DÉPARTEMENT D'ÉTUDES COGNITIVES

PSL UNIVERSITÉ PARIS

Use of monaural auditory models from the AMT toolbox

3rd AMT Workshop

Turin, 8 September 2023

Alejandro Osses, Ph.D.

Posdoctoral researcher, ENS Paris, France
Acoustic researcher at Sorama BV, Eindhoven, the Netherlands

ale.a.osses@gmail.com, alejandro.osses@sorama.eu

What can you expect from this presentation?

- I will show how to use and customise a monaural model of the auditory periphery
- The models:
  - dau1997.m
  - zilany2014.m
  - verhulst2015.m
  - verhulst2018.m
  - bruce2018.m
  - relanoiborra2019_featureextraction.m
  - king2019.m
  - osses2021.m
- In other words you will be able of:
  - Understanding the inputs and outputs of the models
  - How to enable / disable modules of the models (**flags**)
  - How to change parameters of the models (**keyvals**)

Use of **flags** and **keyvals**

There are other models, e.g.,:

breebaart2001
hohmann2002
zilany2007
carney2015
decheveigne2023

Input: Not waveforms but AN responses

Advanced task?

But not all models are parametrised using flags and keyvals
Not all models have the same structure

Useful references to keep in mind

- Comparison of the eight models: Osses et al. (2022, Acta Acustica)

Table 1. List of selected models. The model labels used in this study correspond with the model functions in AMT 1.1.

Label	Reference
dau1997	Dau *et al.* (1997) [31]
zilany2014	Zilany *et al.* (2014) [32] and Carney *et al.* (2015) [33]
verhulst2015	Verhulst *et al.* (2015) [34]
verhulst2018	Verhulst *et al.* (2018) [35]
bruce2018	Bruce *et al.* (2018) [36] and Carney *et al.* (2015) [33]
king2019	King *et al.* (2019) [37]
relanoiborra2019	Relaño-Iborra *et al.* (2019) [38]
osses2021	Osses and Kohlrausch (2021) [39]

(Paper here)

(Paper by Maxwell et al. 2020 here)

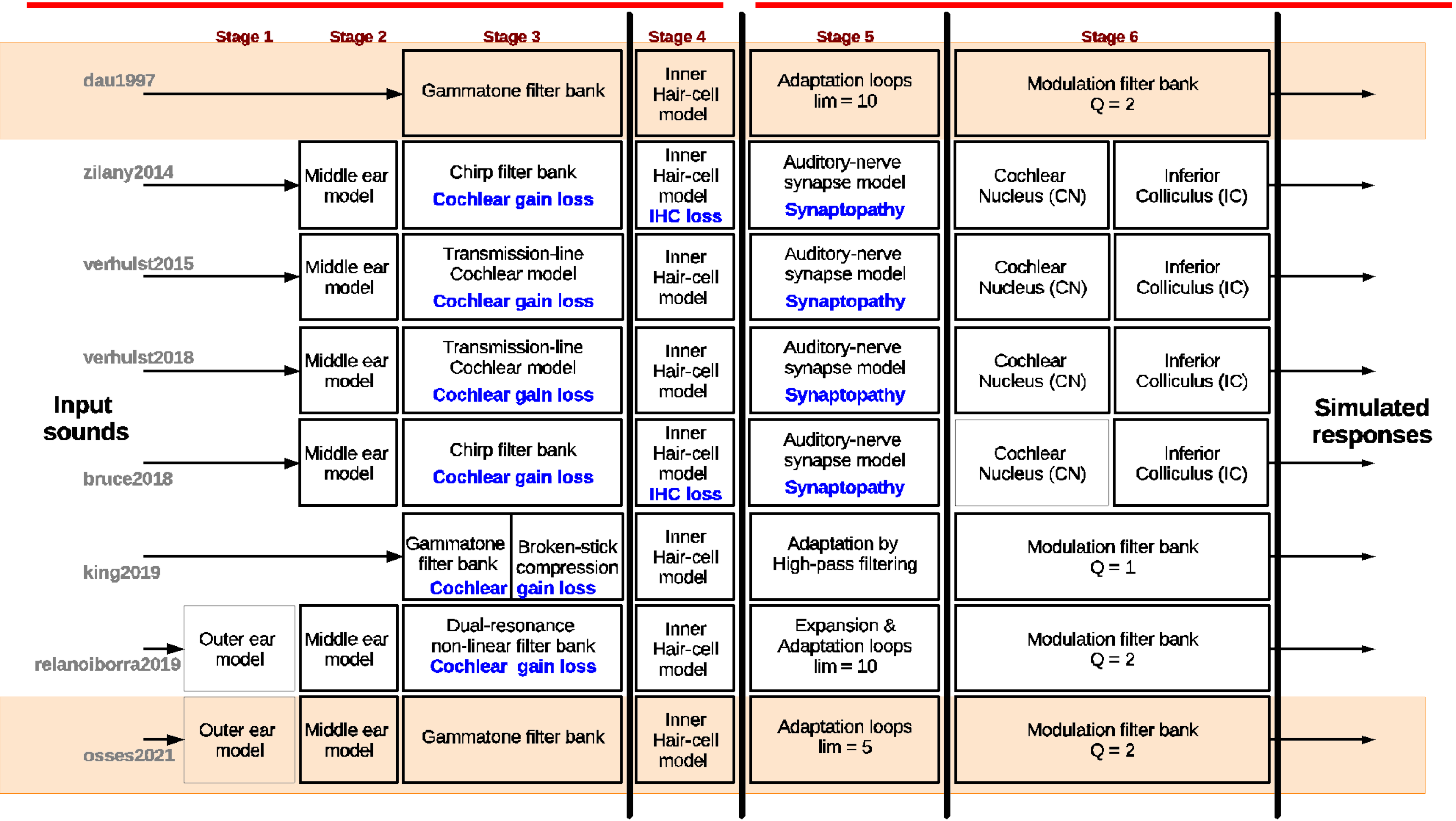
(v1.2 of the model here)

(UR EAR toolbox here)

(Paper here)

(Paper here)

(Paper here)

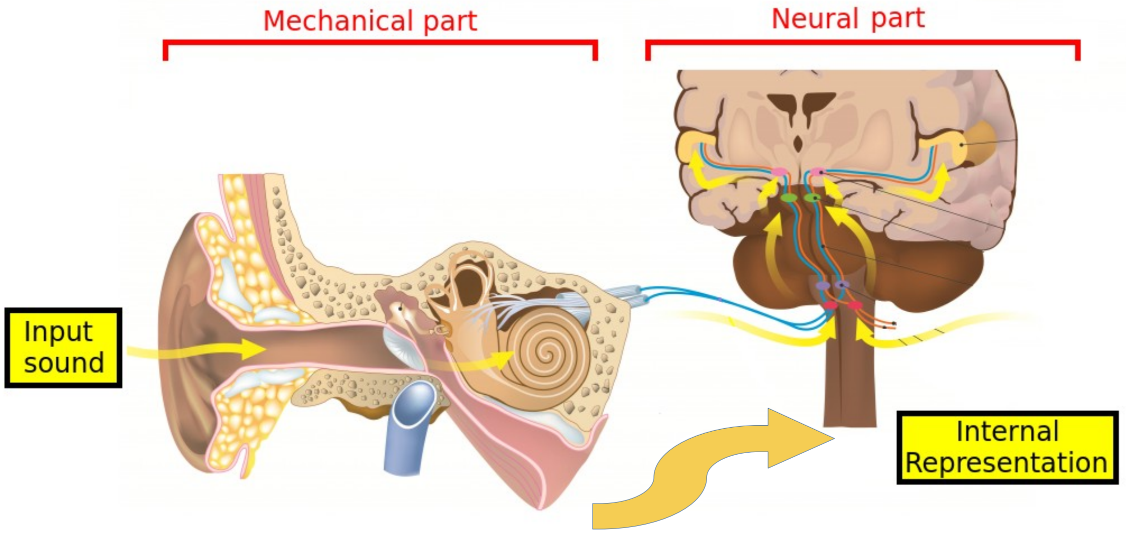


Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Stage 6
dau1997
Gammatone filter bank
Inner Hair-cell model
Adaptation loops lim = 10
Modulation filter bank Q = 2
zilany2014
Middle ear model
Chirp filter bank
Cochlear gain loss
Inner Hair-cell model
IHC loss
Auditory-nerve synapse model
Synaptopathy
Cochlear Nucleus (CN)
Inferior Colliculus (IC)
verhulst2015
Middle ear model
Transmission-line Cochlear model
Cochlear gain loss
Inner Hair-cell model
Auditory-nerve synapse model
Synaptopathy
Cochlear Nucleus (CN)
Inferior Colliculus (IC)
verhulst2018
Middle ear model
Transmission-line Cochlear model
Cochlear gain loss
Inner Hair-cell model
Auditory-nerve synapse model
Synaptopathy
Cochlear Nucleus (CN)
Inferior Colliculus (IC)
Input sounds
Simulated responses
bruce2018
Middle ear model
Chirp filter bank
Cochlear gain loss
Inner Hair-cell model
IHC loss
Auditory-nerve synapse model
Synaptopathy
Cochlear Nucleus (CN)
Inferior Colliculus (IC)
king2019
Gammatone filter bank
Broken-stick compression
Cochlear gain loss
Inner Hair-cell model
Adaptation by High-pass filtering
Modulation filter bank Q = 1
relanoiborra2019
Outer ear model
Middle ear model
Dual-resonance non-linear filter bank
Cochlear gain loss
Inner Hair-cell model
Expansion & Adaptation loops lim = 10
Modulation filter bank Q = 2
osses2021
Outer ear model
Middle ear model
Gammatone filter bank
Inner Hair-cell model
Adaptation loops lim = 5
Modulation filter bank Q = 2

Auditory modelling framework

- Extended Dau et al. model:

We will work today with the model internal representations



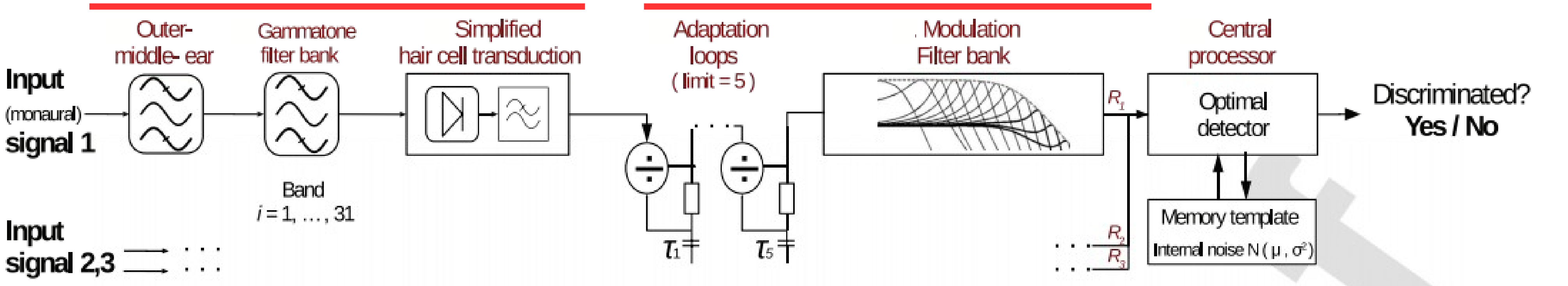


FIG. 1. (Color online) Block diagram of the updated PEMO model. All processing stages and their corresponding set of parameters are explained in the text.

(Figure from Osses & Kohlrausch, 2021, JASA)

See also
auditoryfilterbank
ihcenvelope
adaptloop
modfilterbank
plot_audspecgram
relanoiborra2019
exp_osses2021
exp_osses2022
lopezpoveda2001
dau1996

DAU1997 - Linear filtering for monaural masking (improved)

Usage

```
[outsig, fc] = dau1997(insig,fs);
[outsig, fc] = dau1997(insig,fs,...);
```

Description

Input parameter:

insig : input acoustic signal. fs : sampling rate.

`dau1997(insig,fs)` computes the internal representation of the signal *insig* sampled with a frequency of *fs* Hz.

`[outsig,fc,mfc]=dau1997(...)` additionally returns the center frequencies of the filter bank and the center frequencies of the modulation filterbank.

The model consists of the following stages:

1. a gammatone filter bank with 1-erb spaced filtes.
2. an envelope extraction stage done by half-wave rectification followed by low-pass filtering to 1000 Hz.
3. an adaptation stage modelling nerve adaptation by a cascade of 5 loops.
4. a modulation filterbank

Any of the optinal parameters for auditoryfilterbank, ihcenvelope and adaptloop may be optionally specified for this function. They will be passed to the corresponding functions.

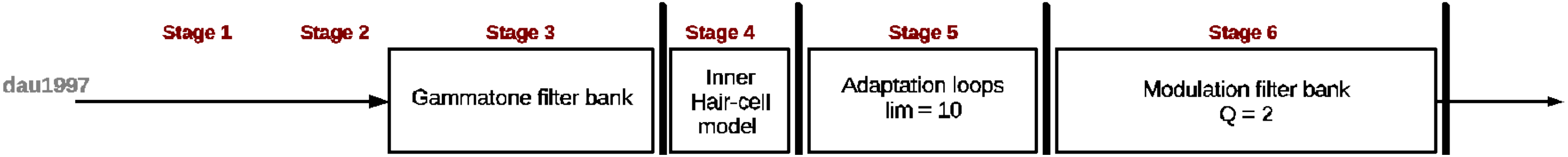
References:

T. Dau, B. Kollmeier, and A. Kohlrausch. Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J. Acoust. Soc. Am.*, 102:2892--2905, 1997a.

T. Dau, B. Kollmeier, and A. Kohlrausch. Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *J. Acoust. Soc. Am.*, 102:2906--2919, 1997b.

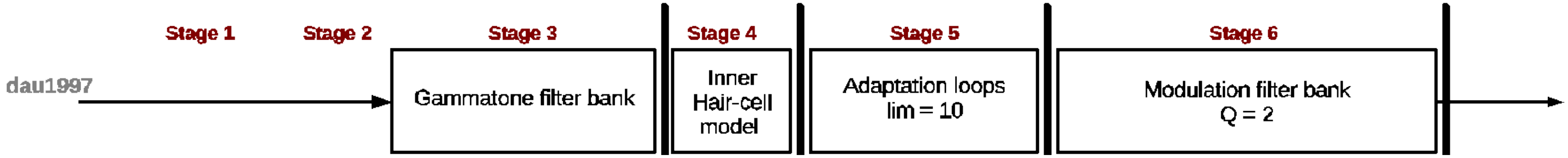
Default parameters for dau1997: Flags

- Outer ear: No
- Middle ear: No
- Gammatone filter bank: Yes (it cannot be by passed)
- Inner hair cell: Yes
- Adaptation loops: Yes
- Modulation filter bank: Yes



Default parameters for dau1997: Keyvals

- Outer ear: No
- Middle ear: No
- Gammatone filter bank: **flow, fhigh, bwmul, basef**
- Inner hair cell: cutofffreq (hard coded), **ihc_filter_order (flag ‘ihc_dau1996’)**
- Adaptation loops: **limit**, **minspl, tau (flag adt_dau1997)**
- Modulation filter bank: **mfc_upper_limit_max (flag mfb_dau1997)**



Where to find this information?

```
definput.import={'auditoryfilterbank','ihcenvelope','adaptloop','modfilterbank'};
definput.importdefaults={'afb_dau1997', 'ihc_dau1996', 'adt_dau1997','mfb_jepsen2008'};
definput.keyvals.subfs=[];

[flags,keyvals]  = ltfatarghelper({'flow','fhigh'},definput,varargin);
```

"Use these configurations" if other input parameters are not specified:

arg_auditoryfilterbank.m	afb_dau1997
arg_ihcenvelope.m	ihc_dau1996
arg_adaptloop.m	adt_dau1997
arg_modfilterbank.m	mfb_jepsen2008

Group of flags / values

home ▸ alejandro ▸ Desktop ▸ AMT-workshop ▸ amtoolbox-full-1.1.0 ▸

dro/Desktop/AMT-workshop/amtoolbox-full-1.1.0/defaults/arg_modfilterbank.m

```
% Attenuation factor applied to mod filters above 10 Hz (only applied if phase_insens
definput.flags.att_factor       = {'att_factor','no_att_factor'};

definput.keyvals.Q_mfb = 2;
definput.keyvals.mfc_upper_limit_max = 1000; % Hz, maximum upper limit

definput.groups.mfb_dau1997 = { 'no_mfc_upper_limit',...
                                'no_LP_150_Hz', ...
                                'no_att_factor'};

definput.groups.mfb_verhey1999 = {'mfc_upper_limit',...
                          'no_LP_150_Hz', ...
                          'no_att_factor'};

definput.groups.mfb_jepsen2008 = {'mfc_upper_limit',...
                          'LP_150_Hz', ...
                          'att_factor'};

definput.groups.mfb_king2019 = {'mfc_upper_limit',...
                          'no_LP_150_Hz', ...
                          'no_att_factor'};

definput.groups.mfb_osses2021_att_gain = {'mfc_upper_limit',...
                          'LP_150_Hz_att', ...
                          'att_factor'}; % corresponds to the "att. gain" condition fr
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

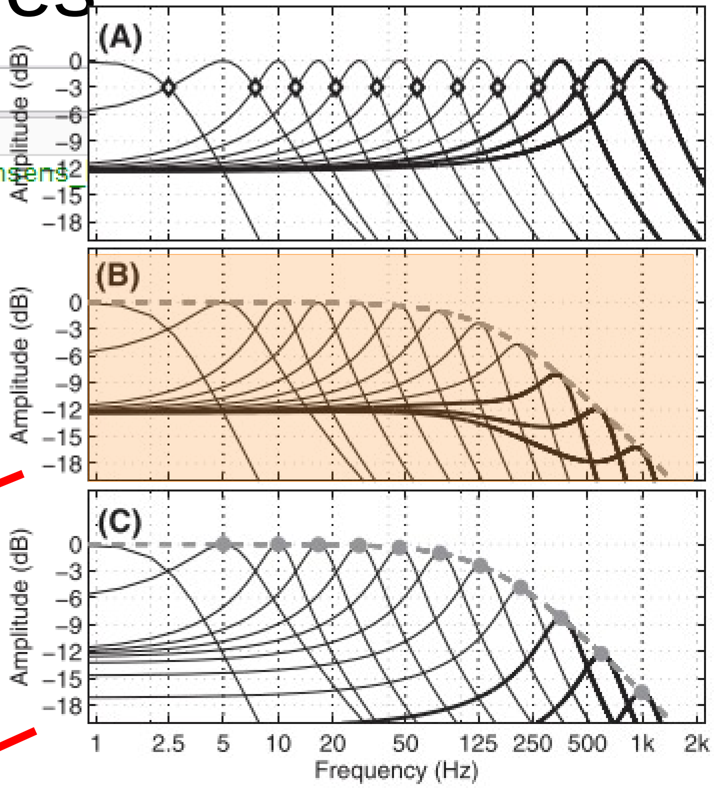


FIG. 14. MTFs of the modulation filter bank without (A) and with (B) the 150-Hz LPF filter. In (C), MTFs using a different implementation of the 150-Hz LPF are shown (see model configuration "mod att. gain," in the text). The MTFs of filters 10–12, which have frequencies $mf_c > 250$ Hz, are drawn in thick black lines. The cut-off frequencies of Table II are indicated by open diamonds in (A). The transfer function of the 150-Hz LPF is shown by the dashed gray line.

Running the model with different configurations

dau1997

Command Window

```
>>
[outsig,fc,mfc] = dau1997(insig,fs); % dau1997 with default parameters
[outsig,fc,mfc] = dau1997(insig,fs,'flow',200,'fhigh',4000); % dau1997 with custom limits for fc between 200 and 4000 H
[outsig,fc,mfc] = dau1997(insig,fs,'mfb_dau1997');
[outsig,fc,mfc] = dau1997(insig,fs,'outerear','middleear');
[outsig,fc,mfc] = dau1997(insig,fs,'basef',2000);
[outsig,fc,mfc] = dau1997(insig,fs,'flow',2000,'fhigh',2000,'basef',2000);
>>
```

The task of today
Process a wave file and obtain its “internal representation”

1) Install the AMT toolbox: https://amtoolbox.org/download.php
2) Pick up one of the eight models
3) Pick up one sound and load it using audioread: suggested names “insig” for the signal, and “fs” for the sampling frequency:

   ababa1.wav (French utterance): Download here

   greasy.wav: type [insig,fs] = greasy;
4) Obtain the “internal representation” by running the model

- Force the model to have a band centred at 200 Hz
- Plot the model output of the band at 200 Hz
- Run the model again but now only up to the adaptation stage (for zilany2014 and bruce 2018 this is already the last stage)

Hands on:
Use of monaural auditory models from the AMT toolbox

3rd AMT Workshop

Turin, 8 September 2023

Alejandro Osses, Ph.D.

Posdoctoral researcher, ENS Paris, France
Acoustic researcher at Sorama BV, Eindhoven, the Netherlands

ale.a.osses@gmail.com, alejandro.osses@sorama.eu

The task of today
Process a wave file and obtain its “internal representation”

1) Install the AMT toolbox: https://amtoolbox.org/download.php
2) Pick up one of the eight models
3) Pick up one sound and load it using audioread, suggested names “insig” for the signal, and “fs” for the sampling frequency:

   ababa1.wav (French utterance): Download here

   greasy.wav: type [insig,fs] = greasy;
4) Obtain the “internal representation” by running the model

- Force the model to have a band centred at 200 Hz
- Plot the model output of the band at 200 Hz
- Run the model again but now only up to the adaptation stage (for zilany2014 and bruce 2018 this is already the last stage)

The task of today
Step by step

1) Install the AMT toolbox: https://amtoolbox.org/download.php

- I recommend “Download the AMT full package” (follow the instructions of the website): i.e., open MATLAB and type amt_start;
- If needed, type amt_mex;

The task of today
Step by step

2) Pick up one of the eight models

- If you pick up verhulst2015 or verhulst2018, note that these models require Python preinstalled and that they are slow models, so at first, use short sounds to be processed.
- zilany2014, bruce2018, verhulst2015, verhulst2018 have many outputs and configurable parameters
- dau1997, king2019, relanoiborra2019_featureextraction, and osses2021 share a similar structure (as I showed earlier in this presentation). **So, if you don’t know which model to use, pick up one of these models.**

The task of today
Step by step

3) Pick up one of sound:

   ababa1.wav (French utterance): Download here

   greasy.wav: type [insig,fs] = greasy; (this option doesn't require a download)

4) Obtain the "internal representation" by running the model

- Force the model to have a band centred at 200 Hz
- Plot the model output of the band at 200 Hz
- Run the model again but now only up to the adaptation stage (for zilany2014 and bruce 2018 this is already the last stage)

The task of today
Step by step

4) Obtain the “internal representation” by running the model

- Force the model to have a band centred at 200 Hz.

  dau1997 or similar: This requires to specify the keyval ‘basef’

  zilany2014, bruce2018: set the input parameter cf to 200 Hz

  verhulst2015, verhulst2018: use either (1) cf_flag=‘abr’ and then get the closest frequency to 200 Hz, or (2) use cf_flag=200

The task of today
Step by step

4) Obtain the “internal representation” by running the model

- Plot the model output of the band at 200 Hz
- **dau1997** or similar: You get the modulation filter bank outputs, plot one of the bands (the one you want, look at the parameter mfc) or several bands superimposed.
- **zilany2014, bruce2018**: You get the auditory-nerve synapse outputs. For simplicity, I suggest to plot the “mean_rate” outputs
- **verhulst2015, verhulst2018**: You get several inferior colliculus (IC) outputs.
- Run the model again but now only up to the adaptation stage (for zilany2014 and bruce 2018 this is already the last stage)

The task of today
Step by step

4) Obtain the “internal representation” by running the model

- Run the model again but now only up to the adaptation stage.
- For this stage you need to identify the name of the last module, so that you can bypass the flag (e.g., ‘mfb’ needs to be specified as ‘no_mfb’)