



Modulation Theory

The Foundation Layer
of Communication

Dafydd Gibbon

Bielefeld University, Germany

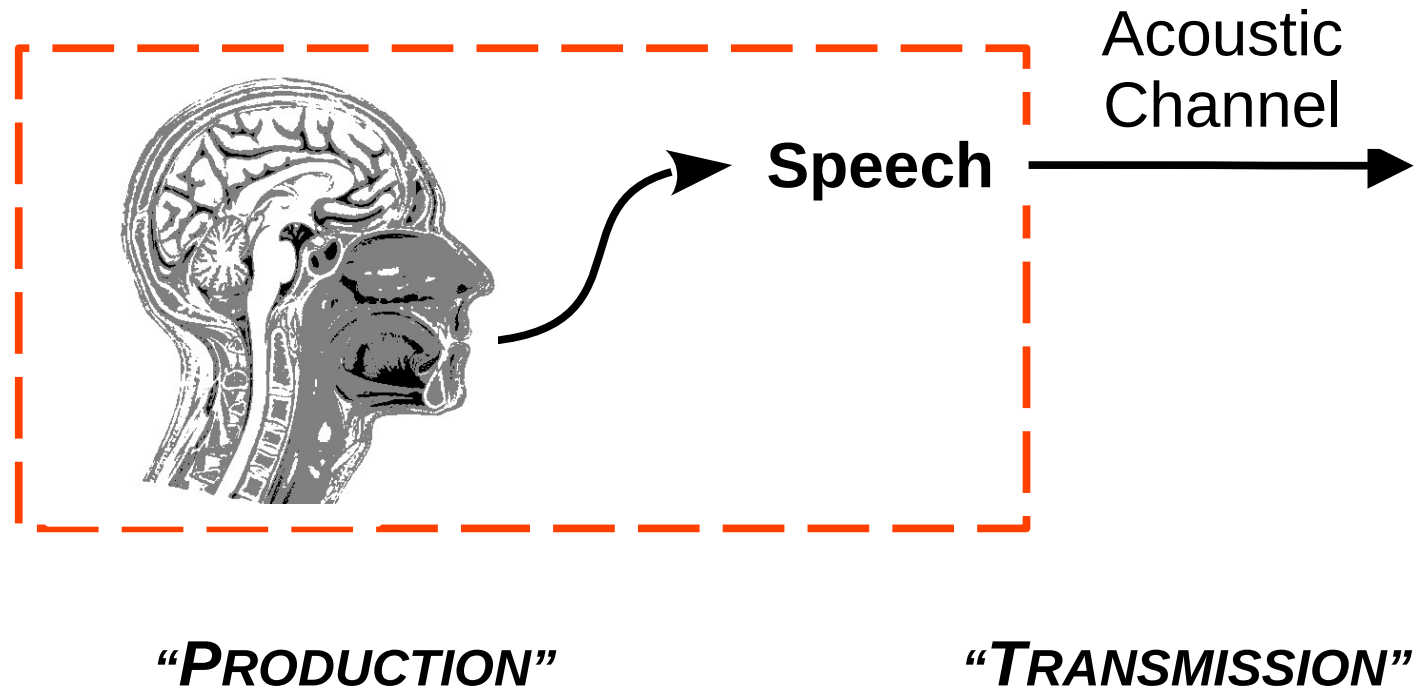
Summer School “Contemporary Phonetics and Phonology”

Xi'an, 2026

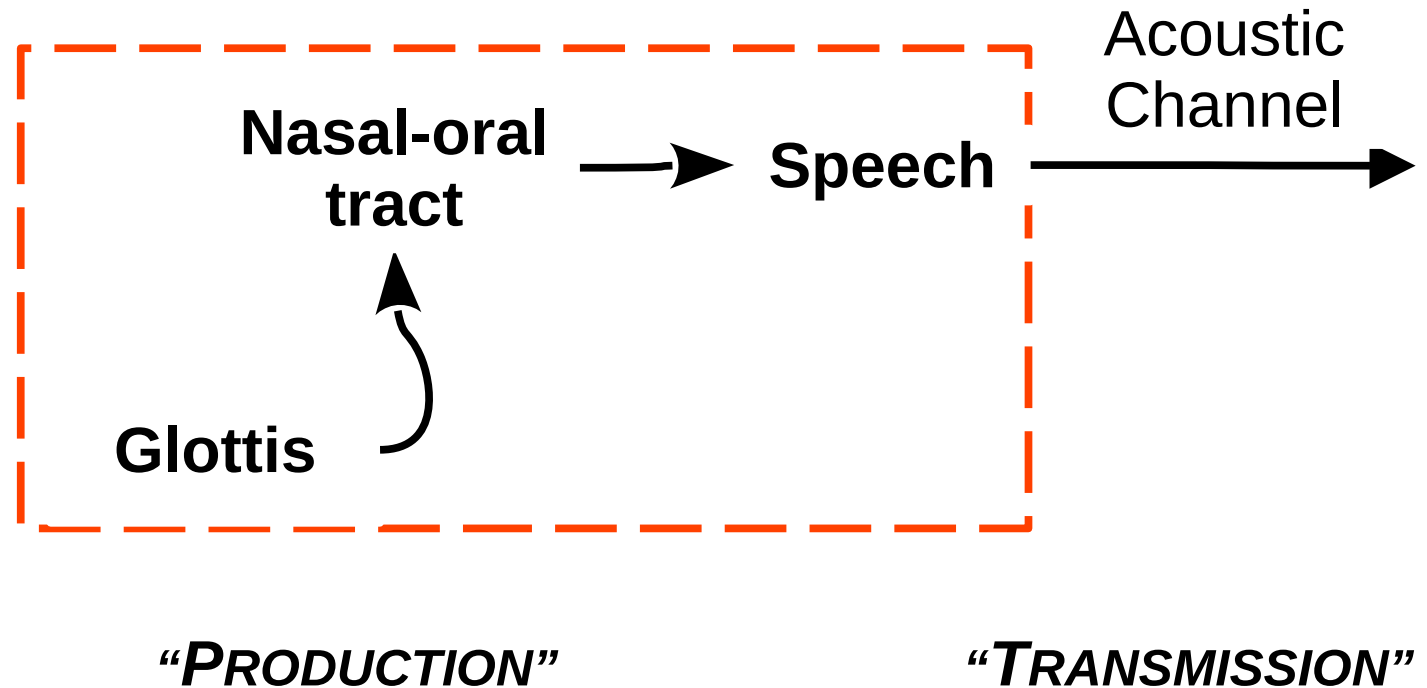


Models of Speech

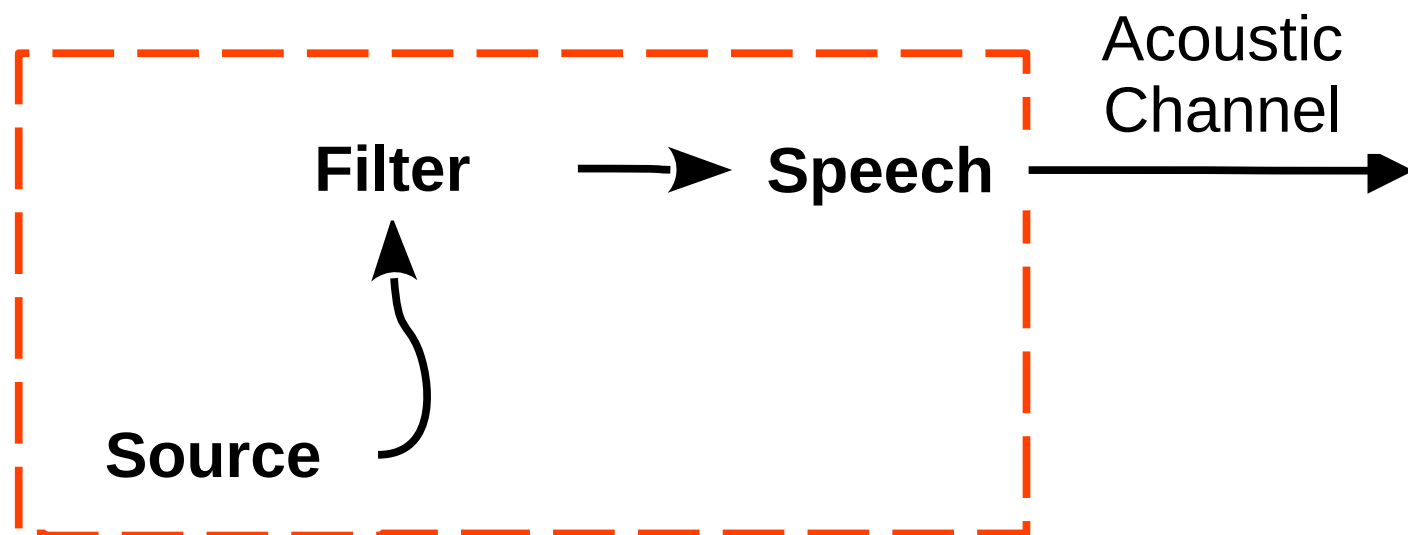
Modelling speech – Model 1



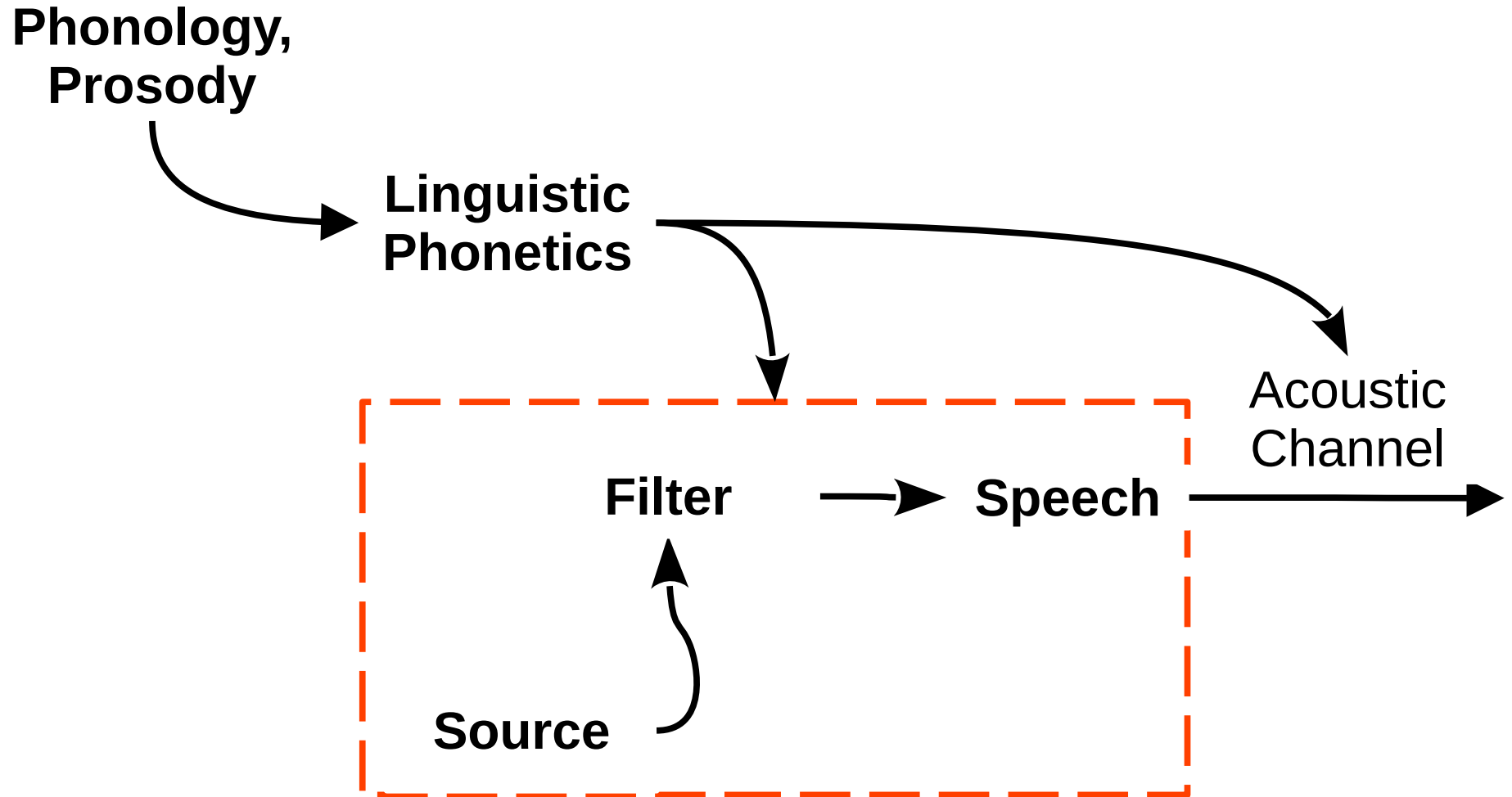
Modelling speech – Model 2



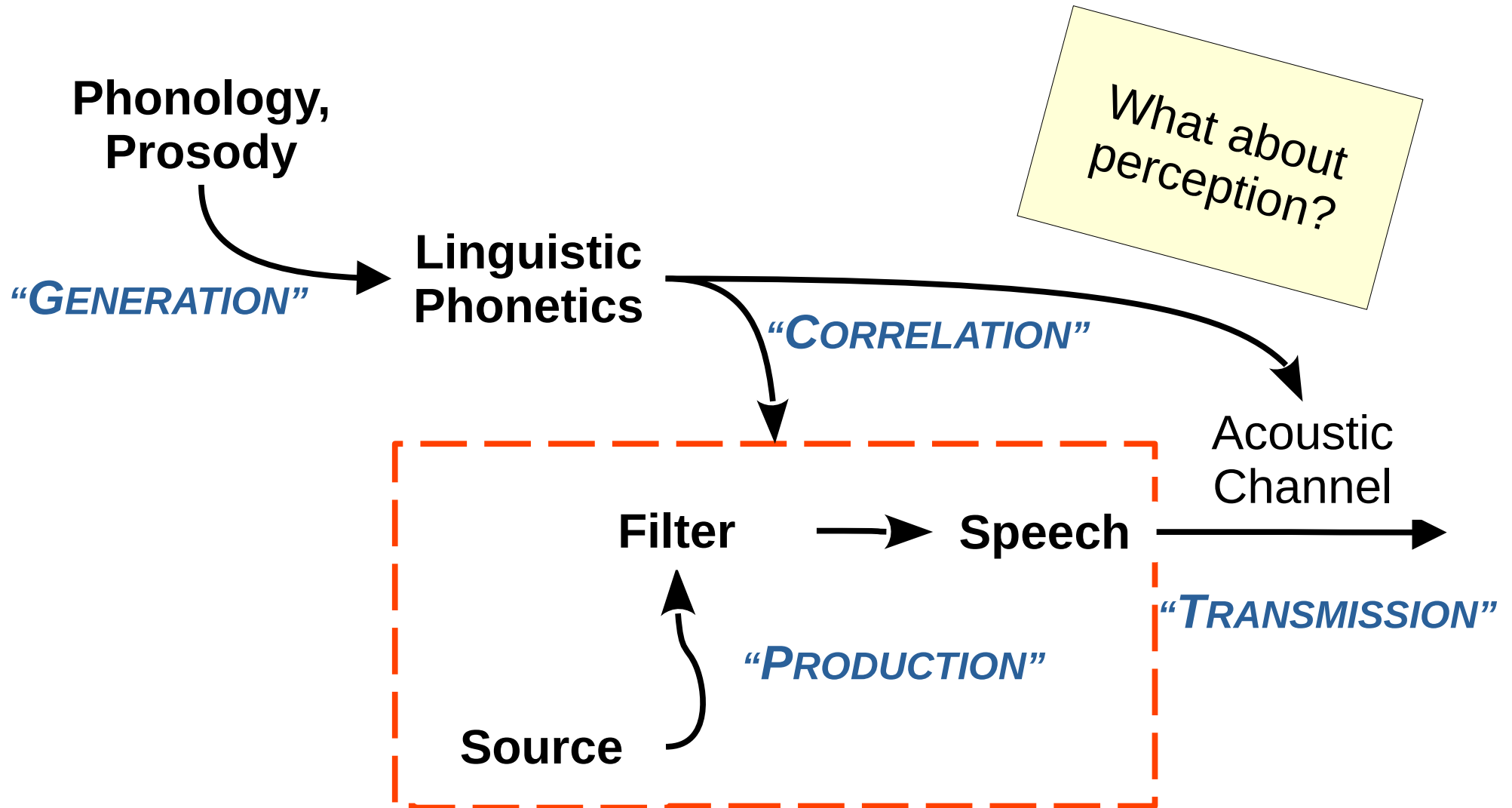
Modelling speech – Model 3



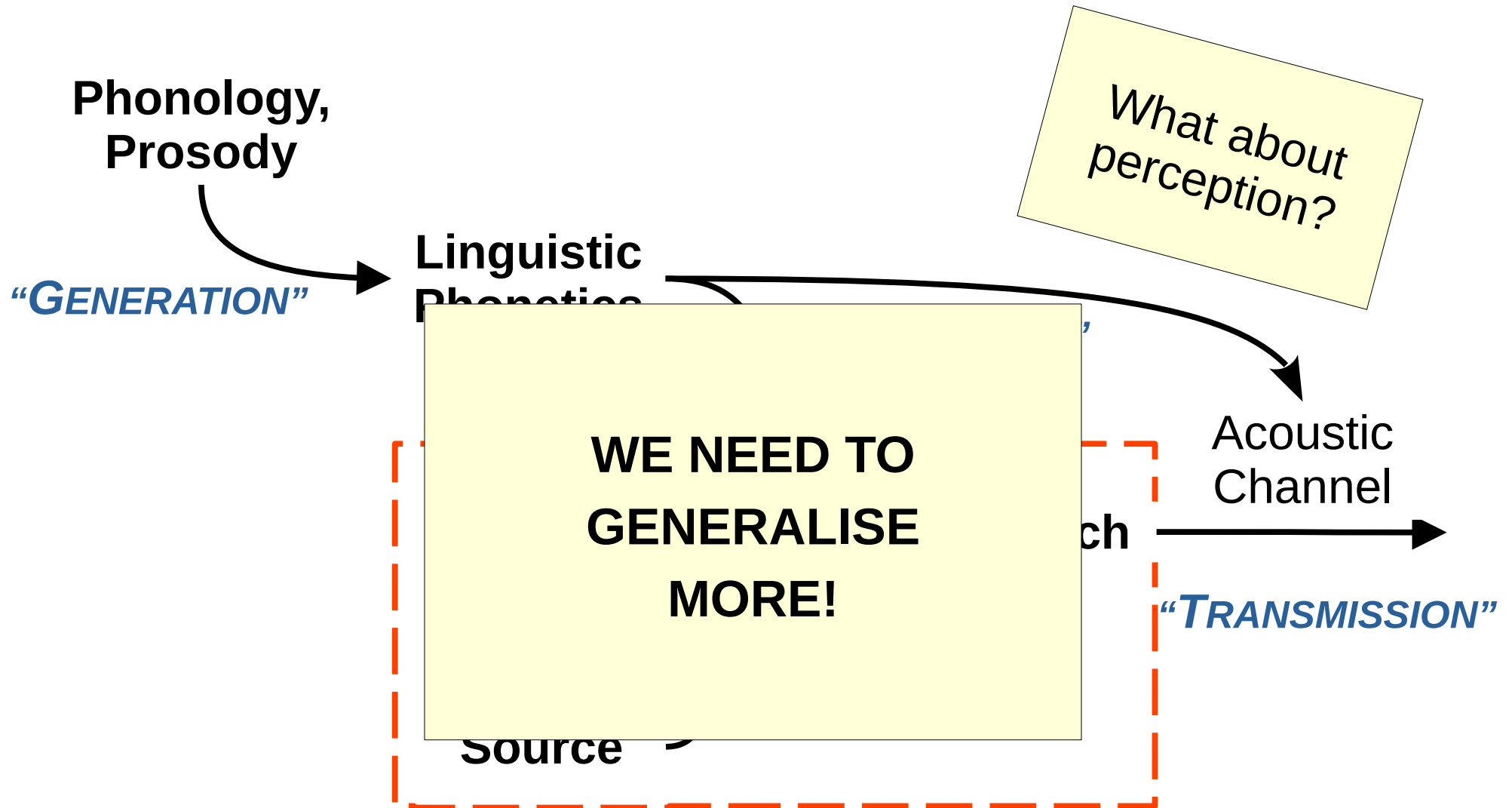
Modelling speech – Model 4



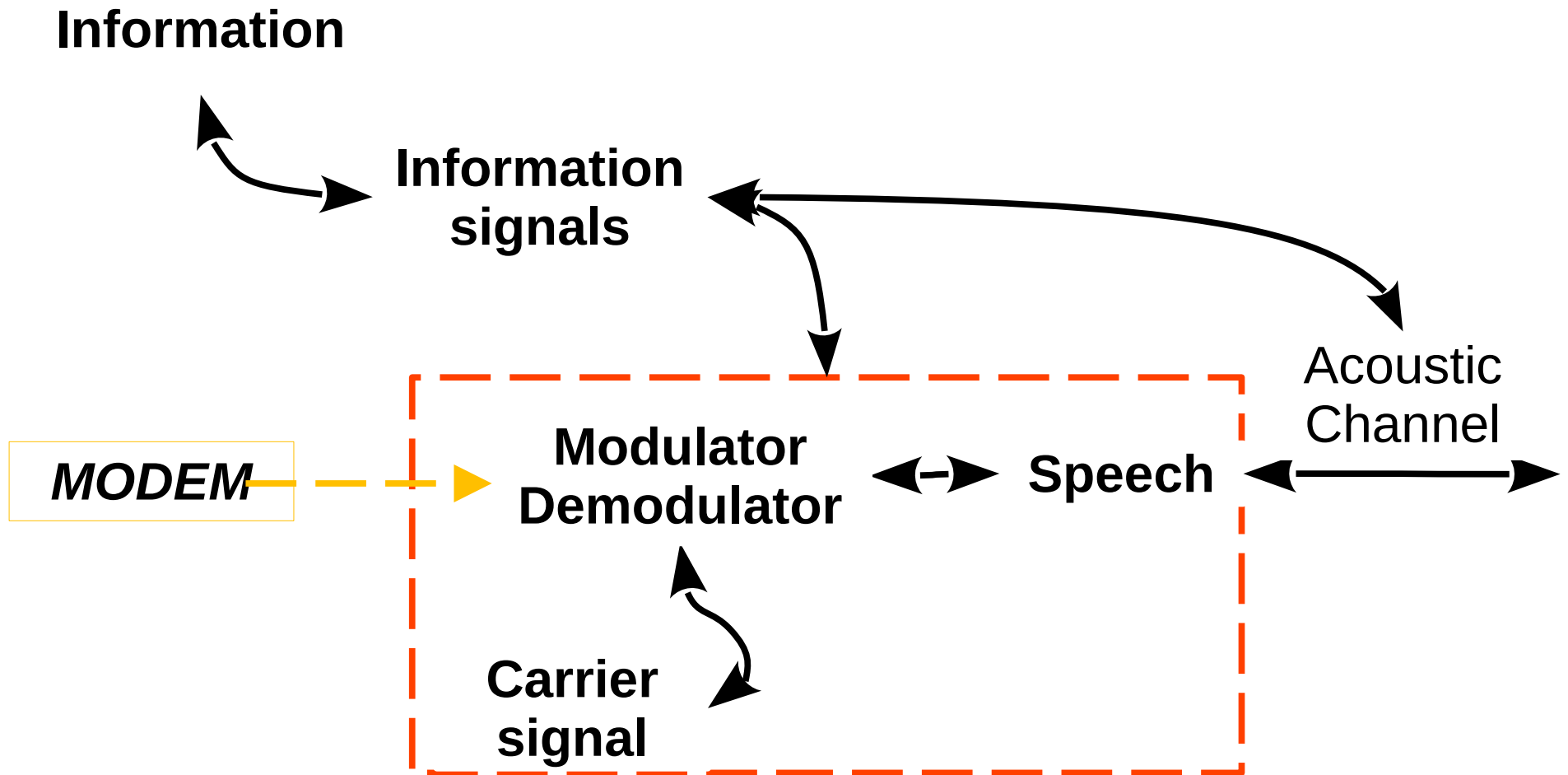
Terminological confusion ...



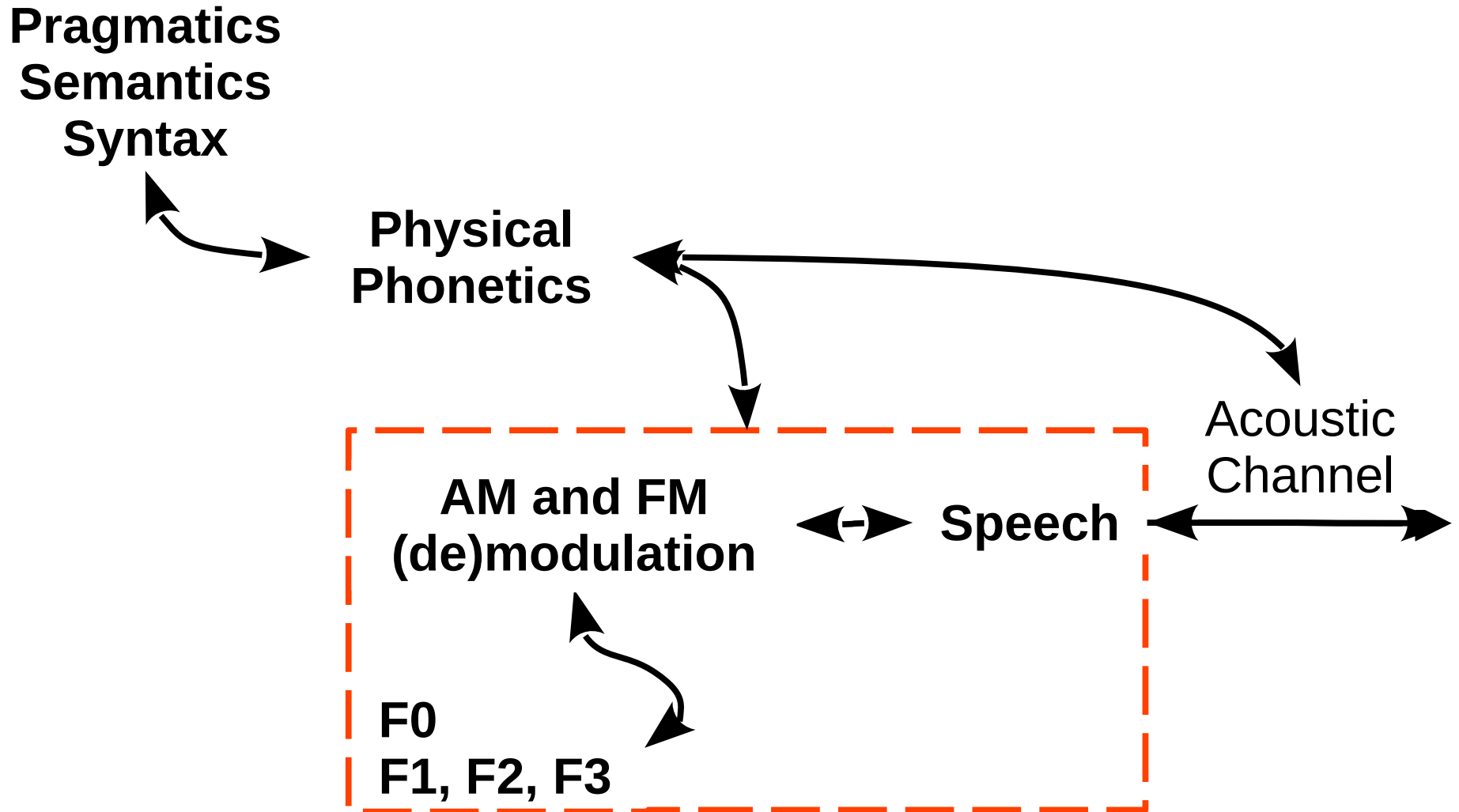
Terminological confusion ...



Modelling speech – Model 5



Modelling speech – Model 6



Linguistics and Phonetics meet Modulation Theory

Information:

Pragmatics

Semantics

Syntax

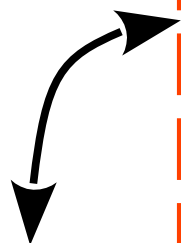


**Information
signal**



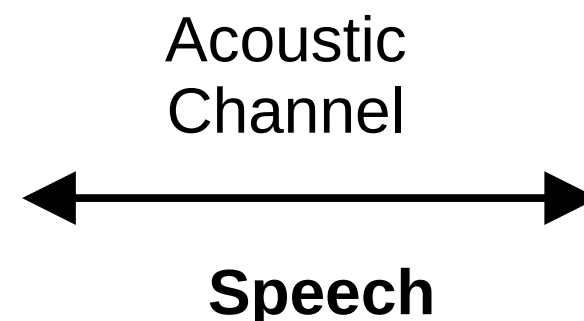
**Vocal Tract + EAR
as MODEM**

**AM and FM
(De) Modulator**



**High frequency
carrier signals:**

F0, F1, F2, F2, ... obstruent noise



AM and FM modulation as you already know it!



AM and FM modulation as you already know it!



Carrier Signals and Modulation Signals

Carrier Signals ($f > 60$ Hz)

have higher frequencies than

Information Signals ($f < 20$ Hz)

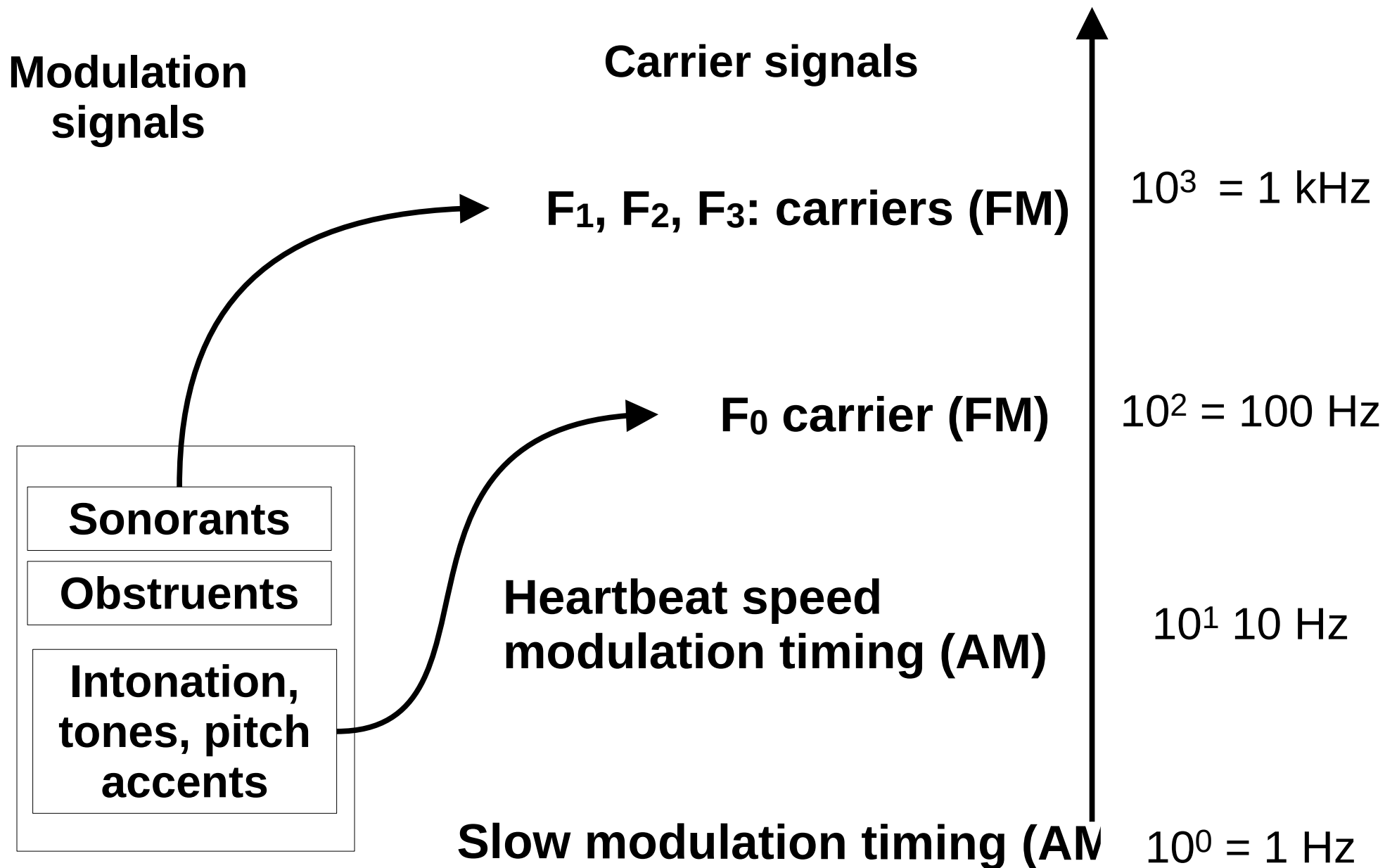
and thus enable us to hear

Speech Modulations

and transmit them across distances ...

... because we hear low frequencies only as rhythms.

Carrier and Modulation Frequencies



Types of Speech Modulation

Carrier and Modulation Frequencies

Carrier signals:

F_1, F_2, F_3 : modulated carriers (FM)
vowel and consonant carriers

F_0 modulated carrier (FM)
intonation, tone, pitch accent carriers

$10^3 = 1 \text{ kHz}$

$10^2 = 100 \text{ Hz}$

Information (modulation) signals:

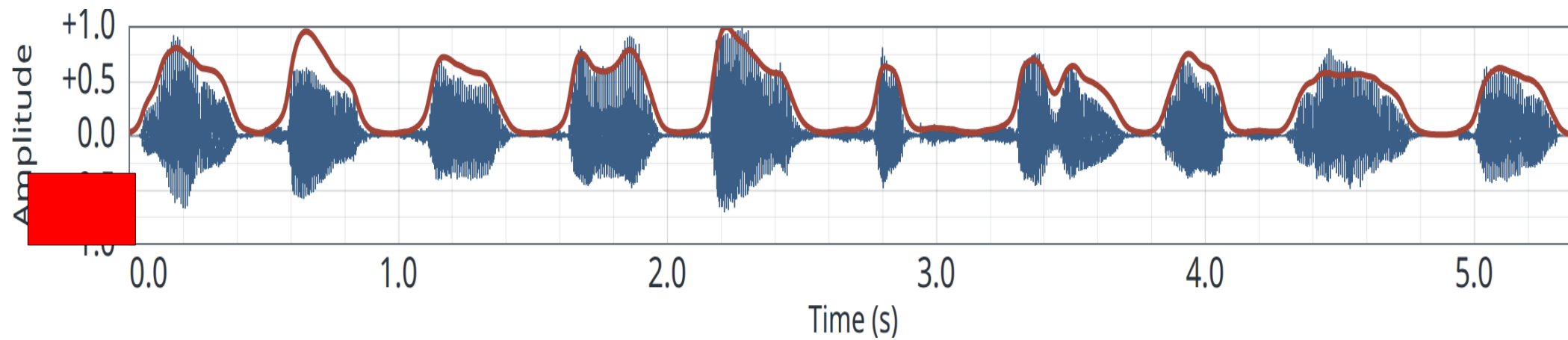
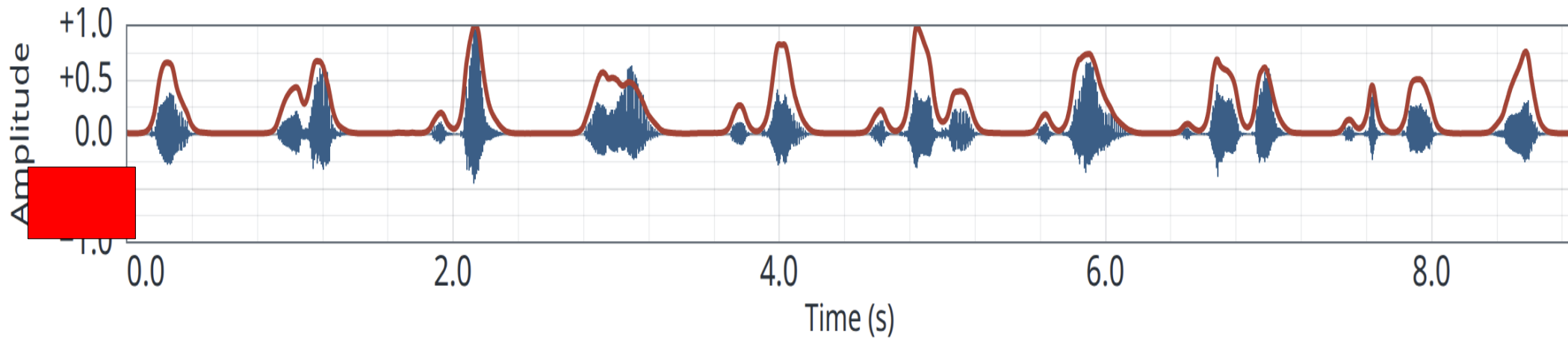
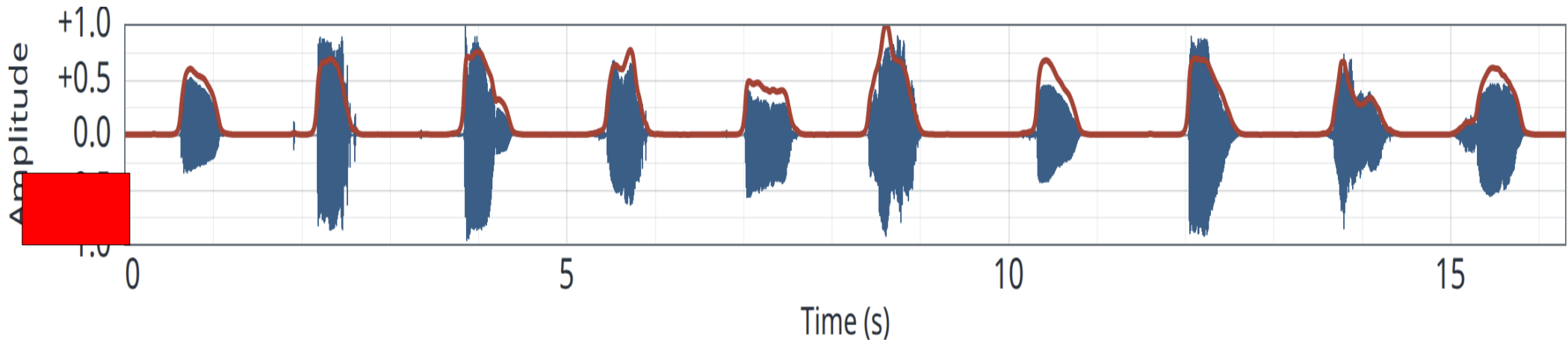
Heartbeat modulation timing (AM)
C-V, CV-CV, words, ..., rhythms
tone, pitch accent, intonation

Slow modulation timing (AM)
phrases, utterances, dialog

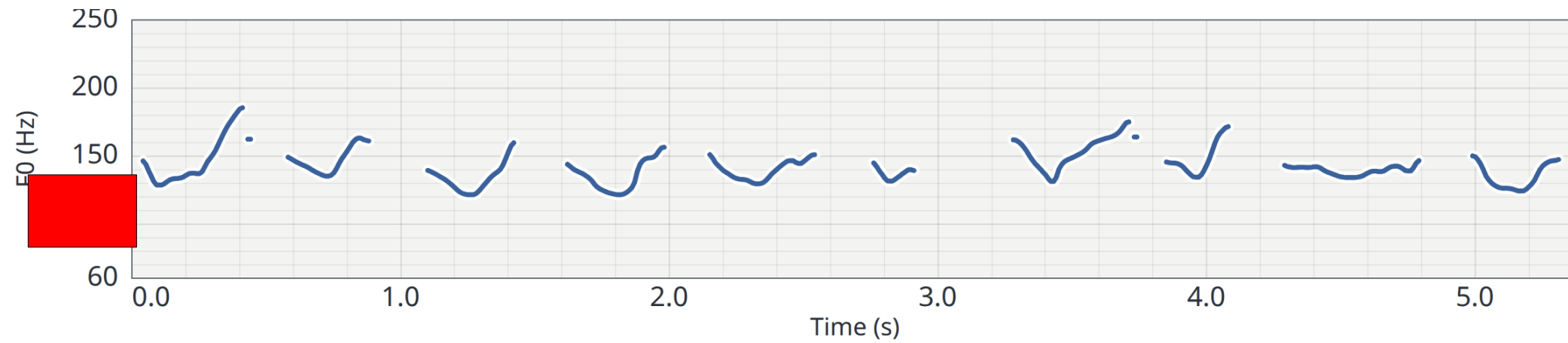
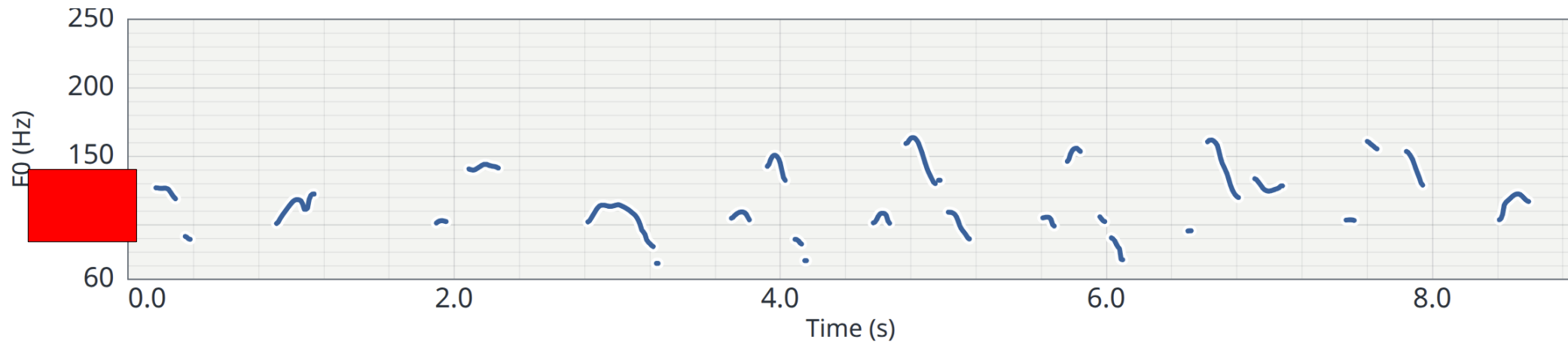
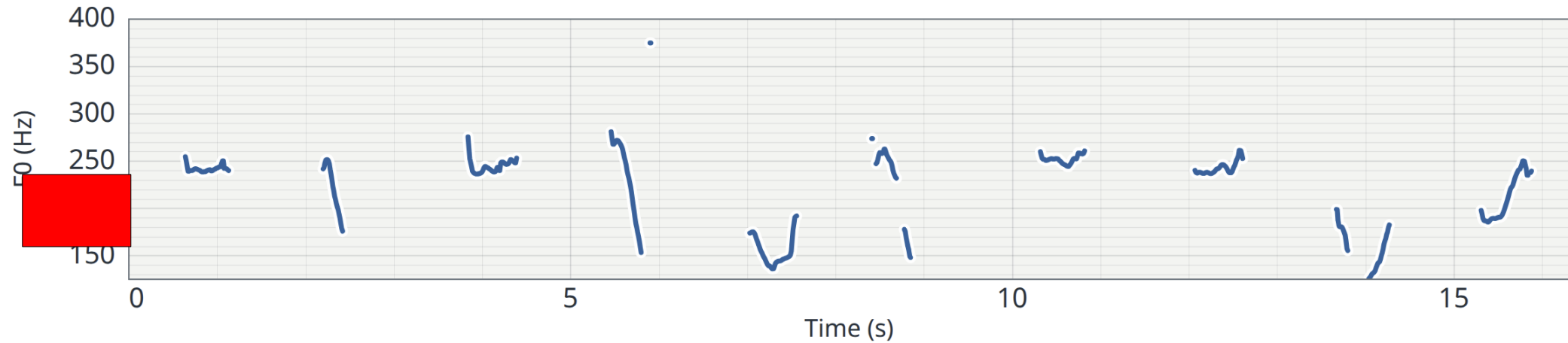
$10^1 = 10 \text{ Hz}$

$10^0 = 1 \text{ Hz}$

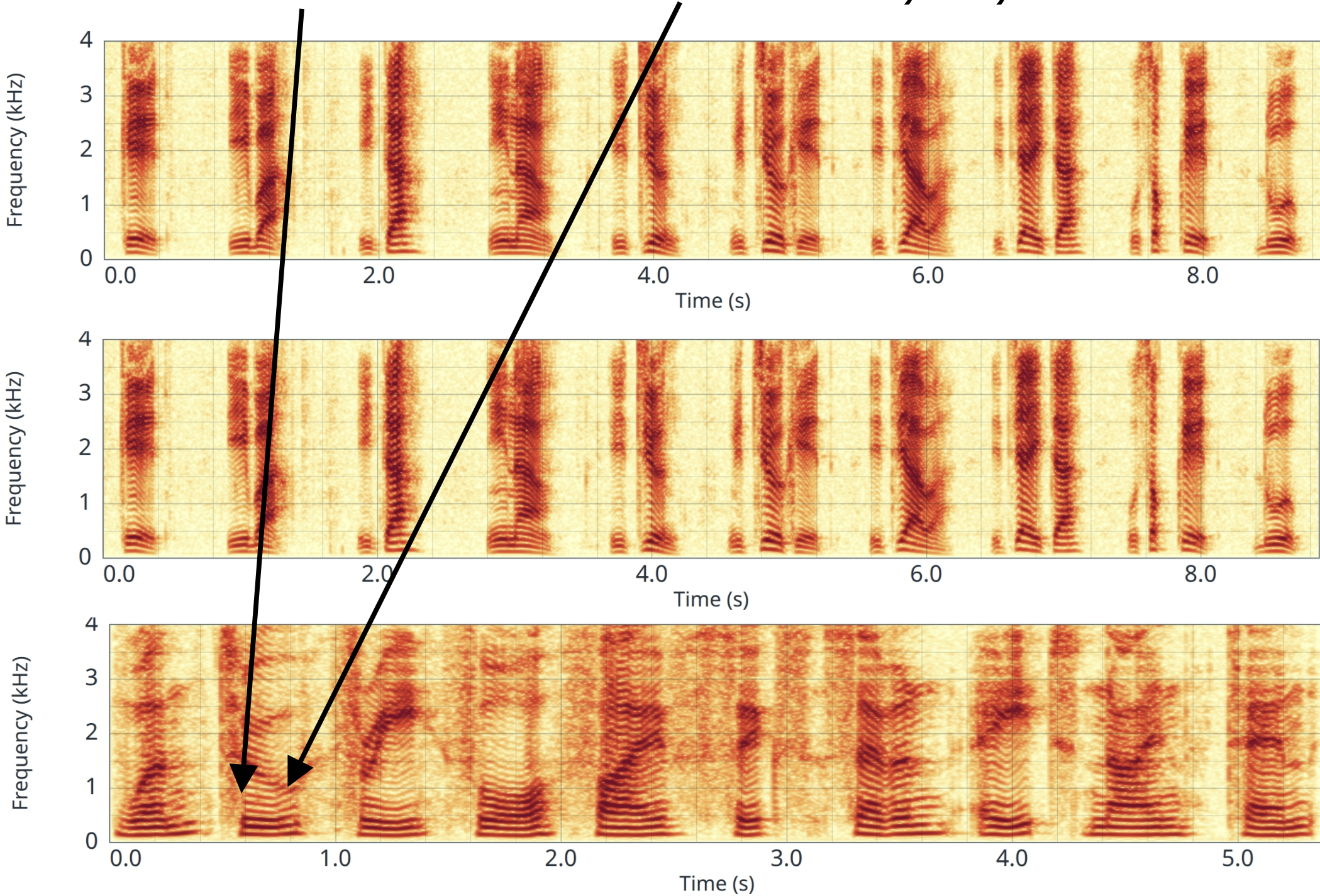
AM envelope carrier: AM (sonority, rhythms ...)



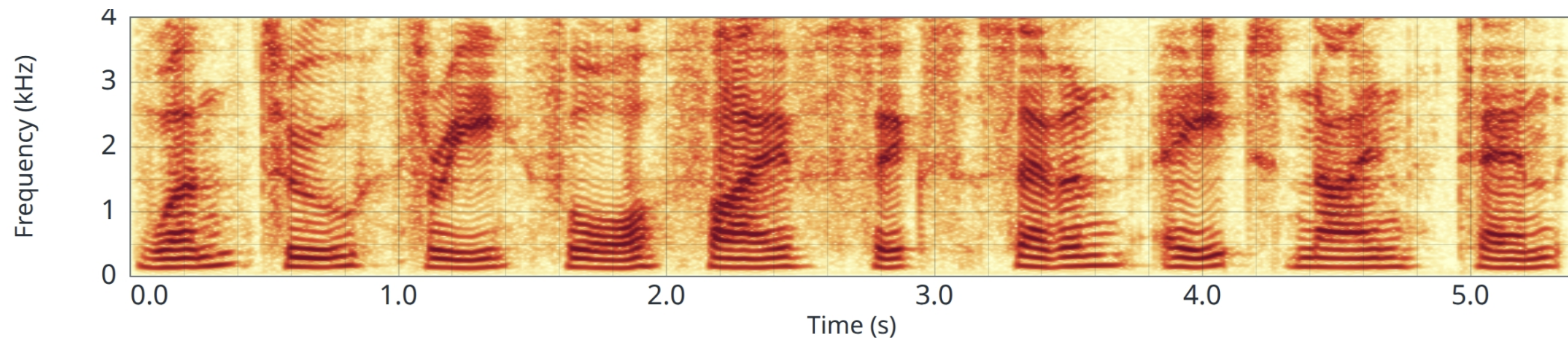
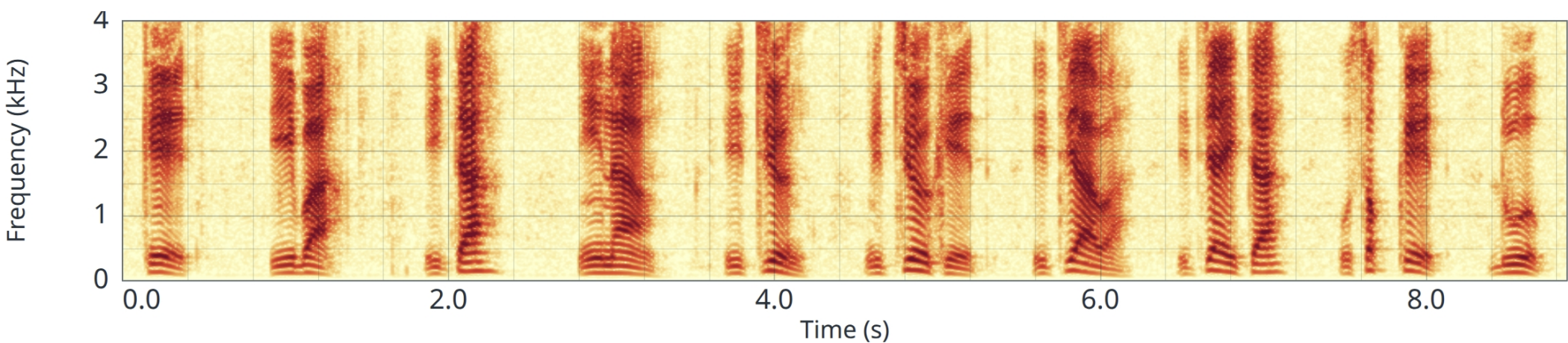
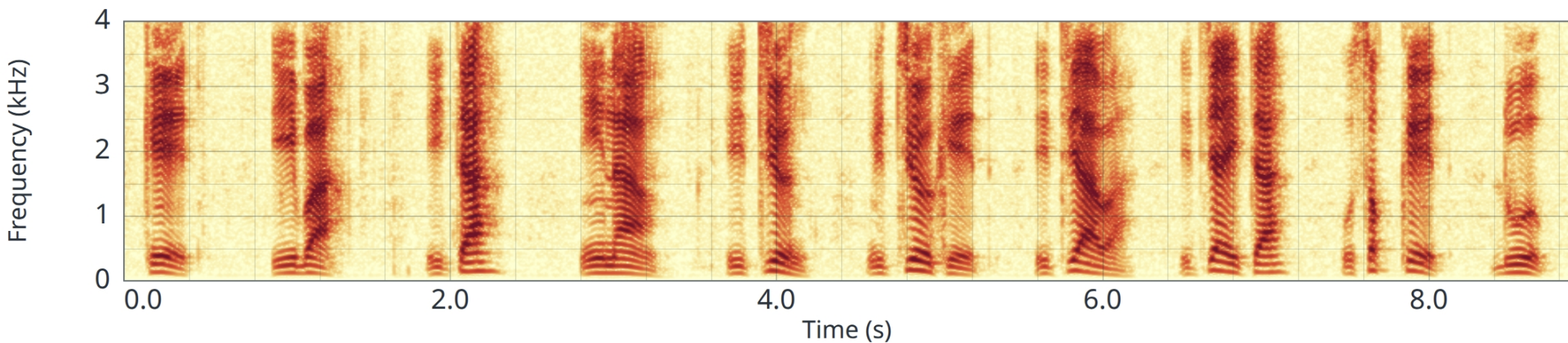
F0 carrier: FM tuning (intonation, tone, accent)



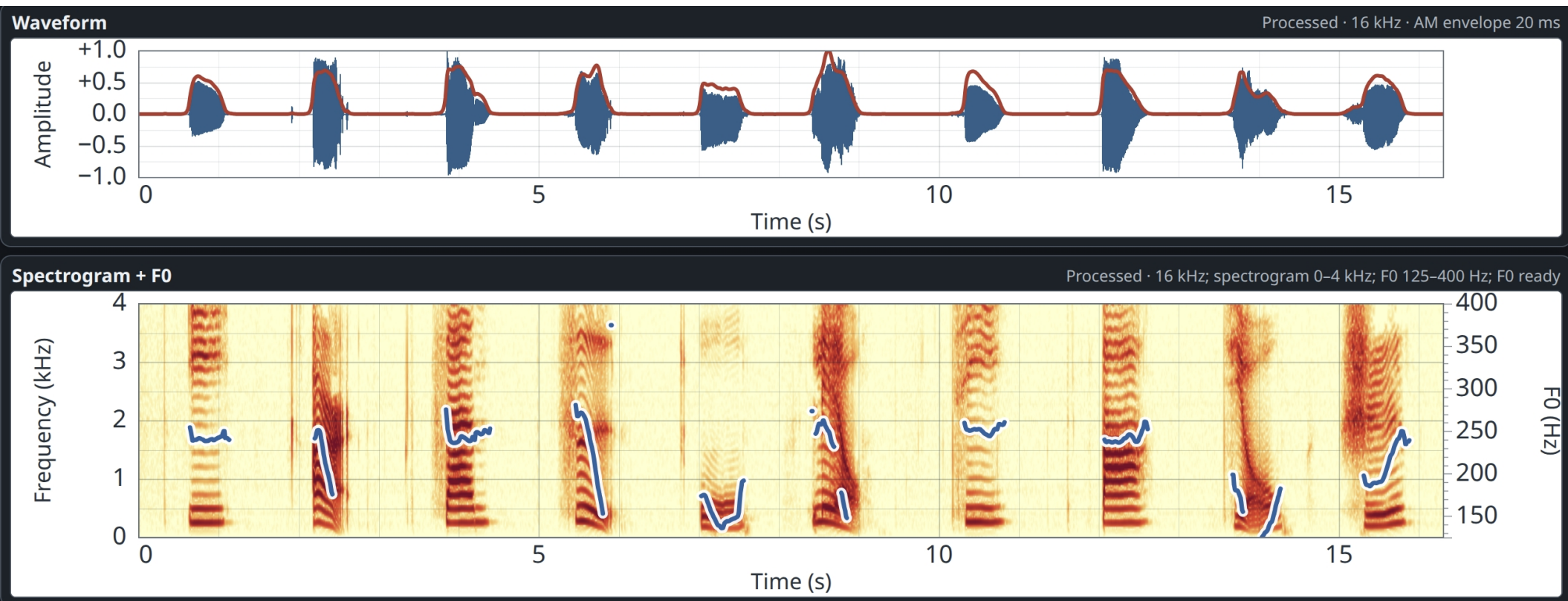
Harmonic carrier: AM creates *F1*, *F2*, *F3* filters



Formants as carriers: *FM by filter tuning*



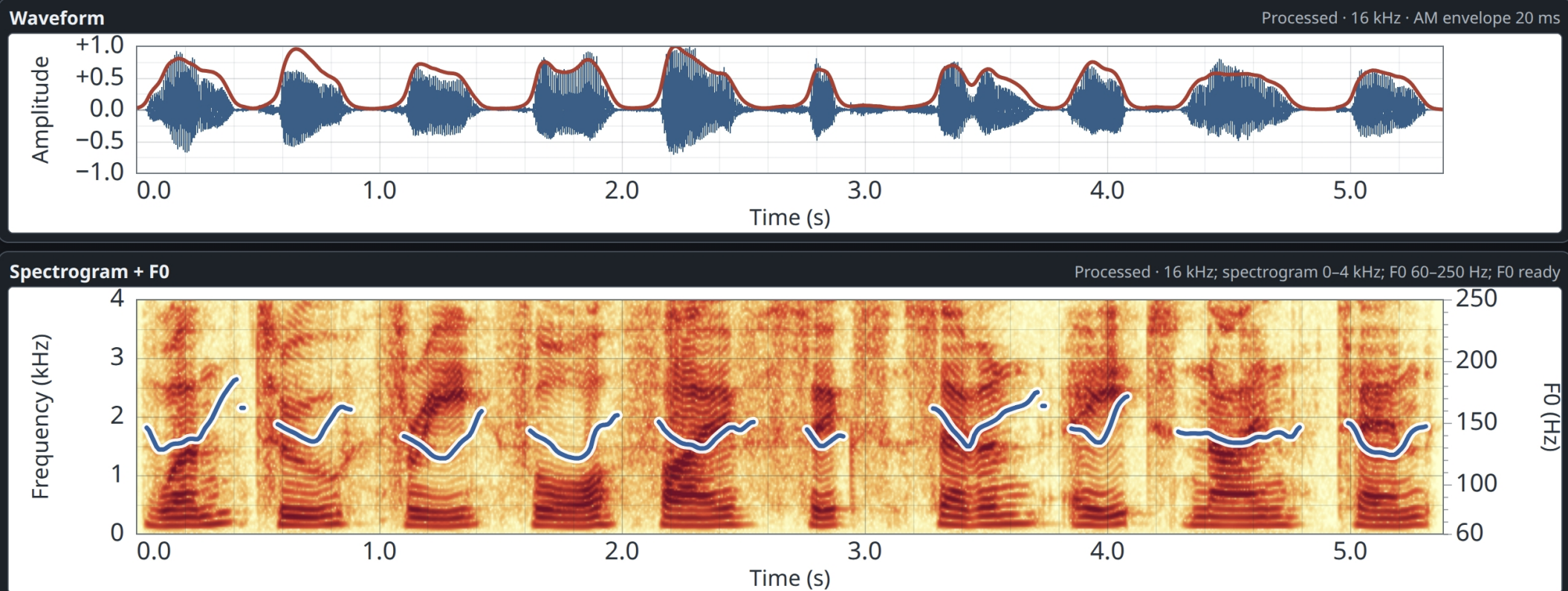
Harmonics: AM → Formants: FM



All four modulations:

- *AM envelope*: **AM** → sonority; syllable, word, phrase, discourse rhythms
- *F0 estimation*: **FM** → intonation, tone, pitch accent
- *F0 harmonics*: **FM** → formant creation by filtering
- Formants F1, F2, F3: **FM** → sonorants (voiced intervals: vowels etc.)

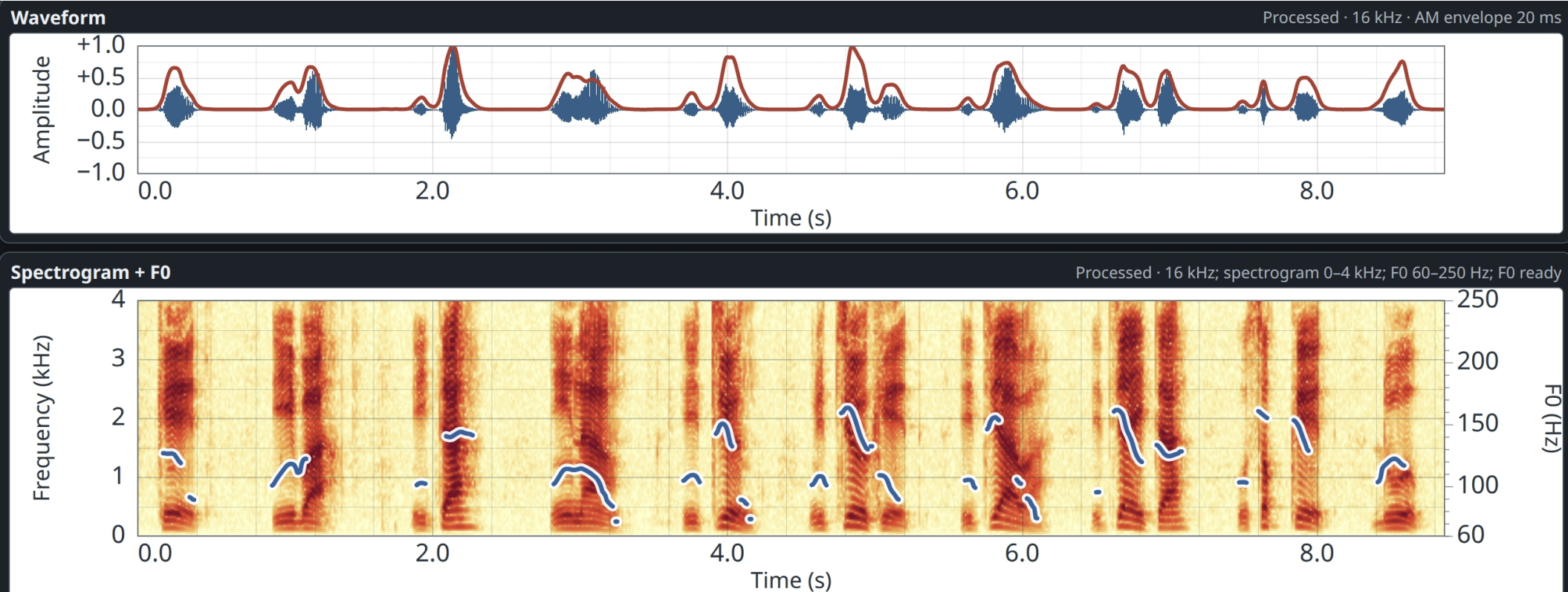
Harmonics: AM → Formants: FM



All four modulations:

- *AM envelope*: **AM** → sonority; syllable, word, phrase, discourse rhythms
- *F0 estimation*: **FM** → intonation, tone, pitch accent
- *F0 harmonics*: **FM** → formant creation by filtering
- Formants F1, F2, F3: **FM** → sonorants (voiced intervals: vowels etc.)

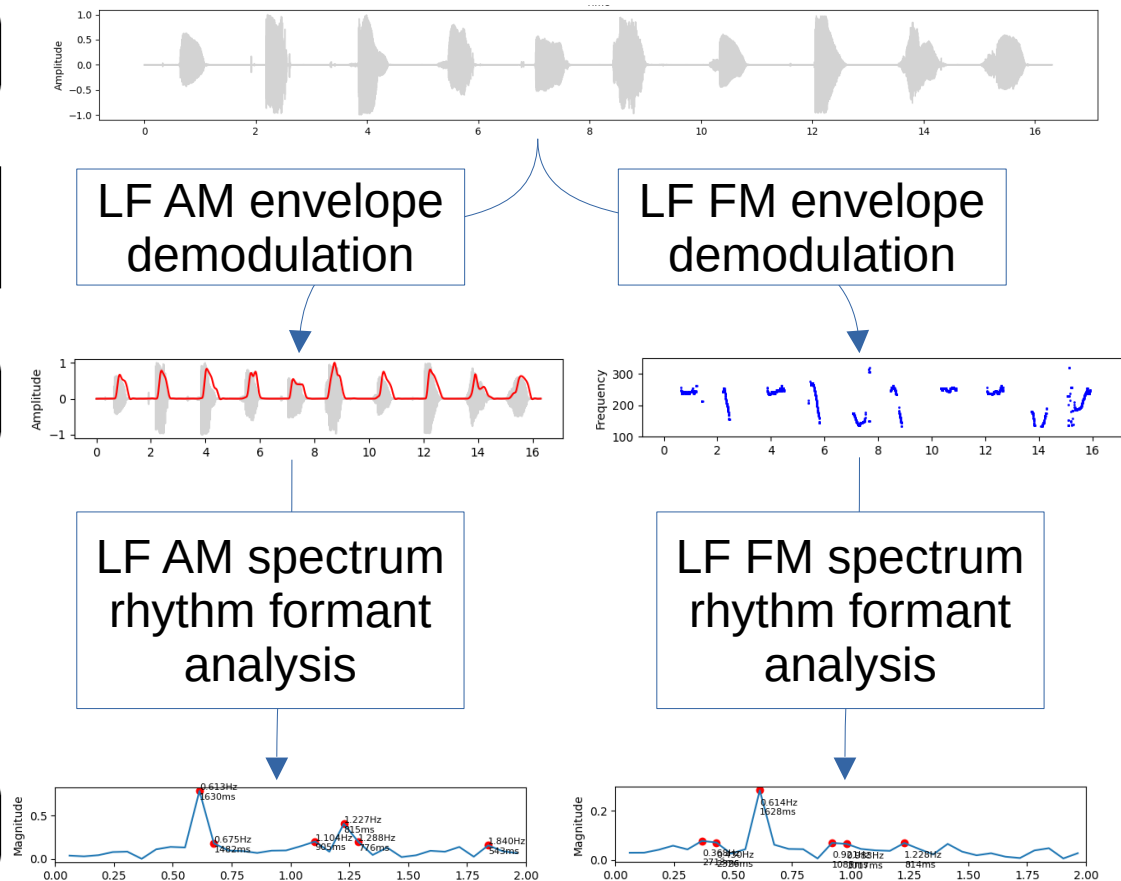
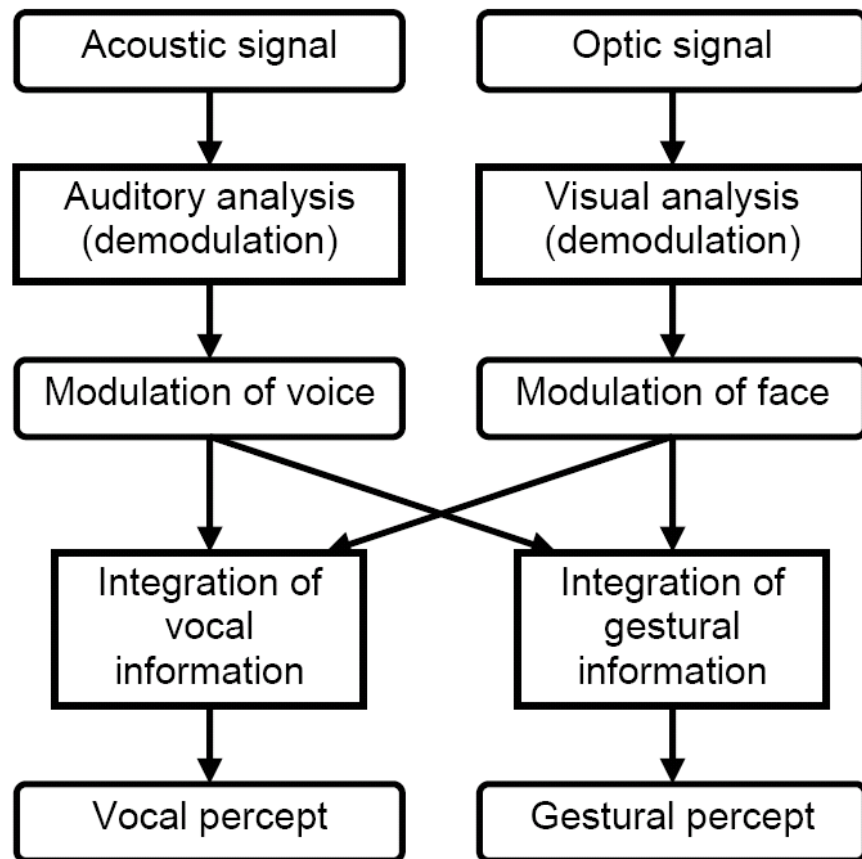
Harmonics: AM → Formants: FM



All four modulations:

- *AM envelope*: **AM** → sonority; syllable, word, phrase, discourse rhythms
- *F0 estimation*: **FM** → intonation, tone, pitch accent
- *F0 harmonics*: **FM** → formant creation by filtering
- Formants F1, F2, F3: **FM** → sonorants (voiced intervals: vowels etc.)

Comparison with Traunmüller's multimodal demodulation model

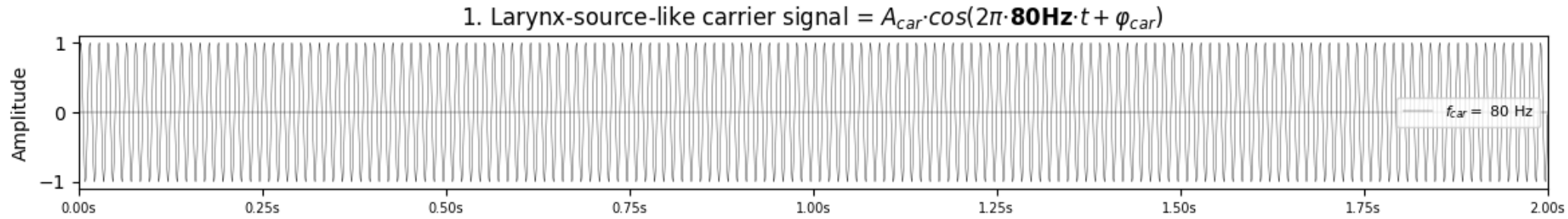


Hartmut Traunmüller (1994) "Conventional, biological, and environmental factors in speech communication: A modulation theory" *Phonetica* 51: 170-183. doi (Also in *PERILUS XVIII*: 92-102.)

Hartmut Traunmüller (2007) "Demodulation, mirror neurons and audiovisual perception nullify the motor theory" *Contr. to Fonetik 2007*, *TMH-QPSR* 50: 17-20. *Detpt. of Speech, Music and Hearing*, Royal Inst. of Technology, Stockholm.

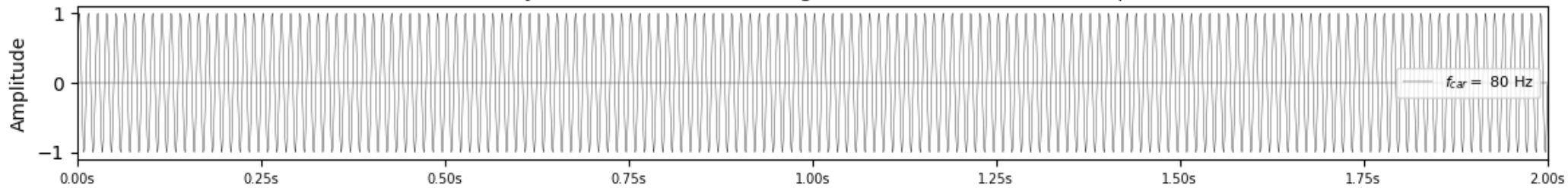
AM and FM modulation step by step

Modulation: carrier signal

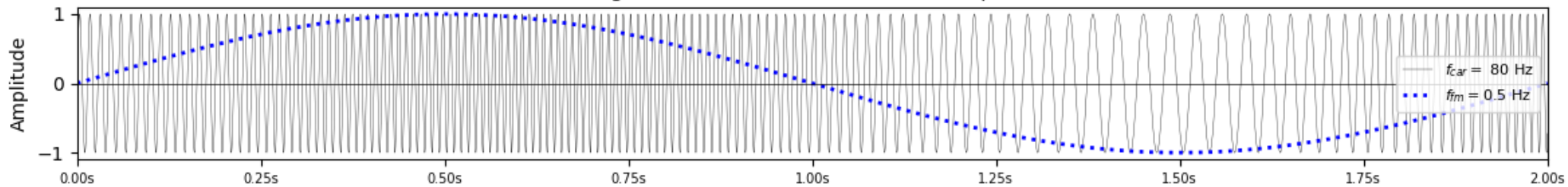


Modulation: FM signal with low frequency information

1. Larynx-source-like carrier signal = $A_{car} \cdot \cos(2\pi \cdot 80\text{Hz} \cdot t + \varphi_{car})$

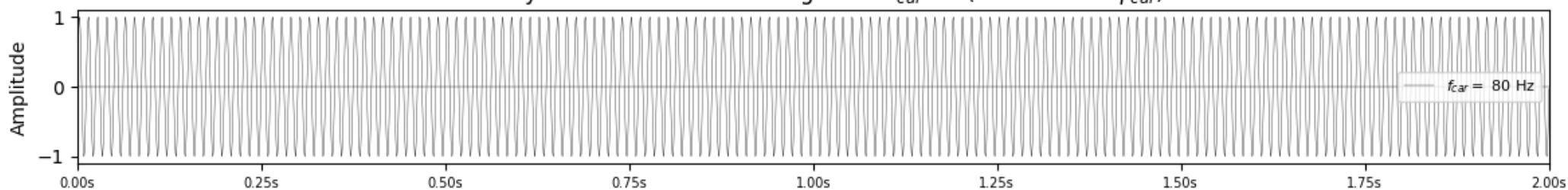


2. Intonation-like FM signal ($\text{FM} = A_{fm} \cdot \cos(2\pi \cdot 0.5\text{Hz} \cdot t + \varphi_{fm})$, modulation overlaid)

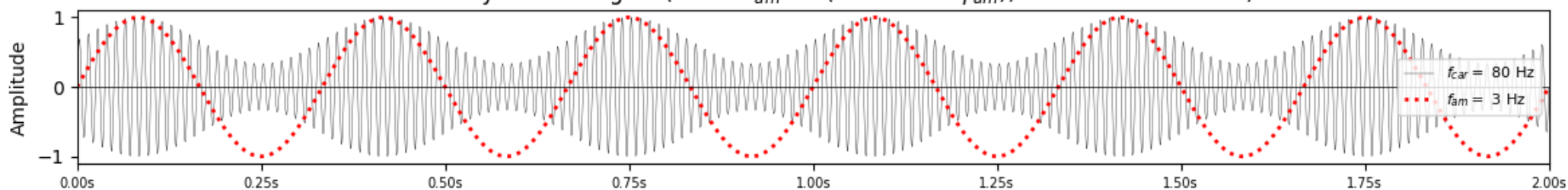


Modulation: AM signal with low frequency information

1. Larynx-source-like carrier signal = $A_{car} \cdot \cos(2\pi \cdot 80\text{Hz} \cdot t + \varphi_{car})$

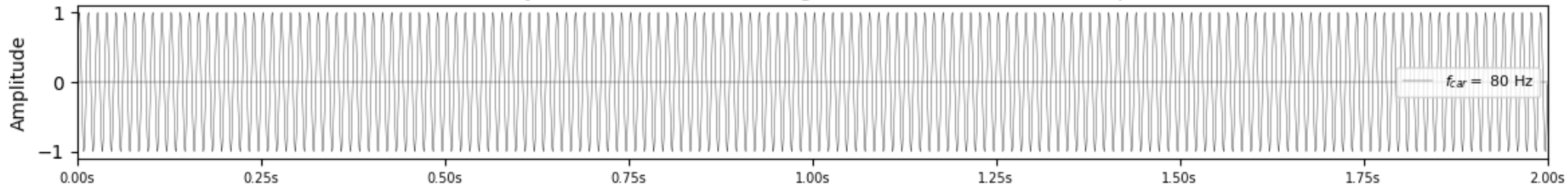


3. Sonority-like AM signal (**AM** = $A_{am} \cdot \cos(2\pi \cdot 3\text{Hz} \cdot t + \varphi_{am})$, modulation overlaid)

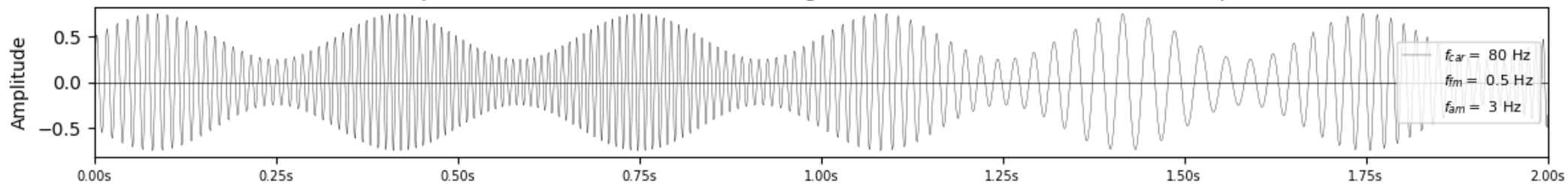


Modulation Theory

1. Larynx-source-like carrier signal = $A_{car} \cdot \cos(2\pi \cdot \mathbf{80Hz} \cdot t + \varphi_{car})$

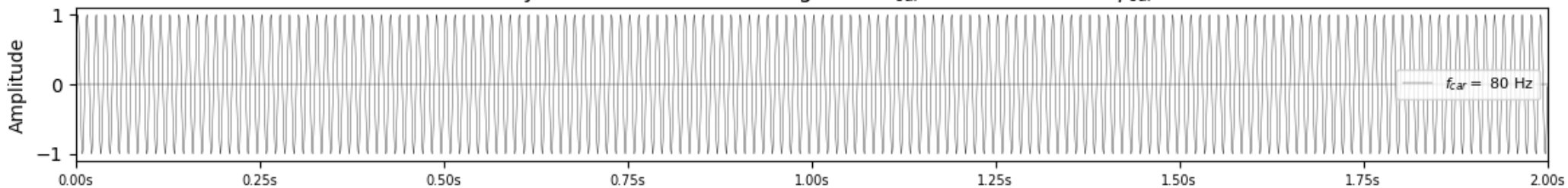


4. Speech-like combined FM and AM signal ($\mathbf{AM} + A_{car} \cdot \cos(2\pi \cdot (f_{car} + \mathbf{FM}) \cdot t + \varphi_{car})$)

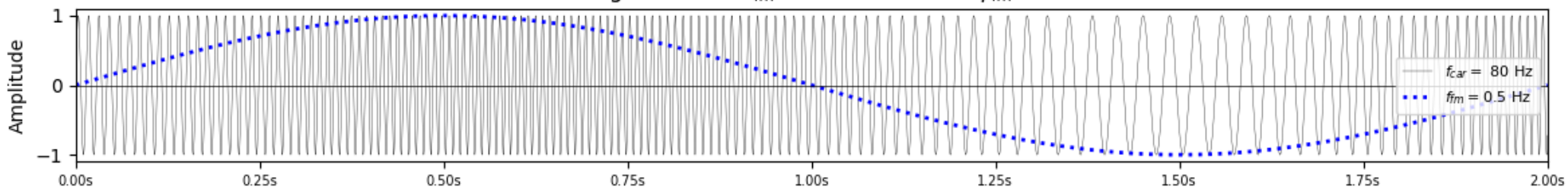


Modulation Theory

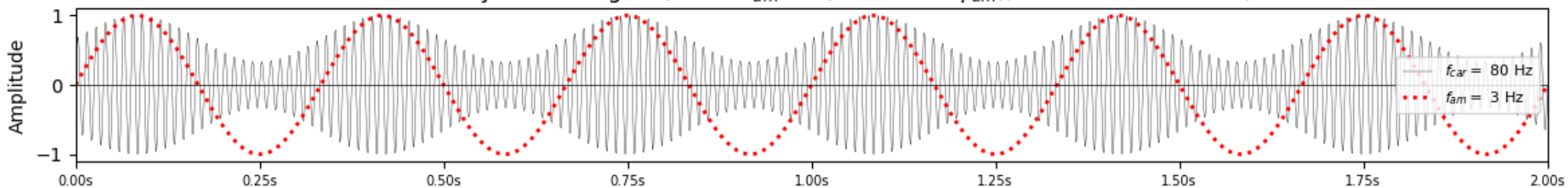
1. Larynx-source-like carrier signal = $A_{car} \cdot \cos(2\pi \cdot 80\text{Hz} \cdot t + \varphi_{car})$



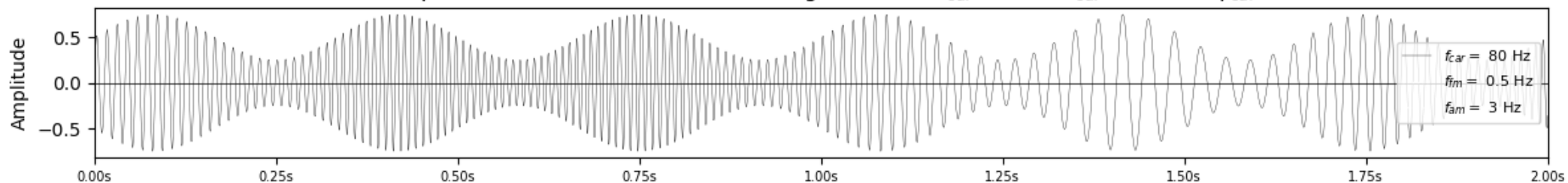
2. Intonation-like FM signal ($\text{FM} = A_{fm} \cdot \cos(2\pi \cdot 0.5\text{Hz} \cdot t + \varphi_{fm})$, modulation overlaid)



3. Sonority-like AM signal ($\text{AM} = A_{am} \cdot \cos(2\pi \cdot 3\text{Hz} \cdot t + \varphi_{am})$, modulation overlaid)



4. Speech-like combined FM and AM signal ($\text{AM} + A_{car} \cdot \cos(2\pi \cdot (f_{car} + \text{FM}) \cdot t + \varphi_{car})$)



Demodulation and analysis procedures

- Time domain procedures:
 - Envelope extraction
 - Fundamental frequency estimation ('pitch' extraction)
- Frequency domain procedures:
 - Spectral analysis
 - Spectrogram analysis
 - there are also frequency domain procedures for F0 estimation
- Comparison using distance metrics
 - distance calculation with different distance metrics
 - hierarchical clustering with distance and different clustering criteria
- Output:
 - Graphical display
 - Numerical files and figure files

Demodulation

AM envelope demodulation:

- Phonetics: syllable, stress-accent
- Phonology: sonority curve, syllable sequences, stress values

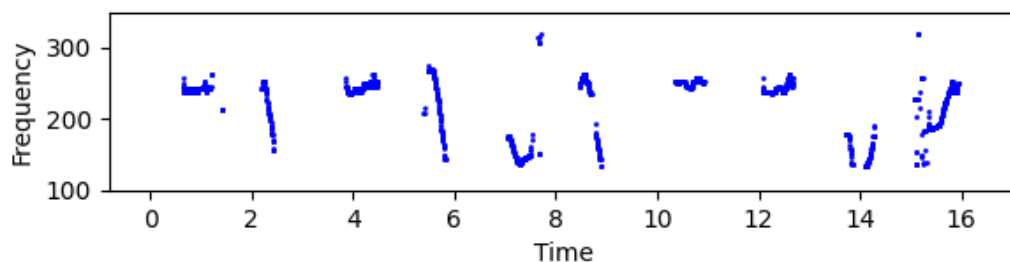
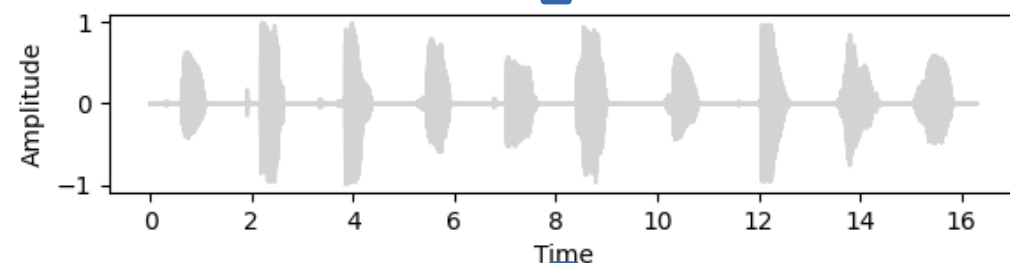
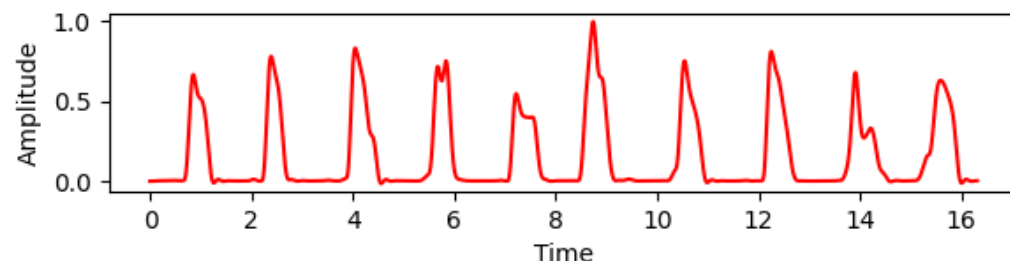


Modulated carrier signal

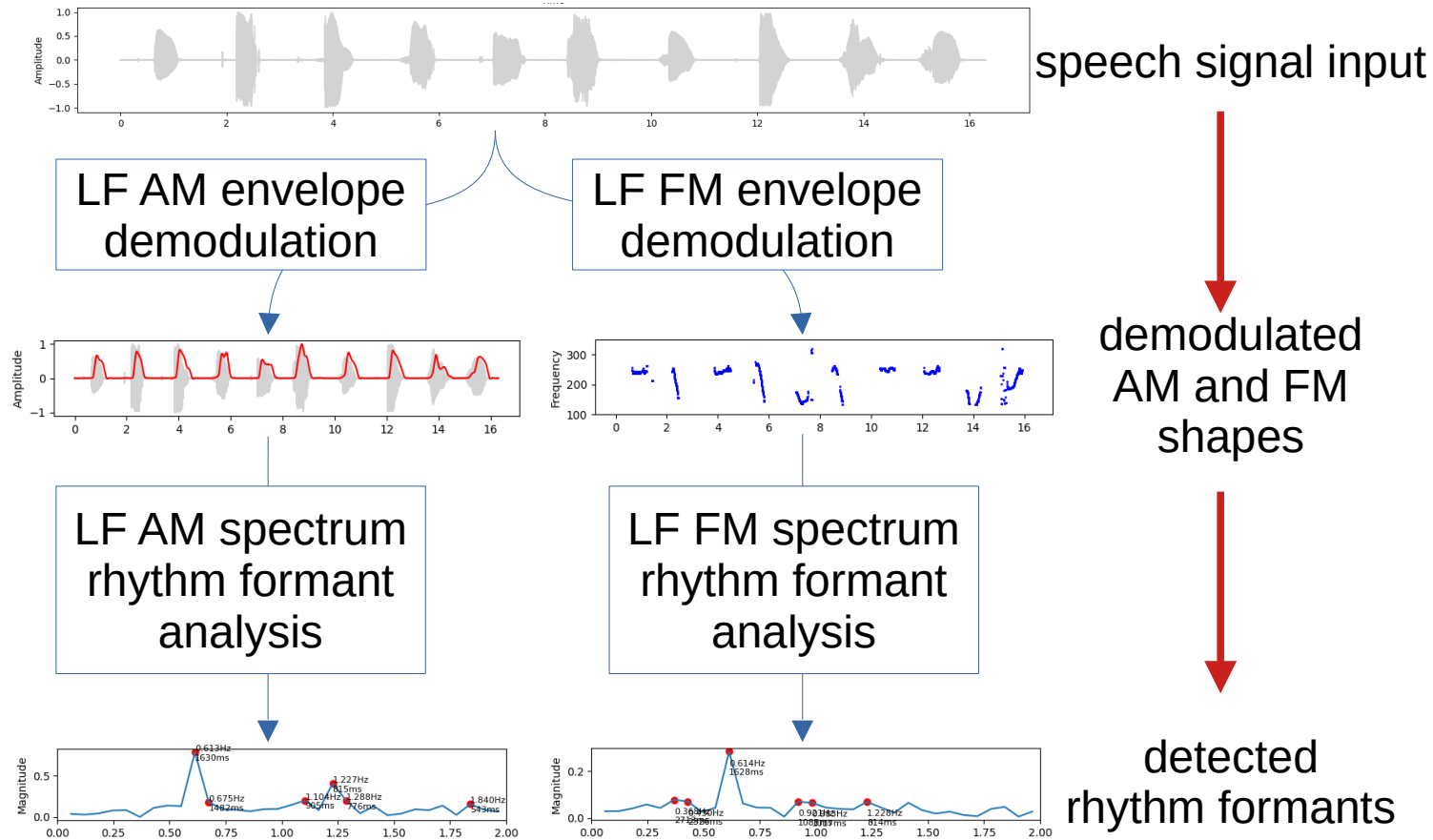


FM envelope demodulation:

- Phonetics: F0, pitch track
- Phonology: tones, pitch accents, intonation



AM and FM demodulation and detection of rhythm



Hartmut Traunmüller (2007) "Demodulation, mirror neurons and audiovisual perception nullify the motor theory" Contr. to Fonetik 2007, TMH-QPSR 50: 17-20. Dept. of Speech, Music and Hearing, Royal Inst. of Technology, Stockholm.