

Die akustische Analyse von
Sprachlauten.

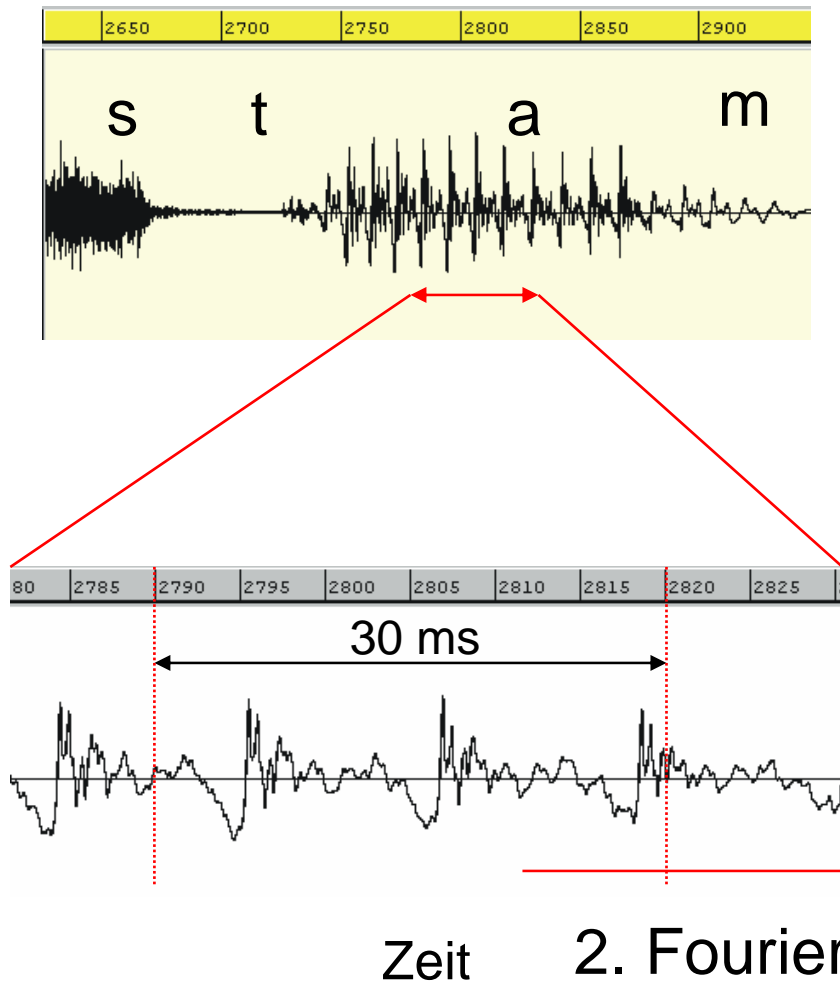
Die Interpretation von Spektrogrammen.

Jonathan Harrington

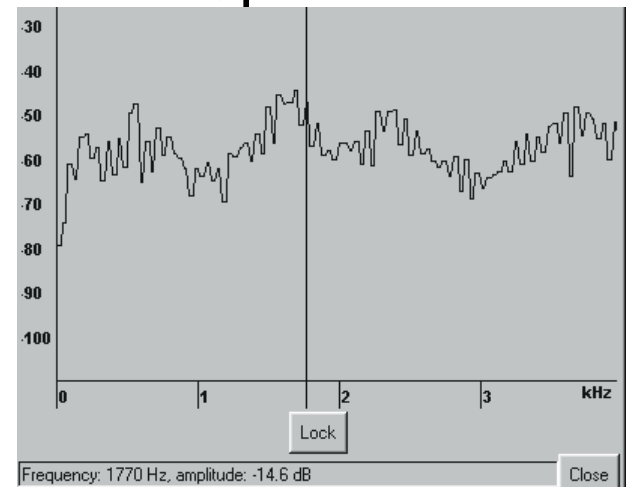
IPDS, Kiel.

Vom Zeitsignal zum Spektrum

1. Ein Teil vom Sprachsignal aussuchen: die **Zeitauflösung** (hier 30 ms) ist die Dauer des Signals die Fourier-analysiert wird



Spektrum



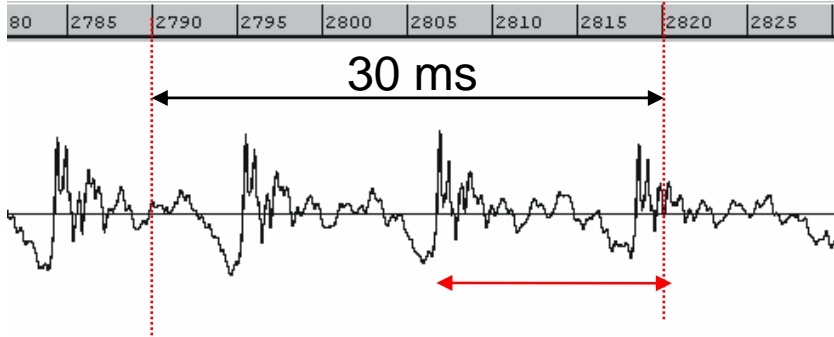
Amplitude

Frequenz

2. Fourier Analyse

Periodendauer, Grundfrequenz, Harmonischen

Zeitsignal

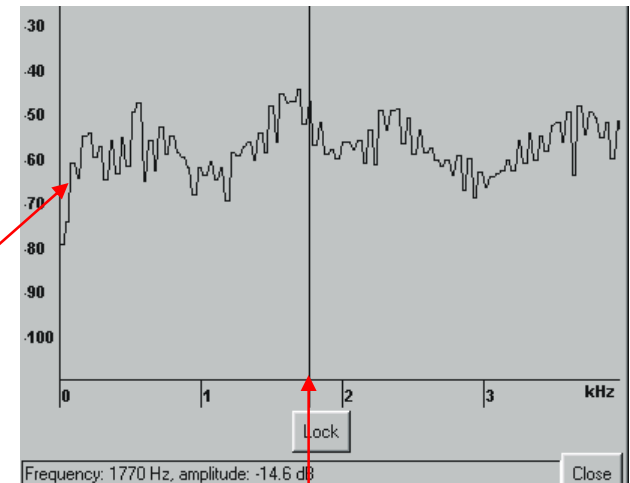


Ca. 2.5 Perioden in 30 ms

Periodendauer $\approx 30/2.5$ ms
 ≈ 12 ms = 0.012 s

Grundfrequenz $\approx 1/0.012$ Hz
 ≈ 85 Hz

Spektrum

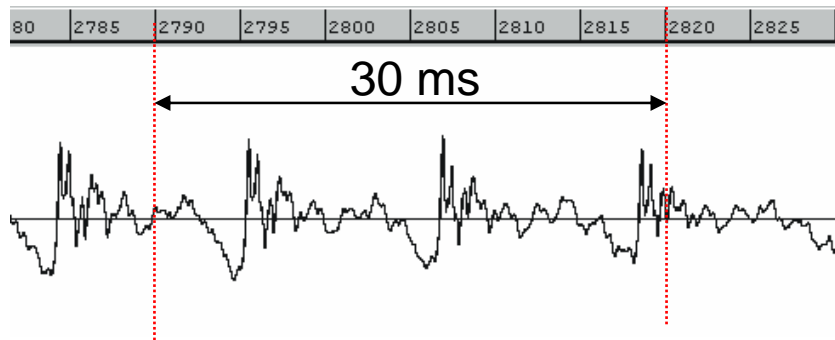


21^e Harmonische \approx
 $85 \times 21 = 1770$ Hz

Frequenz- und Zeitauflösung

Die **Frequenzauflösung** (Hz) ist ca. im umgekehrten Verhältnis zur Dauer des Zeitfensters (s) oder zur **Zeitauflösung**.

Das bedeutet: Je breiter die Zeitauflösung, umso schmaler die Frequenzauflösung und umgekehrt



$$\text{Zeitauflösung} = 30 \text{ ms} = 30/1000 \text{ s}$$

$$\text{Frequenzauflösung} = 1000/30 \text{ Hz}$$

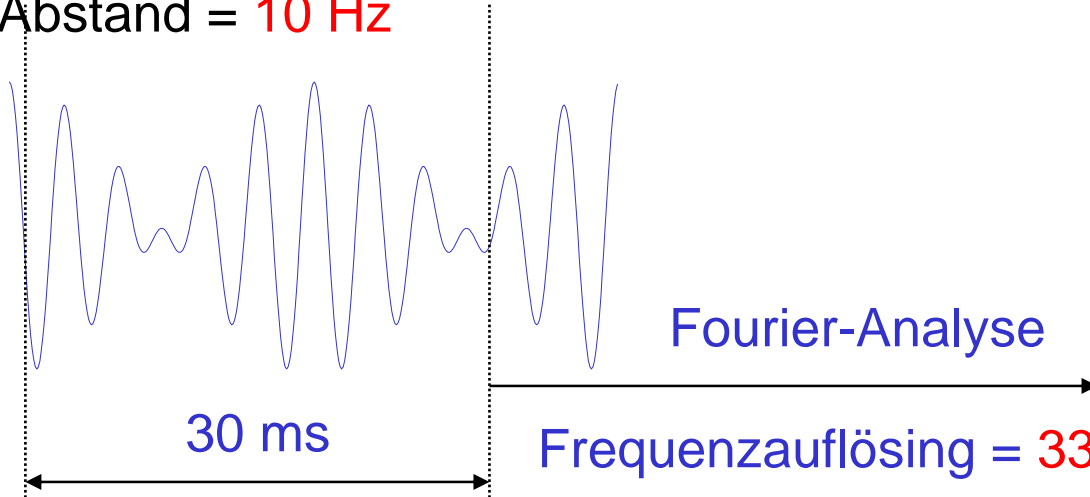
$$= 33.3 \text{ Hz}$$

Im Spektrum werden Frequenzen nur differenziert, **wenn deren Abstand größer als die Frequenzauflösung ist:**

Zeitsignal

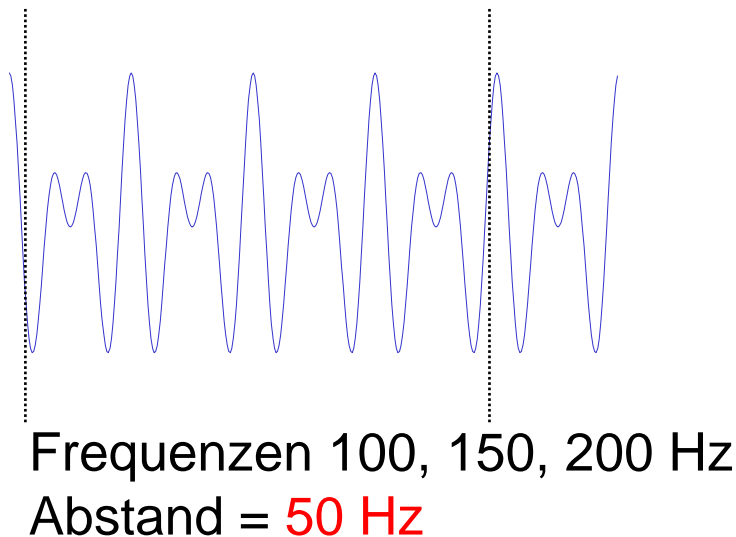
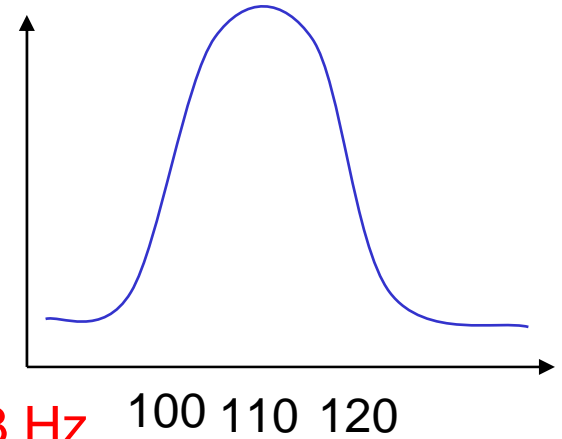
Frequenzen 100, 110, 120 Hz

Abstand = **10 Hz**



Spektrum

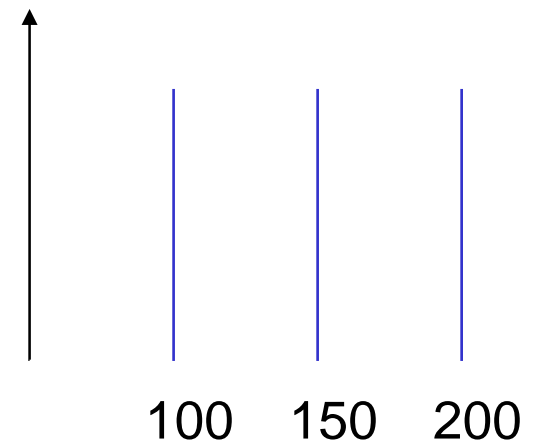
Undifferenzierte Frequenzen



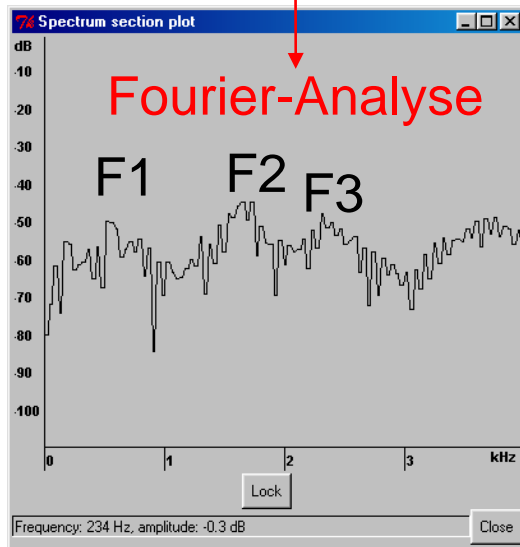
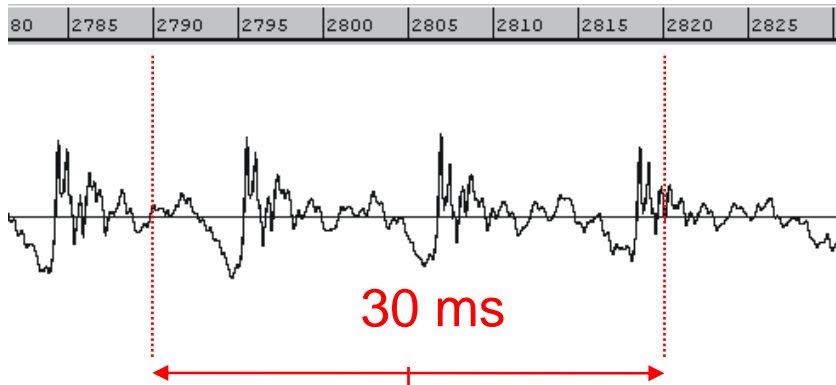
Frequenzen 100, 150, 200 Hz

Abstand = **50 Hz**

Differenzierte Frequenzen

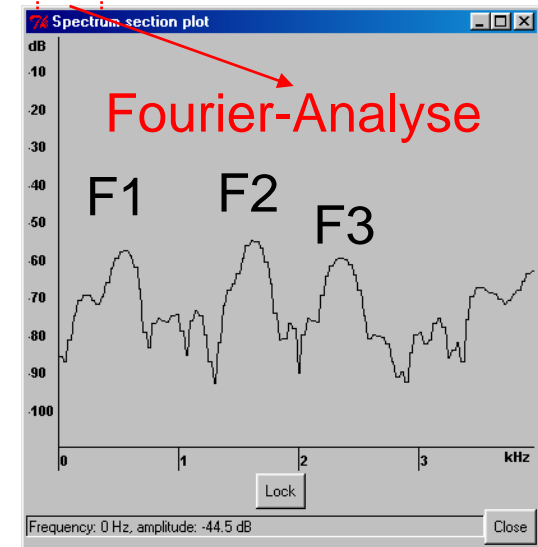
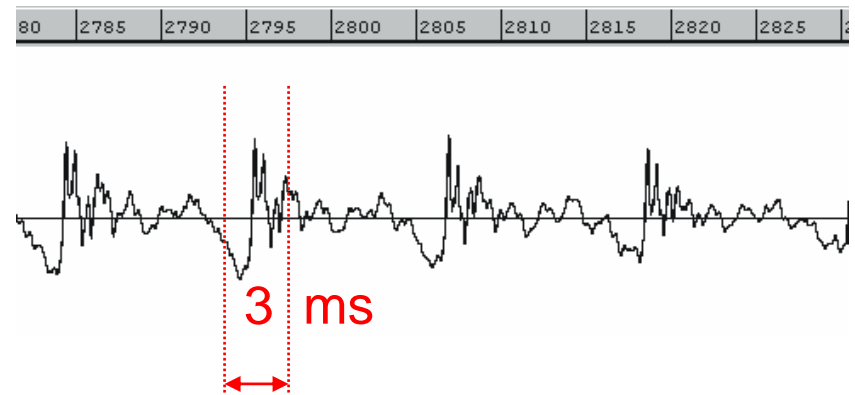


Breite Zeitauflösung: 30 ms
schmale Frequenzauflösung: 33.3 Hz
(Schmalbandanalyse)



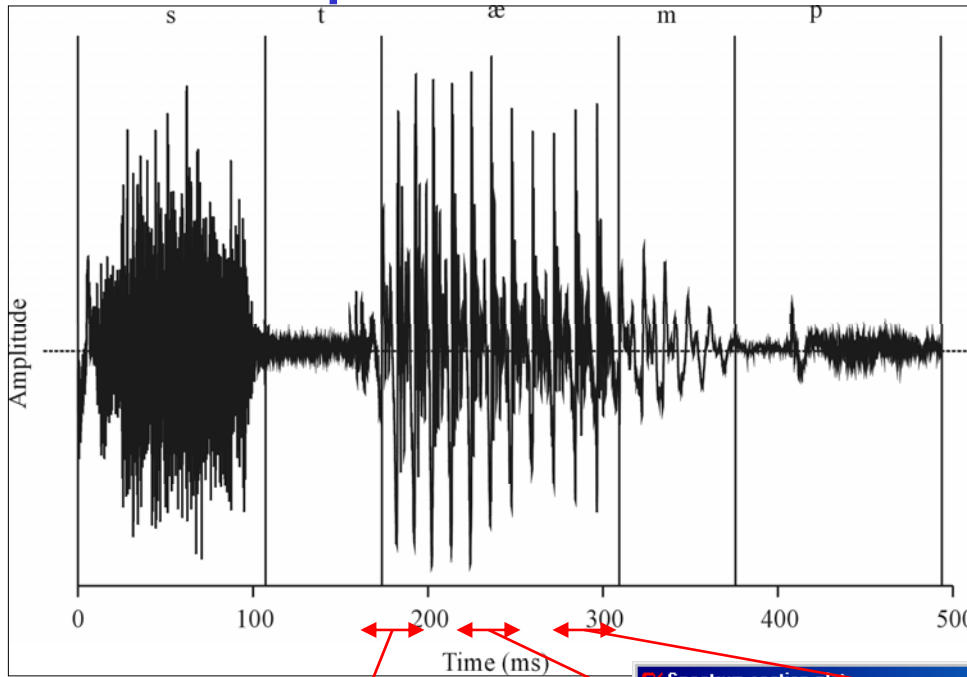
Hier sind die Harmonischen sichtbar, da die Grundfrequenz und daher der Abstand zwischen Harmonischen \approx ca. 85 Hz

Schmale Zeitauflösung: 3 ms
Breite Frequenzauflösung: 333 Hz
(Breitbandanalyse)

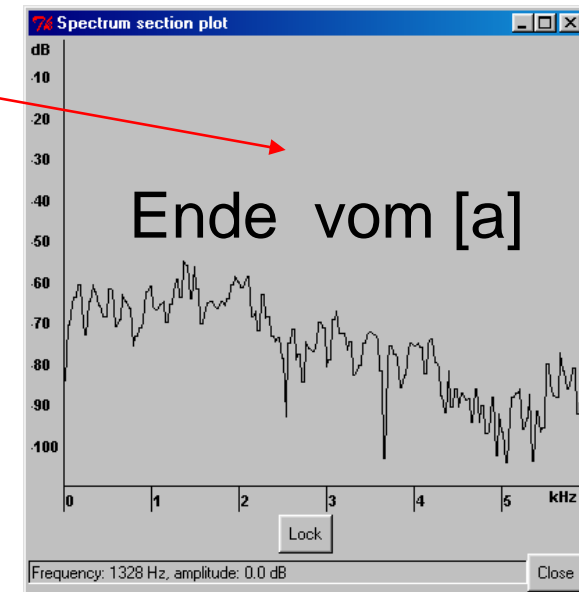


Die Harmonischen sind nicht mehr sichtbar, und die Formanten treten dadurch deutlicher hervor

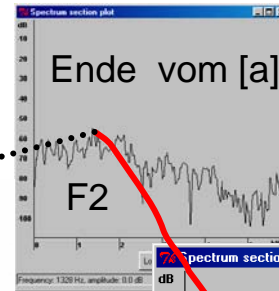
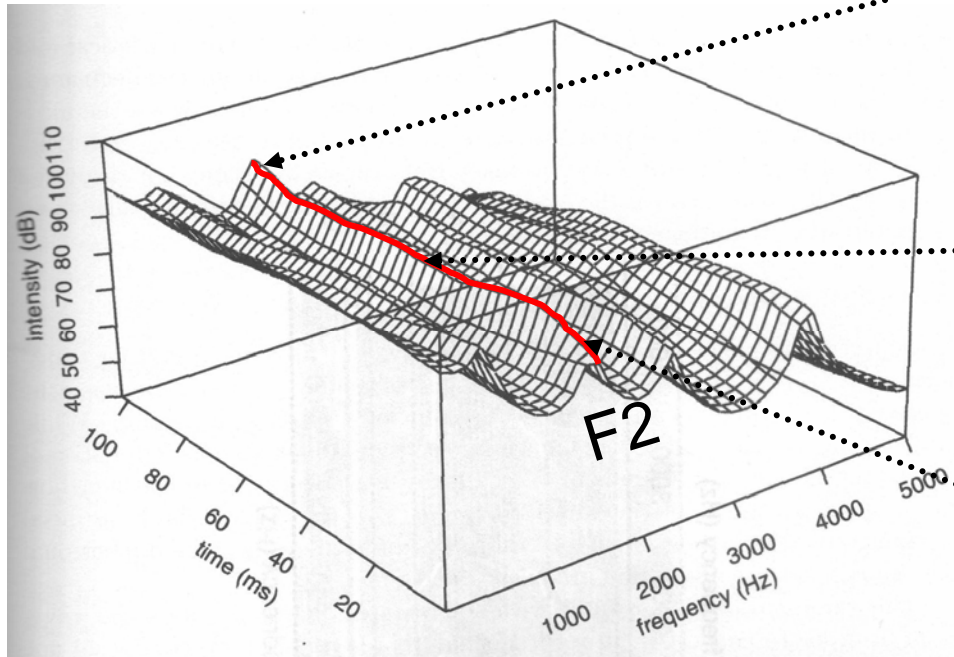
Spektrale Änderungen in der Zeit



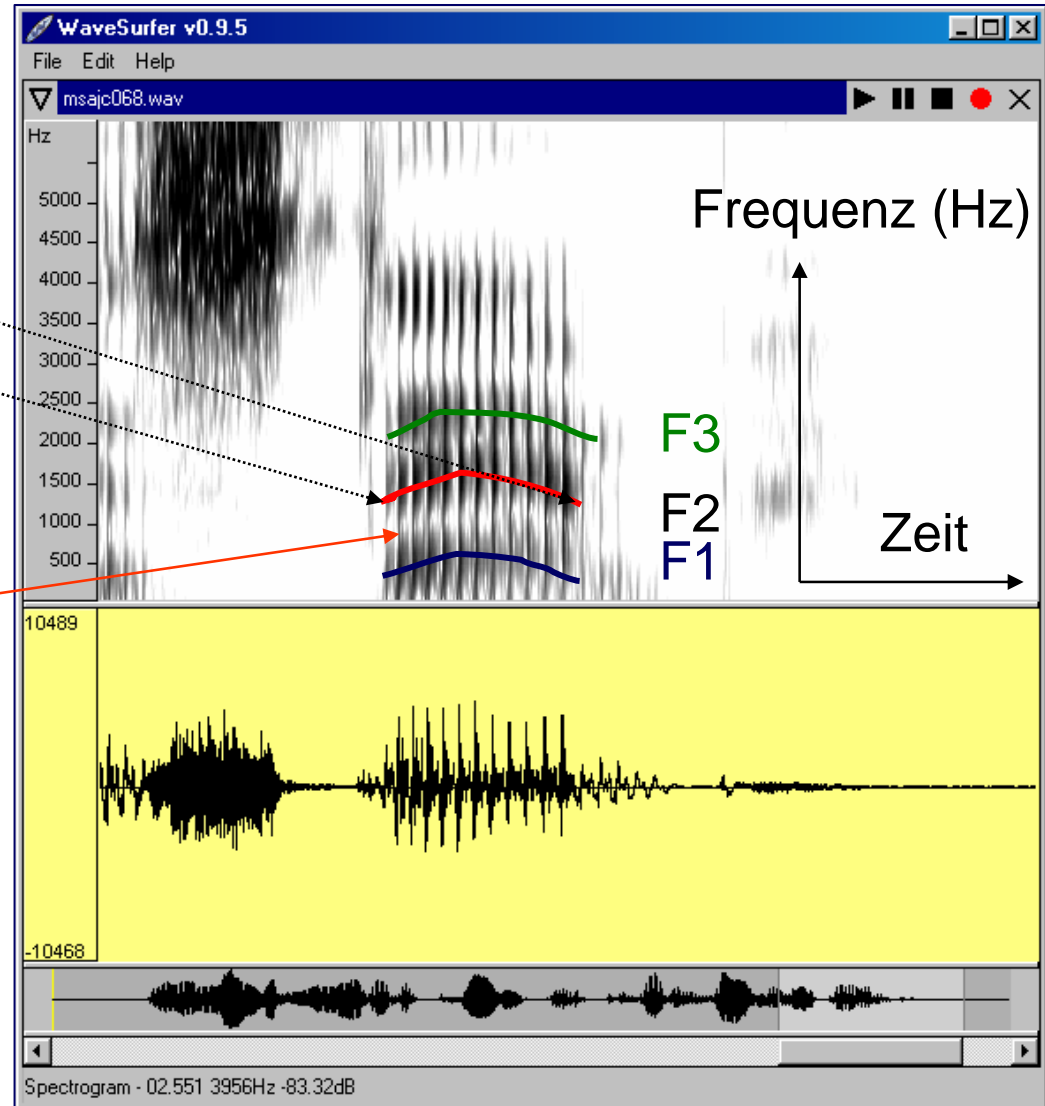
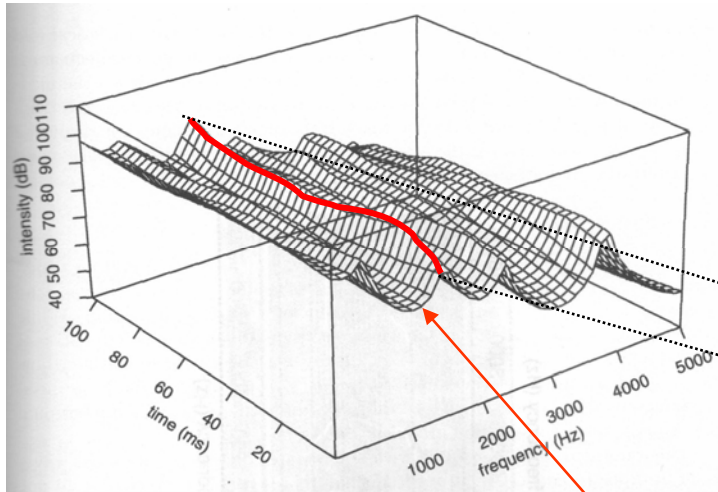
Wenn wir feststellen wollen, wie sich das Spektrum mit der Zeit ändert, müssen Fourier-Analysen zu **sämtlichen** Zeitpunkten berechnet werden



Wasserfall Abbildung



Wasserfall-Abbildung zum Spektrogramm



Tal

In einem Spektrogramm wird die Amplitude in **Dunkelheit** umgesetzt – je dunkler umso höher die Amplitude

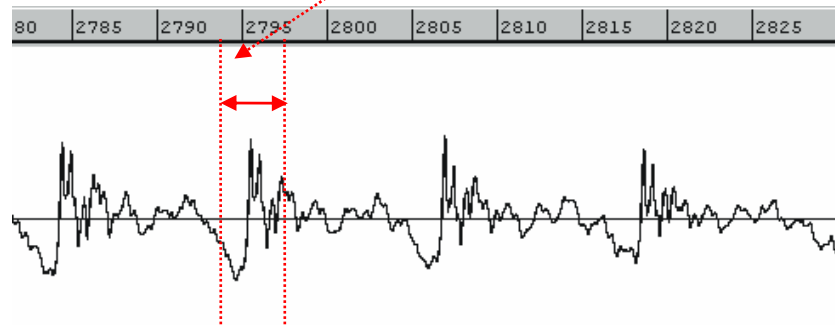
In einem Breitbandspektrogramm ist:

Die Frequenzauflösung
mindestens **300 Hz**

Die Zeitauflösung ist daher
 $1/300 \approx$ **3 ms**

Für dieses
Signal:
 $f_0 \approx$ ca. **85 Hz**

Periodendauer, T ,
 $\approx 1/85 \text{ s} \approx$ **12 ms**

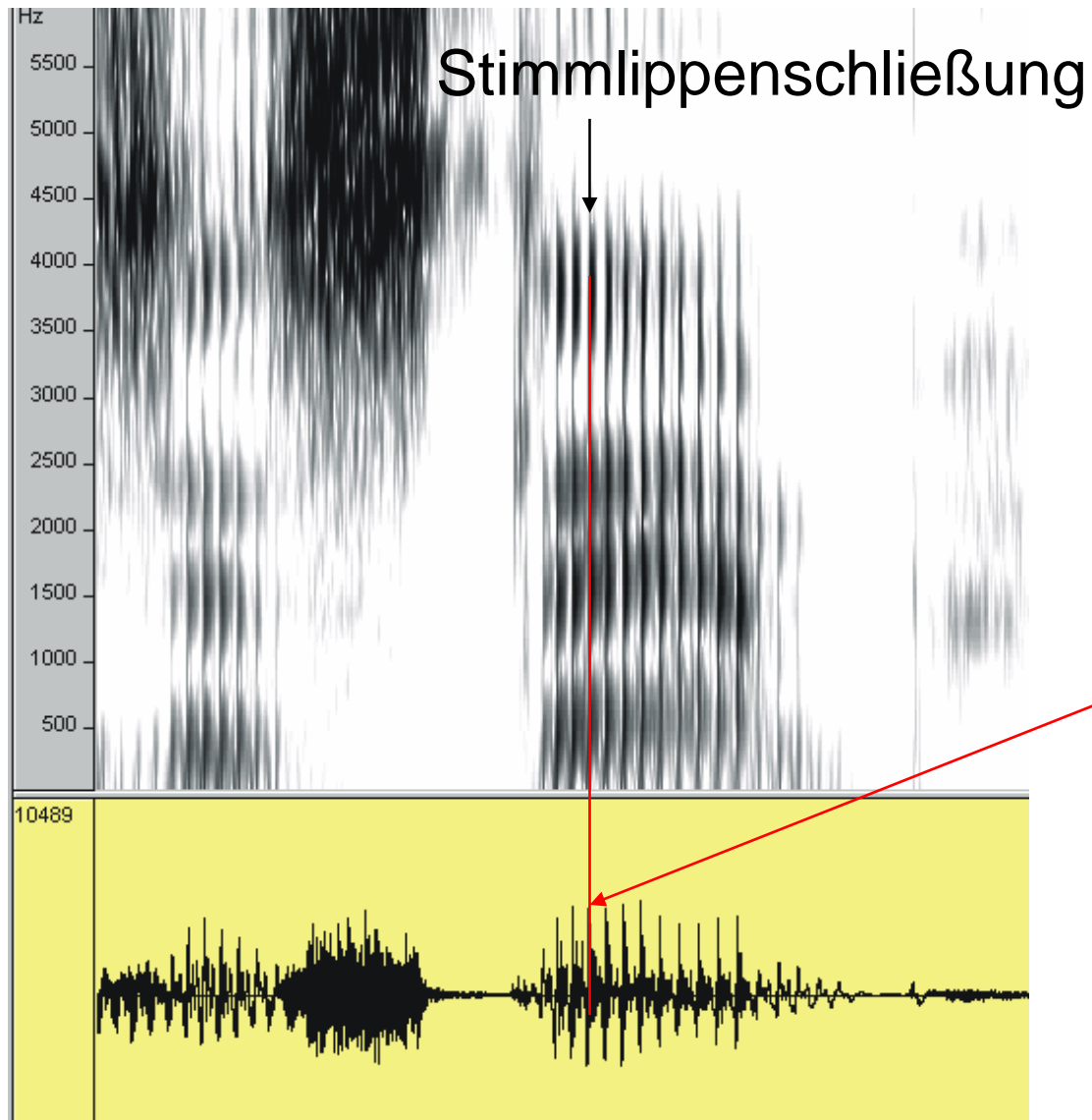


Die Harmonischen:
nicht sichtbar
(da $f_0 <$ **300 Hz**)

Die
Stimmlippenschließung:
sichtbar
(Da $T >$ **3 ms**)

Breitbandspektrogramm

Frequenzauflösung = 333.3 Hz, Zeitauflösung = 3 ms



Jede Schließung erzeugt eine gerade Linie im Breitband-Spektrogramm und ist mit einem Höhepunkt im Zeitsignal synchronisiert

Schmalbandspektrogramm

Frequenzauflösung ist meistens 45 Hz

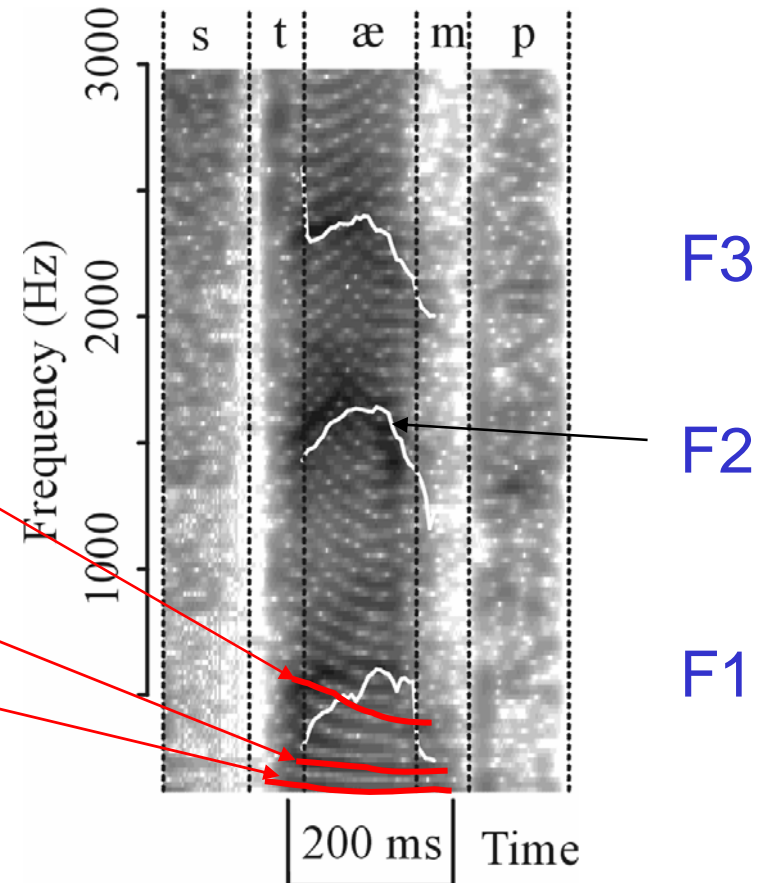
Harmonische sind sichtbar,
da $f_0 \approx 85$ Hz

Zeitauflösung = $1/45 \approx 22$ ms
Stimmlippenschließungen
sind **nicht** sichtbar, da die
Periodendauer ≈ 12 ms

5^e Harmonische

2^e Harmonische

Grundfrequenz



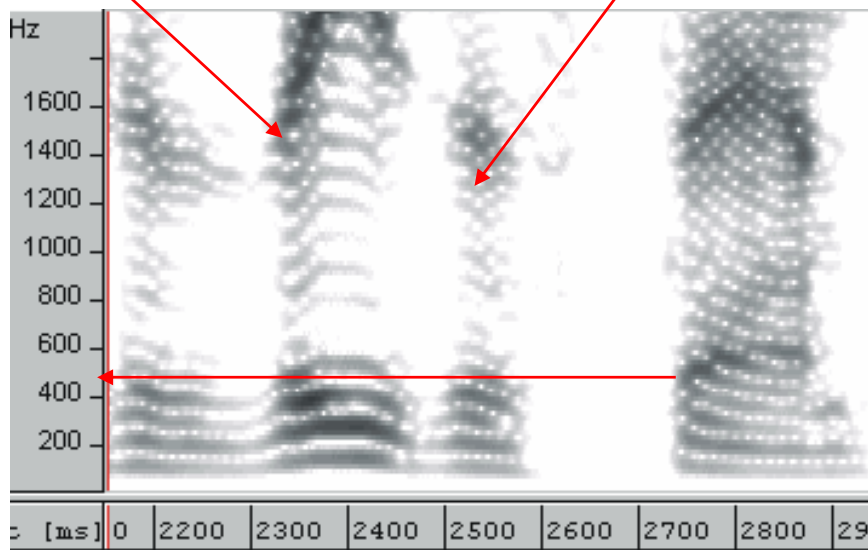
Schmalbandspektrogramm und die Tonhöhe

Mit einem SB-Spektrogramm kann man oft an Hand der höheren Harmonischen

- Die Änderungen der Tonhöhe sehen
- Die Grundfrequenz einschätzen

Tonhöhe steigt

fällt



5^e Harmonische \approx 450 Hz, daher $f_0 \approx$ 90 Hz

Zusammenfassung

Ein Spektrum wird durch die Anwendung einer Fourier-Analyse auf einen Teil oder **ein Fenster** vom einem Zeitsignal angewendet

In einem Spektrogramm werden Spektre in regelmäßigen Zeitabständen berechnet: die Amplitude entspricht der Dunkelheit

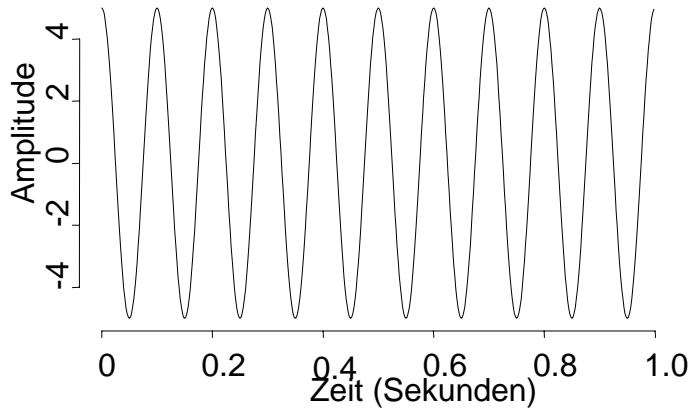
Die Frequenzauflösung ist im umgekehrten Verhältnis zur Dauer oder Zeitauflösung dieses Fensters: je schmaler das Fenster, umso grober (breiter) die Frequenzauflösung

Zusammenfassung (fortgesetzt)

Um Harmonischen im Spektrum/Spektrogramm zu sehen, muss **die Grundfrequenz größer als die Frequenzauflösung** sein. Daher ist ein Schmalbandspektrogramm (Frequenzauflösung typischerweise 45 Hz) für die Analyse von Harmonischen geeignet

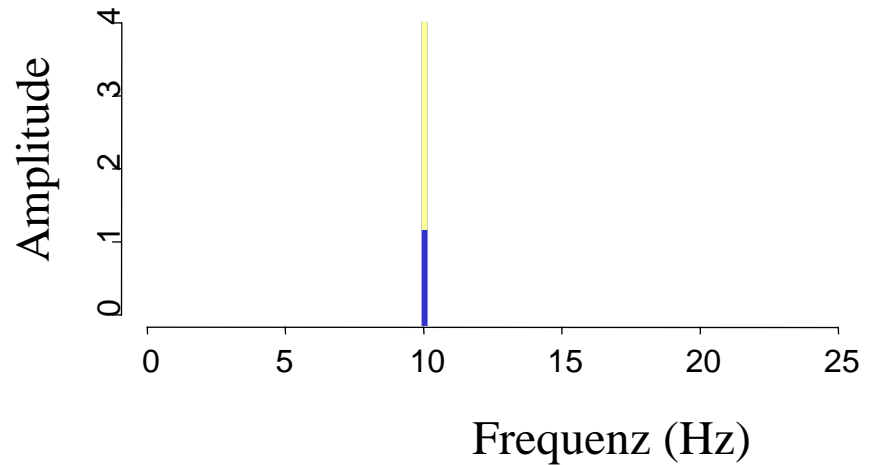
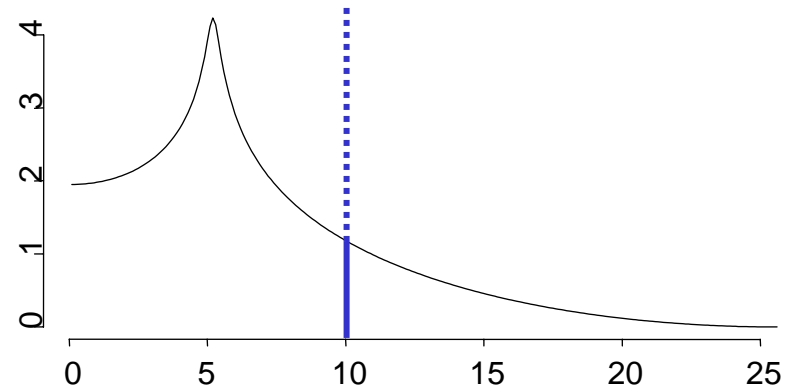
Um Stimmlippen-Schließungen im Spektrogramm zu sehen, muss die **Periodendauer größer als die Zeitauflösung** sein. Dafür ist ein Breitbandspektrogramm (Zeitaufklärung ca. 3 ms, Frequenzauflösung über 300 Hz) geeignet.

7. Der Sinusoid links wird mit der Resonanzkurve rechts gefiltert. Erzeugen Sie eine Abbildung des Spektrums der Ausgabe.

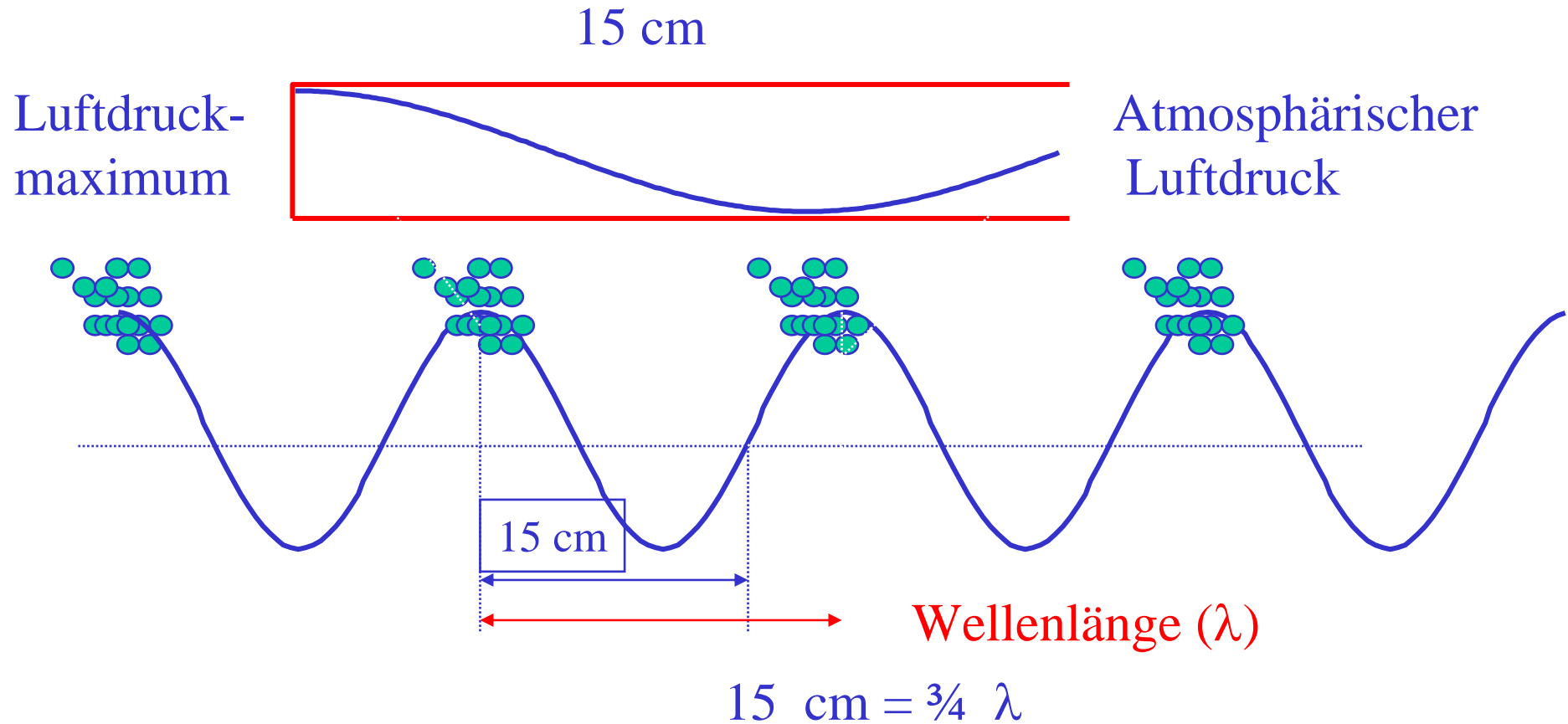


10 Perioden in einer Sekunde.

Frequenz = 10 Hz



8. Die typische Länge von einem Vokaltrakt einer Frau ist 15 cm. Berechnen Sie den zweiten Resonanz eines [ə] Vokals an Hand eines Rohres von der selben Länge.



Resonanz für eine Wellenlänge von:

$$\lambda = (4 \times 15) / 3 \text{ cm} = 20 \text{ cm}$$

Und für eine Frequenz von $f = c / \lambda = 35000 / 20 = 1750 \text{ Hz}$

9. Beim Tief-Tauchen besteht die Luft oft aus einer Mischung von Sauerstoff und Helium, wodurch die Schallgeschwindigkeit erheblich höher wird. Welcher Einfluss übt dies auf die Resonanzen vom Vokaltrakt des Tauchers aus?

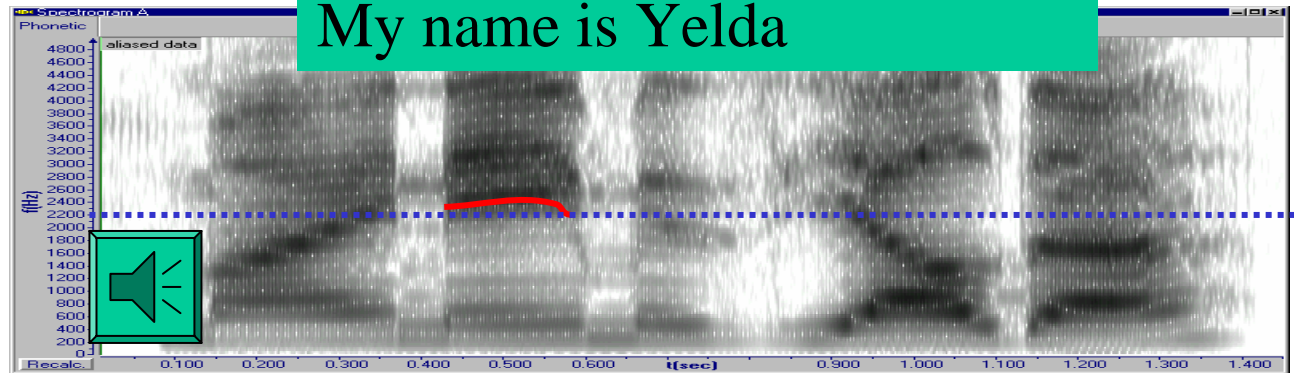
Frequenz = Schallgeschwindigkeit / Wellenlänge

$$f = c / \lambda$$

Je grösser c , umso größer c / λ , also umso höher die Frequenz

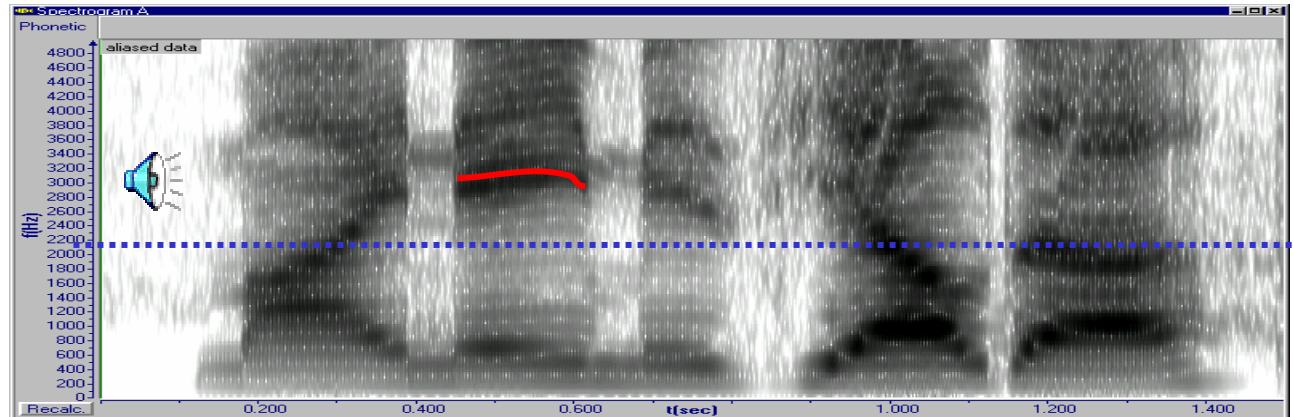
My name is Yelda

Normal



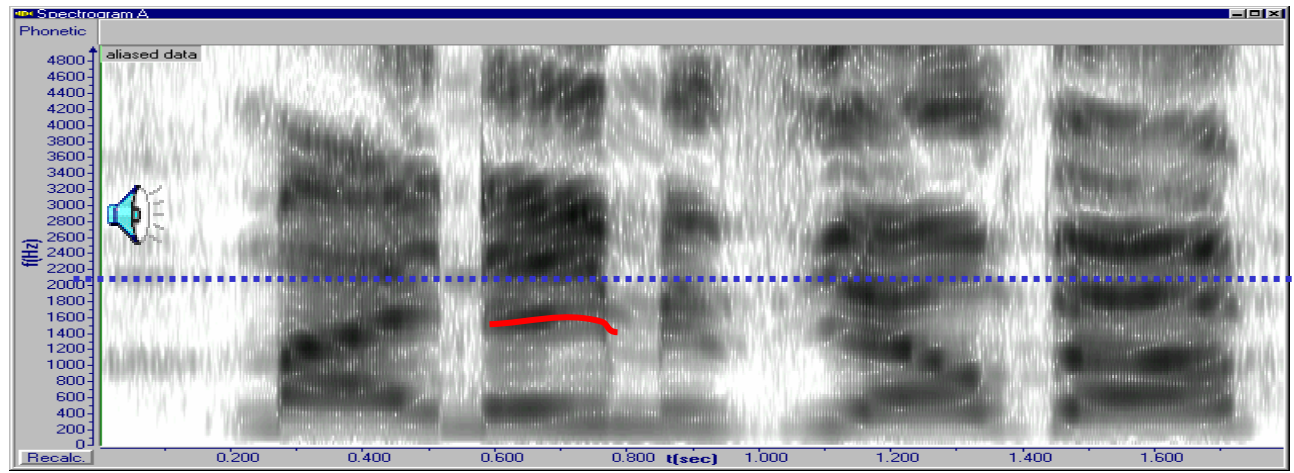
2 kHz

Einatmung
von Heliox
(*c* ist höher)



2 kHz

Einatmung
von Schwefel-
Hexafluorid (*c*
ist niedriger)



2 kHz

Beispiele: John Ohala, Phonetics Laboratory San Francisco

10. Wenn Vokale gerundet werden, wird der Vokaltrakt länger.
Welcher Einfluss übt dies auf die Resonanzen aus?

Je länger der Vokaltrakt, umso grösser wird λ
(die Wellenlänge)

f (die Frequenz der Resonanz) wird *kleiner* wenn λ
grösser wird, da:

$$f = c / \lambda$$

Daher führt die Lippenrundung zu einer *Senkung* der
Formantfrequenzen

11. Berechnen Sie den zweiten Formant für die Frau in 8. bei einer Grundfrequenz von (i) 200 Hz
(ii) 280 Hz.

Welches Vielfach von 200 Hz liegt 1750 Hz am nächsten?

Antwort = die 9^e Harmonische = 1800 Hz = F2

Für $F_0 = 280$ Hz,

Antwort = die 6^e Harmonische = 1680 Hz

Bitte Fragen 12-28 (Seiten 12-16) zum nächsten Mal beantworten.