

EIN COMPUTERPROGRAMM ZUR UNTERSUCHUNG VON EINZELPERIODEN PHONETISCHER SIGNALE

DIETER STOCK

Stimmhafte Sprachschallsignale werden in der phonetischen Literatur nicht als periodisch, sondern zumeist als QUASIPERIODISCH bezeichnet. Diese Einschränkung trägt dem empirischen Tatbestand Rechnung, daß die aufeinanderfolgenden Perioden eines phonetischen Signals in mehr oder weniger starkem Maß unähnlich sind. Die Unähnlichkeit dokumentiert sich sowohl in der Periodenlänge als auch in der Amplitude korrespondierender Signalextrema. Schließlich besitzt auch die innerperiodische Feinstruktur keineswegs exakt wiederkehrende Regularitäten. Der Grund dieses im strengen Sinne aperiodischen Phänomengehalts liegt sowohl auf Seiten der Schallerzeugung, nämlich der Phonation, als auch auf Seiten der Schallmodifikation, d.h. der Artikulation. Beide Kenngrößen, also die Parameter des glottalen Erregungssignals wie die Systemeigenschaften der vom Ansatzrohr gebildeten akustischen Übertragungsfunktion, weisen unabhängig voneinander zeitliche Fluktuationen auf. Sie sind naturgemäß dort besonders groß, wo sich das Signal in einem segmentalen oder suprasegmentalen Übergangsstadium befindet, d.h. an den Lautgrenzen, bzw. beim Wechsel von Tonhöhe oder Dynamik. Aber auch im Lautinneren lassen sich -trotz neutraler Affektlage während des Sprechens- mehr oder weniger deutliche Variationen der Signalstruktur nachweisen. Ihrem Einfluß ist es zuzuschreiben, daß wir synthetischen Sprachschall auditiv spontan von natürlichem zu unterscheiden vermögen.

Um den Strukturgehalt eines phonetischen Signals phänomenologisch deskriptiv zu erfassen, um also beispielsweise Einzelperioden visuell miteinander zu vergleichen, bietet das Oszillogramm eine hervorragende Grundlage. Vorauszusetzen ist dabei eine hinreichend breitbandige Auflösung bezüglich des Schreibsystems und ein genügend schneller Papiertransport von mindestens 2-3 m/s. Mit zunehmender zeitlicher Auflösung wird aber ein Nachteil spürbar, der den visuellen Vergleich außerordentlich erschwert, und zwar ist dies die fortschreitende räumliche Distanz phasengleicher Punkte. Insbesondere gelingt es kaum, den Ort und den Grad zwischenperiodischer Diskrepanzen auf einen Blick zu erkennen oder, was uns noch wichtiger erscheint, den Verlauf bestimmter Signalmuster über mehrere Perioden hinweg figurativ zu verfolgen, wie dies etwa beim Sonagramm hinsichtlich spektraler

Intensitätskonzentrationen in sehr einfacher und einprägsamer Weise möglich ist.

Alle genannten Nachteile lassen sich trotz beliebig hoher zeitlicher Auflösung vermeiden, wenn man dazu übergeht, die einzelnen Perioden nicht nebeneinander, sondern UNTEREINANDER zu registrieren, indem die Aufzeichnung am Periodenende abgebrochen und mit dem Anfang der nächsten Periode erneut am linken Bildrand beginnend fortgesetzt wird. Dadurch geraten phasengleiche Punkte mit einem bestimmten vertikalen Versatz untereinander und lassen sich, auch über längere Segmente hinweg, leicht miteinander vergleichen. Sorgt man außerdem dafür, daß der vertikale Versatz automatisch vom Grad der Unähnlichkeit gesteuert wird, dann kommt man zu graphischen Gebilden wie beispielsweise dem in Abb. 1 gezeigten.

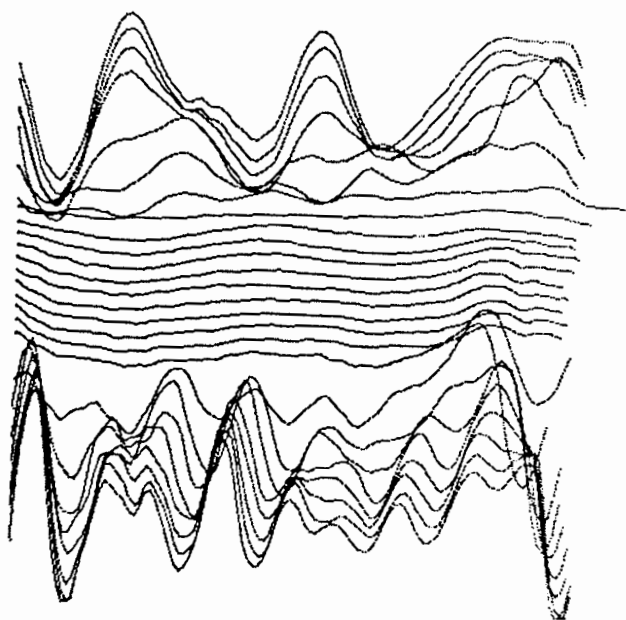


Abb. 1

Es handelt sich um eine vollständige Registrierung des Kunstwortes 'oma', das eine Gesamtdauer von ca. 300 ms besitzt und von einer Männerstimme gesprochen wurde. Man erkennt deutlich die drei Segmente [o], [m] und [a] sowie die dazwischenliegenden Übergangsstücke. Außerdem spiegelt die durch die Endpunkte der einzelnen Perioden gebildete senkrechte Berandungskurve den Tonhöhenverlauf der Äußerung wieder.

Das nächste Bild (Abb. 2) demonstriert den Periodenverlauf des phonetischen Diphthongs [ae]. Hier ist die sehr gleichmäßige Veränderung des Schwingungsmusters von der ersten bis zur letzten Periode, also der Übergang von [a] nach [e] hervorzuheben, was sich auch in der relativ konstanten mittleren Distanz der Kurven dokumentiert.

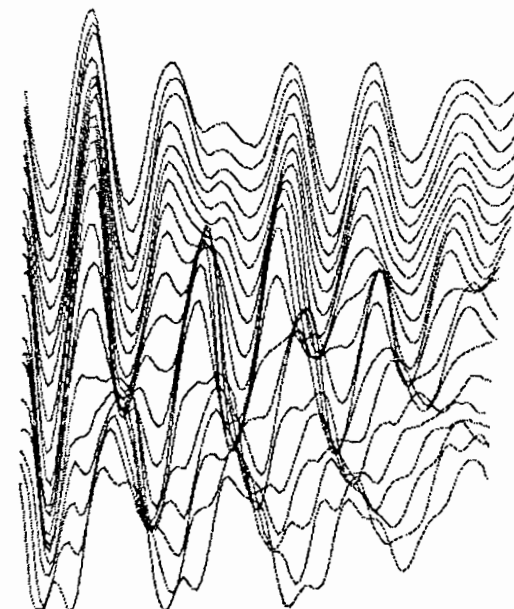


Abb. 2

Die praktische Realisierung dieser speziellen Registriermethodik war mit konventionellen Verfahren, wie beispielsweise mittels Oszillograph und Periodizitätsanalysator nicht zu bewältigen. Stattdessen haben wir auf die digitale Signalverarbeitung mit einem Computer vom Typ PDP 15/20 zurückgegriffen. Das elektrische Sprachsignal wird zunächst mit einer Samplefrequenz von 14 kHz digital konvertiert und als diskrete Zahlenfolge in den Kernspeicher eingelesen. Sodann wird am Signalbeginn auf einer Länge von 12,5 ms (der größten potentiellen Periodendauer) die Lage und der Betrag des Maximums, der benachbarten Minima sowie des linken Nebenmaximums bestimmt. Diese Informationen dienen der Festlegung des 'physiologischen' Periodenbeginns, d.h. desjenigen Signalnulldurchgangs, der die Grenze zwischen den energetisch fallenden und wieder ansteigenden Formantschwingungen bildet. Die Hauptschwierigkeit bestand darin, das Ende der einzelnen Perioden, das ja gleichzeitig über den Beginn der nächsten entscheidet, in jeder segmentalen Phase sicher zu bestimmen. Hier war eine außerordentliche Präzision vonnöten, da aufgrund der vertikalen Darstellungsform bereits Meßfehler in der Größenordnung von einer Sampleperiode, also von ca. $\pm 100 \mu\text{s}$ visuell erkannt werden.

Wir haben uns im vorliegenden Fall für die Anwendung der Autokorrelationsanalyse entschlossen, die, auch im phonetischen Bereich, als eins der zuverlässigsten Hilfsmittel zur Periodenforschung gilt. Um jeweils eine Periode zu isolieren, werden rund 3.000 Signalaufnahmen miteinander korreliert, wobei die Sampleperiode ($T_s = 70 \mu\text{s}$) den Verzögerungsparameter τ repräsentiert. Die Beobachtungszeit beträgt ca. $2/3$ der letzten Periodendauer. Der größte gemessene Korrelationskoeffi-

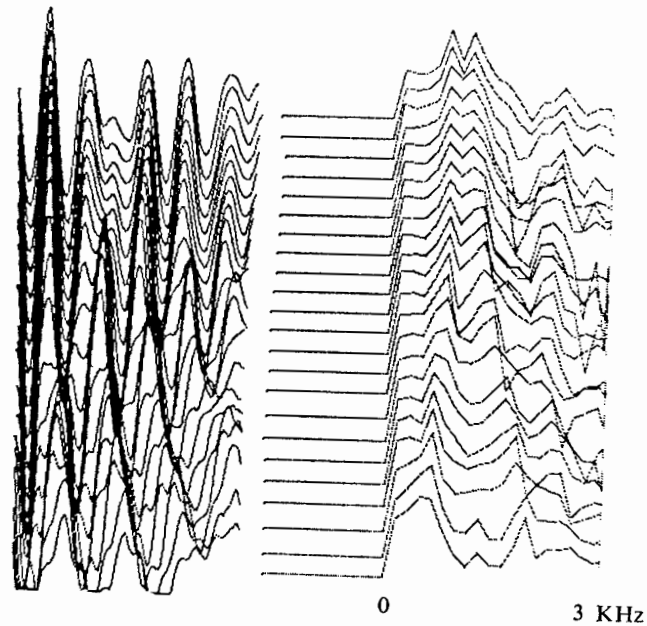


Abb. 3

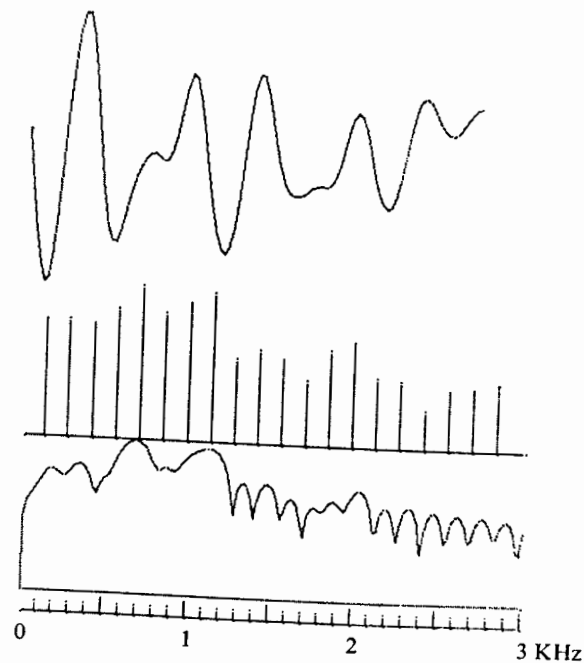


Abb. 4

zient entscheidet nicht nur über die Periodengrenze, sondern (mit seinem Kehrwert) auch über den vertikalen Versatz bei der Darstellung der Kurven. Für die Registrierung haben wir einen Speicheroszillographen benutzt und das vollständige Bild schließlich mit einer Kamera vom Bildschirm abfotografiert.

Zum Schluß sei noch anhand zweier weiterer Schirmbilder eine zusätzliche Anwendungsmöglichkeit der geschilderten Methode angedeutet, nämlich die ebenfalls vom Rechner durchgeführte Fouriertransformation einzelner Signalperioden. Abb. 3 zeigt links wiederum den Zeitverlauf der Einzelperioden von [æ] und rechts daneben das logarithmierte Amplitudenspektrum jeder einzelnen Periode bis 3 kHz. Hier ist das Auseinanderstreben der beiden unteren Formanten zum Lautende hin besonders evident. Das letzte Bild schließlich (Abb. 4) dokumentiert oben den Zeitverlauf einer Periode von [a], in der Mitte das diskrete Amplitudenspektrum für den (unrealistischen) Fall einer periodischen Fortsetzung dieser Signalkonfiguration und unten das kontinuierliche Spektrum des einmaligen Vorgangs.

*Institut für Kommunikationsforschung
und Phonetik
Universität Bonn*

DISCUSSION

IIVONEN (Oulu)

Ist es möglich, automatisch eine beliebige Stelle am Sprachsignal für eine weitere Analyse zu lokalisieren? Z.B., kann der Mittelpunkt eines [a] in irgendeinem Wort automatisch lokalisiert und spektral dargestellt werden?

STOCK

Im augenblicklichen Versuchsstadium können nur stimmhafte Sprachschallsignale verarbeitet werden. Abgesehen von dieser Restriktion, lässt sich jede Periode oder Gruppe von Perioden automatisch lokalisieren, die sich durch bestimmte objektive Kriterien von ihrer Umgebung unterscheidet, Z.B. also durch Angabe der Periodennummer, eines bestimmten Zeitpunktes oder einer hinreichend signifikanten zeitabhängigen oder frequenzabhängigen Amplitude.