

1.1.1.3. Fazit

Durch die z.T. sehr großen Unterschiede zwischen der tatsächlichen Aussprache und den "Zitierformen" der Wörter und Wortverbindungen ergeben sich für die Erkennung gesprochener Sprache gravierende Probleme (in weitaus geringerem Maße auch für die Synthese). Insbesondere ist der Status dieser Formen zu bestimmen. Wenngleich - aus welchen Gründen auch immer - die Sprechwirklichkeit von ihnen abweicht und sie daher als Konstrukte erscheinen, so behalten sie doch ihren Wert als Referenzmuster und Ausgangspunkt von Veränderungen. Es ist notwendig, einerseits eine Abstraktion von der lautlichen Realisierung vorzunehmen und andererseits diejenigen lautlichen Unterschiede auszumachen, die in einer konkreten natürlichen Sprache zu "wirklichen" Unterschieden führen. Dies ist die Aufgabe der Phonologie (s. 1.2.).

1.1.2. Akustische Phonetik

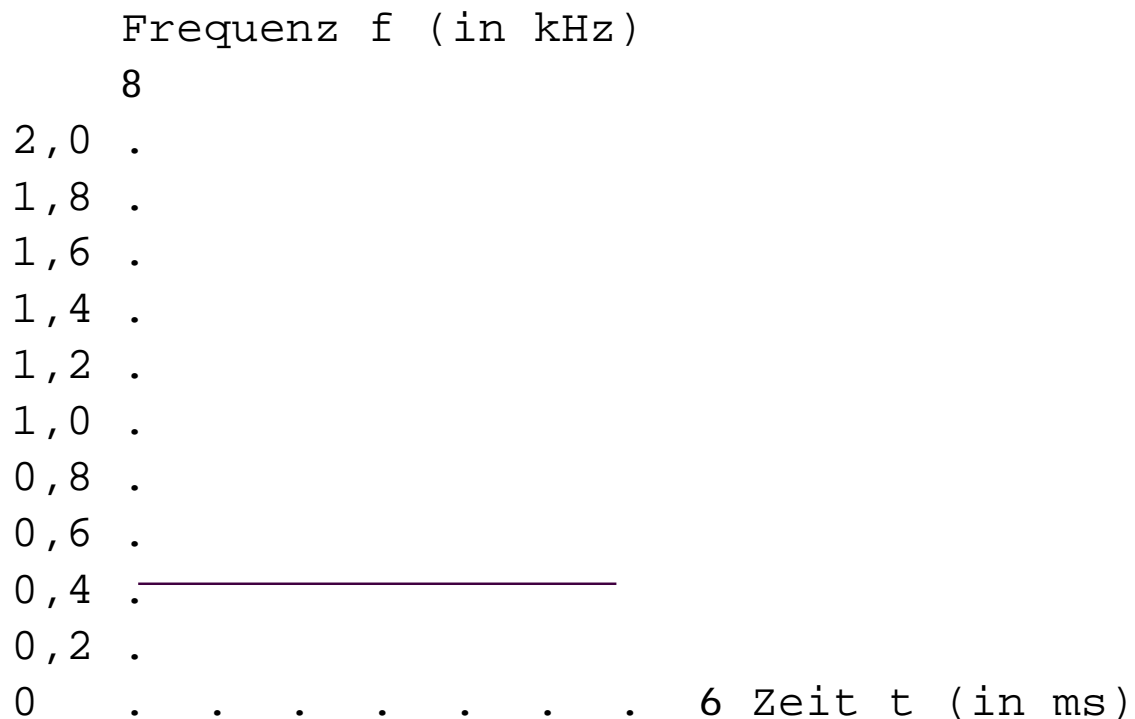
Die akustische Phonetik untersucht das Sprachsignal (weitgehend) unabhängig von der Artikulation und der Perzeption (1.1.3.). Es wird als physikalisches Phänomen betrachtet.

Ein (Sprach)Signal ist ein akustisches Er-

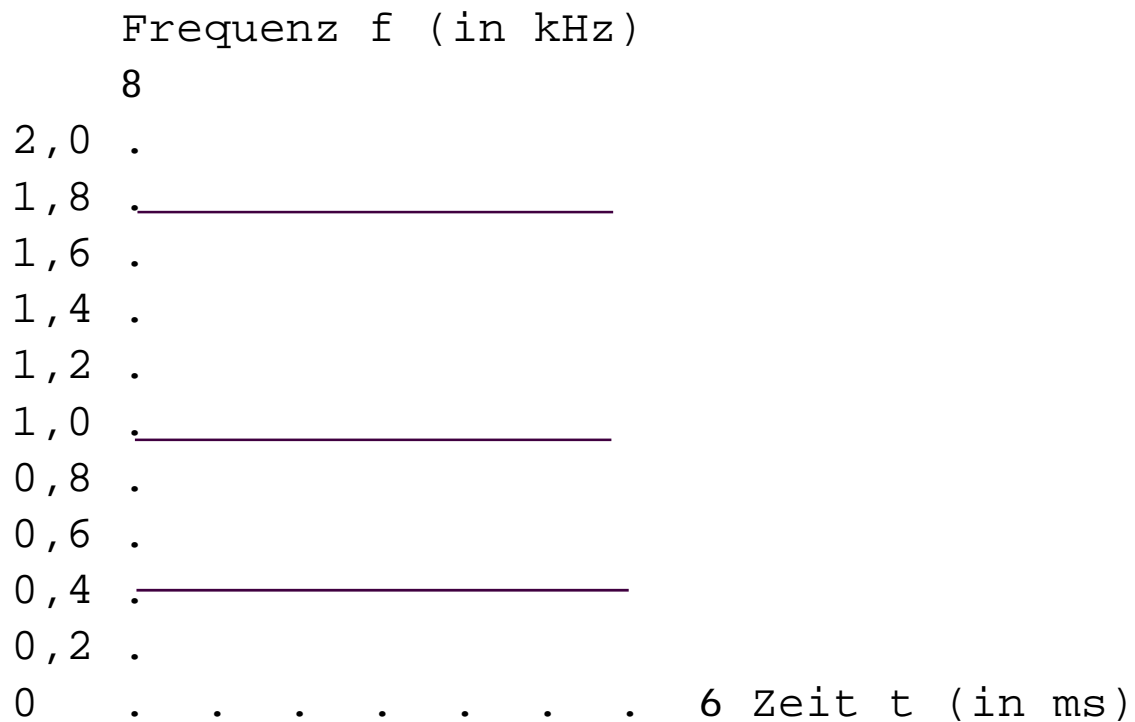
eignis, d.h. eine mit der Zeit t variierende Verteilung von Amplituden (Intensitäten = Energien) A auf die einzelnen Frequenzen f :

$$A = F(f, t)$$

Jede Frequenz f ist in dem Gesamt ereignis zum Zeitpunkt t gerade so stark vertreten, wie der Wert von A ist. Der Kammerton "a" würde sich so darstellen lassen:



Erzeugt ein Instrument diesen Ton, so kommen charakteristische Obertöne hinzu, die gerade seine Klangfarbe ausmachen (außer den Okta-ven sind weitere möglich). Diejenigen Fre- quenzen, die besonders hohe Werte von A haben, heißen die Formanten des Signals.



Bei einer einstimmigen Melodie ist die Funktion F keine Konstante mehr, ein mehrstimmiges Instrumentalstück sieht schon sehr verschwommen aus, die Formanten sind nur schwer zu erkennen (aber natürlich zu hören!).

Man definiert die Formanten eines Sprachsignals als (relativ enge) Frequenzbänder, da sie von Sprecher zu Sprecher(in) selbstverständlich variieren.

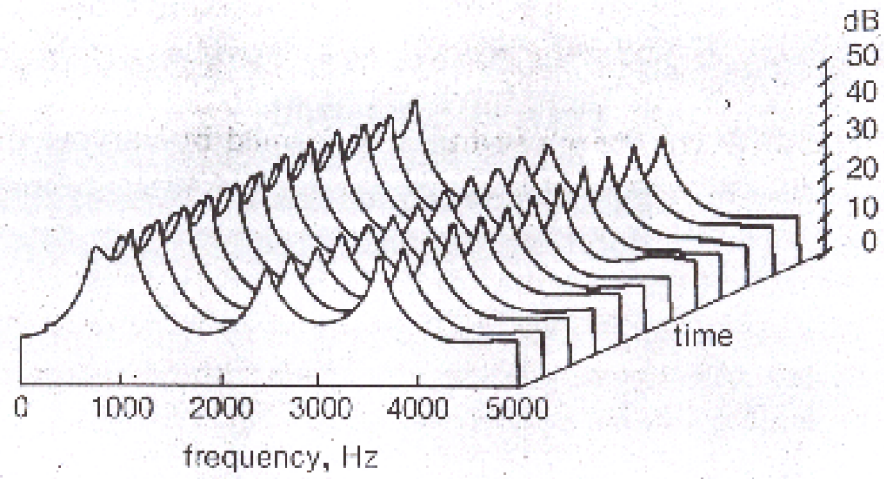


Fig 2 · Spectrum of vowel segment.

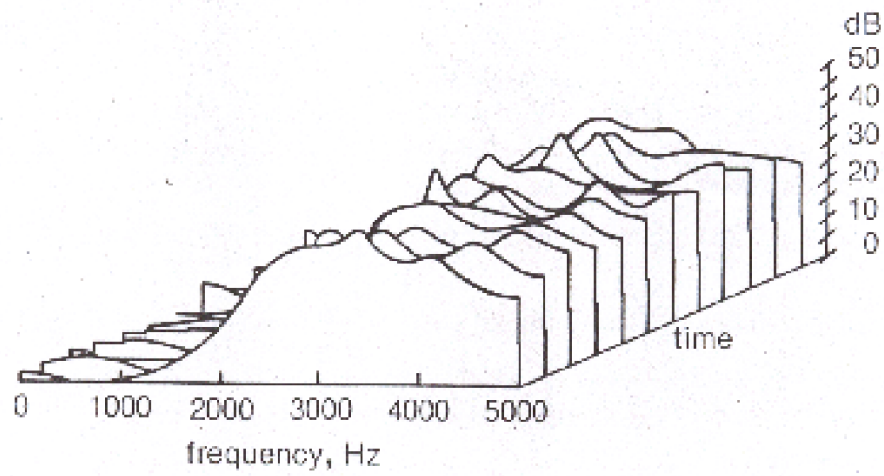
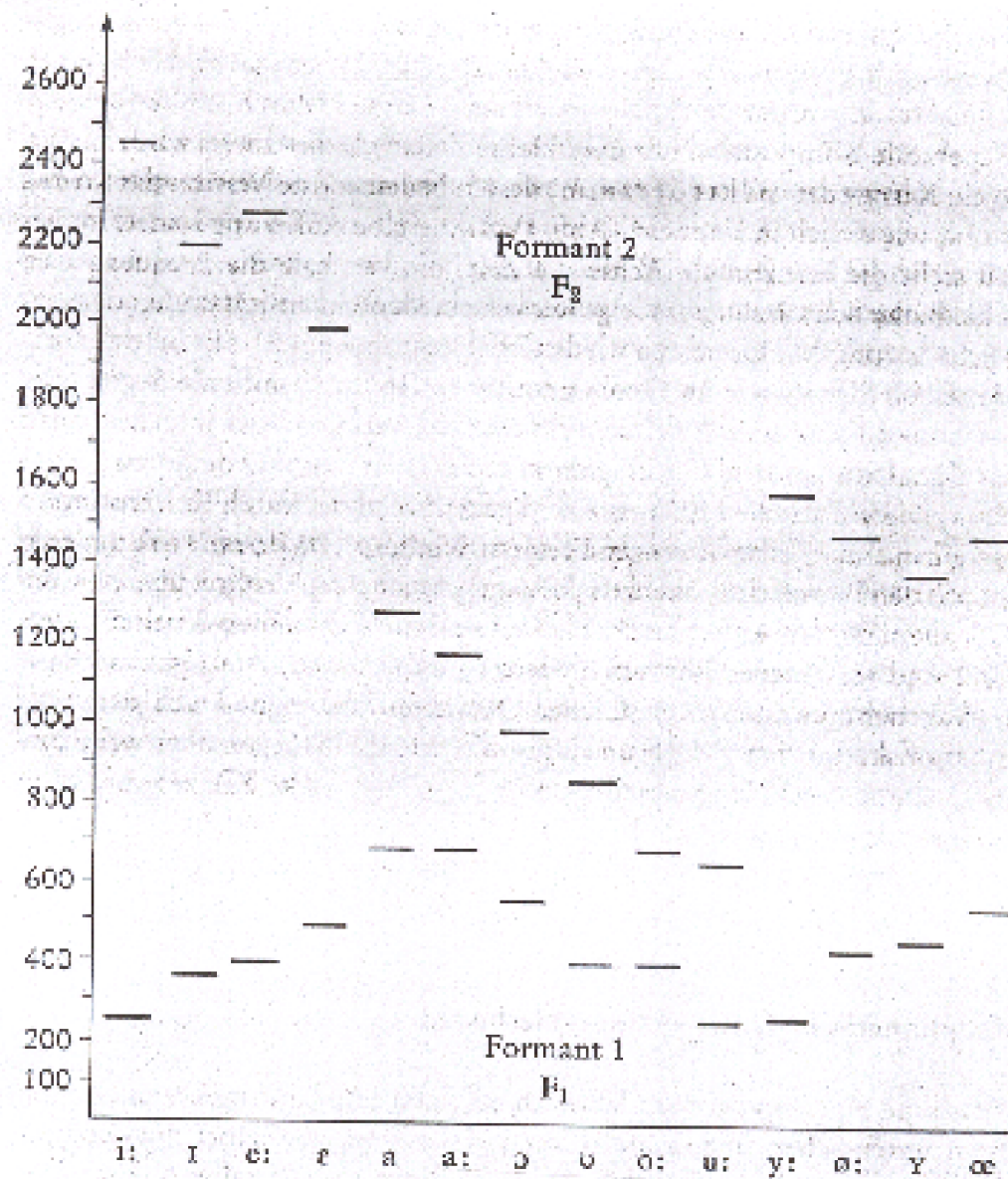


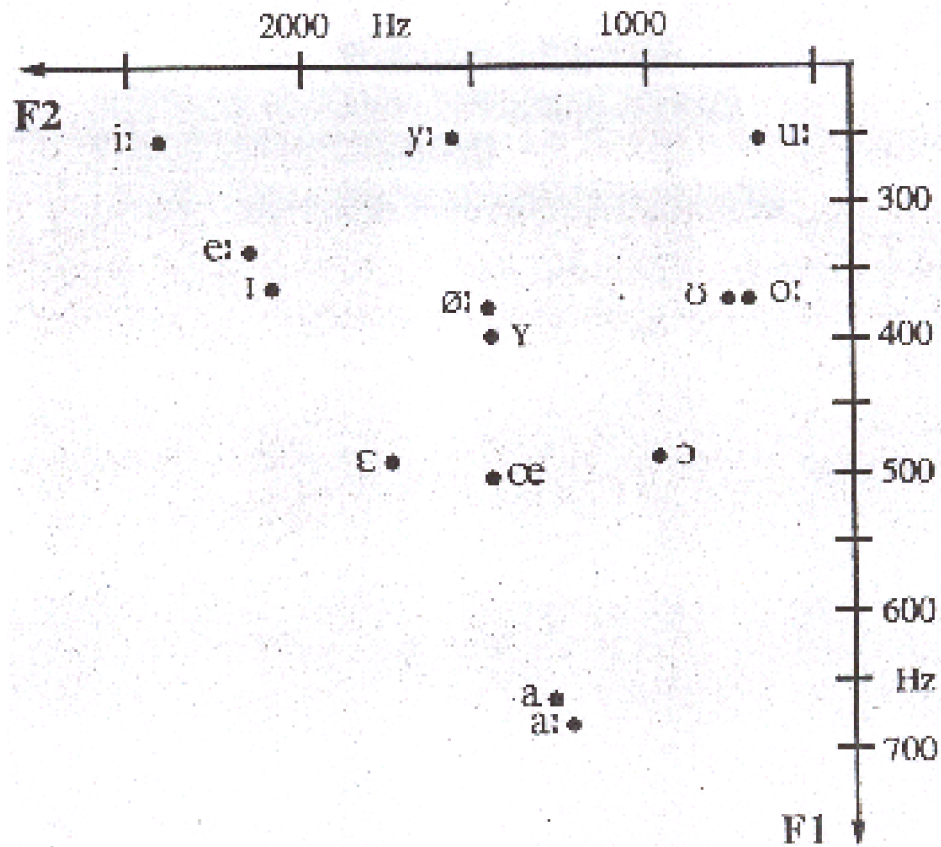
Fig 3 · Spectrum of voiceless sound.



Breitband-Sonagramme der Vokalaufierungen [ur:], [o:], [a:], [e:], [i:].



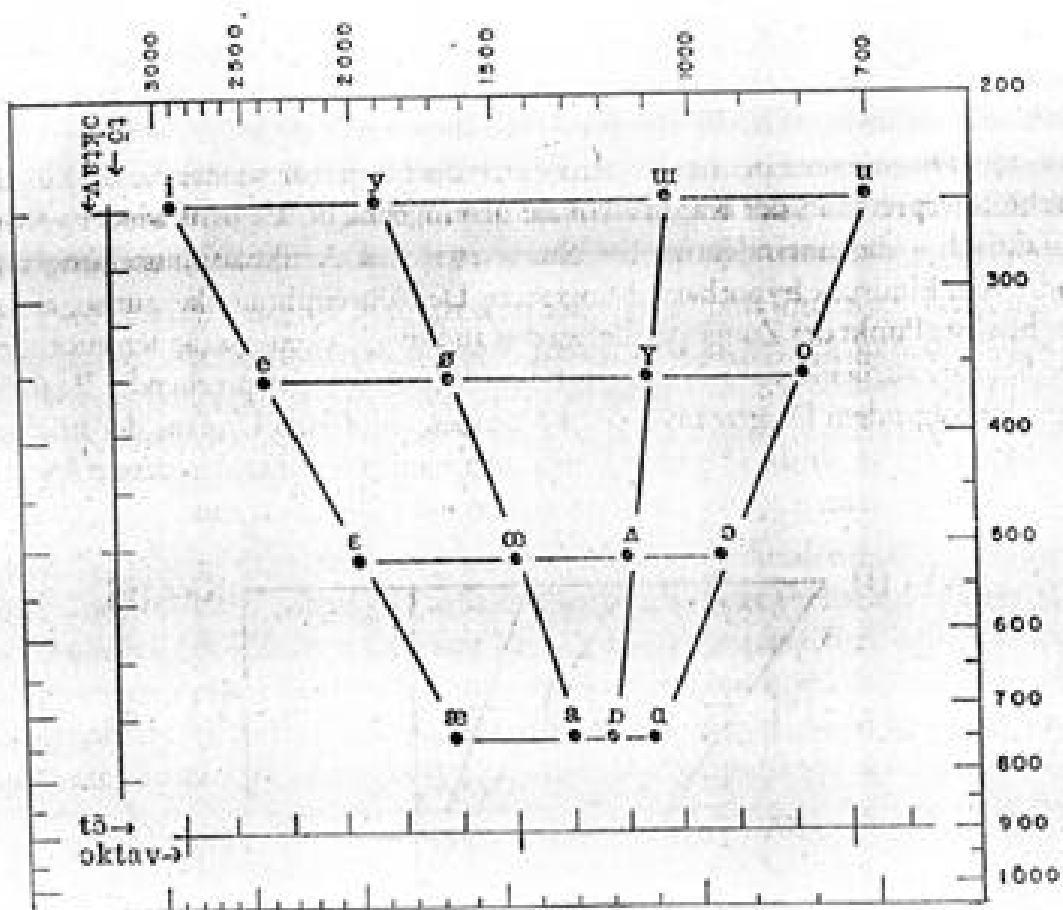
Durchschnittswerte der zwei untersten Formanten deutscher Vokale



Durchschnittliche Formantwerte deutscher Vokale (nach Rausch 1972).

Man beachte auf der letzten Folie die Lage von [i:] und [e:] zu [ɪ] und [ɛ]!

Ein Zusammenhang zwischen der artikulatorischen und der akustischen Seite der Vokale ergibt sich dadurch, daß die Darstellung mit zwei Formanten das Vokalviereck reproduziert (es handelt sich um synthetische Vokale!):



Breitband-Sonagramme erfassen einen größeren Frequenzbereich (hier: 0 bis 3 ... 10 kHz) als Schmalband-Sonagramme (hier 0 bis 1,5 ... 2,5 kHz).

Oszillogramme dagegen geben die Schwingungen als zeitabhängige Veränderung des Schalldrucks wieder:

