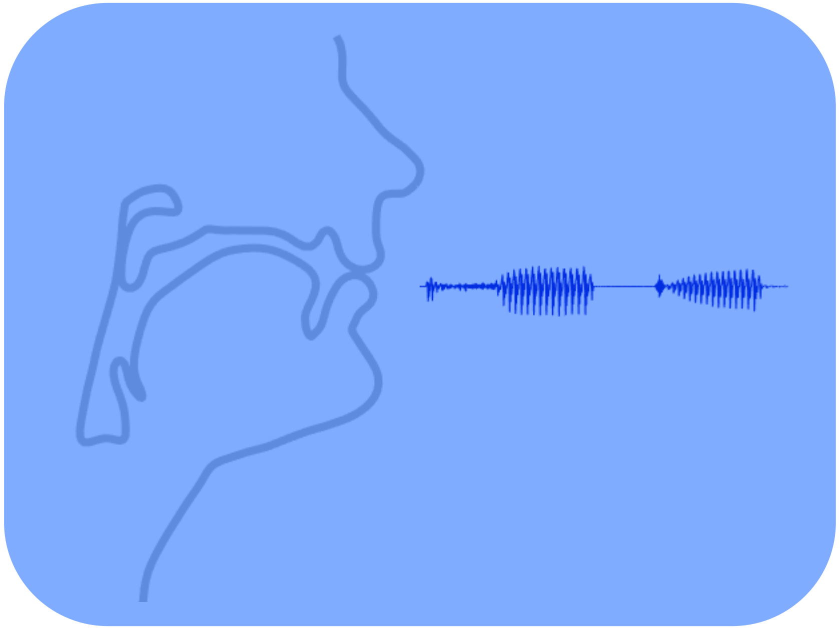
DIVA model

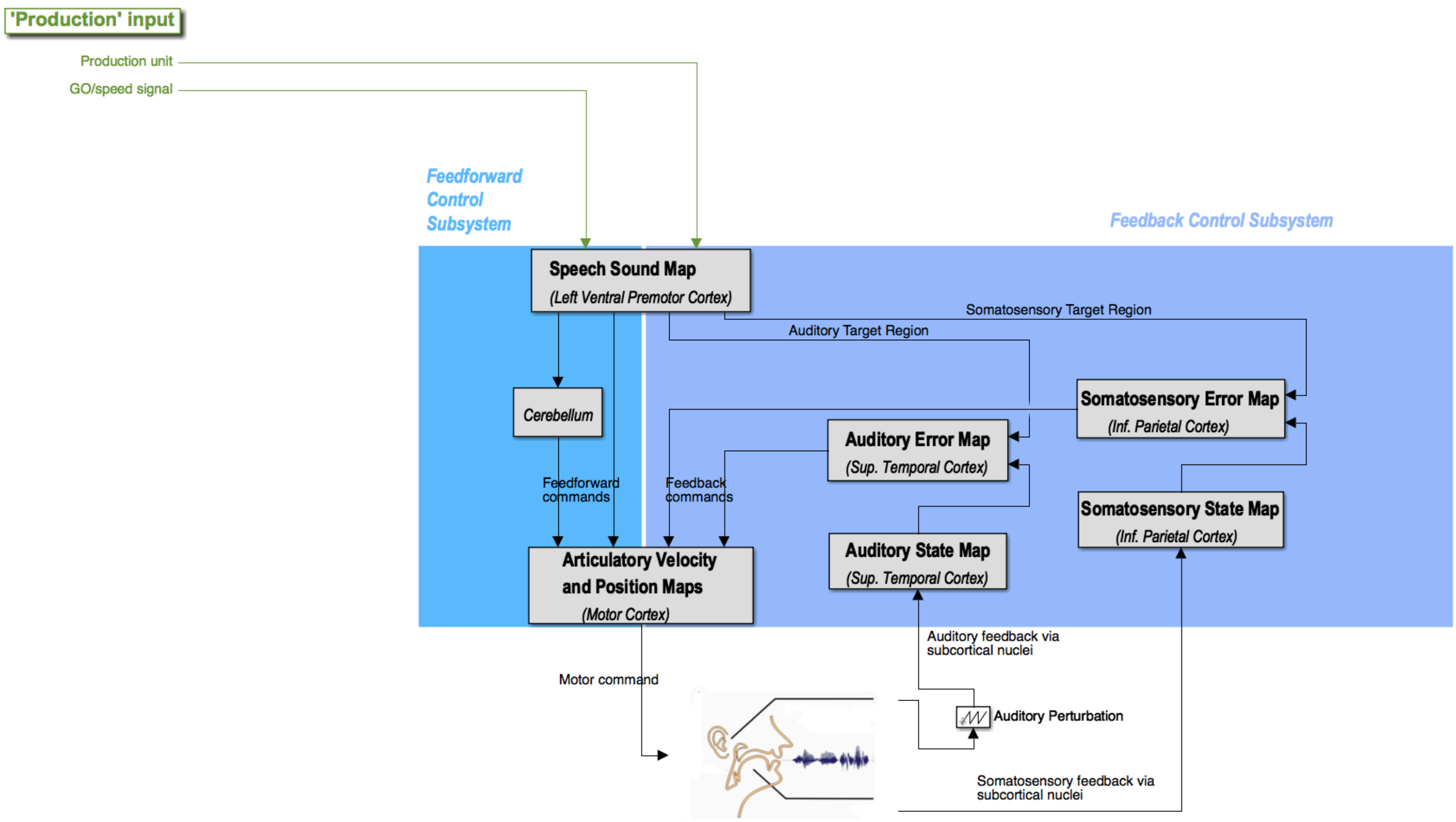
Matlab/Simulink implementation

SH680 Lecture 5



DIVA model overview

DIVA model overview

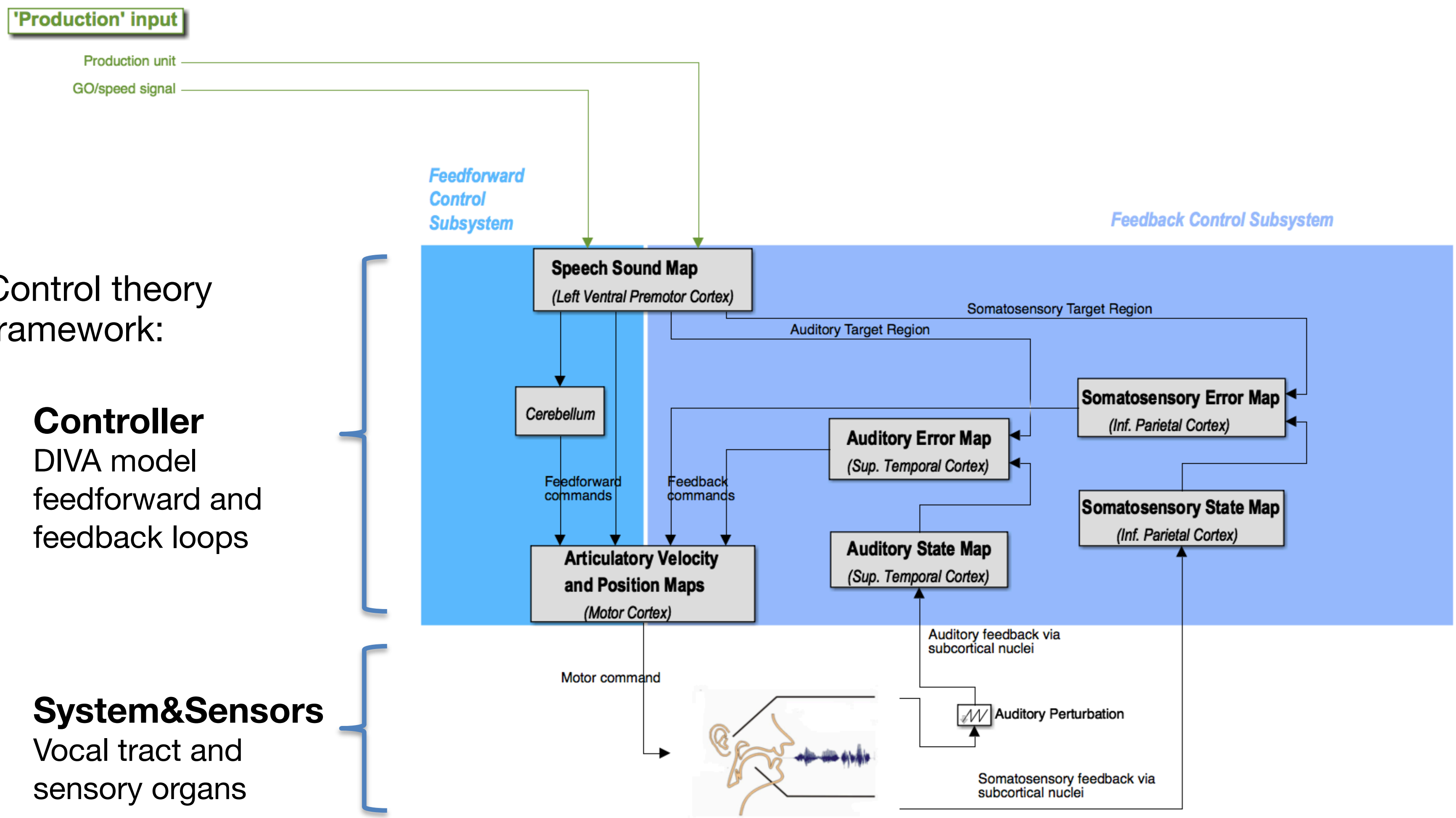


DIVA model overview

Control theory framework:

Controller
DIVA model feedforward and feedback loops

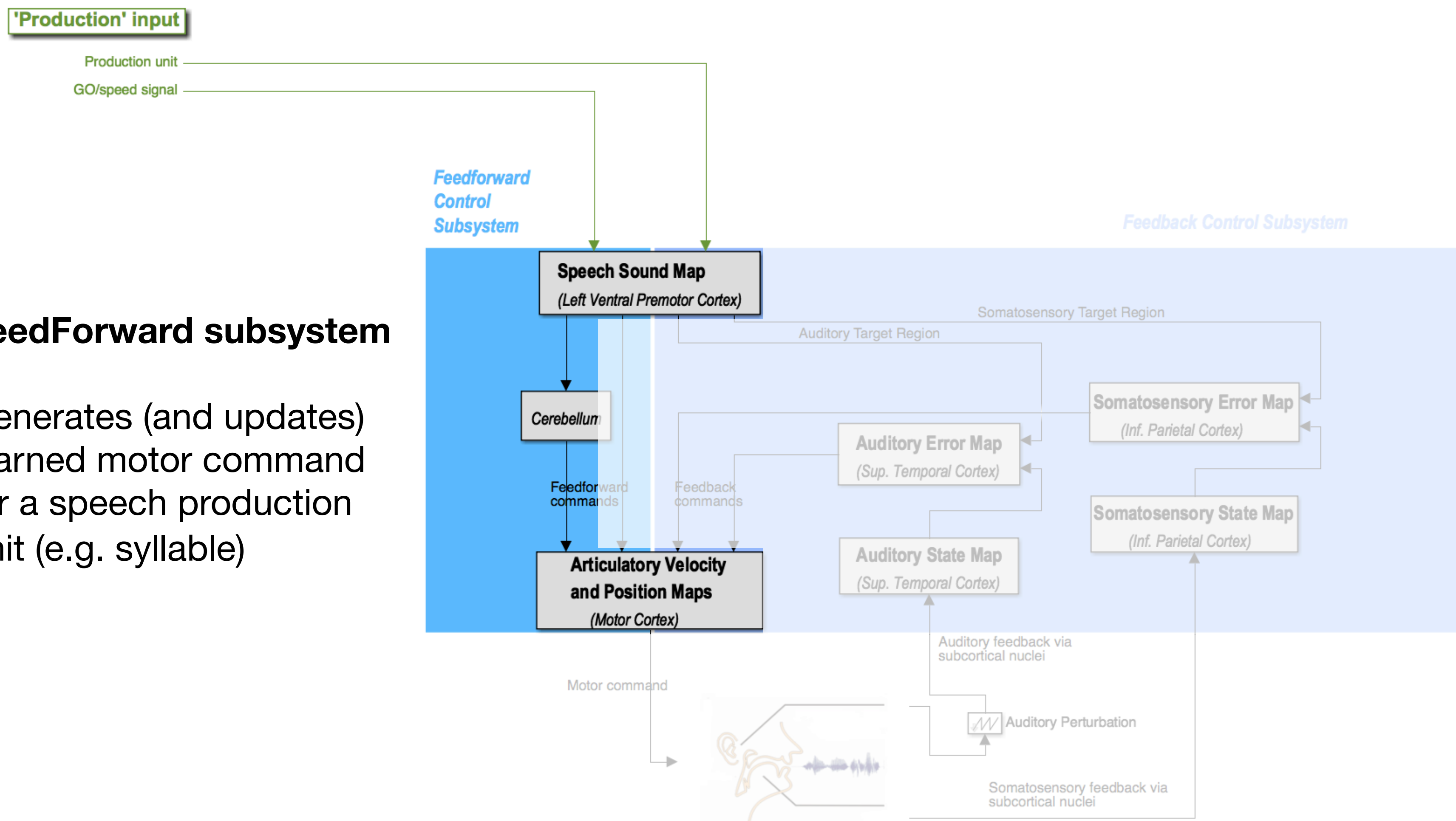
System&Sensors
Vocal tract and sensory organs



DIVA model overview

FeedForward subsystem

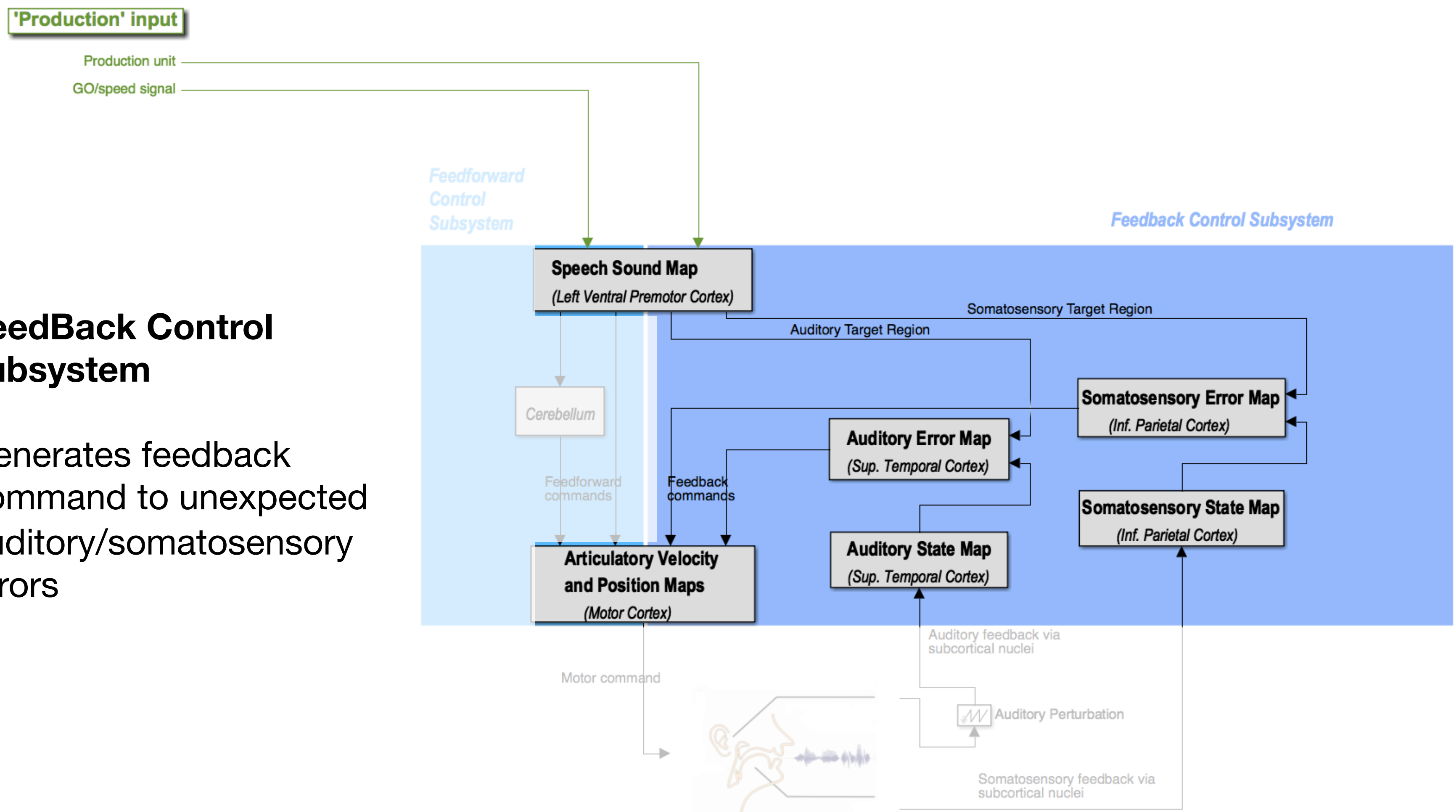
Generates (and updates) learned motor command for a speech production unit (e.g. syllable)



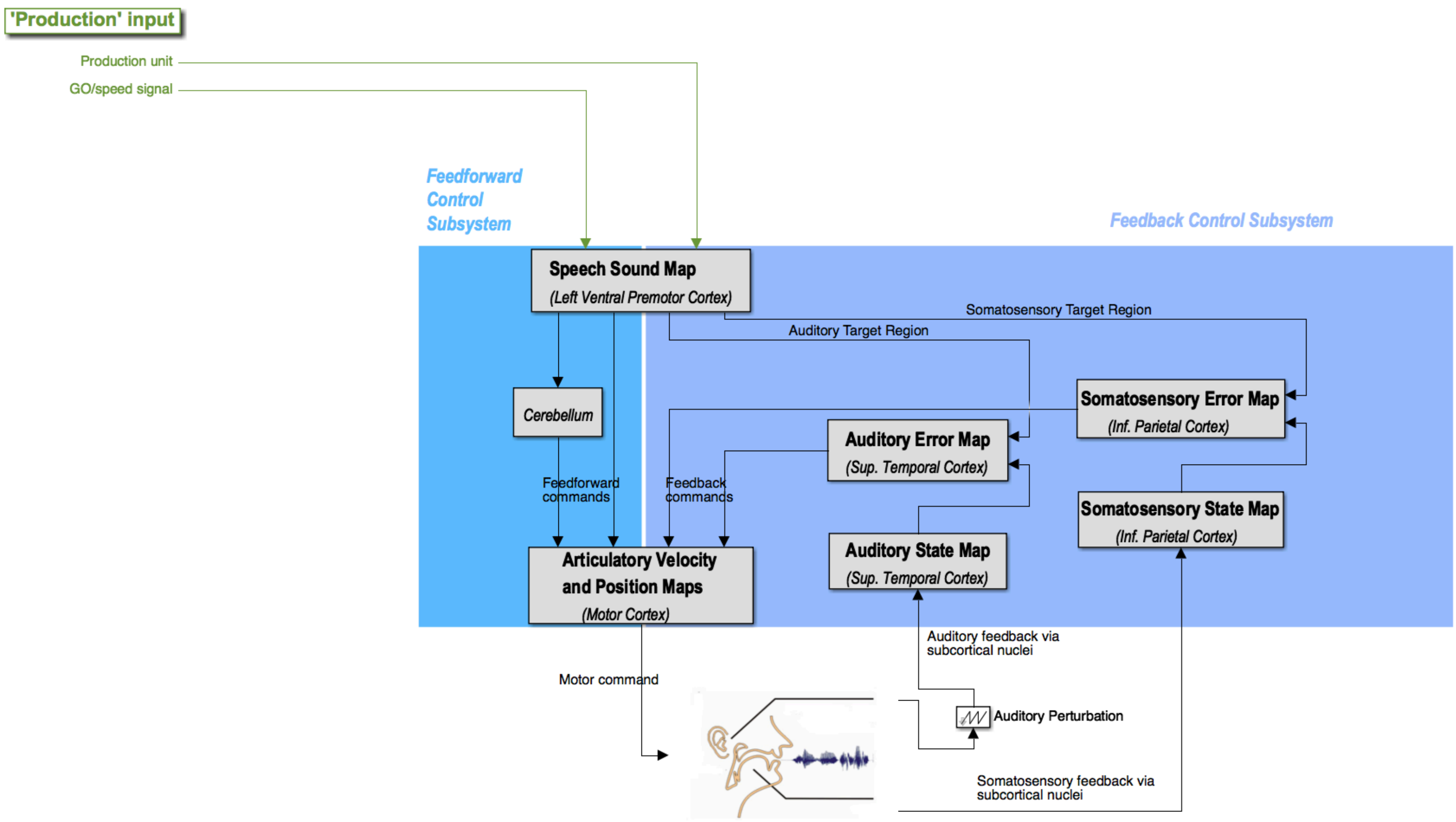
DIVA model overview

FeedBack Control subsystem

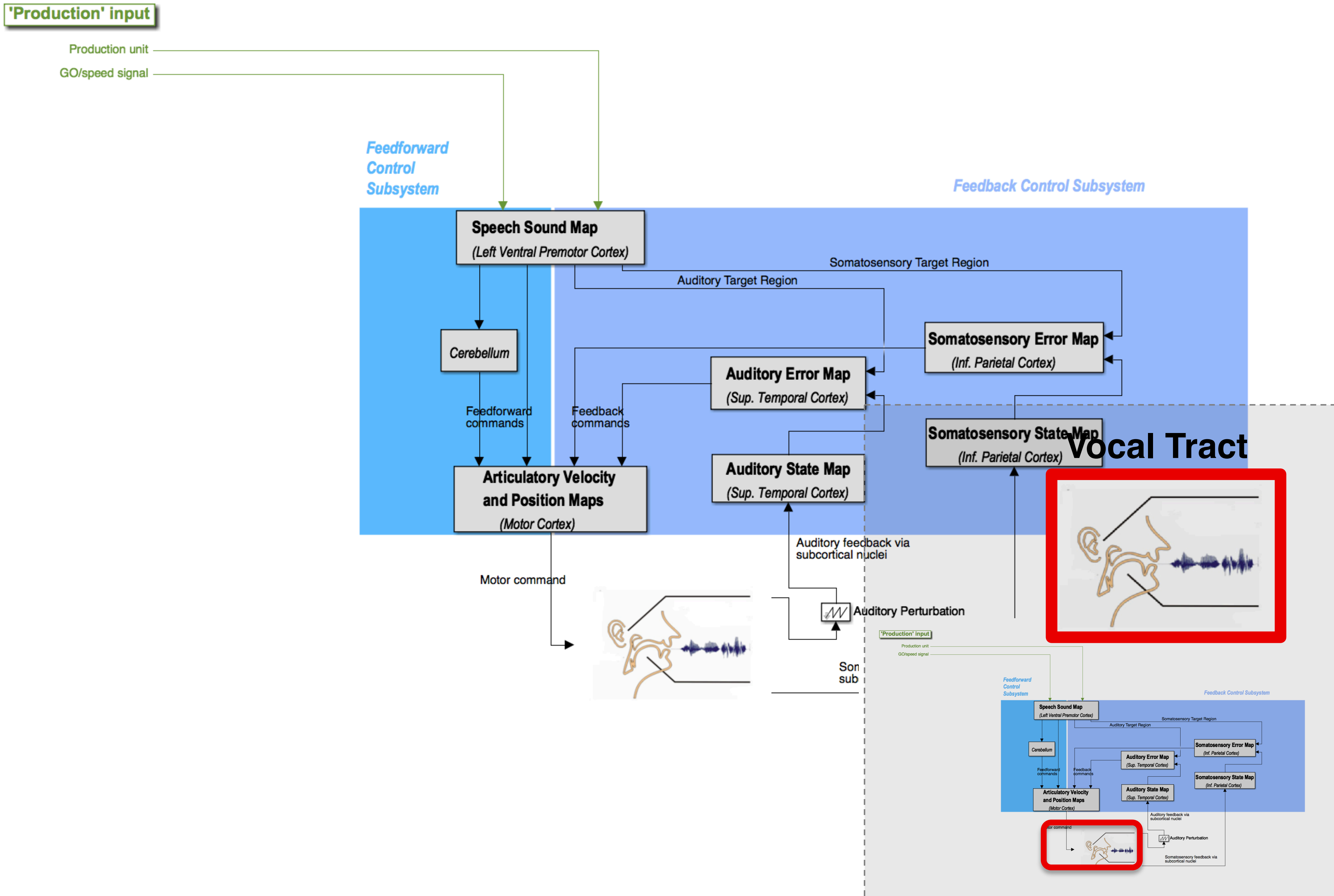
Generates feedback command to unexpected auditory/somatosensory errors



DIVA model components

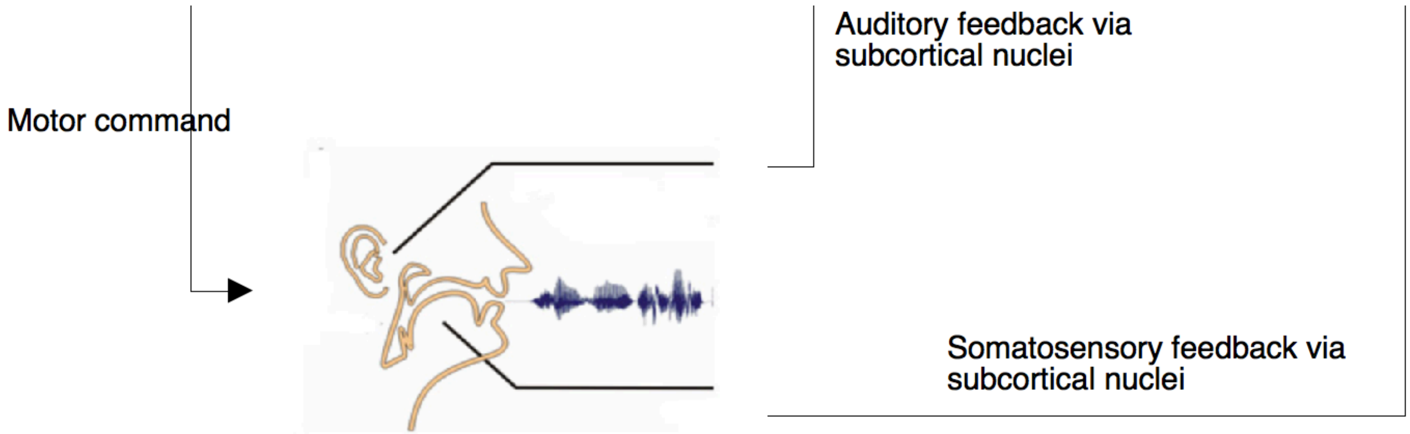


DIVA model components
Vocal Tract



DIVA model components

Vocal Tract



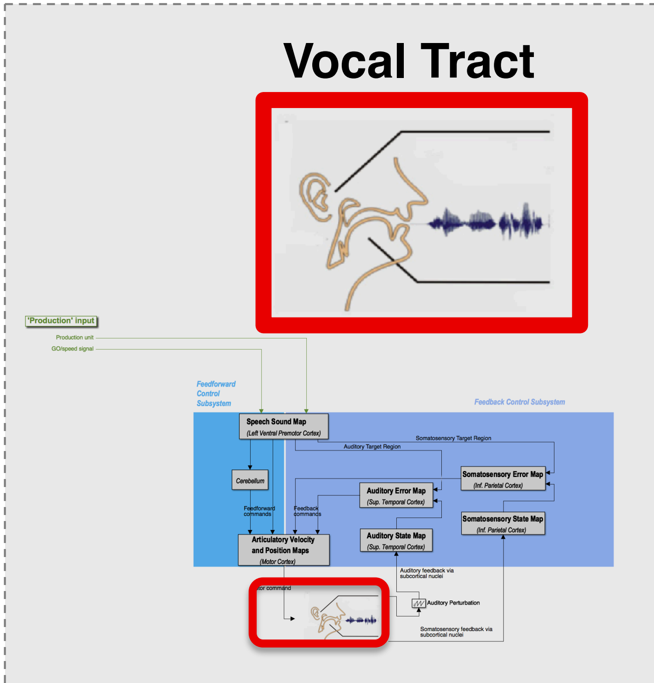
Vocal tract model

Dynamic model relating vocal tract articulatory position commands and associated soundwave and perceptual signals (auditory/somatosensory)

Input: motor command, defining vocal tract shape and glottal source characteristics

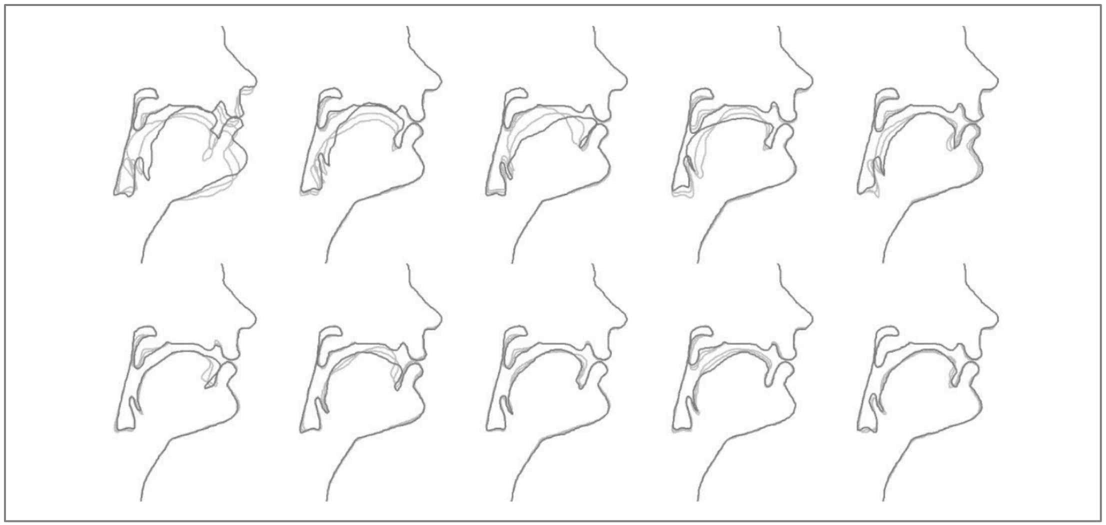
[area function tube model of vocal tract cavity + glottal model = radiated soundwave filter]

Output: acoustic and somatosensory representation of produced sound



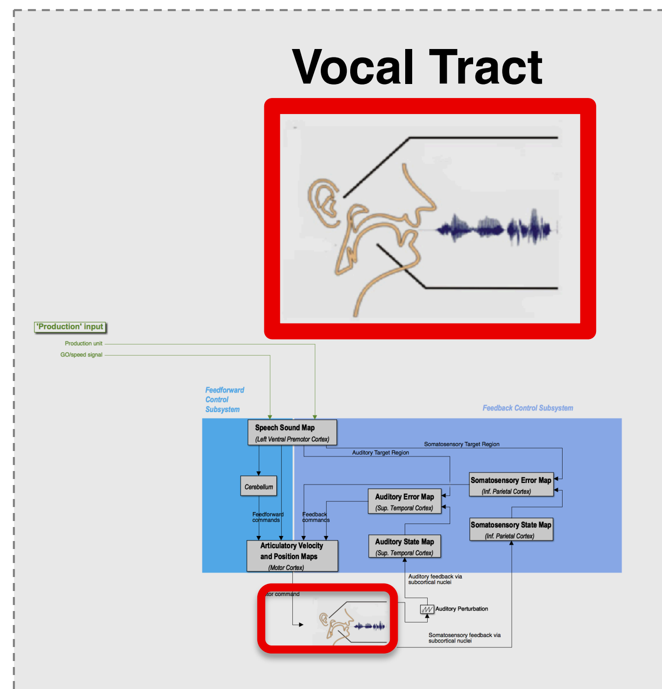
DIVA model components

Vocal Tract

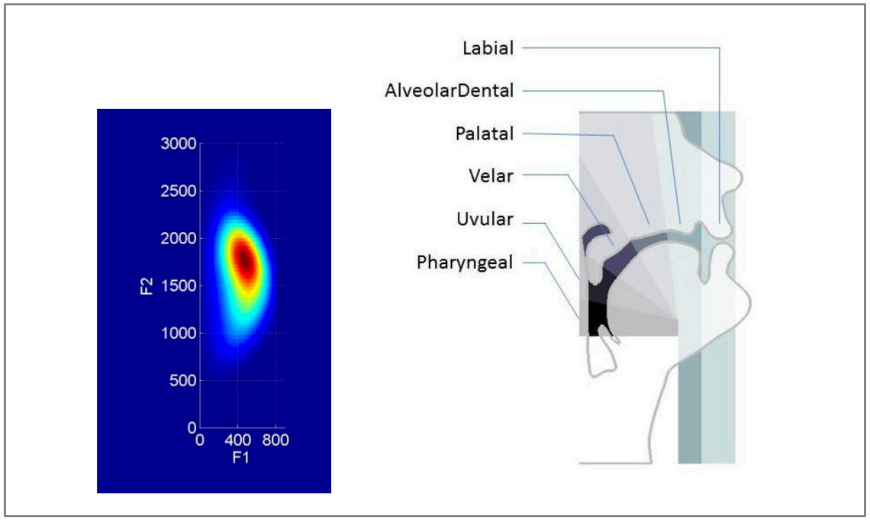


Motor representation (input to Vocal Tract module; 13 dimensions):

Ten *vocal tract shape* dimensions (shown above), plus **three** *source* dimensions (not shown: *tension*, *pressure*, and *voicing*)

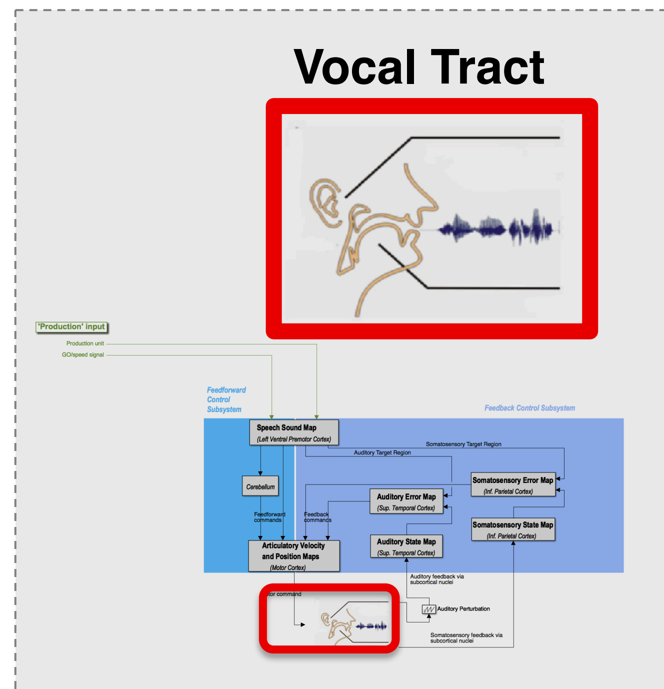


DIVA model components
Vocal Tract



Somatosensory/Auditory representation
(output of Vocal Tract module, 12 dimensions):

Four *formant* dimensions (F0, F1, F2, and F3), **six** *place of articulation* dimensions (shown above), and **two** *source* dimensions (not shown: *pressure*, and *voicing*)



DIVA model components

Vocal Tract

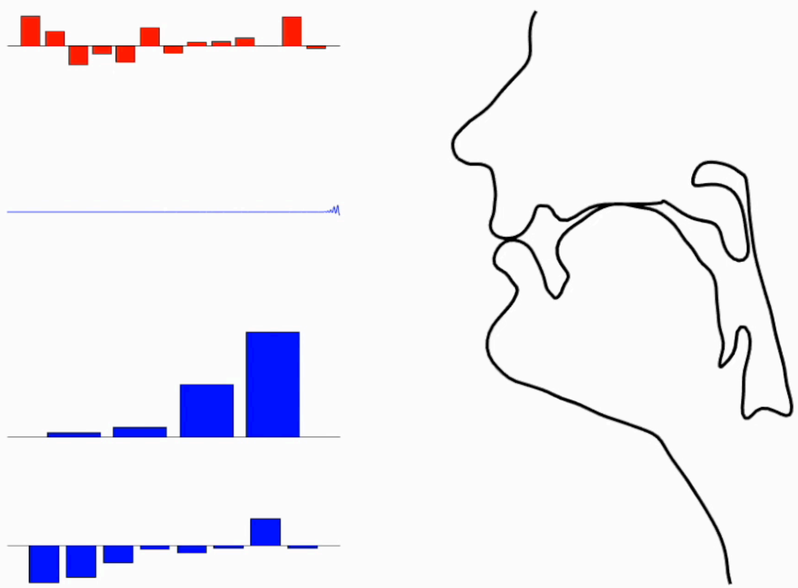
Motor input

Outputs

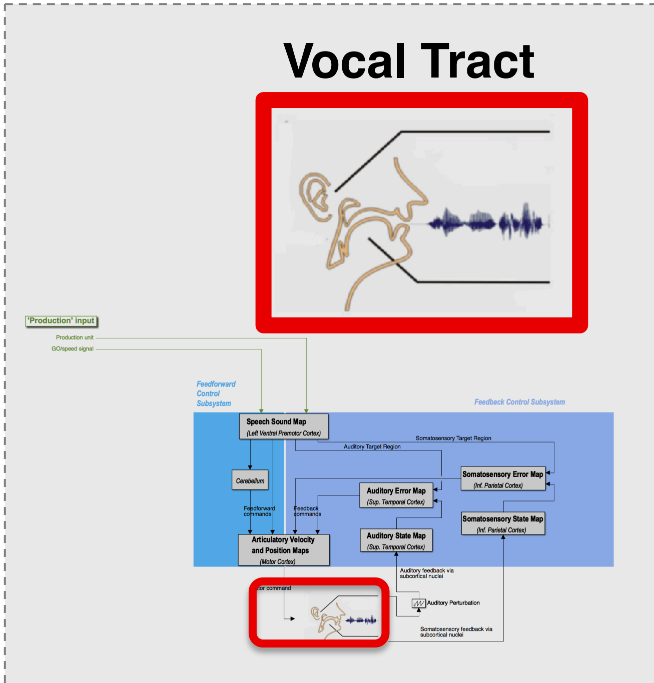
Soundwave

Auditory
(formants)

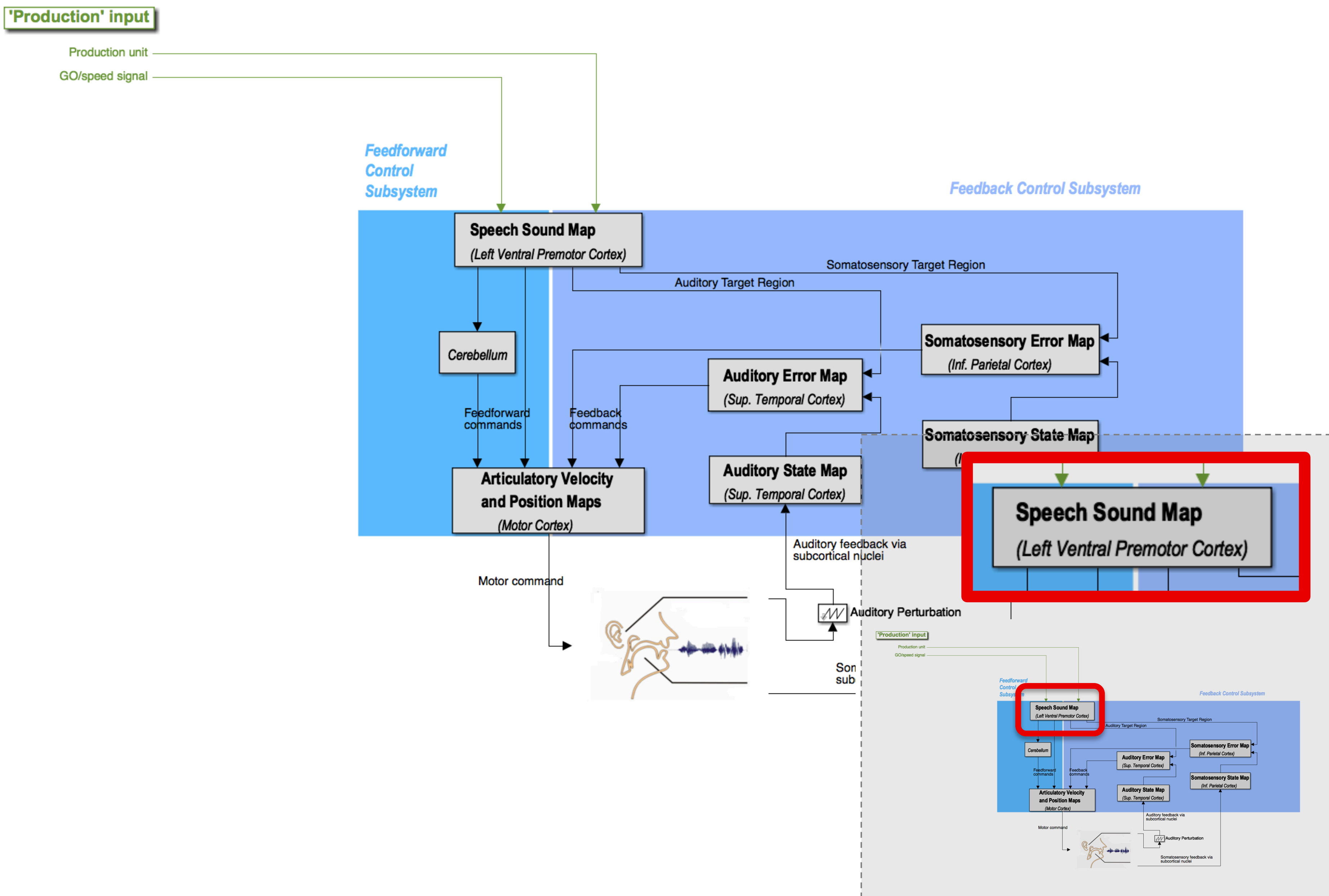
Somatosensory



Example of random articulatory movements

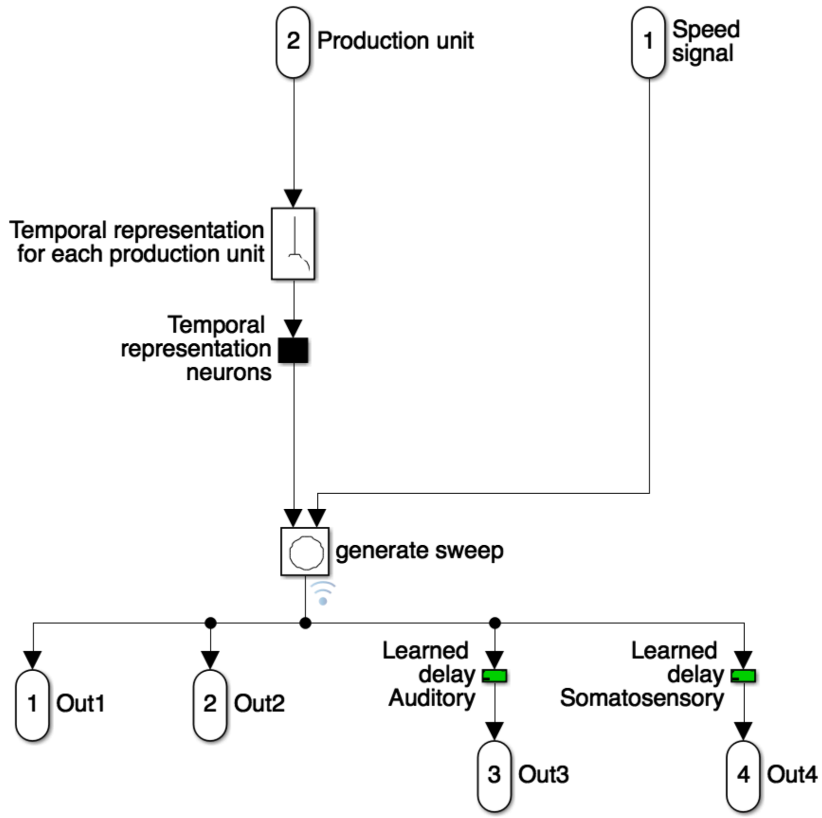


DIVA model components
Speech Sound Map



DIVA model components

Speech Sound Map



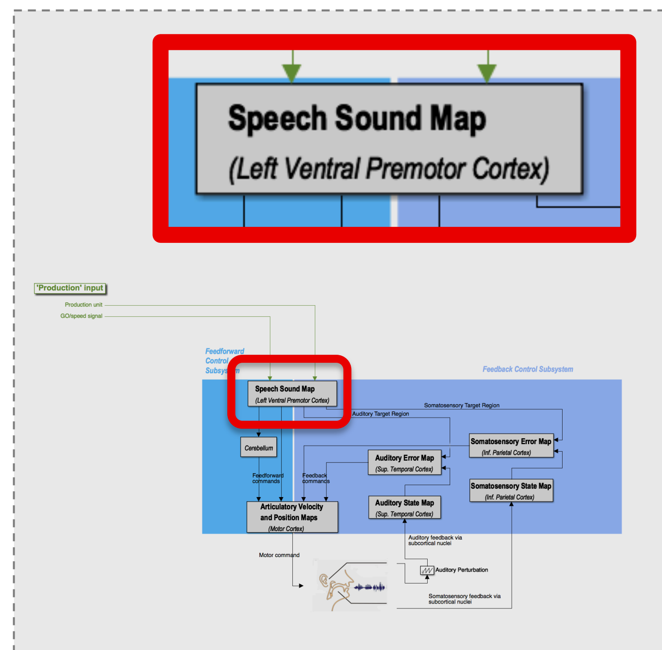
Key symbols in block diagrams

Input/Output of block module

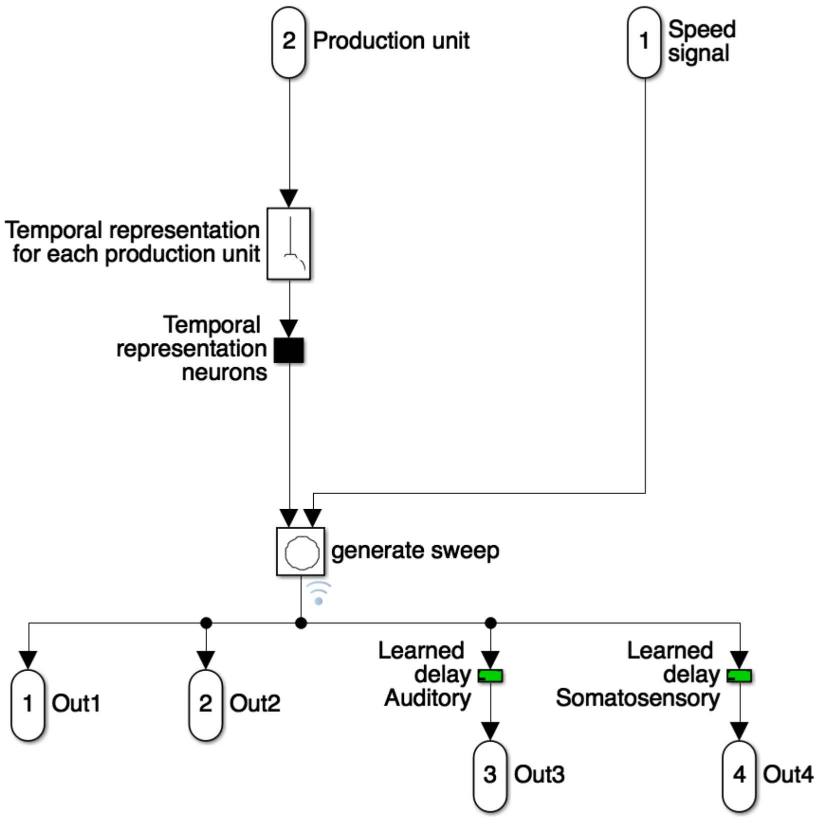
Weights/projections between two layers of neurons

Layer of neurons (user-editable activation level)

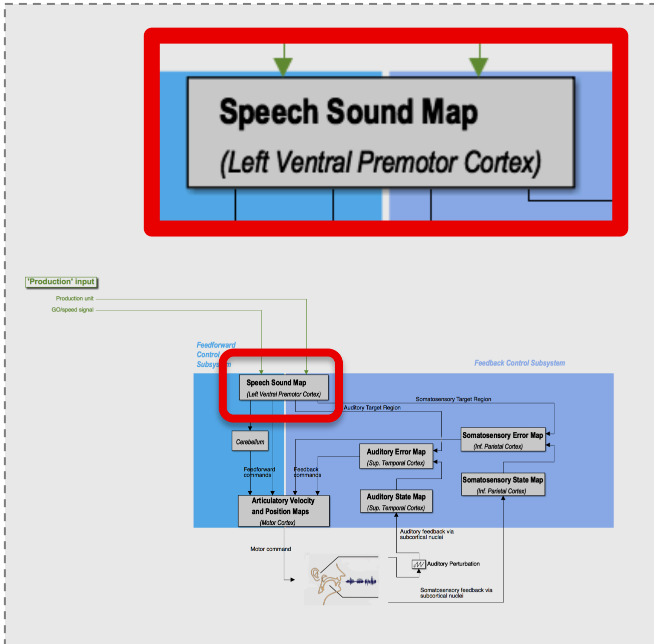
Delays (user-editable delay amount)



DIVA model components
Speech Sound Map

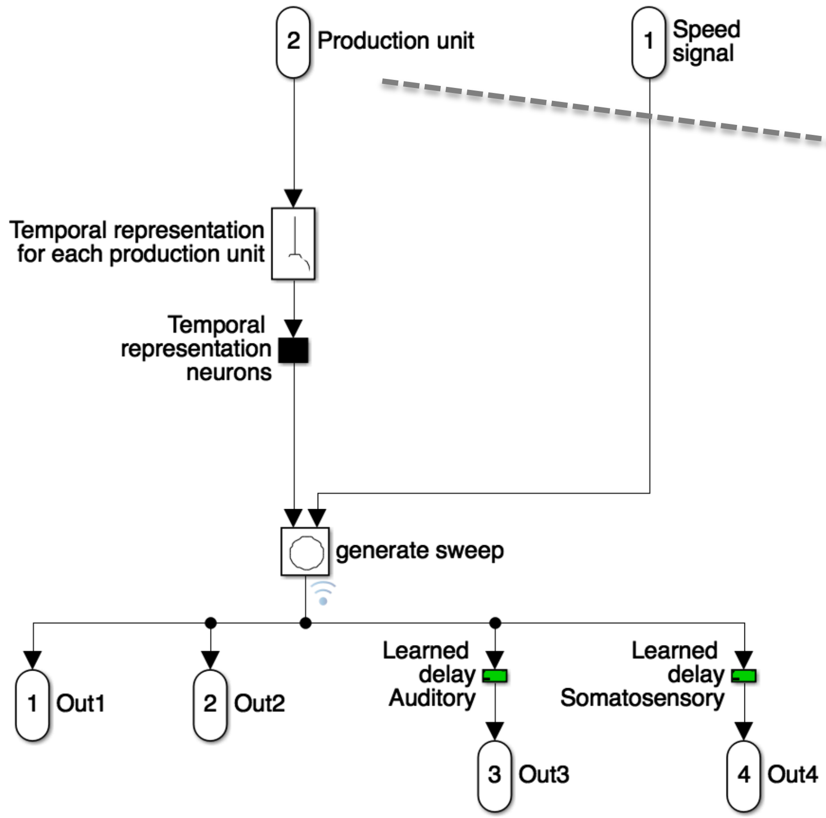


Neurons in **Premotor Cortex (Speech Sound Map)** encode learned speech productions. They initiate the production of a learned sequence, and they project to motor and sensory cortices

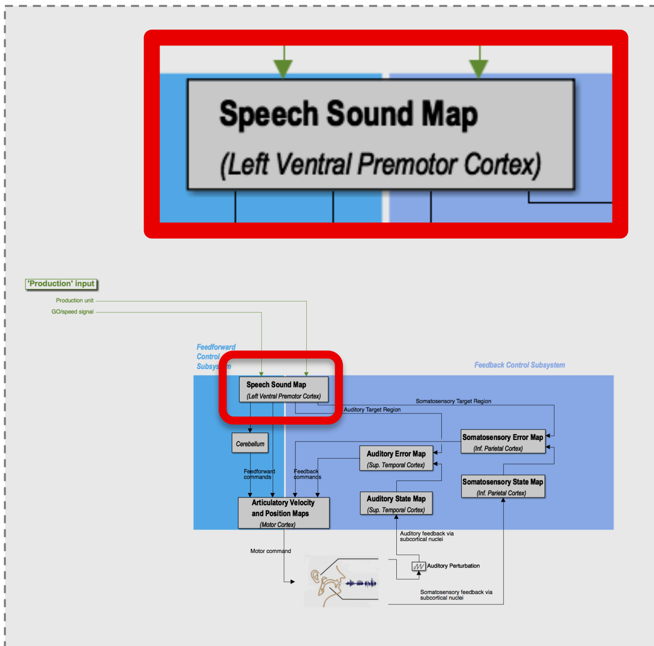


DIVA model components

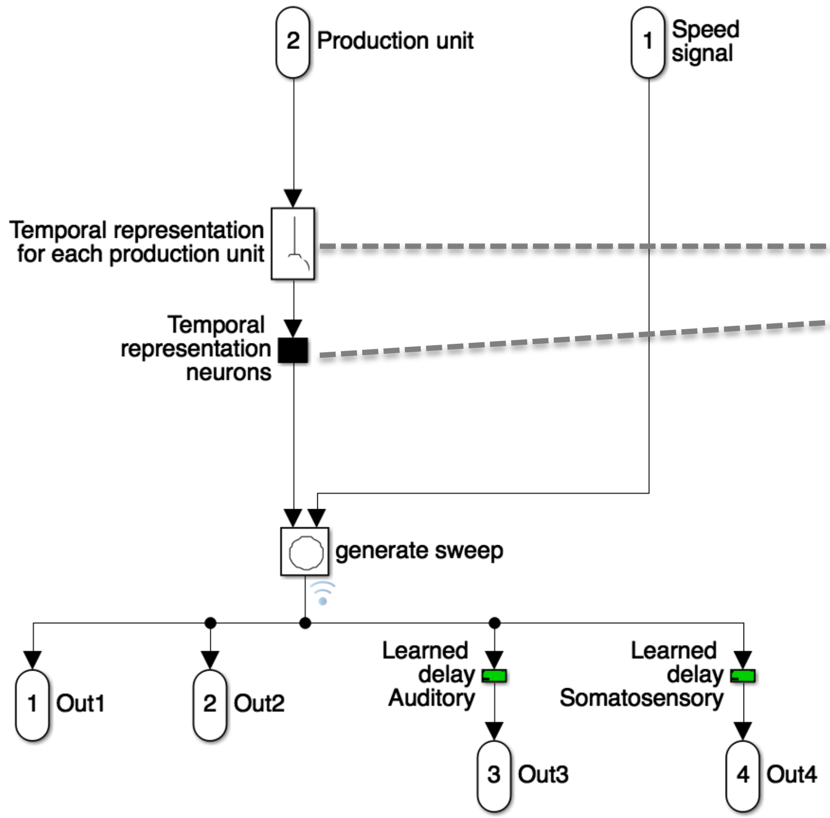
Speech Sound Map



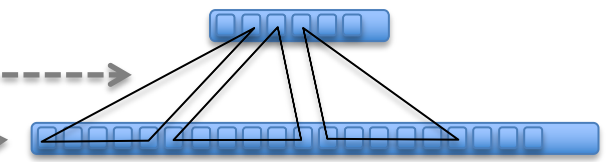
Production units: each neuron represents a different learned production



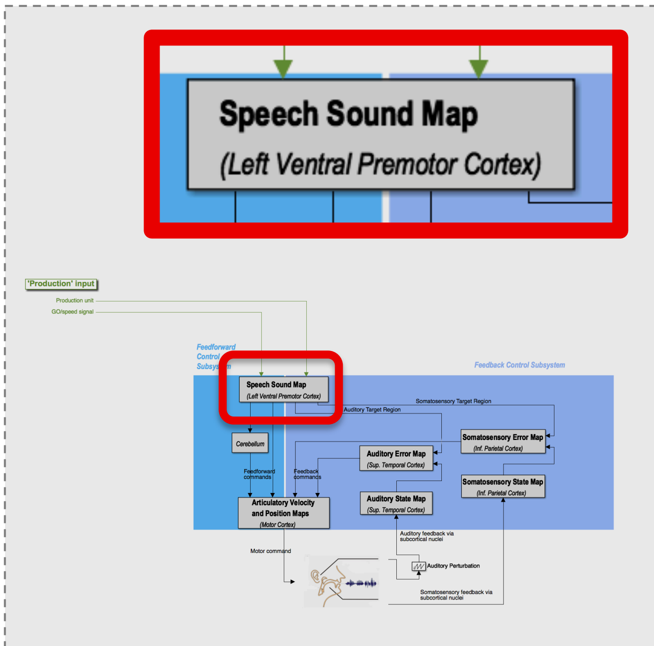
DIVA model components
Speech Sound Map



Production units: each neuron represents a different learned production

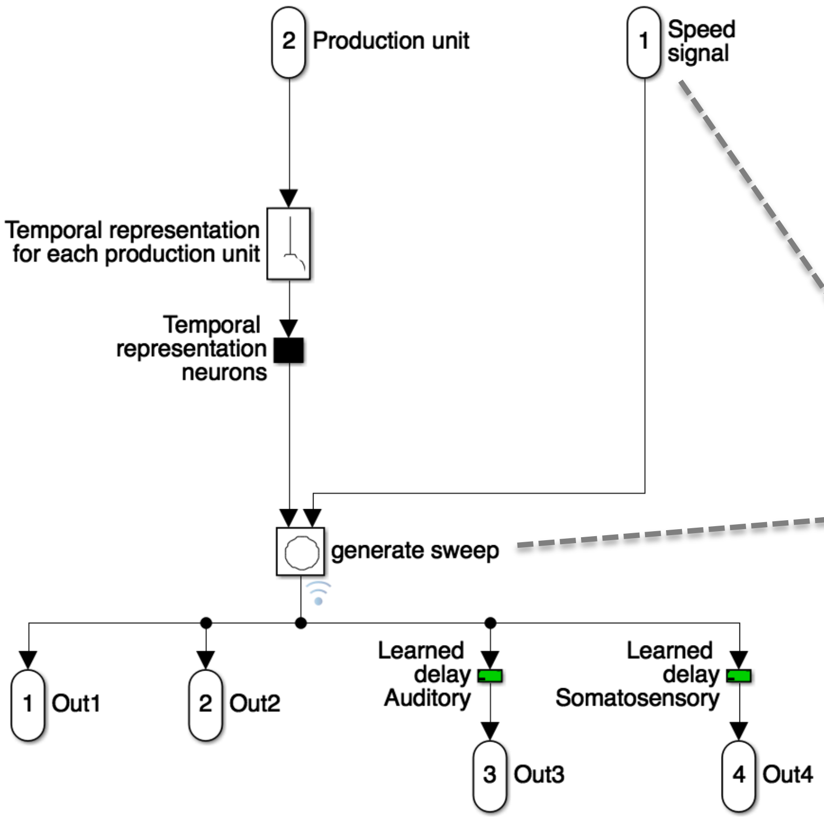


Temporal representation units: each neuron represents a different timepoint of a learned production (spatial encoding of temporal information)

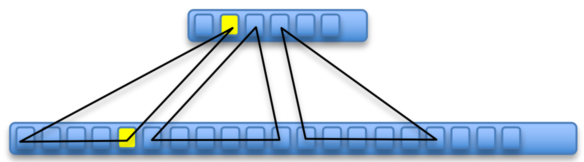


DIVA model components

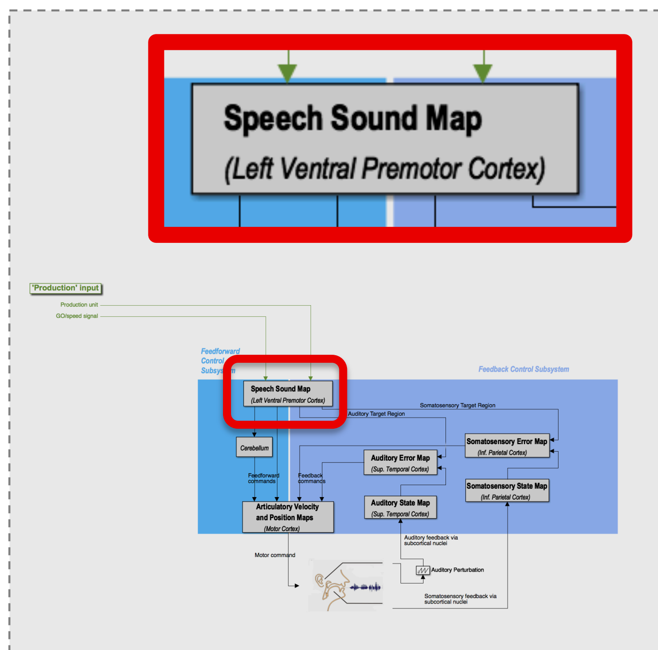
Speech Sound Map



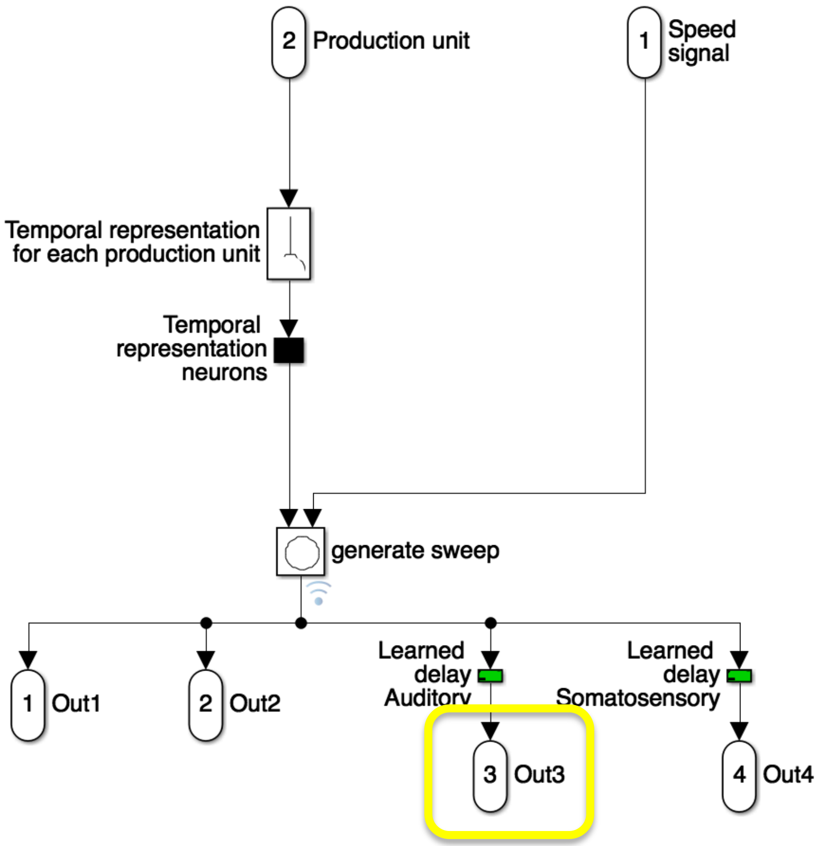
Production units: each neuron represents a different learned production



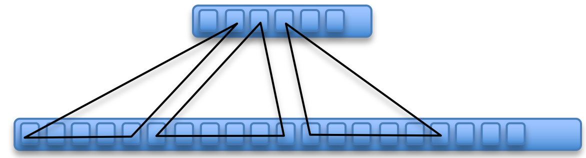
Temporal representation units: each neuron represents a different timepoint of a learned production (spatial encoding of temporal information). Temporal representation units are activated sequentially, at a speed controlled by the *speed signal* model input



DIVA model components
Speech Sound Map

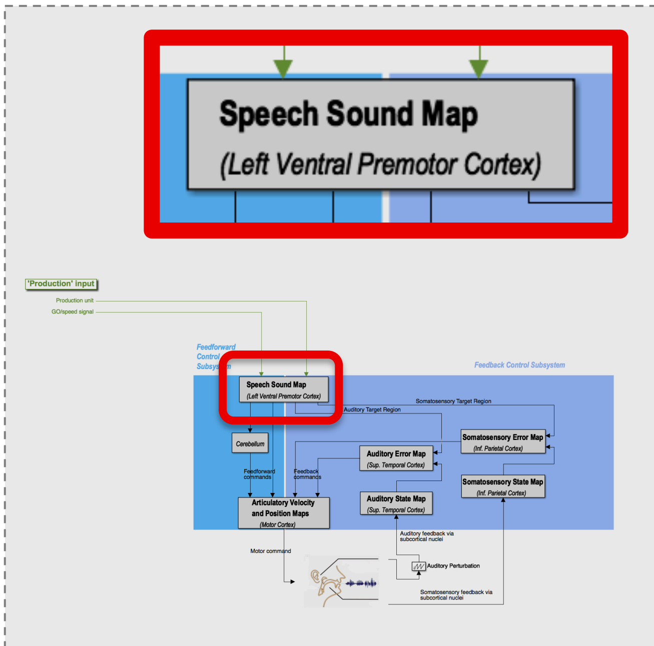


Production units: each neuron represents a different learned production



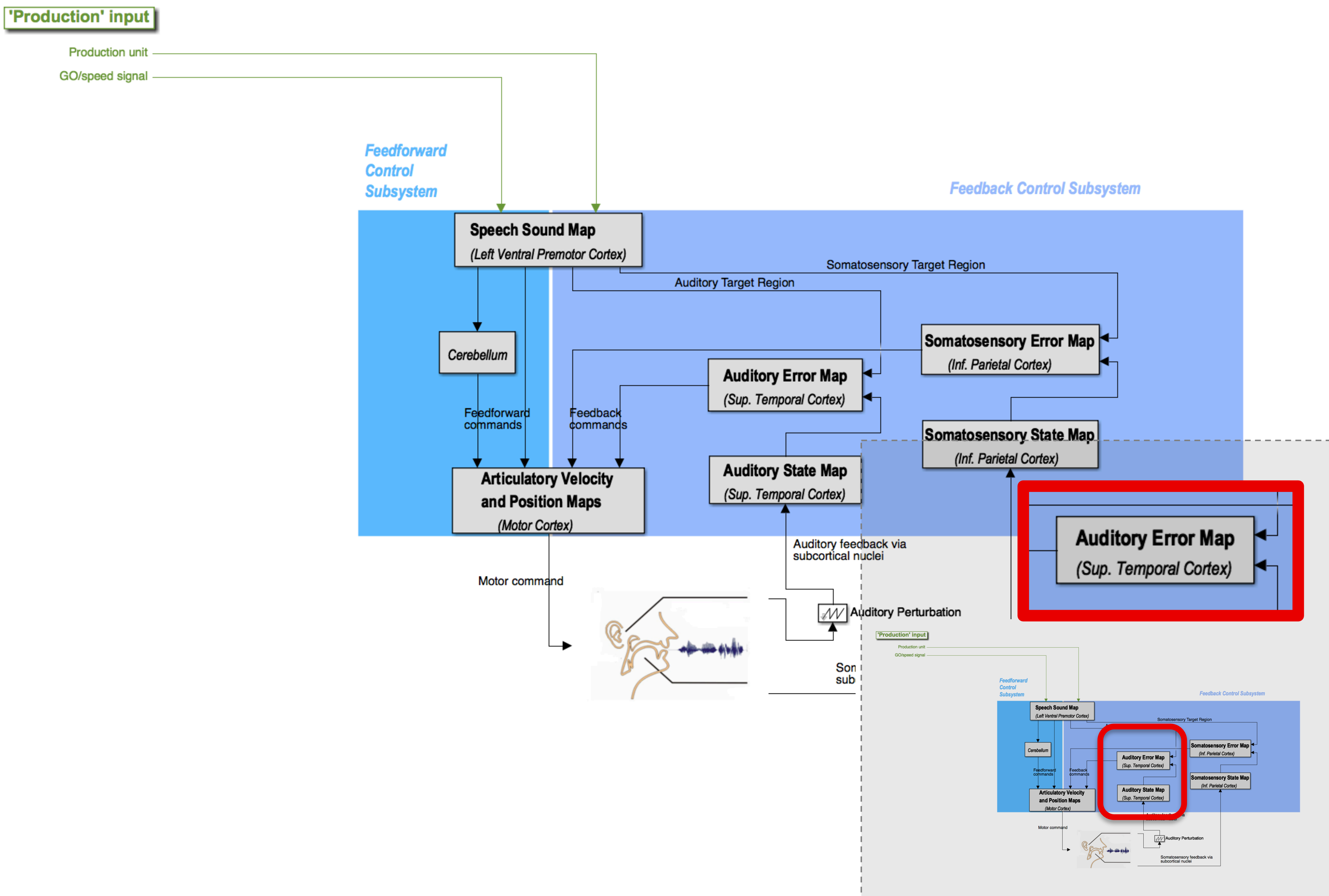
Temporal representation units: each neuron represents a different timepoint of a learned production (spatial encoding of temporal information). Temporal representation units are activated sequentially, at a speed controlled by the *speed signal* model input

Output of SSM is projected to motor and sensory cortices to encode (through a series of weights defined in the following DIVA modules) the expected feedforward command at each timepoint for a given production target, as well as the expected auditory and somatosensory signals associated with this target



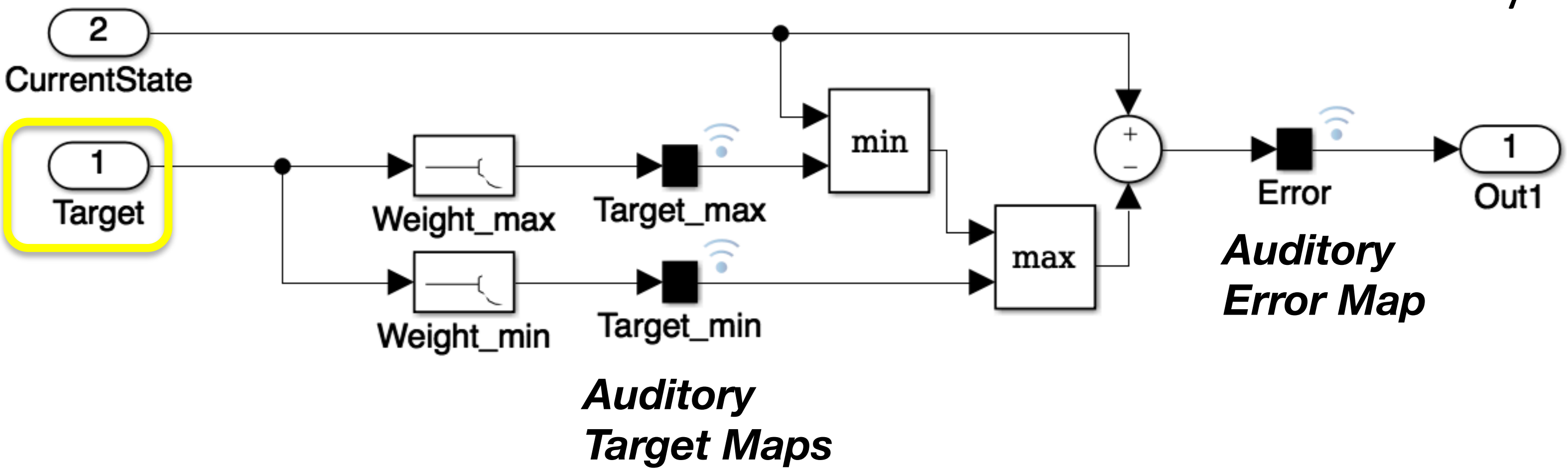
DIVA model components

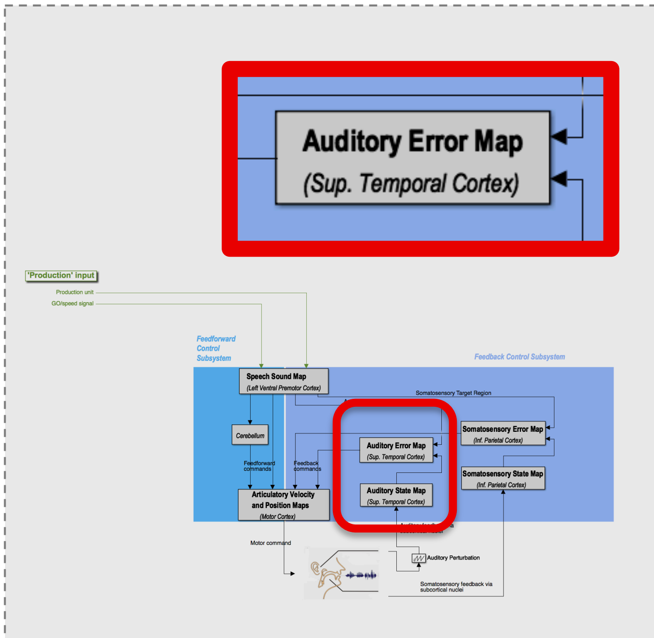
Auditory Error Map



DIVA model components

Auditory Error Map

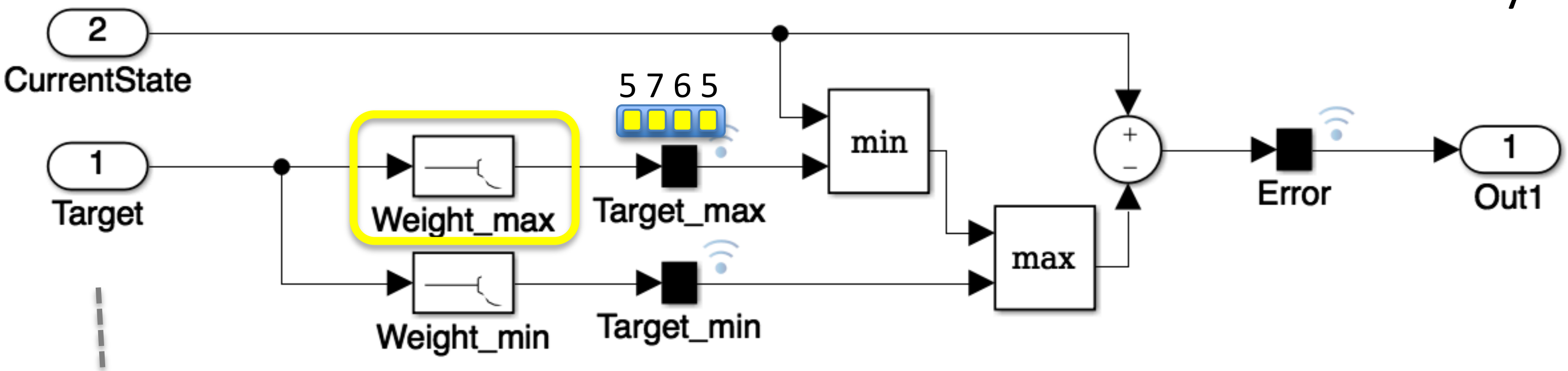




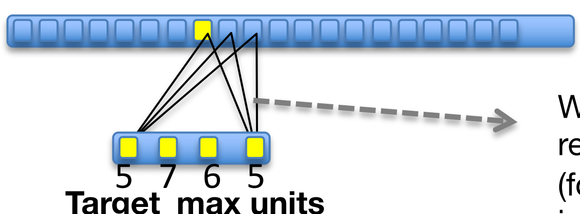
Neurons in **Superior Temporal Cortex (Auditory Error Maps)** compute the auditory error signals: distance between the current auditory signal and the desired auditory target region

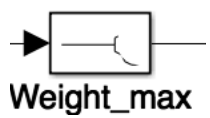
DIVA model components

Auditory Error Map

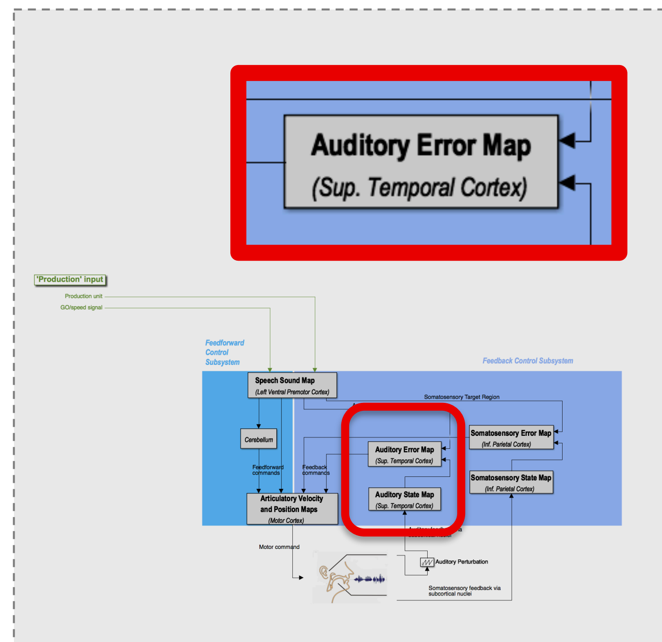


Temporal representation units



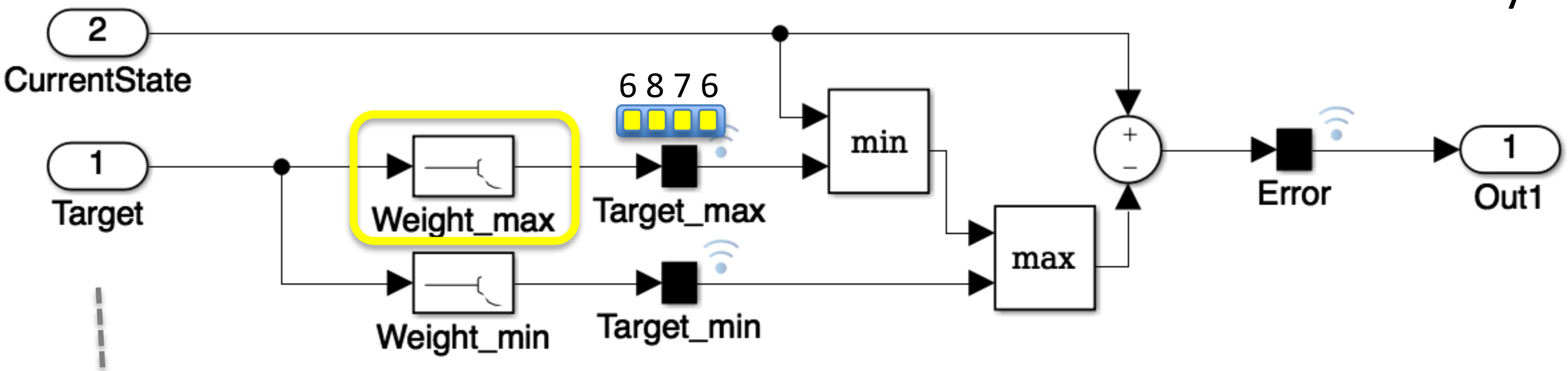


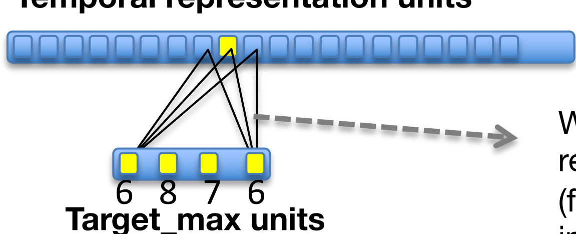
Weights encode auditory target region *higher/maximum* range (for a given timepoint of an individual learned production)

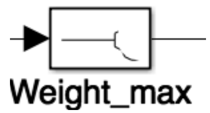


DIVA model components

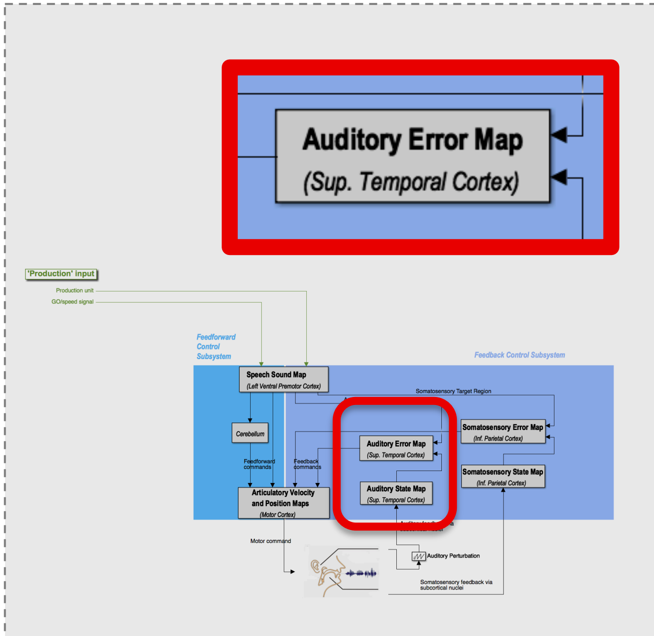
Auditory Error Map





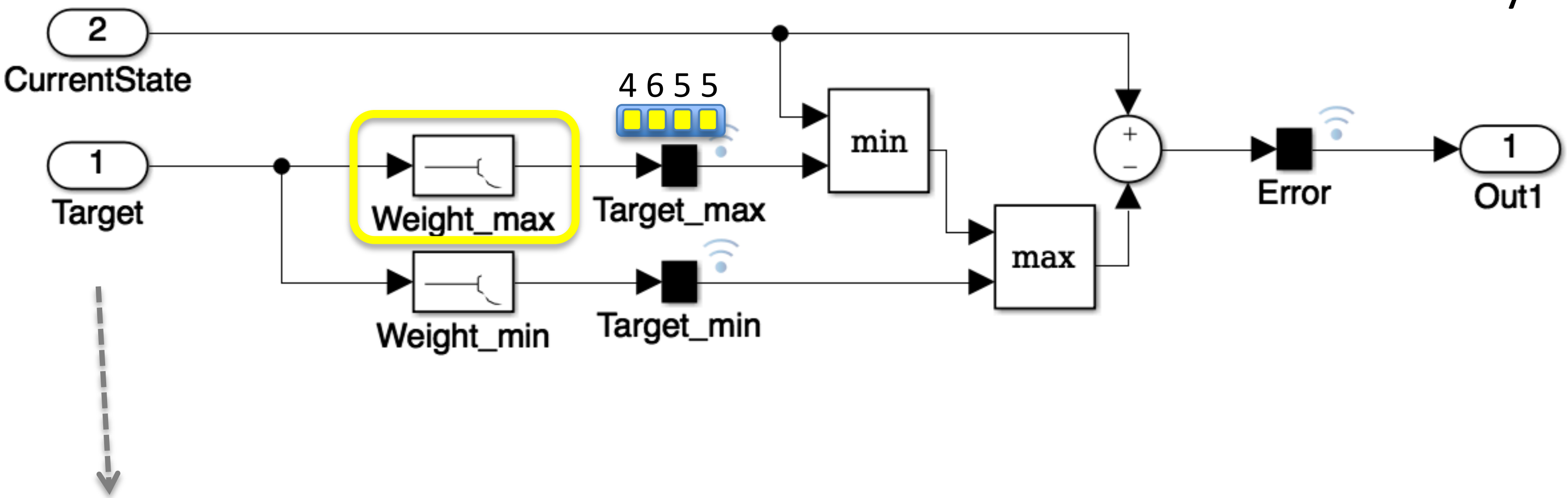


Weights encode auditory target region *higher/maximum* range (for a given timepoint of an individual learned production)

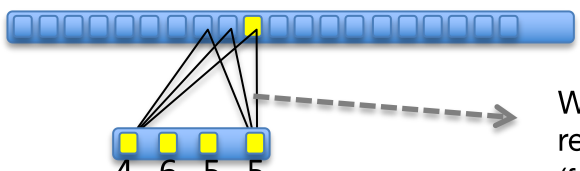


DIVA model components

Auditory Error Map



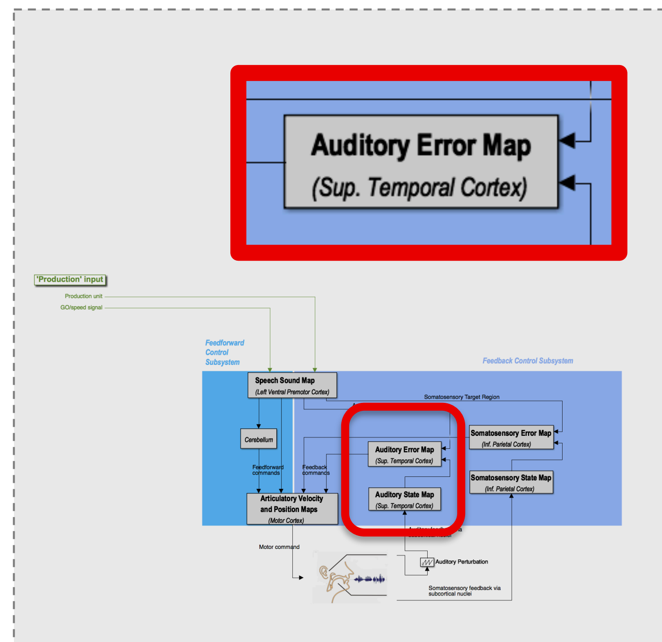
Temporal representation units



Target_max units

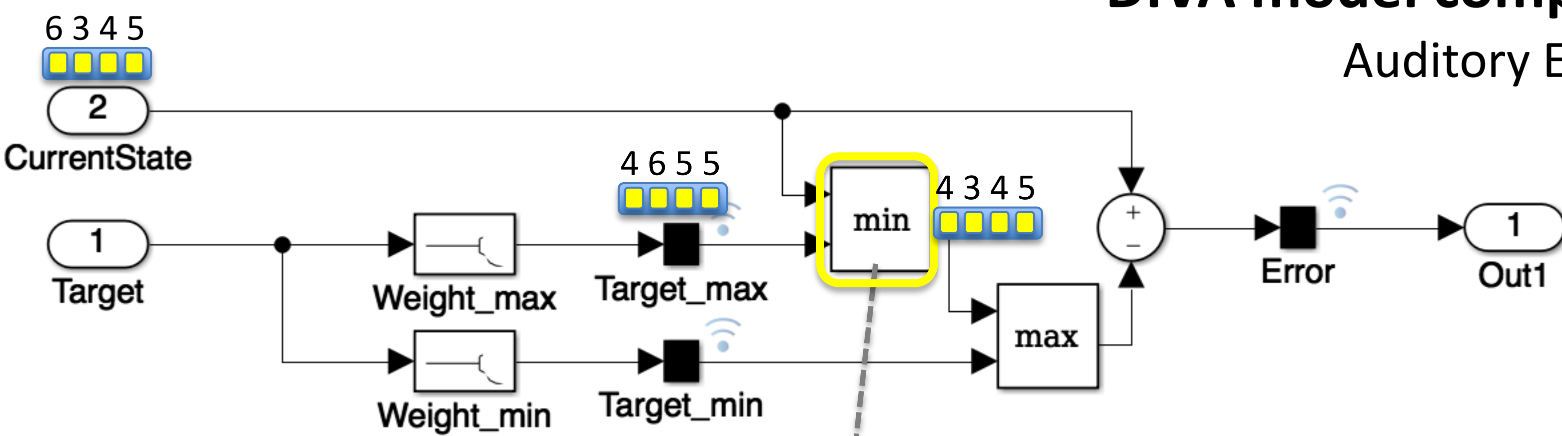
Weight_max

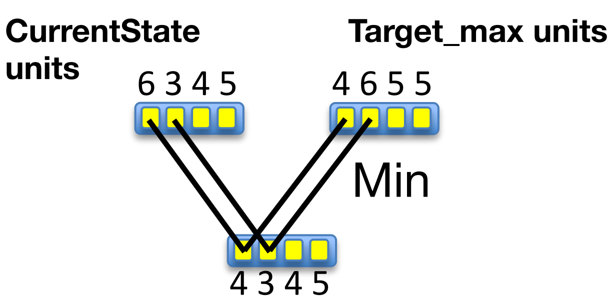
Weights encode auditory target region *higher/maximum* range (for a given timepoint of an individual learned production)



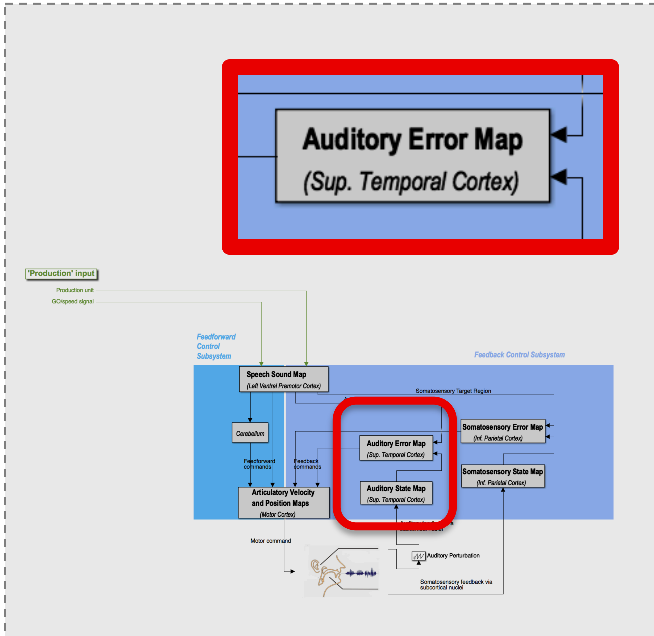
DIVA model components

Auditory Error Map



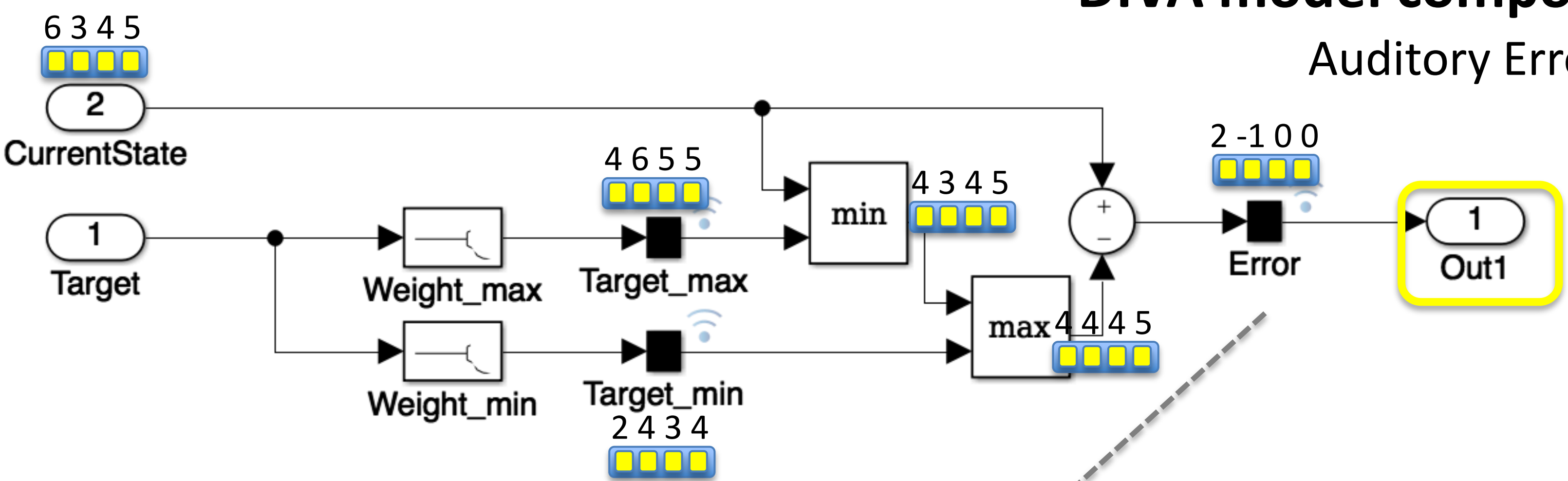


min block computes minimum value among pairs of inputs

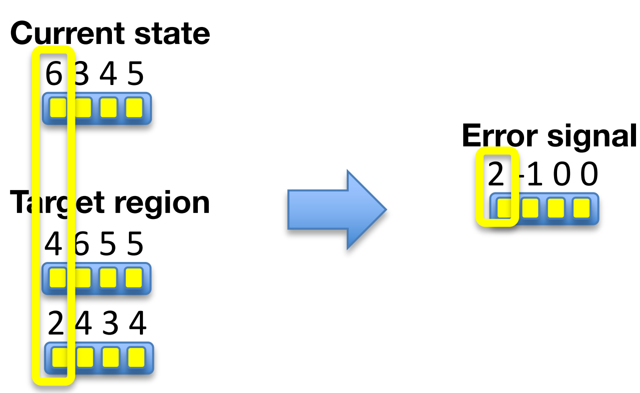


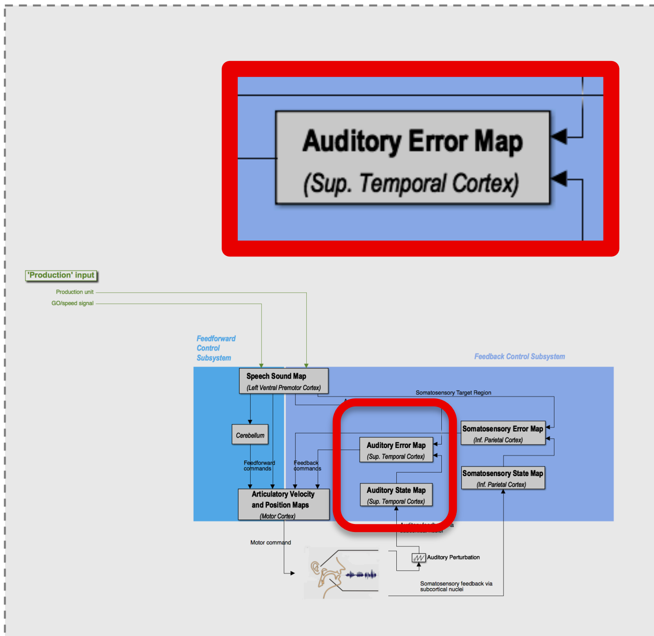
DIVA model components

Auditory Error Map



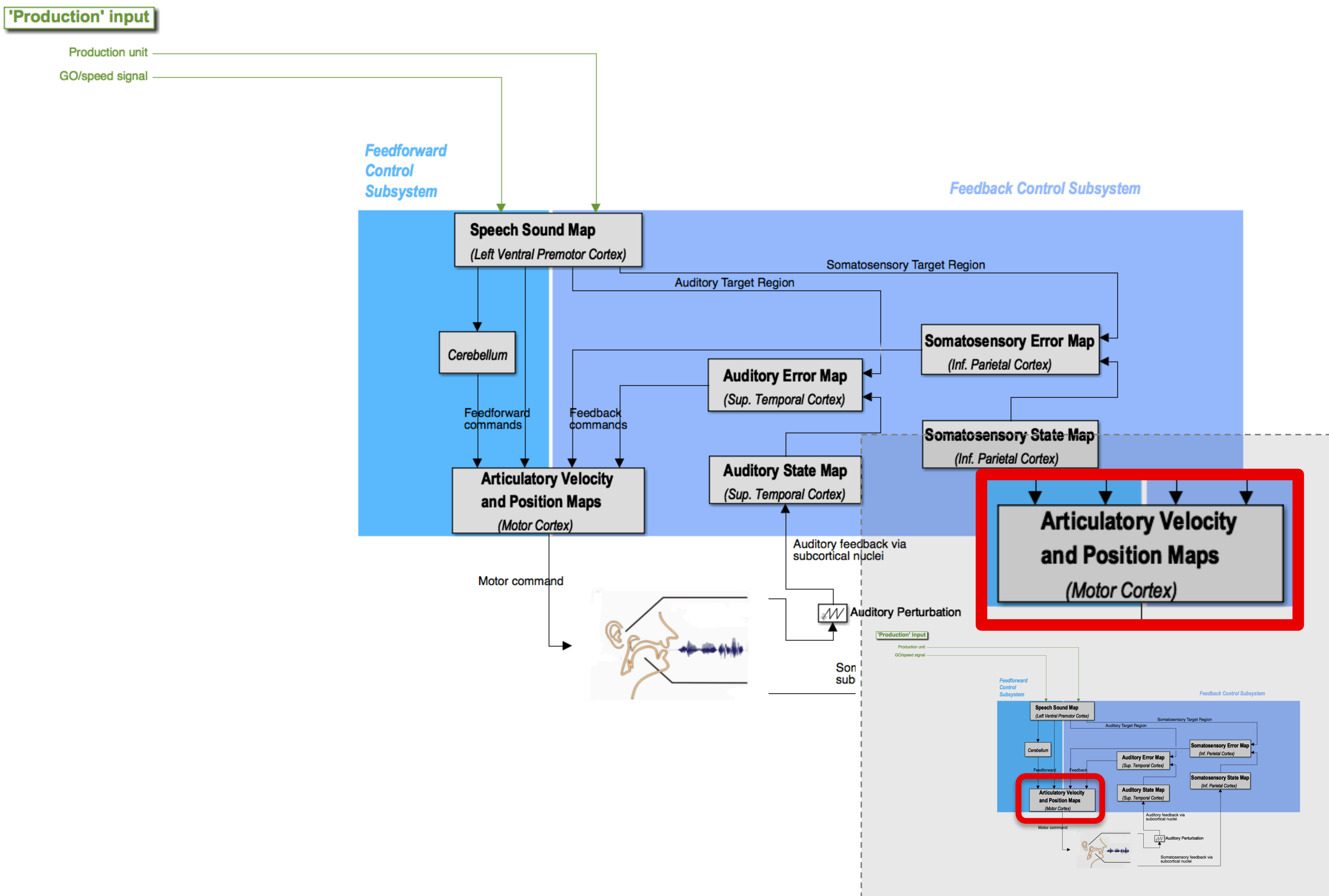
Min/Max/Diff operations jointly compute **error signal** (distance between current state and target region)





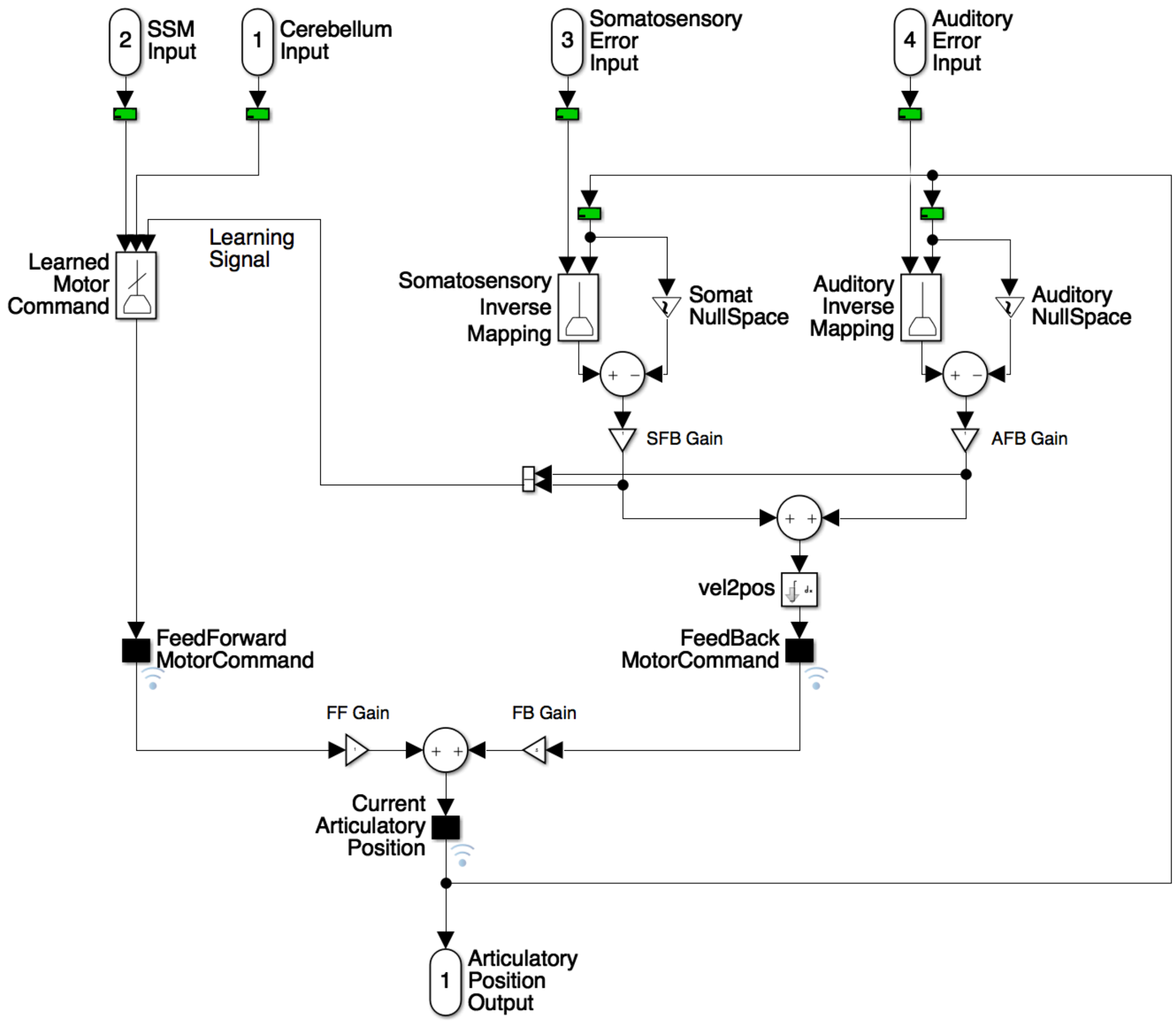
DIVA model components

Articulatory Velocity and Position Maps

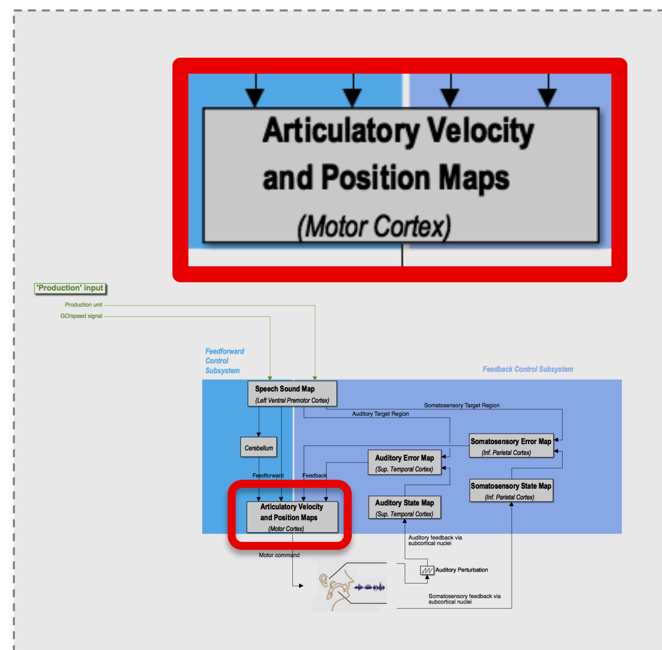


DIVA model components

Articulatory Velocity and Position Maps

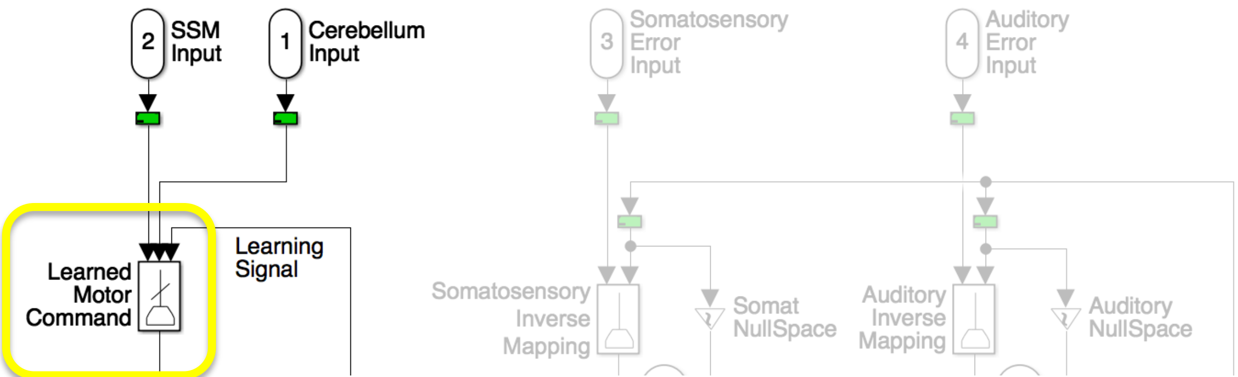


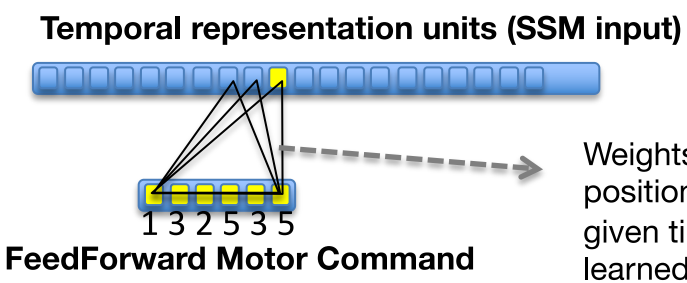
Neurons in **Motor Cortex (Articulatory Velocity and Position Maps)** compute the motor commands, combining feedforward and feedback components

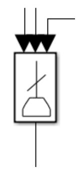


DIVA model components

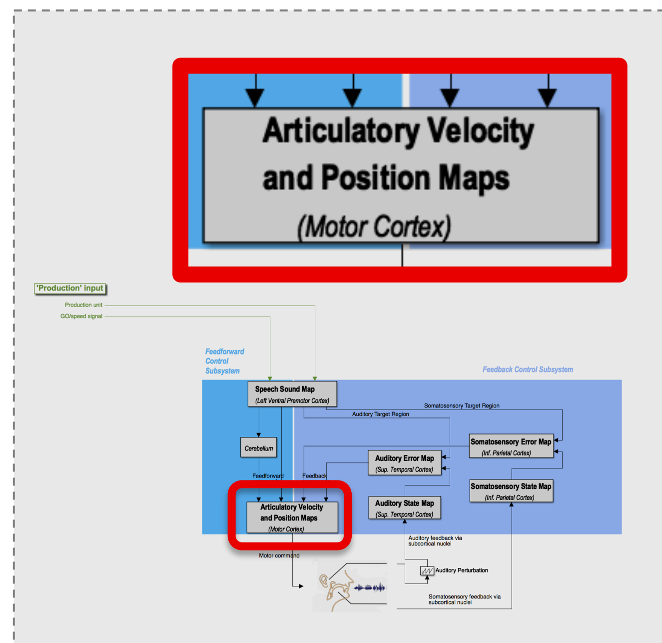
Articulatory Velocity and Position Maps





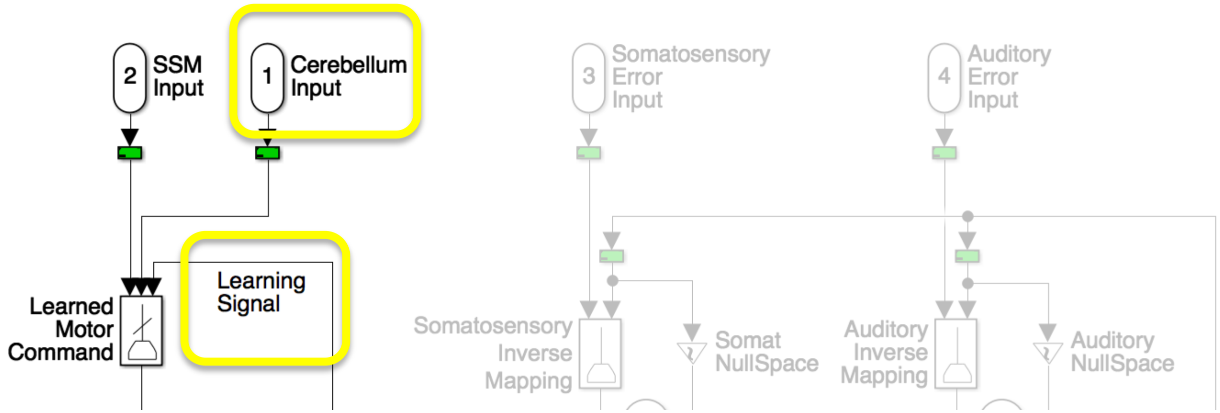


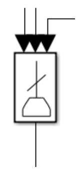
Weights encode learned motor positions commands (for a given timepoint of an individual learned production)



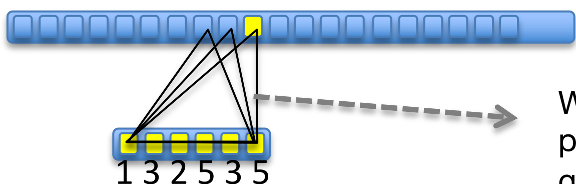
DIVA model components

Articulatory Velocity and Position Maps





Temporal representation units (SSM input)



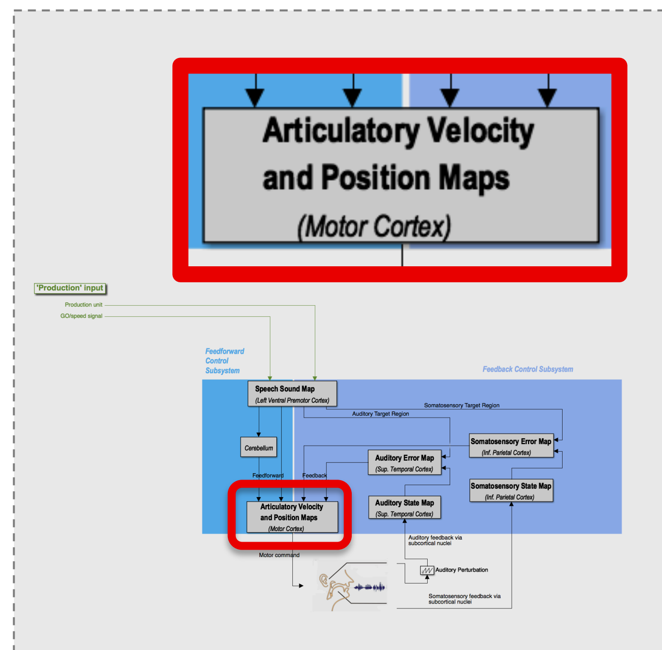
FeedForward Motor Command

Weights encode learned motor positions commands (for a given timepoint of an individual learned production)

Weights are adaptive. Change in weights is determined by the combination (outer product) of *Cerebellum Input* and *Learning Signal* inputs

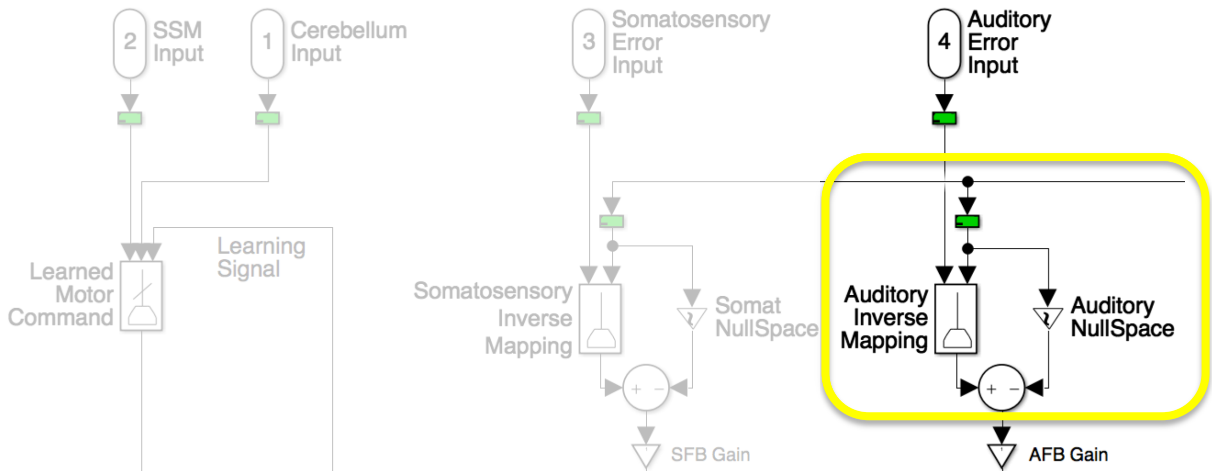
Cerebellum Input

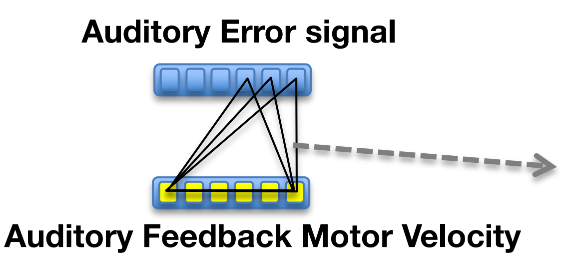
Learning Signal

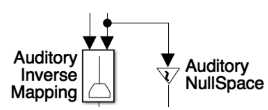


DIVA model components

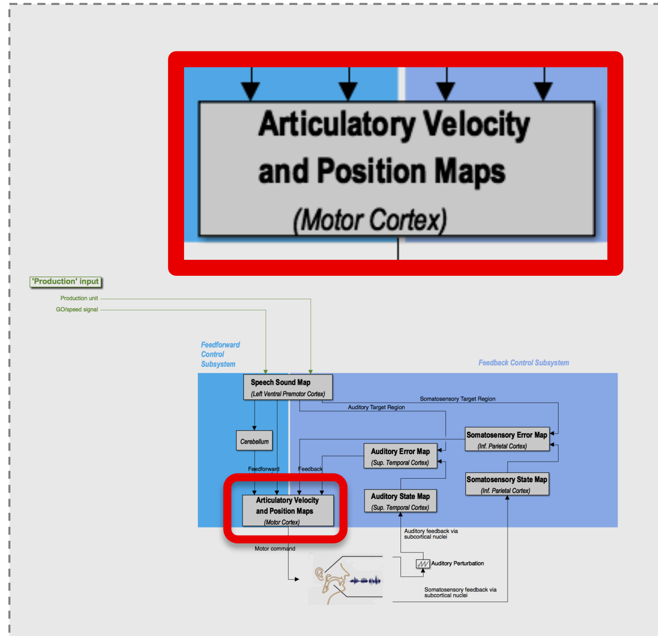
Articulatory Velocity and Position Maps







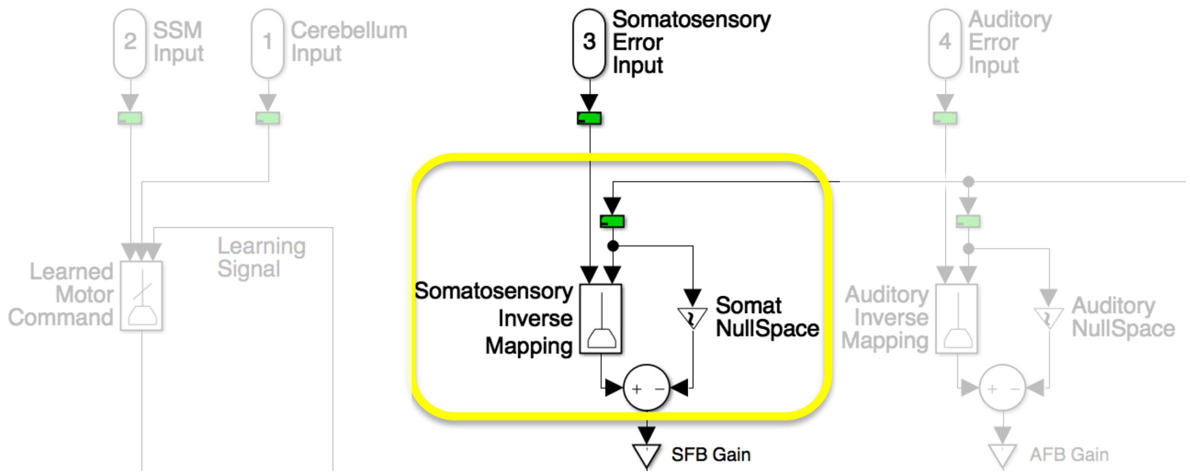
Weights encode directional auditory-to-articulatory mapping. Compute direction of change in articulatory space that acts to reduce the observed auditory error



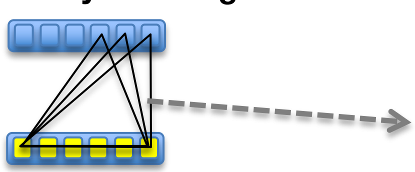
Note: current implementation uses explicit Jacobian pseudoinverse (proposed learned map)

DIVA model components

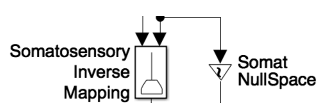
Articulatory Velocity and Position Maps



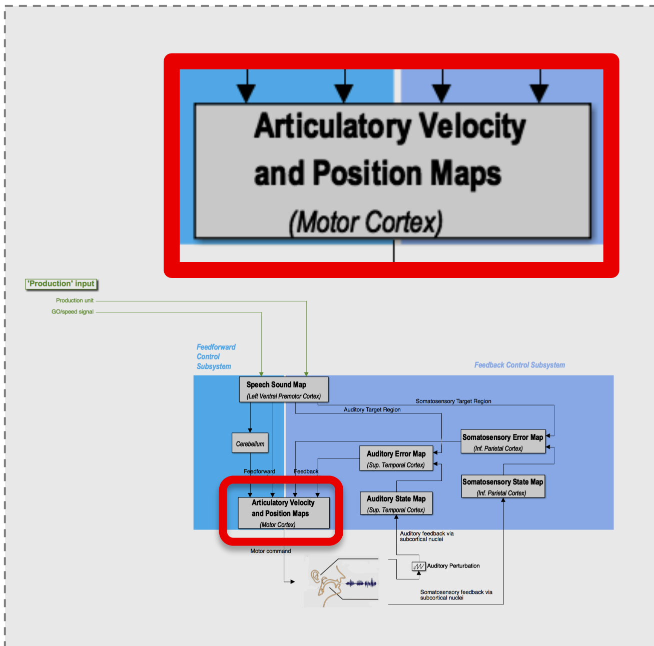
Somatosensory Error signal



Somatosensory Feedback Motor Velocity



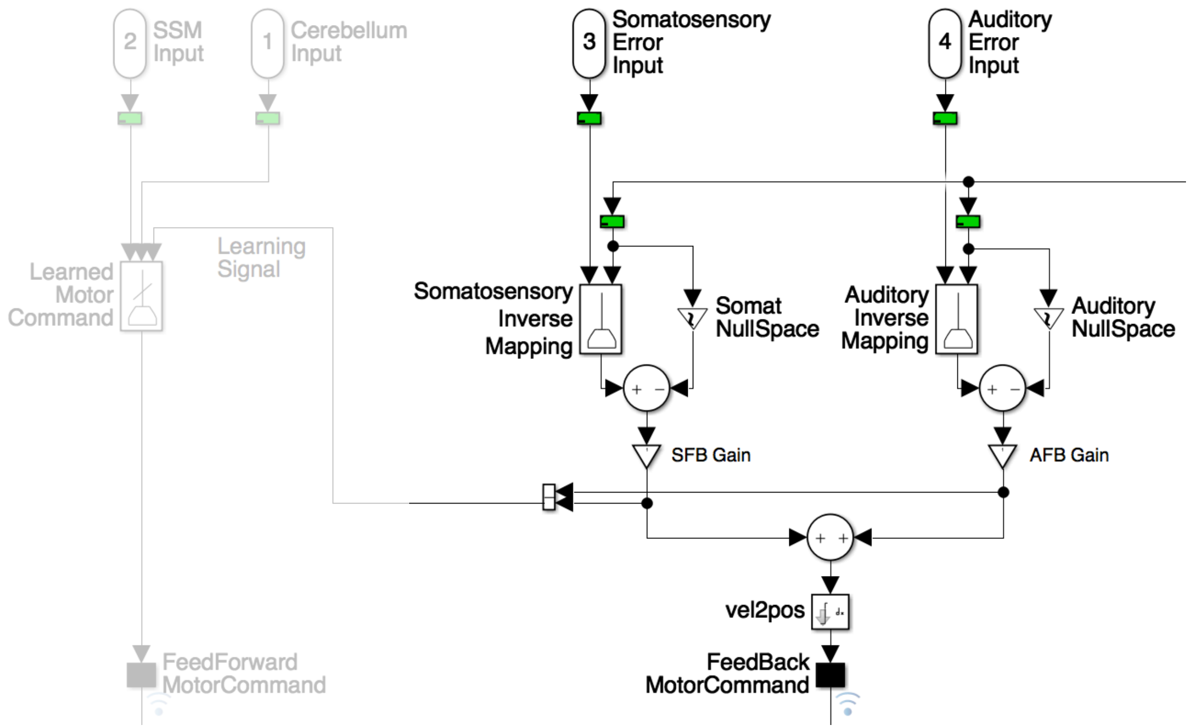
Weights encode directional somatosensory-to-articulatory mapping. Compute direction of change in articulatory space that acts to reduce the observed somat. error

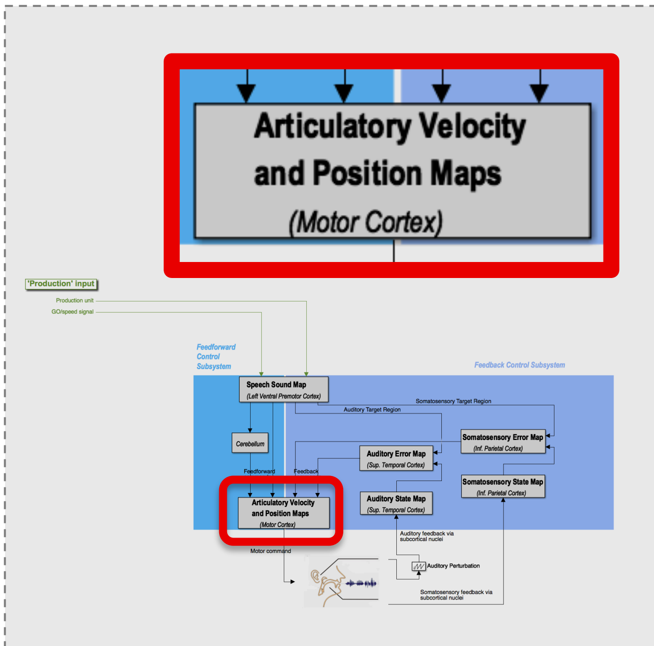


Note: current implementation uses explicit Jacobian pseudoinverse (proposed learned map)

DIVA model components

Articulatory Velocity and Position Maps

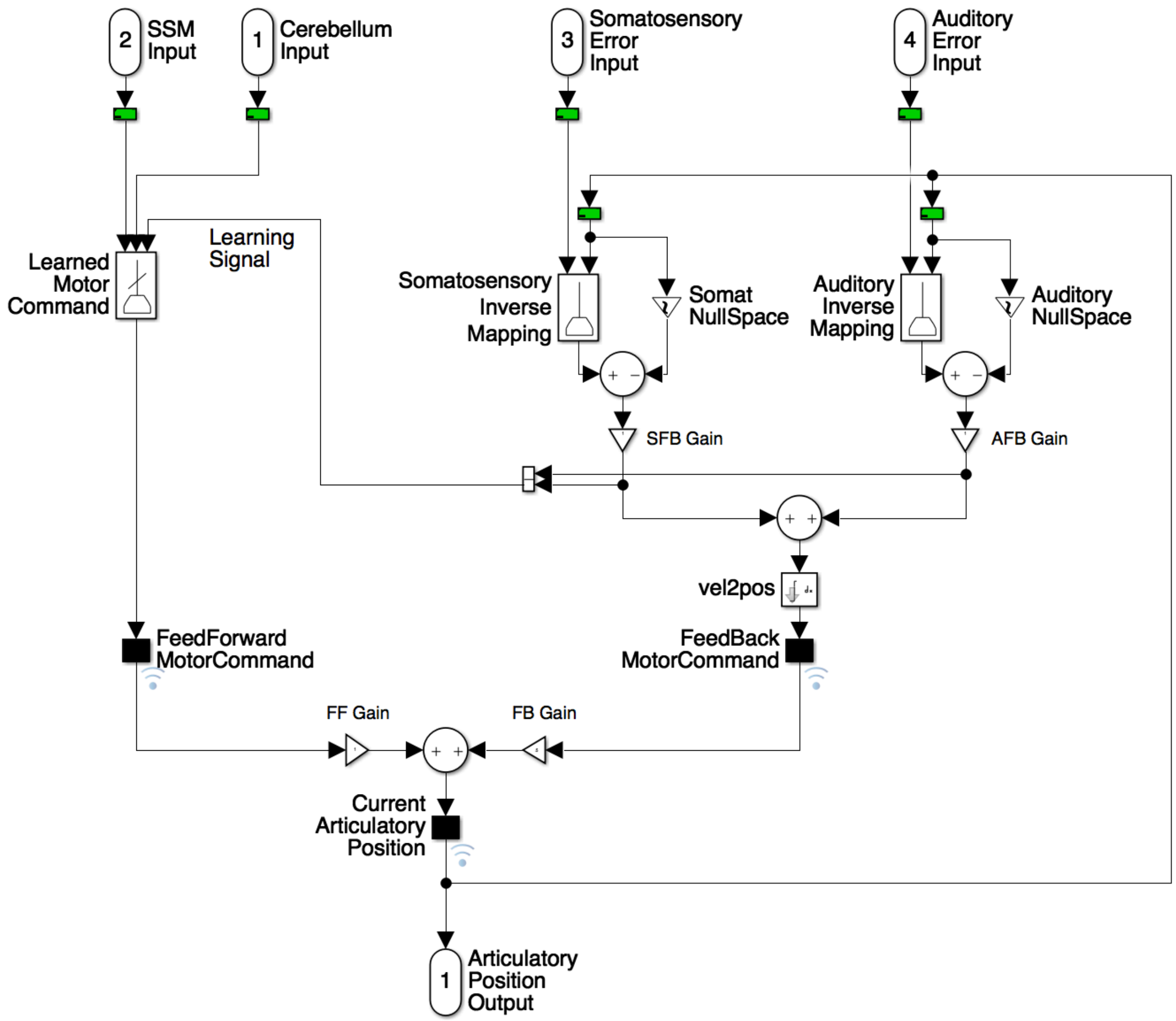




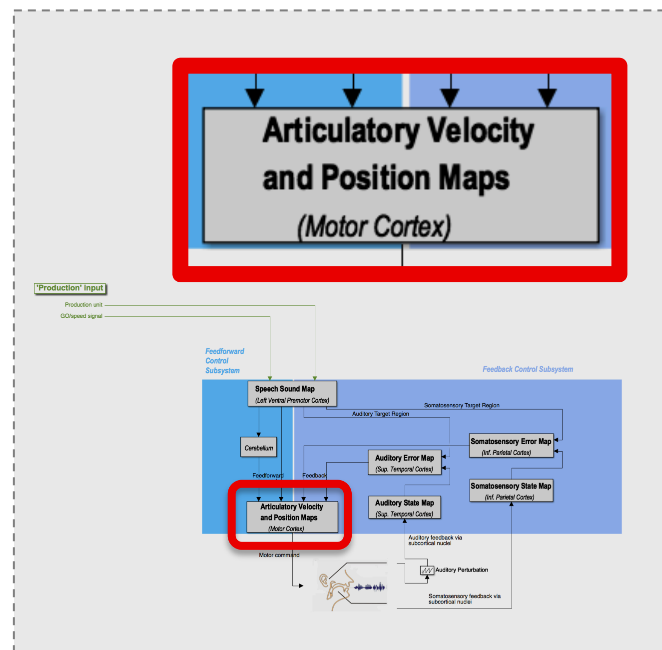
Somatosensory and Auditory feedback components are combined and integrated to compute the **FeedbBack Motor Command**

DIVA model components

Articulatory Velocity and Position Maps

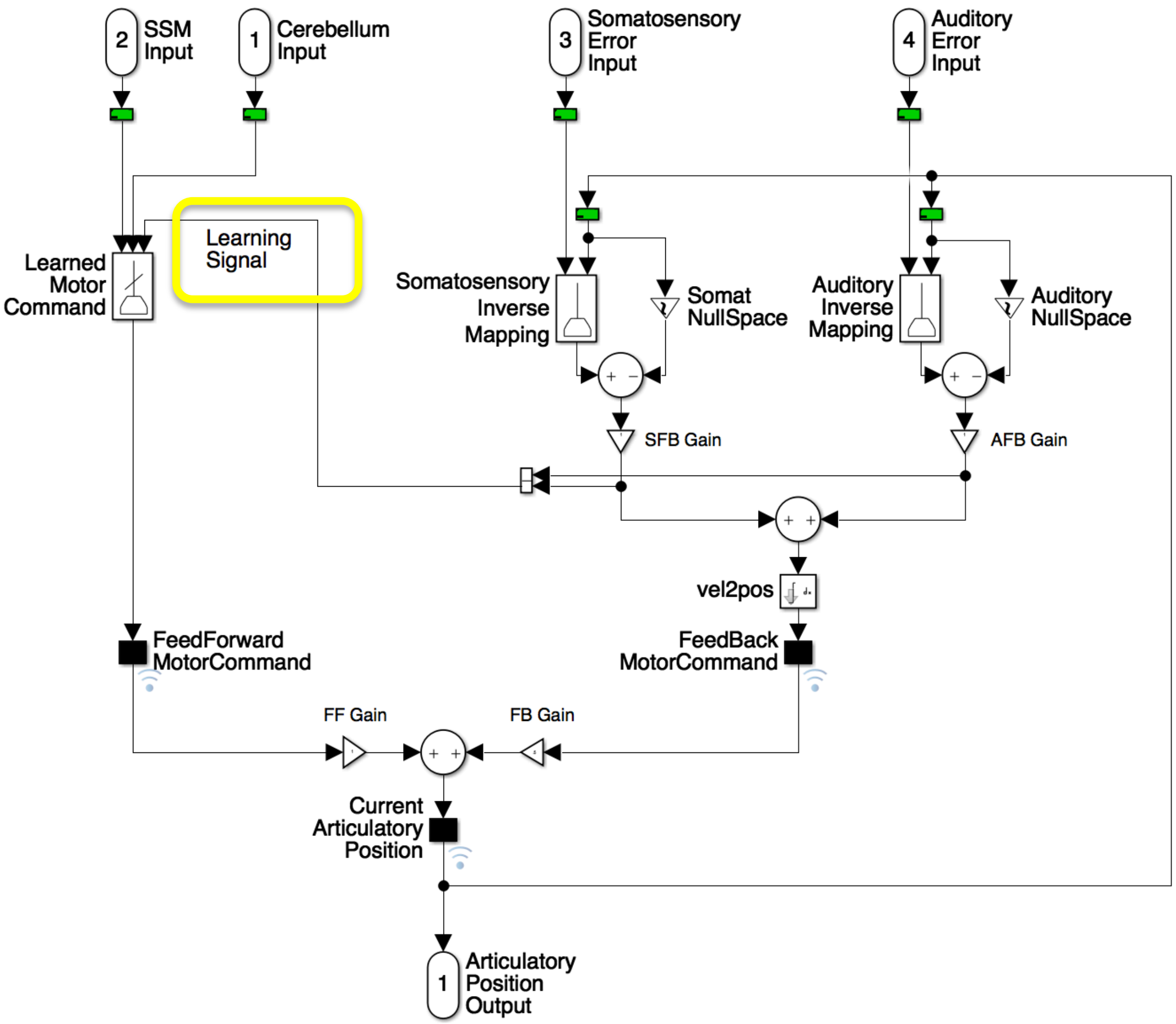


And the FeedForward and FeedBack components are combined to form the desired **Motor Command**



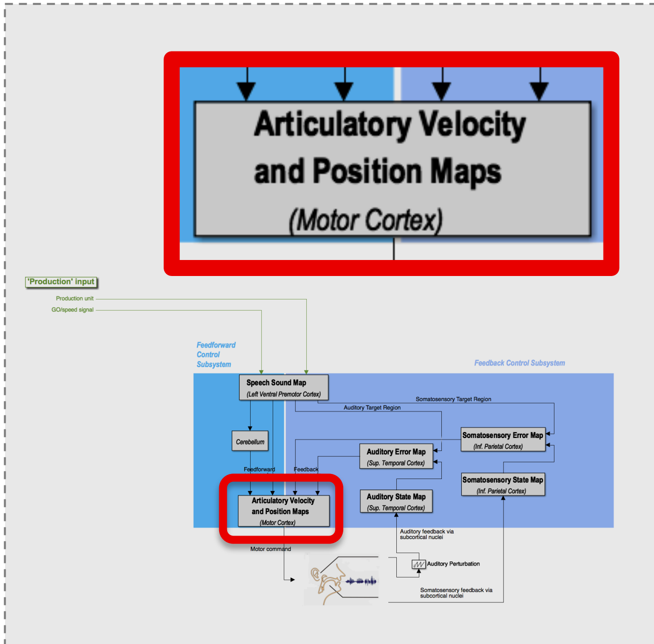
DIVA model components

Articulatory Velocity and Position Maps

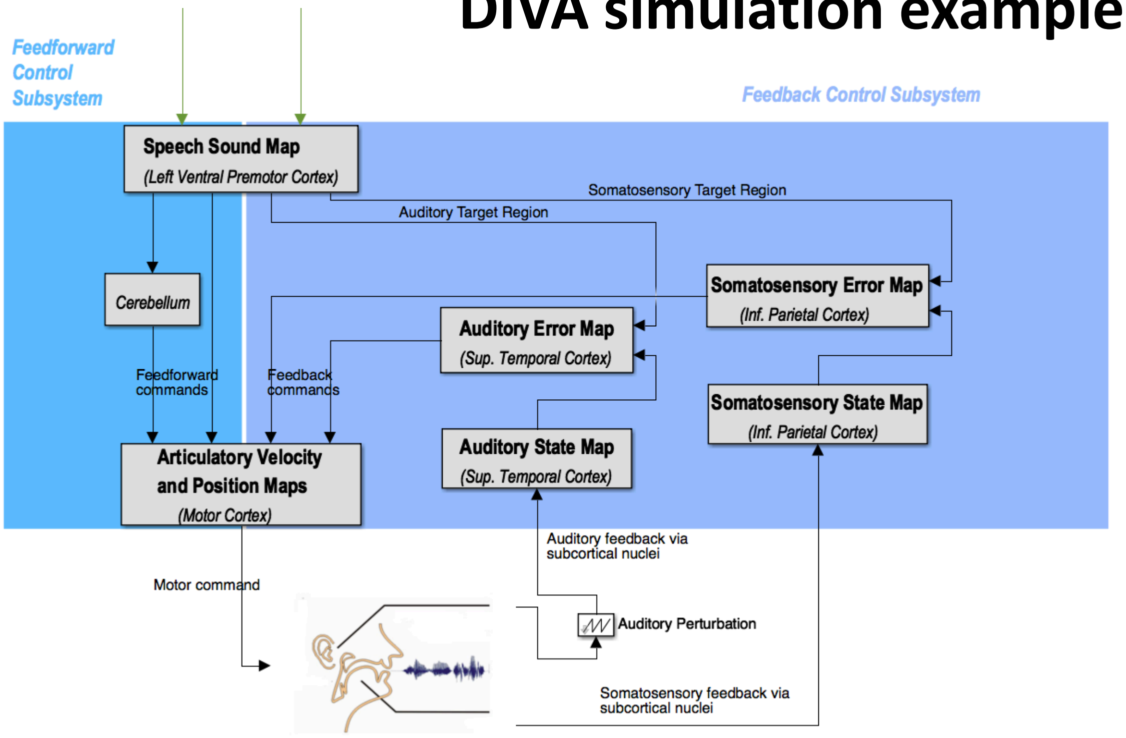


And the FeedForward and FeedBack components are combined to form the desired **Motor Command**

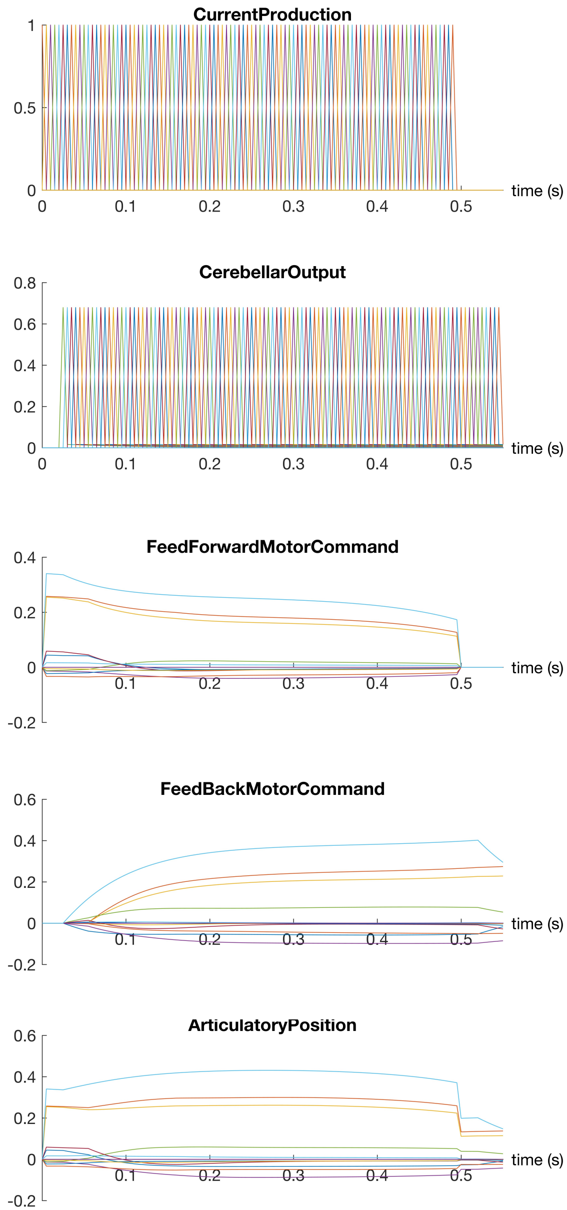
Also note how the computed SomatoSensory and Auditory FeedBack signals are used to drive learning of the adaptive FeedForward weights

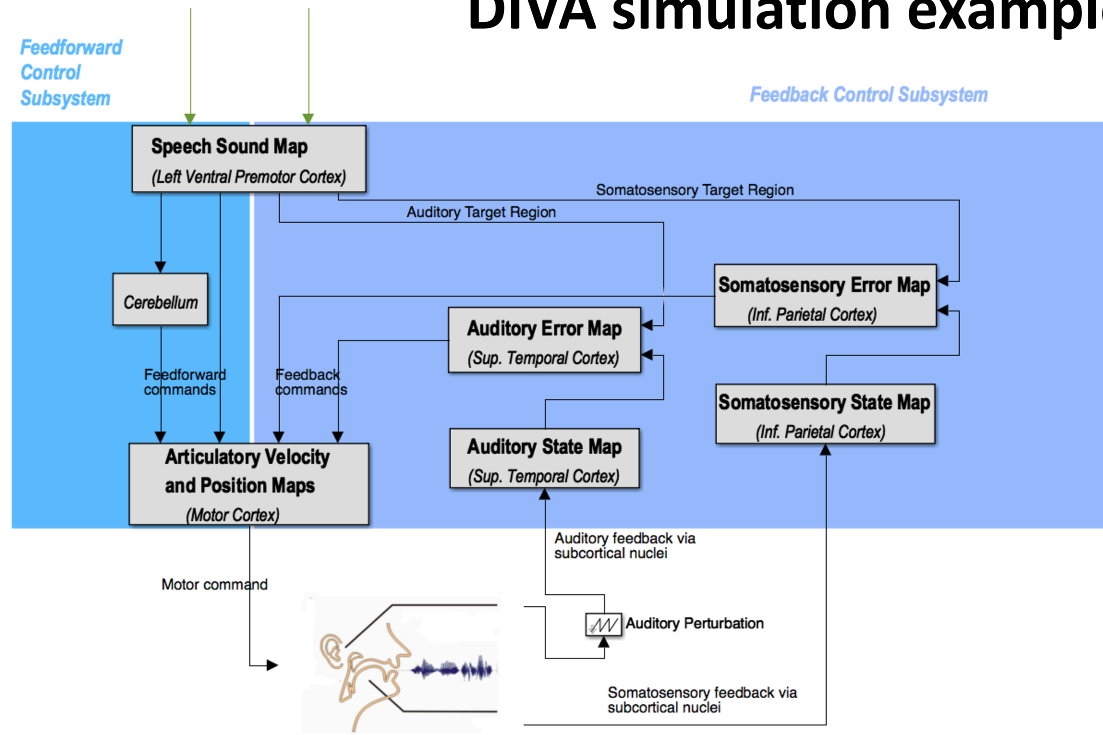


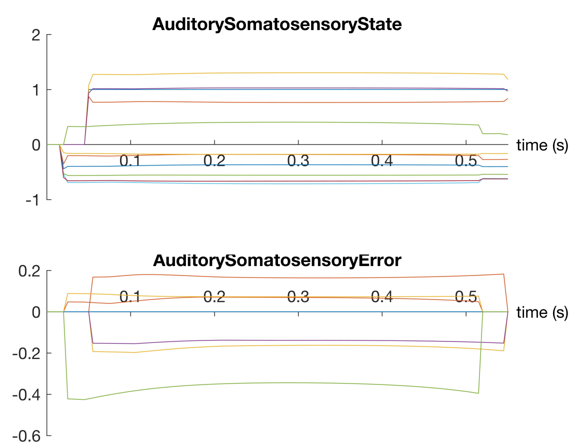
DIVA simulation example



DIVA simulation example



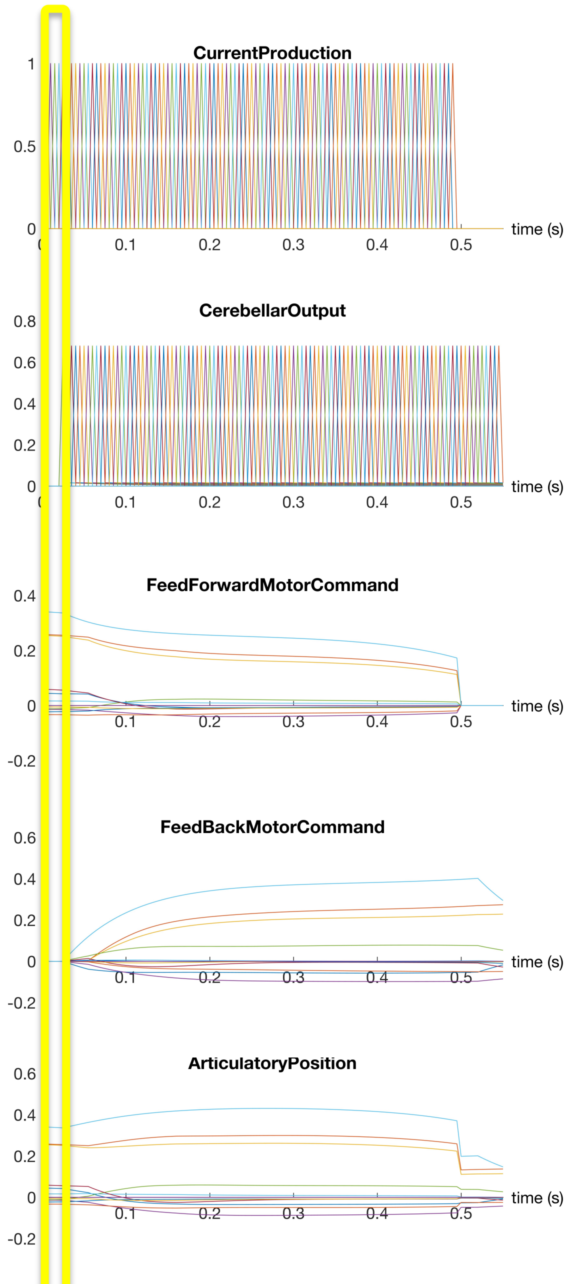


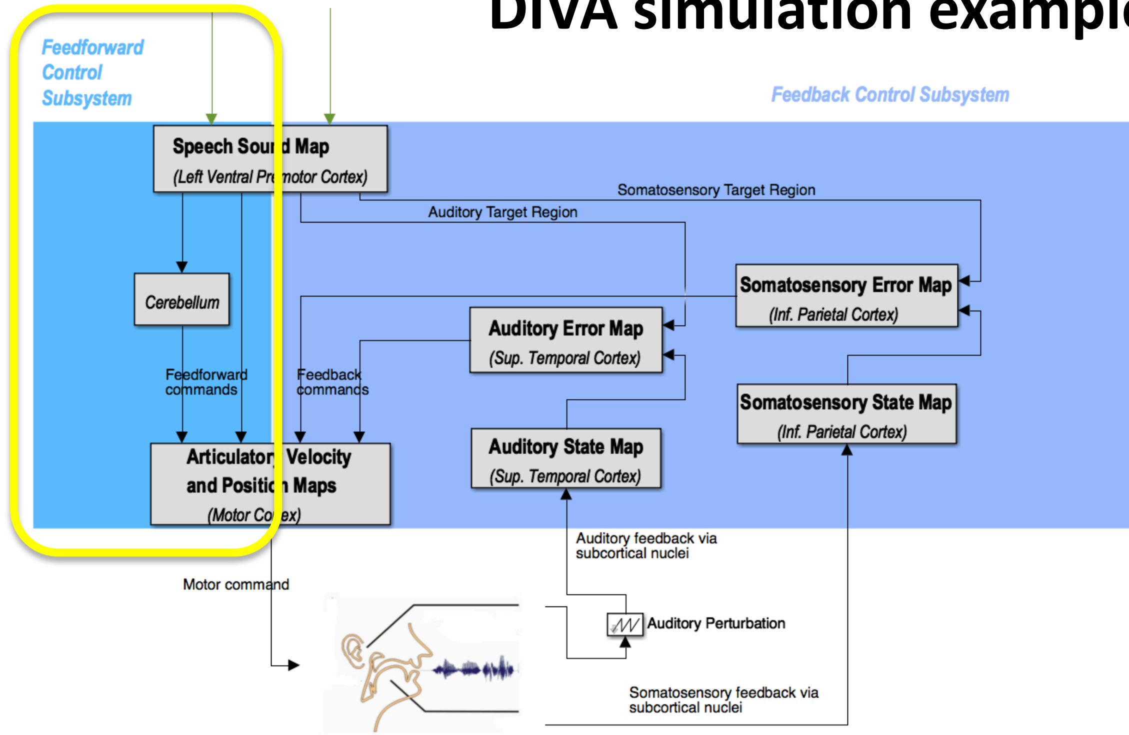


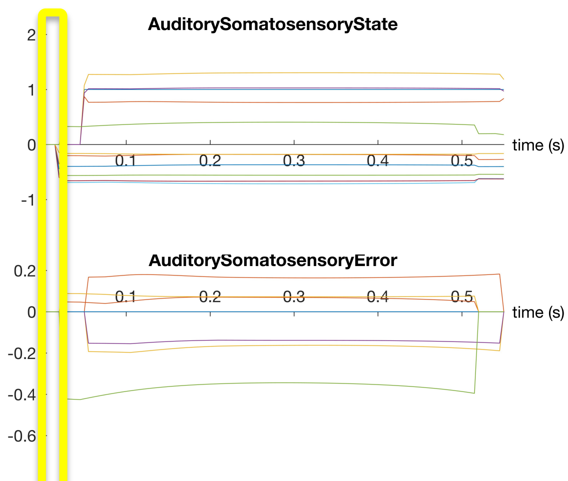
Model simulations

Plots describe the activation timeseries of selected components of the DIVA model (e.g. feedback and feedforward motor commands, auditory errors, etc.)

DIVA simulation example

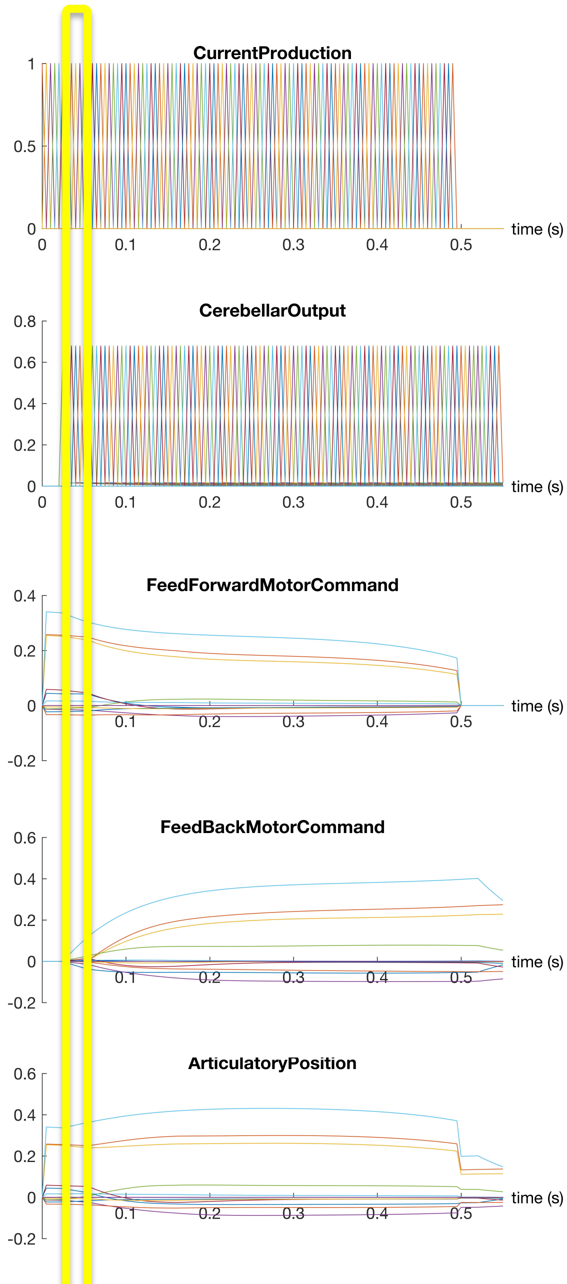


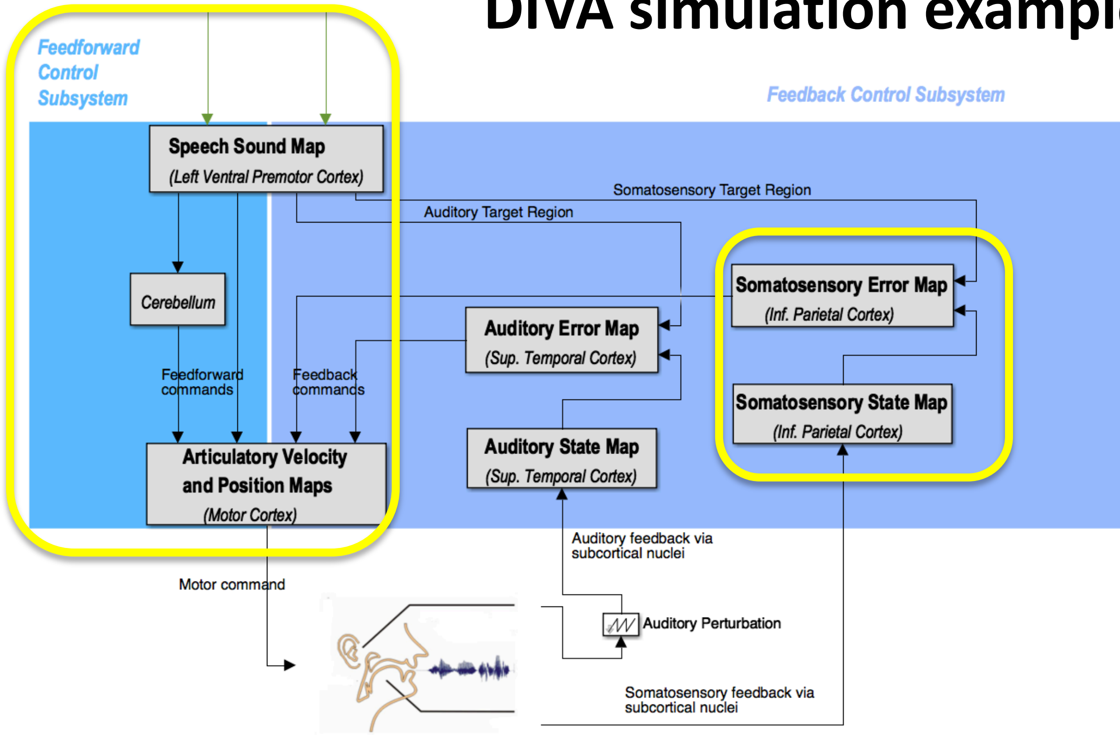


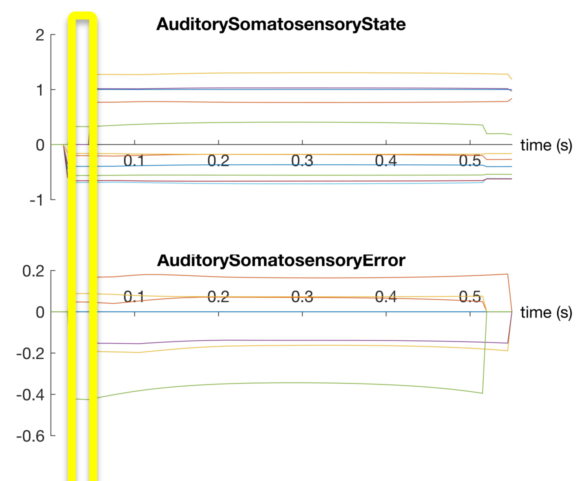


The first few milliseconds (0-20ms) are marked by feedforward responses, as somatosensory and auditory signals have longer delays

DIVA simulation example

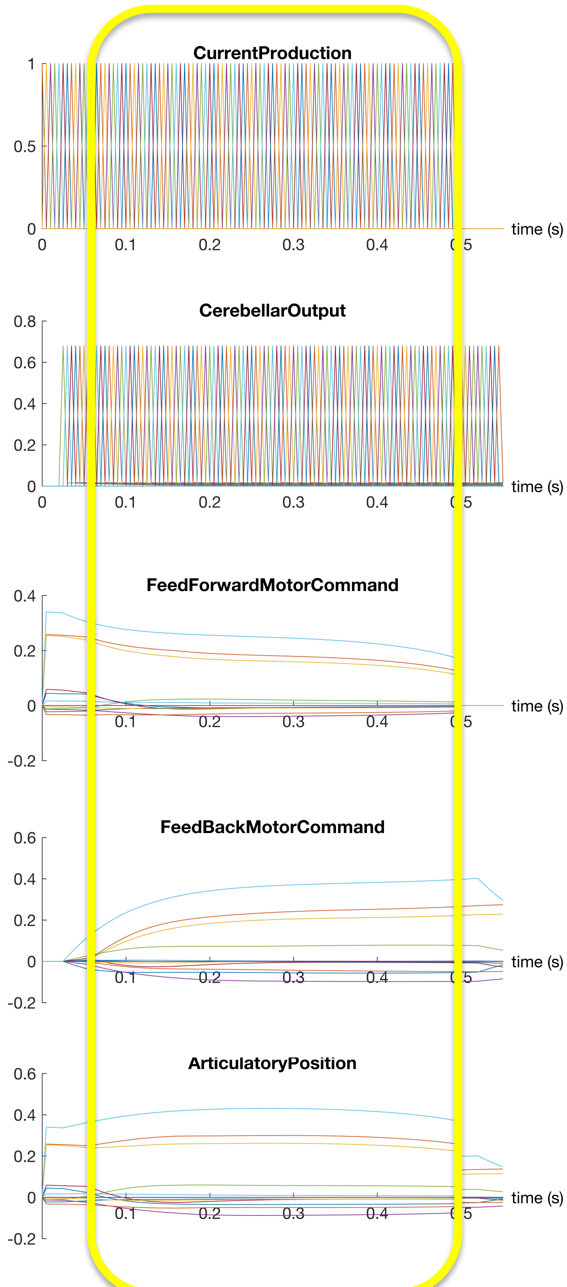


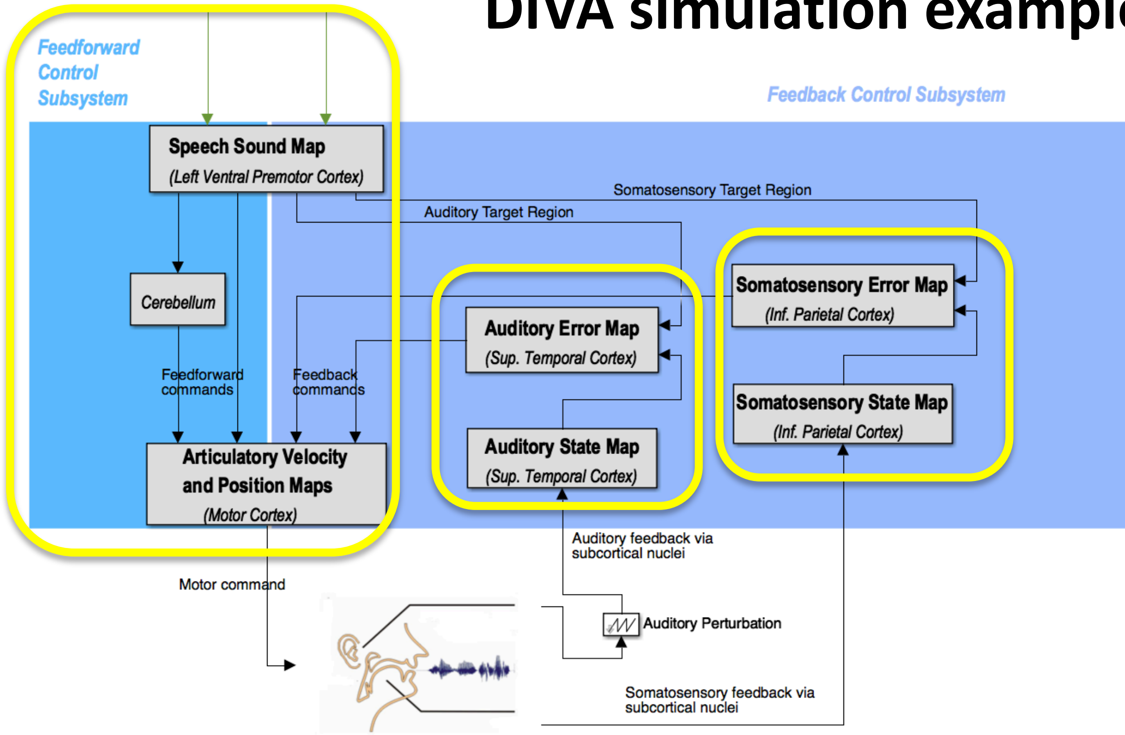


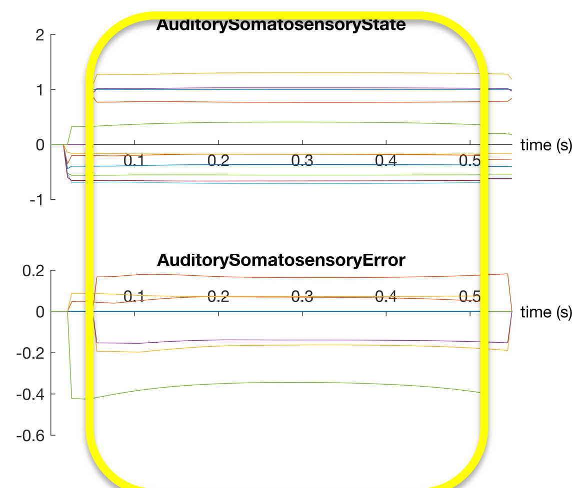


After that (>20ms) feedback responses start to be incorporated into the observed motor response, driven first by **somatosensory** inputs and associated errors…

DIVA simulation example



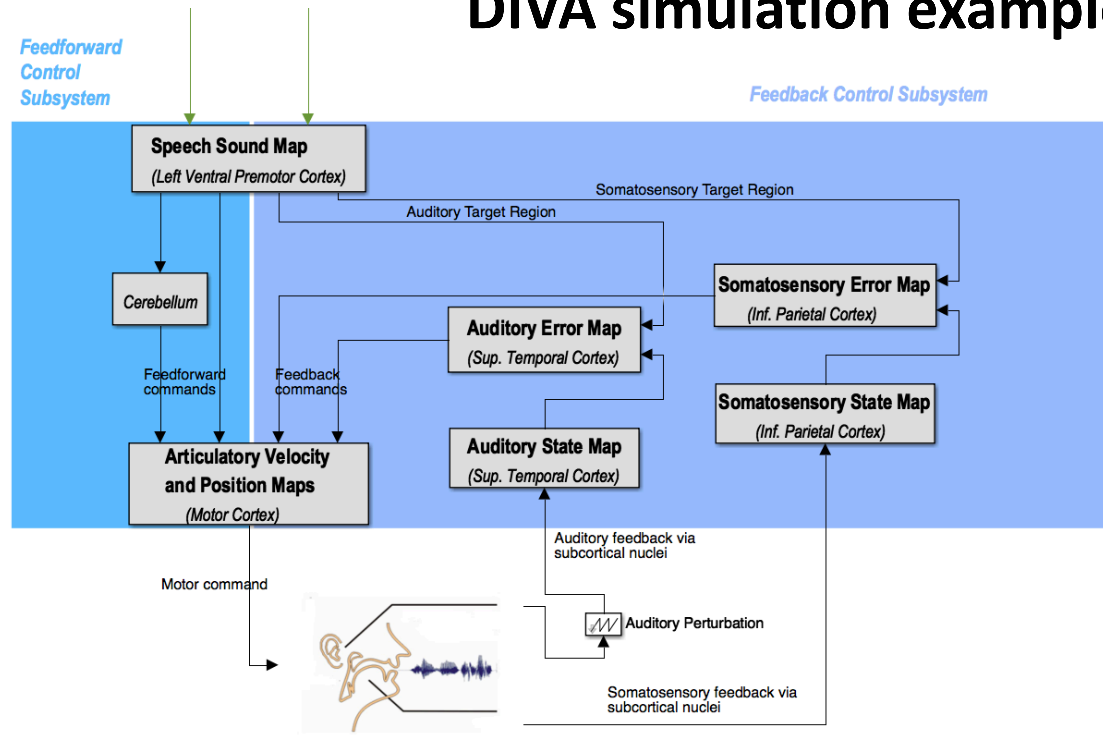


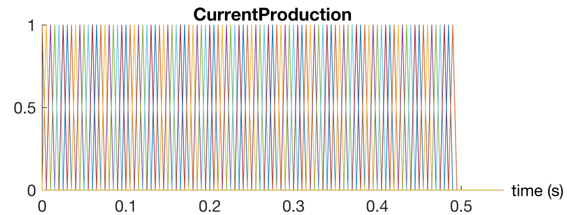


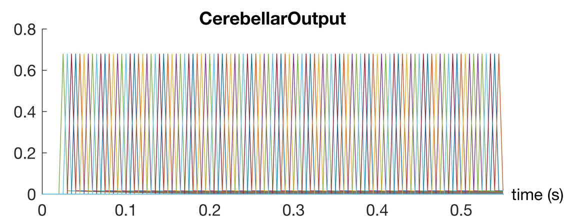
… followed (>40ms) by **auditory** signals and associated errors

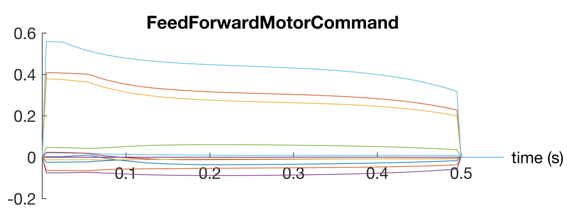
At this point if there is an observable error in either somatosensory or auditory targets, the feedback component attempts to produce a compensatory command to reduce this error

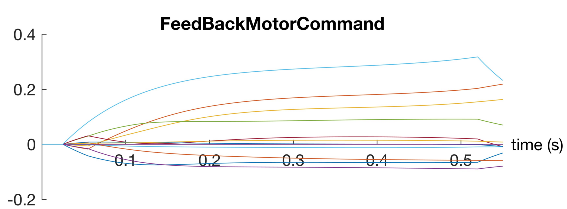
DIVA simulation example

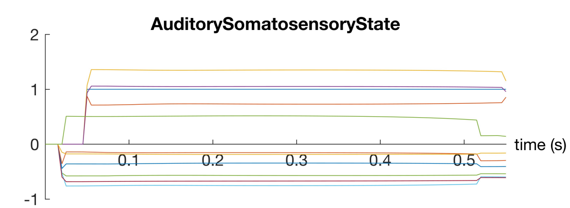




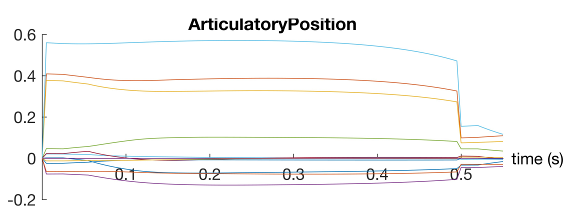


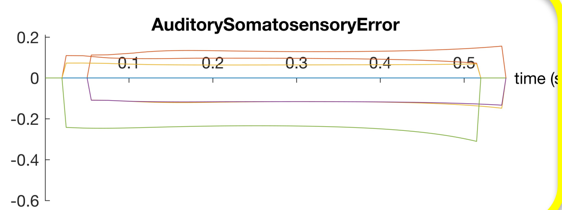




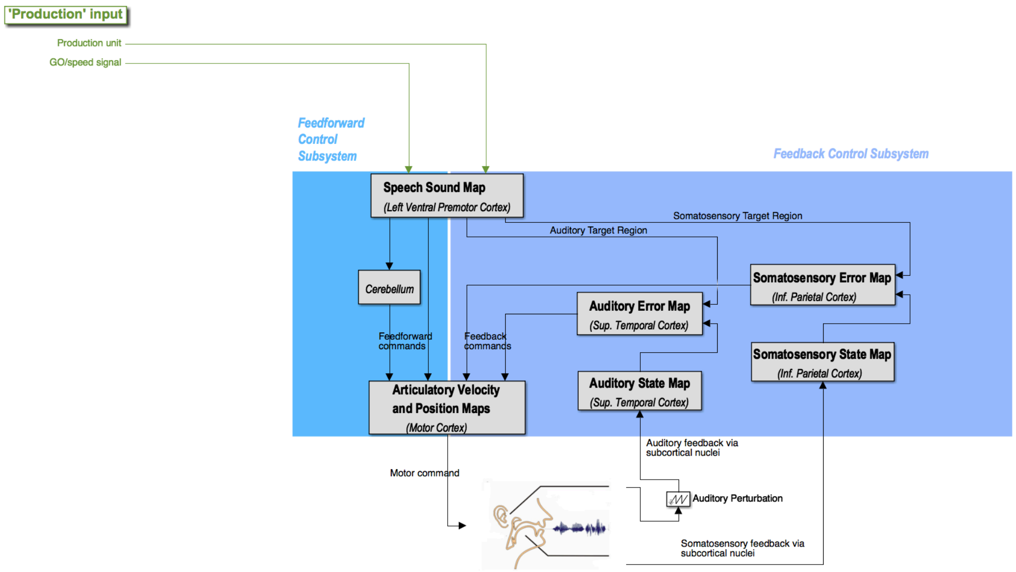


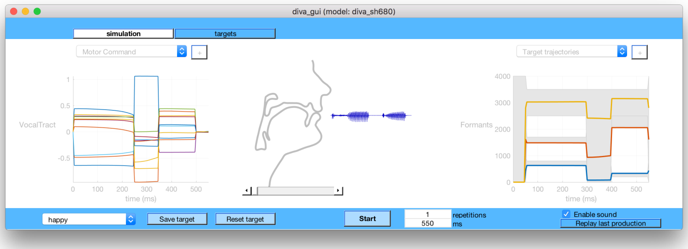
This feedback command is also used to drive learning, so the next repetition of this same production will produce smaller errors

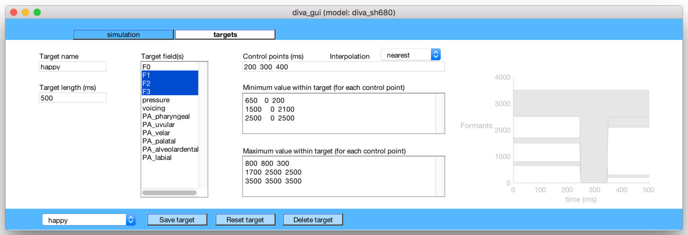




DIVA GUI
hands-on practice







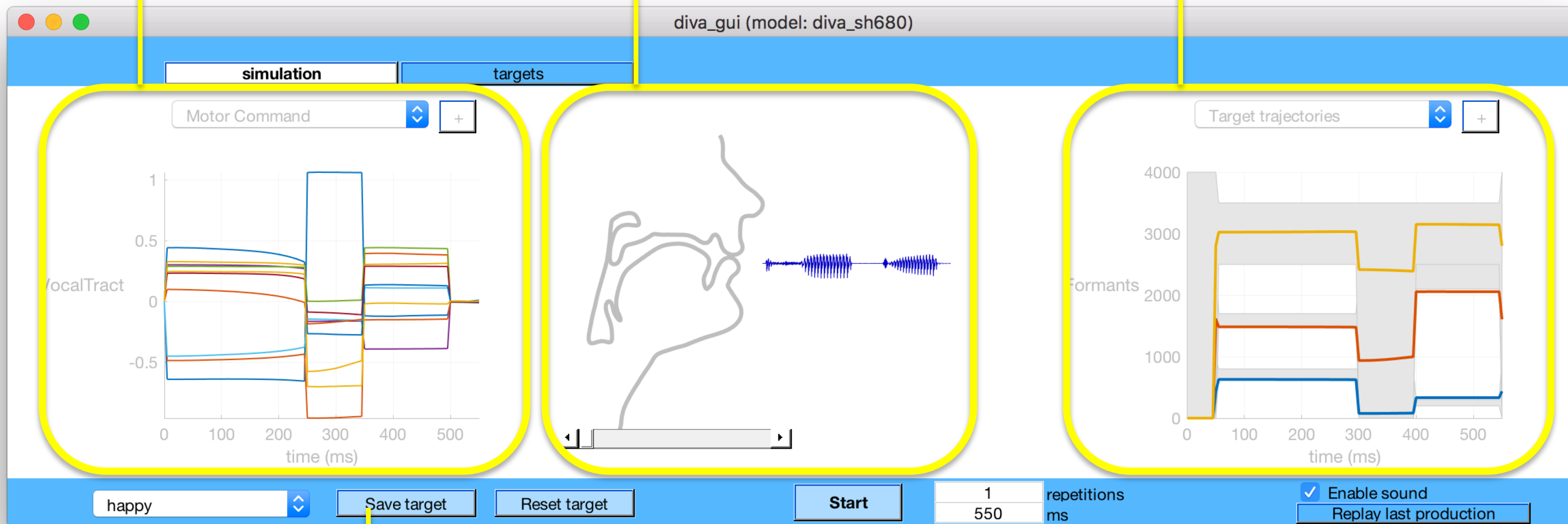
In addition to the main DIVA Simulink model window, which is fully editable by users (e.g. to modify the behavior of the DIVA model or the general experimental design), there is an additional DIVA graphical interface that allows simple initialization of the DIVA Simulink model simulations as well as exploration of the simulation results

DIVA GUI
hands-on practice

Motor signals display

Vocal tract and soundwave display

Auditory/Somatosensory signals display



Select production target

Resetting the target “forgets” the learned feedforward motor command for this production target

Saving changes (feedforward adaptation/learning) after a simulation allows you to use the new learned feedforward command as initial state for future simulations

Start a simulation

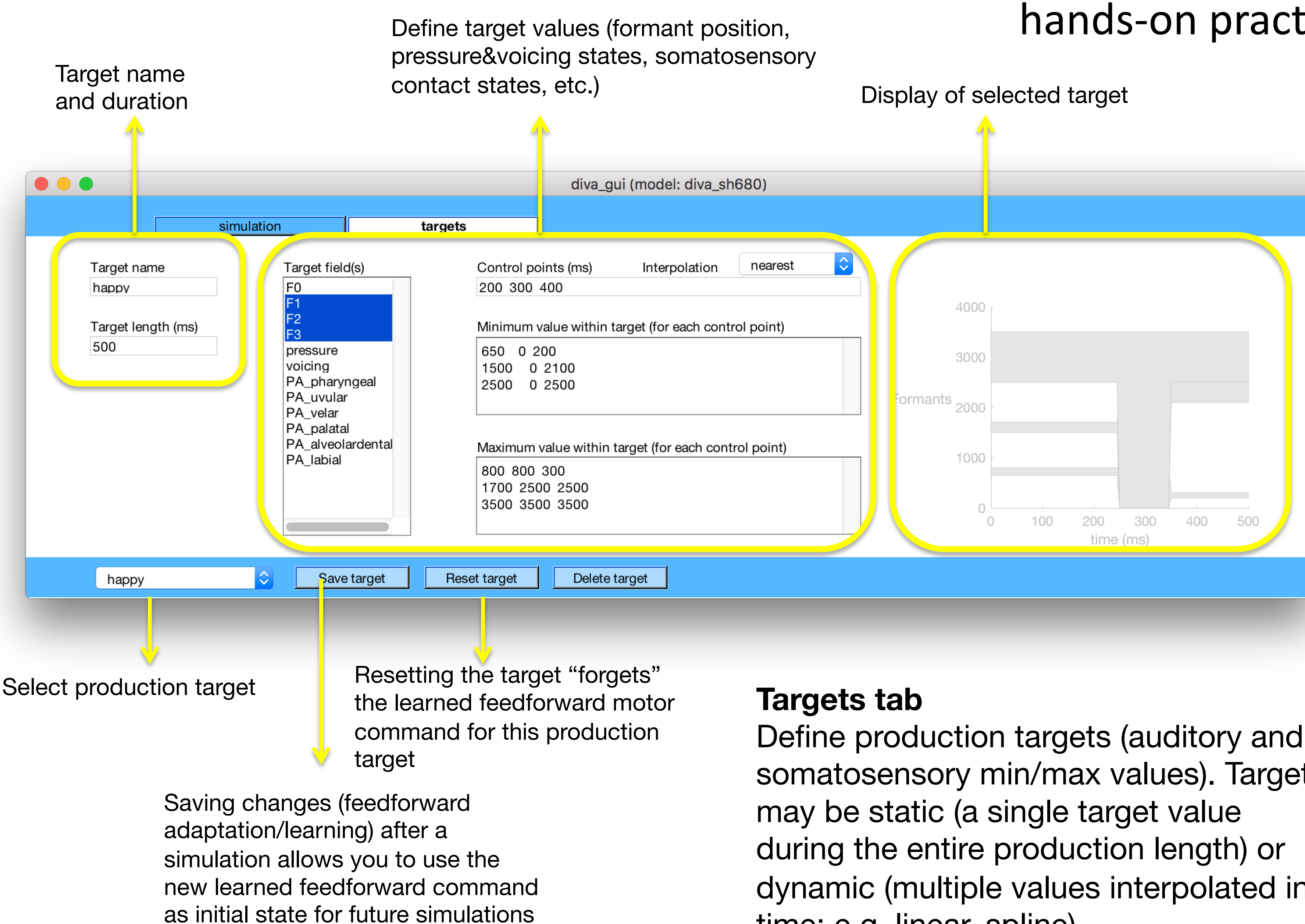
Define number of repetitions (sequential productions of the selected target) and simulation length (ms)

Enable/disable sound output, repeat last production sound

Simulation tab

Select production target, run simulation, and display DIVA component activation timeseries

DIVA GUI
hands-on practice



Targets tab

Define production targets (auditory and somatosensory min/max values). Targets may be static (a single target value during the entire production length) or dynamic (multiple values interpolated in time; e.g. linear, spline)