

Erläuterungen zur Sprachakustik

Diese Erläuterungen sind nicht als Ersatz für die Lektüre von Phonetik-Lehrbüchern, sondern als Ergänzung zu dieser Lektüre gedacht. Schauen Sie für zentrale Begriffe auch in die Bücher, die Sie auf der Literaturliste finden (speziell in das Buch von Henning Reetz), und in die Vorlesungsskripte von Jonathan Harrington.

Die Akustik wird in einer eigenen Lehrveranstaltung noch genauer und ausführlicher behandelt, Sie sollten sich jetzt vor allem um das prinzipielle Verständnis zentraler Grundbegriffe bemühen. Das Verständnis von DFT und FFT ist z. B. zu diesem Zeitpunkt noch nicht so wichtig.

Einleitendes:

- Schallquellen verursachen durch Bewegung Druckänderungen im Übertragungsmedium. Druckänderungen werden im Ohr wieder in Bewegung umgesetzt; diese führt zu neuronalen Signalen und letztendlich zu einem Höreindruck.
- Akustisches Medium ist typischerweise Luft, aber auch Wasser oder Festkörper sind als Medium möglich. Jedes Medium hat eine eigene Übertragungs-(Schall-) Geschwindigkeit c , bei Luft liegt c bei etwa 330-350 m/s.
- Im Ruhezustand befinden sich die Luftmoleküle etwa im gleichen Abstand zueinander (abhängig vom atmosphärischen Druck). Wenn eine Schallquelle sich bewegt, werden die Moleküle aufeinandergedrückt bzw. auseinandergezogen, neigen aber dazu, wieder auf gleichen Abstand zu kommen. → Druckwelle bewegt sich von der Quelle weg. (vgl. Stein im Teich.)
- Die Luftmoleküle selbst bewegen sich nur minimal, die Druckänderung (Schallwelle) pflanzt sich fort. Es handelt sich in allen Fällen um Hin- und Herbewegungen der Moleküle, also Schwingungen.
- Diese Schwingungen können auch als akustische Signale betrachtet werden:
 - rein periodisch (Sinustöne, Klänge)
 - quasiperiodisch (Vokale, Nasale)
 - aperiodisch (Frikative, aber auch die Verschlusslösungen von Plosiven)

Darstellungsformen von akustischen Signalen:

- Akustische Signale können im Zeitbereich oder im Frequenzbereich betrachtet werden.
- Oszillogramm/Zeitsignal zeigt die Luftdruckschwankungen im zeitlichen Verlauf: A (Amplitude) über t (Zeit)
- Spektrum: Zeigt die Frequenzbestandteile zu einem ausgewählten Zeitpunkt. Verschiedene Frequenzen kommen dadurch zustande, dass menschliche Sprache nicht aus reinperiodischen Sinusoiden besteht, sondern aus quasiperiodischen bzw. aperiodischen Schwingungen zusammengesetzt ist.
 - Amplitudenspektrum: A über f (Frequenz)
In Praat sind unter *Spectrum/View spectral slice* Hüllkurven zu sehen:
Hüllkurve: Mittlung eines Spektrums, abhängig von Bandbreite und Fensterfunktion (s. u.)
- Sonagramm (Spektrogramm):
 - zeigt f über t, A als Graustufen. Je höher die Amplitude, umso dunkler der Bereich.

Ausgangspunkt von Sonagrammen:

Fourier-Theorem: Jedes streng periodische Signal ist eine Summe von unendlich vielen Sinusoidalsignalen, deren Frequenzen ganzzahlige Vielfache (Harmonische) eines Ausgangssignals (Grundfrequenz) sind. Sinussignale sind durch die drei Parameter Frequenz, Amplitude und Phase eindeutig bestimmt.

Sprachlaute sind nicht streng periodisch (zeitbegrenzt, manchmal extrem kurz, variabel in Amplitude und Frequenz, bei Aperiodizität gar keine Perioden). Daher können aus einzelnen Perioden keine weiteren Perioden vorhergesagt werden (wie z. B. bei Sinussignalen oder Klängen, die durch Fouriertransformation aus einer Periode heraus rekonstruiert werden können).

Lösung des Problems: Analysefenster, die einen Ausschnitt des Signals von bestimmter Länge betrachten. Multiplikation von Fenster mit Werten des Signals (z. B. in der Mitte des Fensters Multiplikation mit 1, gegen Ende des Fensters mit 0). Die Fenster haben eine bestimmte Länge (gemessen in Abtastwerten), zeitliche Dauer lässt sich in Abhängigkeit von der Abtastrate herausfinden. Da es sich um einzelne (diskrete) Abtastwerte handelt, spricht man auch von Diskreter Fouriertransformation (DFT). Eine rechnerische Vereinfachung durch Verwendung von Zweierpotenzen stellt die Fast Fourier Transformation (FFT) dar.

Der Fensterinhalt wird als streng periodisch wiederkehrend betrachtet, und das Fenster wird über das gesamte Sprachsignal verschoben. Auf diese Weise erhält man eine Abfolge von Spektren (für jeden Abtastwert eins), so dass die Sprachlaute im Sonogramm nun sowohl im Zeitbereich als auch im Frequenzbereich betrachtet werden können.

- Die Bandbreite bestimmt die Frequenz-Auflösung des Signals und hängt mit der Länge des Analysefensters zusammen.

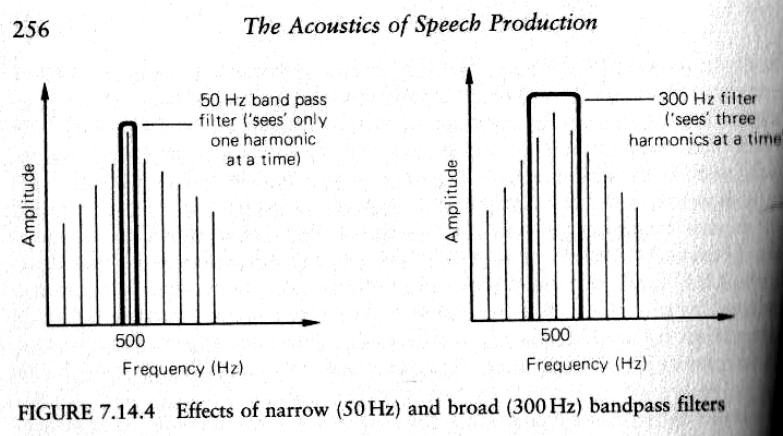


Abb. aus Clark and Yallop (1995), *An introduction to Phonetics and Phonology*, S. 256

- Breitband (Bandbreite 300 Hz, also größer als die höchste Grundfrequenz eines männlichen Sprechers). Das Breitbandsonogramm zeigt Frequenzbereiche mit viel Energie (hohe Amplituden der Signalkomponente) als dunkle, relativ breite Streifen, die als Formanten bezeichnet werden. Ein Breitbandsonogramm hat eine gute zeitliche Auflösung, dafür aber eine schlechtere Frequenzauflösung (z. B. sind einzelne Harmonische nicht erkennbar). Geeignet für Formantanalysen.
- Schmalband (Bandbreite 50 Hz, kleiner als die niedrigste Grundfrequenz) zeigt die einzelnen Signalkomponenten als schmale dunkle Streifen. Hier

ist die Frequenzauflösung gut (z. B. Harmonische), dafür ist die zeitliche Auflösung schlechter (Plosivlösungen oder einzelne Glottisschläge sind dadurch nicht so gut erkennbar wie im Breitbandsonagramm). Geeignet für F0-Analysen.

- Formanten¹ sind im Sonagramm als dunkle Balken zu erkennen. Sie sind das Resultat von Resonanzen (Verstärkungen) im Ansatzrohr: Die Luftsäule im Ansatzrohr wird durch das Glottissignal angeregt, an den durch die individuelle Geometrie vorbestimmten Frequenzen „mitschwingen“. Dadurch verformt sich das gleichmäßig abfallende Glottisspektrum zu Gipfeln, die als die typischen Formanten für bestimmte Laute bekannt sind.
 - Vokale werden in der Regel stimmhaft produziert. Durch die Schwingungen der Stimmlippen entsteht eine periodische Druckschwankung, die sich akustisch (d.h. als Schallwelle) durch den weiteren Vokaltrakt fortpflanzt.
 - Bei der Vokalproduktion kann der menschliche Vokaltrakt (Ansatzrohr) analog zu einem einseitig offenen Rohr betrachtet werden (geschlossenes Ende: Schallquelle Glottis, offenes Ende: Mundöffnung).
 - Die Glottis erzeugt eine Schallwelle, die von den Wänden des Ansatzrohres reflektiert werden, und die Reflektionen überlagern sich.
 - Bei bestimmten Resonanzfrequenzen des Rohres (= Formanten, abhängig von der Form des Ansatzrohres) entstehen stehende Wellen: geschlossenes Ende Bauch (Druckmaximum/-minimum), offenes Ende Knoten (Umgebungsluftdruck). Nun muss berechnet werden, bei welchen Frequenzen Bäuche am geschlossenen Ende (Glottis) bzw. Knoten am offenen Ende (Mundöffnung) auftreten. Für die Berechnung der Frequenzen gilt die Formel $F=c/\lambda$ (Lambda: Wellenlänge). Ein Berechnungsbeispiel befindet sich in Reetz (2003), *Artikulatorische und akustische Phonetik*, S. 34 f.
 - Formanten sind keine Harmonischen, aber ohne Harmonische im Glottisspektrum wäre auch keine Resonanz möglich (ohne Anregung durch die Glottis keine Resonanz). Formanten sind keine Eigenschwingungen wie z. B. bei einer angeschlagenen Glocke, aber wie bei einer Glocke werden die Formanten durch die Resonanzfähigkeit des Ansatzrohres und damit durch die Geometrie entscheidend bestimmt.
 - Formanten sind strenggenommen keine einzelnen Resonanzfrequenzen wie bei einem einseitig offenen Rohr, weil es sich beim menschlichen Ansatzrohr nicht um einen verlustfrei reflektierenden Körper handelt. Es handelt sich eher um Frequenzbänder (mehrere benachbarte Frequenzen, an denen Resonanzen entstehen), wodurch die dunklen Balken (und nicht nur Striche wie bei einer Frequenz) im Sonagramm entstehen. (nach Pompino-Marschall (2003), S. 108).
 - Eine nützliche Übersicht über Formantwerte deutscher Vokale befindet sich in Kohler (1995), *Einführung in die Phonetik des Deutschen*, 2. Auflage, S. 50.

¹ Die Ausführungen über Formanten basieren auf www.phonetik.uni-muenchen.de/AP.