlich-phoniatrischer Routineuntersuchungen bei Studienbewerbern berichtet. Die Aufstellung der Vergleichsnorm und die Durchführung des Vergleichs werden diskutiert.

PROBLEMSTELLUNG

In der DDR bewerben sich in jedem Jahr zehntausende von Abiturienten an den Universitäten und Hochschulen mit dem Ziel, eine Ausbildung für einen sprechintensiven Beruf (z.B. als Lehrer, als Schauspieler, als Kindergärtnerin) zu absolvieren. Alle diese Bewerber müssen sich einer Tauglichkeitsprüfung unterziehen, die durch eine "Gemeinsame Anweisung der Ministerien für Volksbildung und Gesundheitswesen zur Beurteilung der Tauglichkeit für Berufe mit besonderer Stimm- und Sprechbelastung" aus dem Jahre 1974 geregelt wird. Nach dieser Anweisung sind folgende Untersuchungen durchzuführen:

- (1) Sorgfältige Anamneseerhebung zur Einschätzung der Stimm- und Sprechleistung;
- (2) Erhebung des HNO-fachärztlichen Status, gegebenenfalls mit audiologischen und röntgenologischen Überprüfungen;
- (3) Ermittlung des Stimmstatus, d.h. des Stimmklanges, des Stimmesatzes, der Sprechstimmlage und der Steigerungsfähigkeit, gegebenenfalls mit Hilfe der Stroboskopie und der Pneumographie;
- (4) Ermittlung des Sprechstatus, d.h. der Artikulation und des Sprechablaufs.

Als untauglich müssen Bewerber unter anderem dann eingestuft werden, wenn sie Stimm- und Sprechstörungen haben, die einer Behandlung nicht oder nur schwer zugänglich sind.

Diese Routineuntersuchungen sind außerordentlich arbeitsintensiv und zeitaufwendig, sie müssen dringend rationalisiert

werden, damit die Behandlungskapazität der entsprechenden Einrichtungen nicht unzumutbar eingeschränkt wird. Unsere Arbeiten zielen darauf ab, für die Ermittlung des Stimm- und Sprechstatus ein möglichst einfaches und möglichst billiges computergestütztes Verfahren zu entwickeln, das es ermöglicht, einen Teil der Untersuchungen Hilfskräften zu überantworten, ohne daß die Qualität der Beurteilungen dadurch eingeschränkt wird.

Wir haben aus zwei Gründen mit Arbeiten zur objektiven Bewertung von Sigmatismen begonnen: 1. Sigmatismen machen nicht nur im Kindesalter den größten Teil aller Lautbildungsstörungen aus, sondern sie sind auch bei Erwachsenen, wenigstens im deutschen Sprachbereich, als nahezu einzige funktionelle Stammelfehler weit verbreitet und werden von Kindern leicht imitiert. Studienbewerber mit auffälligem Sigmatismus sind deshalb für ein Lehrerstudium nicht tauglich.

2. Die subjektiv-auditive Bewertung der /S/-Realisationen ist vom Hörvermögen abhängig, das bekanntlich mit wachsendem Alter in der Höhe abnimmt, wodurch die Diskriminationsfähigkeit für korrekte bzw. unkorrekte /S/-Allophone eingeschränkt wird. Da die Zahl der Sigmatiker unter Lehrer- und Schauspielstudenten trotz Tauglichkeitsuntersuchung sehr groß ist und möglicherweise anwächst (in manchen Statistiken der letzten zehn Jahre werden bis zu 35 % der Studenten eines Matrikels als Sigmatiker ausgewiesen), sind wir der Beeinträchtigung des Hörvermögens und der Diskriminationsfähigkeit gesondert nachgegangen. Mit einem von U. Hollmach entwickelten Hochfrequenzaudiometer wurden in drei verschiedenen Altersgruppen Hörschwellenuntersuchungen durchgeführt. Es handelte sich (1) um zwölf Lehrerstudenten mit einem Durchschnittsalter von 21,2, (2) um zwölf Sprechwissenschaftler mit einem Durchschnittsalter von 31,7 und (3) um acht Lehrkräfte für Musik, alle über 50 Jahre alt mit einem Durchschnittsalter von 53,9. Das Audiometer kann in der Frequenz stufenlos zwischen 20 Hz und 22 kHz eingestellt werden, die Intensität

ist in 1-dB-Stufen regelbar. Die Umwertung der dB-Angaben in Phonzahlen und die Korrektur, die wegen der Eigenfrequenz des verwendeten orthodynamischen Kopfhörers und des für die Prüfung eingesetzten künstlichen Ohr erforderlich war, wurden mit einem speziellen Programm durch den Kleincomputer KC 85/3 des VEB Mikroelektronik "Wilhelm Pieck" Mühlhausen vorgenommen. Für die Prüfung wurden 15 Frequenzen ausgewählt. Die Mittel der Schwellenintensitäten (\bar{x}) und die jeweilige Standardabweichung (s) sehen abgerundet wie folgt aus:

kHz	Gruppe 1		Gruppe 2		Gruppe 3	
	\bar{x}	s	\bar{x}	s	\bar{x}	s
0.4	6	4	7.8	4	3.8	6
0.7	8	4	6.4	7	3.4	3
1	4	4	4	6	1.5	3
1.2	0.9	3	2.4	5	2.1	4
1.5	4	7	6.2	5	4.9	6
2	5.9	4	4.9	6	12	8
2.5	6.7	5	7	8	13.8	11
3	6.3	6	9	9	16.6	8
4	5.6	9	8.8	7	19.3	12
6.3	9.6	6	12.7	9	27.9	23
7.6	5	6	10.9	10	38	66
9.5	6.6	6	14.8	8	48	63
10.8	9.2	5	12.3	14	51	64
13	2	9	10.2	12	101	82
16	16	11	48.8	81	-	-

Verständlicherweise wurde für die Gruppe der Studenten beim Mittel und bei den Extremwerten das beste Hörvermögen festgestellt. Bei der folgenden S-Bewertung urteilte sie am homogensten. Die zweite Gruppe zeigt die zweitbeste Mittelwertkurve; die Extremwerte liegen weiter vom Mittelwert entfernt, als in der ersten Gruppe. Die stärksten Ausfälle, besonders bei hohen Frequenzen, finden sich in der letzten Gruppe, in der aber eine Besonderheit zu beobachten ist. Bis 1,2 kHz haben die älteren Probanden eine um durchschnittlich 4 dB niedrigere Hörschwelle, als die beiden anderen Gruppen. Außerdem ist die hier nicht zusätzlich ausgewiesene Varianz in diesem Bereich für die Gruppe 3 deutlich geringer als bei den anderen Gruppen. Besonders bei den Studenten ist die Varianz bei 1,5 und 4 kHz sehr groß. Eine Erklärung hierfür steht noch aus.

Der Hörschwellenaudiometrie schloß sich sofort ohne Veränderung der Stellung des Kopfhörers ein Diskriminationstest mit

Satzpaaren an, in dem gegenübergestellte /S/-Allophone auf Geräuschscharfe, auf Geräuschfarbe, auf Ähnlichkeit der Geräusche und auf Korrektheit der /S/-Realisation zu beurteilen waren. Um das Verhältnis zwischen den diskreten, von einem computergestützten Analysator ausgegebenen Spektraldaten der einzelnen /S/-Allophone und den ermittelten Hörschwellenwerten bestimmen zu können, wurde für jeden Probanden die Differenz zwischen Hörschwellenwerten und Spektraldaten ermittelt. Die entstehenden Differenzkurven ermöglichen Aussagen über die Wahrnehmungswahrscheinlichkeit für die einzelnen /S/-Allophone. Je positiver die Kurvenwerte sind, desto größer ist die wahrgenommene Lautstärke für die einzelnen Frequenzen und desto höher die Wahrscheinlichkeit, daß die wahrgenommenen Frequenzbereiche signalgerecht identifiziert werden. Erwartungsgemäß ergab sich, daß die Gruppe der Studenten die Sigmatismen am ehesten erkannte und am besten unterschied, daß die mittlere Altersgruppe auch eine Reihe von stumpfen S-Geräuschen als scharf beurteilte, daß diese sowie die letzte Gruppe den Sigmatisimus stridens nicht erkannte und daß die über 50 Jahre alten Hörer die sigmatischen Allophone nicht sicher unterscheiden konnten. Diese Aussage bezieht sich selbstverständlich nur auf die auditive Beurteilung. In der Diagnose der sprechwissenschaftlichen oder phoniatischen Praxis werden natürlich auch visuell aufgenommene Informationen genutzt. Trotzdem fordern diese Untersuchungsergebnisse, daß die erfahrenen älteren Kollegen, die in der Regel die für die Tauglichkeitsuntersuchungen zuständigen Einrichtungen leiten und dort aktiv tätig sind, nicht nur von Arbeitsaufwand entlastet, sondern auch hinsichtlich der Beurteilung der /S/-Realisationen unterstützt werden müssen. Nach unserer Vorstellung kann das auf sehr ökonomische Weise mit einem entsprechend programmierten Kleincomputer geschehen, der mit einem Analog-Digital-Umwandler bestückt ist und bei der Artikulation von Testwörtern oder kurzen Sätzen eine objektive Bewertung der /S/-Allophone, aber z.B. auch der Vokalartikulation im Fremdsprachenunterricht vornimmt. Der benutzte Kleincomputer KC 85/3 wird von uns vor allem deshalb für den Einsatz in der sprechwissenschaftlich-phoniatrischen Praxis empfohlen, weil er vielfältig nutzbar ist und z.B. durch RAM-Erweiterung und Zusatzmodul mit Textverarbeitungsprogramm auch die Rationalisierung der Anamnese und Diagnose gestattet.

AUFBAU DES SPEKTRALANALYSESYSTEMS

Für die Forschungsarbeiten mußte zunächst ein Spektralanalysesystem entwickelt werden, das es möglich macht, zusammenhängende natürliche Äußerungen mit einer be-

stimmten Mindestdauer, also "fließende Lautsprache", zu untersuchen. Dieses System besteht aus 4 Komponenten: einer Filterbank, einem Analog-Digital-Umwandler, einem Kleincomputer und der erforderlichen Software.

Mit der von U. Hollmach aufgebauten Filterbank wird die gesprochene Sprache in frequenzdiskrete Signale zerlegt, wobei der Abstand der einzelnen Frequenzkomponenten eine Viertel Oktave beträgt. Die Filterbank umfaßt 32 Kanäle, die parallel geschaltet sind und einen Frequenzgang von 80 Hz bis 18 kHz haben. Diese 32 Kanäle werden durch einen Analog-Multiplexer (AMUX) nacheinander durchgeschaltet, der das gleichgerichtete Signal seriell zum Ausgang führt. Die Integrationszeit bzw. das Zeitfenster ist zweistufig ausgelegt; Signale, die sich schnell verändern (z.B. Sprache), können mit einem Zeitfenster von 6 ms erfaßt werden, für quasi-konstante Signale (z.B. Stimmklang) steht ein Zeitfenster von 138 ms zur Verfügung. Während der AMUX an einem beliebigen Filterkanal das Analogsignal abfragt, würde an allen anderen Filterkanälen das Signal zeitlich versetzt weiterhin integriert werden. Das hätte eine Verfälschung der Signale zur Folge. Durch den Einsatz eines Analogspeichers wird die Ungenauigkeit verhindert. Im Analogspeicher steht der integrierte Filterausgangswert bis zur Abfrage zur Verfügung. In der Zwischenzeit kann an jedem Filter der zeitlich äquivalente Wert abgerufen werden.

Der Ausgang der Filterbank führt auf den Analog-Digital-Umwandler (ADU), mit dessen Hilfe die analogen Gleichspannungssignale in digitale, für den Computer verarbeitbare Signale umgesetzt werden. Der von uns benutzte ADU ist ein Zusatzmodul (M 010) für den KC 85/3. Er besitzt eine Auflösung von 10 bit, das entspricht 1024 Stufen.

Der KC 85/3 ist ein 8-bit-Rechner; er hat Vollgraphik und eine 16-farbige Displaystrukturierung. Ein besonderer Vorteil dieses Computers sind seine Erweiterungsmöglichkeiten. So kann selbstverständlich ein Drucker angeschlossen werden, und der RAM-Speicher kann bis zu 4 Mbyte aufgestockt werden.

Die Software wurde zur Hälfte in BASIC und für schnelle Abläufe in Maschinensprache (U 880) geschrieben. Das Programm ist ein Grundprogramm für die Spektralisierung und Auswertung gesprochener Sprache und kann für jede spezielle Anwendung leicht erweitert oder umgestellt werden. Mittels Menütechnik können die graphischen Darstellungs- und Auswertungsvarianten benutzerfreundlich aufgelistet werden. Das Programm sieht eine farbige

sonographische Darstellung mit 12 Farbwerten für die unterschiedlichen Intensitätswerte vor. Im Gegensatz zu den klassischen Sonagrammen mit ihrer Grauwertdarstellung ist die farbige Repräsentation übersichtlicher und genauer. Aus dem als Sonagramm erscheinenden Abschnitt der gesprochenen Sprache läßt sich insbesondere das stimmlose /S/-Allophon leicht herausfinden. Die optische Segmentierung erfolgt mit zwei Leuchtbalen. Aus den eingelesenen Daten kann sowohl ein Frequenz-Intensitätsdiagramm als auch ein Intensitätszeitdiagramm aufgebaut werden. Für die Beurteilung der /S/-Allophone ist das Frequenz-Intensitätsdiagramm vorteilhaft.

REFERENZMUSTER UND VERGLEICH

Die objektive Bewertung beliebiger idiolektaler /S/-Allophone geschieht durch den Vergleich mit einem Referenzmuster, das der Norm des Aussprachestandards entspricht. Als Norm wird hier die von den Sprachbenutzern als dialektfrei und korrekt beurteilte Realisierungsvariante verstanden. Die Herstellung des Referenzmusters und der Vergleich orientieren sich an den Prozessen der Spracherkennung /1/. Die Feststellung und Beschreibung der Norm ist problematisch, so daß spezielle und noch nicht abgeschlossene Untersuchungen angesetzt werden mußten. Mit diesen Untersuchungen wurde in Kommunikationsexperimenten die Wirkung der einzelnen /S/-Allophone bzw. Sigmatismen als subjektive Bewertung durch den Hörer erfaßt. Methodisches Instrument waren standardisierte Polaritätsprofile, mit denen üblicherweise soziale Bezüge charakterisiert werden können. Die bisherigen Ergebnisse besagen, daß bei Kommunikation über Mikrofon (z.B. Rundfunk und Fernsehen) höhere Anforderungen gestellt werden und ein zu scharfes oder zu stumpfes S-Geräusch als auffällig und störend abgelehnt wird. In der face-to-face-Kommunikation und vor allem im Alltagsgespräch wird dagegen ein Sigmatisimus stridens und selbst ein schwacher Sigmatisimus addentalis sehr viel eher toleriert; hier ist die Norm also weniger streng. Je nach dem Berufsziel des Studienbewerbers wird man deshalb möglicherweise mit zwei Referenzmustern arbeiten müssen, einem Muster für die strenge Norm bei Mikrofonsprechern und einem zweiten Muster für die liberalisierte Norm bei Lehrern und Schauspielern. Das Referenzmuster entsteht, indem /S/-Allophone von mehreren Sprechern spektralisiert und die Spektraldaten gemittelt werden. Die einzelnen Allophone entstammen unterschiedlichen Phonenverbindungen in unterschiedlichen Äußerungen; jeder Sprecher wird durch eine Expertengruppe auditiv und visuell auf die Korrektheit seiner /S/-Realisation über-

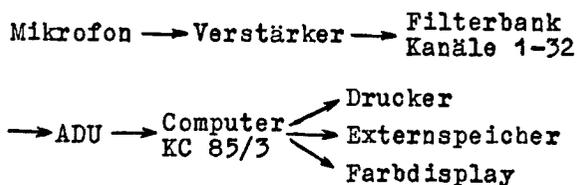
prüft.

Bei der Diagnose bzw. der Erhebung des Sprechstatus werden /S/-Allophone aus Testsätzen mit dem gespeicherten Referenzmuster nach der Minimum-Distanz-Strategie /2/ verglichen. Dieses statistische Verfahren nutzt die Beziehung zwischen dem Mittelwert der korrekten /S/-Realisationen und der durch die idiolektale Variation der Normsprecher entstehenden Streuung der Spektraldaten, wobei der Abstand zwischen dem jeweiligen Mittelwert und dem Streuungswert als die zulässige minimale Distanz bewertet wird. Überschreiten die Spektraldaten der /S/-Realisation eines zu diagnostizierenden Sprechers diese minimale Distanz, so wird diese /S/-Realisation zurückgewiesen. In unserer Forschung muß also praktisch die mögliche Unschärfe der /S/-Realisation bestimmt werden, die noch als korrekte Artikulation bewertet wird. Dabei ist ein Maß für die Ähnlichkeit aus den Distanzzahlen zu ermitteln.

Es besteht auch die Möglichkeit, die Wörter bzw. /S/-Allophone bereits im Sonagramm zu verrechnen. Ein entsprechendes Programm hierzu existiert an der Sektion Informationstechnik der TU Dresden. Hierbei werden mittels dynamischer Programmierung zwei Sonagramme zeitäquivalent übereinander geschoben und dadurch vergleichbar. Diese Methode erfordert jedoch einen hohen Rechenaufwand.

Die Spektralanalyse mit Hilfe der Filterbank dient der Grundlagenforschung und ist nicht für den Einsatz in der Praxis gedacht. Auf der Grundlage der gewonnenen Ergebnisse wird ein System mit Fast-Fourier-Transformation entwickelt, das lediglich den Kleincomputer mit dem ADU-Zusatzmodul nutzt. Die Arbeit mit der Filterbank war jedoch unabdingbar, weil sie gegenüber der Fast-Fourier-Transformation zwei Vorteile bietet: (1) Spektralisierung mit geringem Rechenaufwand und kurzer Rechenzeit, (2) speicherplatzsparendes Zerlegen des Signals bis zu einer Frequenz von 18 kHz (bei 10 bit Auflösung). Bei 64 kbyte des Computers hat die Filterbank eine Grenzfrequenz mit einem Analysebereich von 1.2 s, die bei 20 kHz liegt. Das Zeitfenster bietet einen Ausschnitt von 6 ms.

Das Spektralanalysesystem hat folgenden Aufbau:



LITERATUR

- /1/ Paul, V.: Modelle für die Verarbeitung fließender Sprache. Nachrichten/Elektronik 37 (1987), 22-23
- /2/ Berg, H.: Statistische Untersuchung an einem Spracherkennungssystem. In: 20. Fachkolloquium Informationstechnik. TU Dresden 1987, 94-98