USING NENGO.AI FOR DEVELOPING MODELS OF SPEECH PRODUCTION AND SPEECH PERCEPTION

Prof. Dr. Bernd J. Kröger
Neurophonetics Group, Phoniatrics Department, Medical Faculty
RWTH-Aachen University, Aachen, Germany
www.speechtrainer.eu

Introduction

Download whole talk from YouTube Playlist (6 videos):
https://www.youtube.com/watch?v=4Q1xtjNxefU&list=PLKha1vAjpAIgS6nWVY75OZ1m3D1Dc9xHd

- How can experimental sciences benefit from models?
- The NENGO.ai framework
- Our speech processing model (production and perception)
  - architecture of the modeled brain
  - simulated tasks -> simulated behavior (modeled test person)
- Concluding remarks
  - On the realism of brain models
  - On the benefit of brain models for medical research

Download Video1:
https://www.youtube.com/watch?v=4Q1xtjNxefU

Motivation for models

- Models do not necessarily lead to new (detailed) knowledge
  - -> … can not answer too specific (quantitative) questions, like:
  How strong is a specific effect?
- But: Models help to understand the basics
  - Functional organization of the neural processes of speech production and speech perception
  - Interactions between different processing stages (modules or levels of a model)
  - … can answer the question: an effect occurs because of …

Motivation for models

- Models can generate new knowledge:
  - e.g.,: relationship between neural structure and behavior:
  - i.e.,: how does a neural dysfunction change behavior (-> speech)?
  - Thus: Neural models may help to refine medical screening methods or to develop new screening methods

Motivation for models

- Neural models
  - define structure (neurons and connections)
  - define neural functioning (spiking, forwarding and spreading of neural activity) -> neural processing
  - are able to simulate macroscopic behavior (in case of large-scale models = a avatar)
  - give clear definitions for neural dysfunctions (resulting in specific speech disorders)

The NENGO.ai framework

- “Neural ENGineering Objects” framework:
- www.nengo.ai
- Nengo is free -> open source, but professionally developed!
- Python based: -> runs on: Windows, Mac OS X and Linux
- Embedded in Python Development Environments (e.g., Anaconda)
- Python Math-Libraries, plotting libraries (-> comparable to MatLab) can be used
- Professional scientific background:
  - Group of Prof. Chris Eliasmith, Leader of Centre of Theoretical Neurosciences, University of Waterloo, Waterloo, Canada
  - Eliasmith, C. (2013). How to build a brain: A neural architecture for biological cognition. *Oxford University Press*.

Download Video2:
https://www.youtube.com/watch?v=5lWhF1b4SB4

NENGO comprises

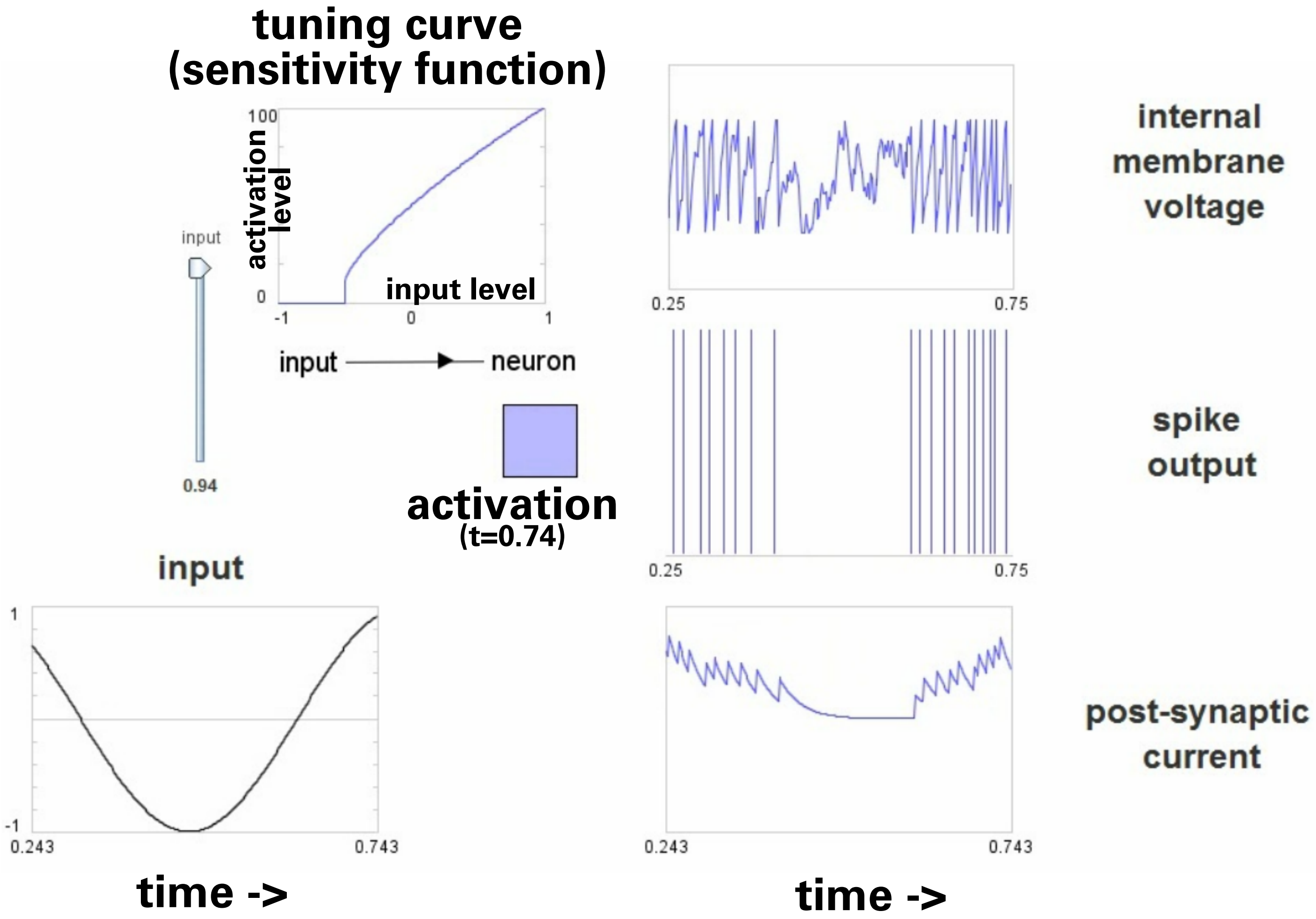
- basic neuron model for cortex, for subcortical regions (LIF-model)
- concepts or strategies for connecting neurons:
  - neuron ensembles, neuron buffers, short-term memories, associative memories, binding buffers, …
- strategies for coding higher level items (concepts <-> S-pointers)
- fully developed model for for neural process control over time (BG-Thalamus-complex: cortico-cortical control loop)
- strategy for connecting the periphery: -> model avatar
  - Sensory input: eyes and/or ears
  - Motor output: arms and/or the speech articulators

NENGO: single neuron and spikes

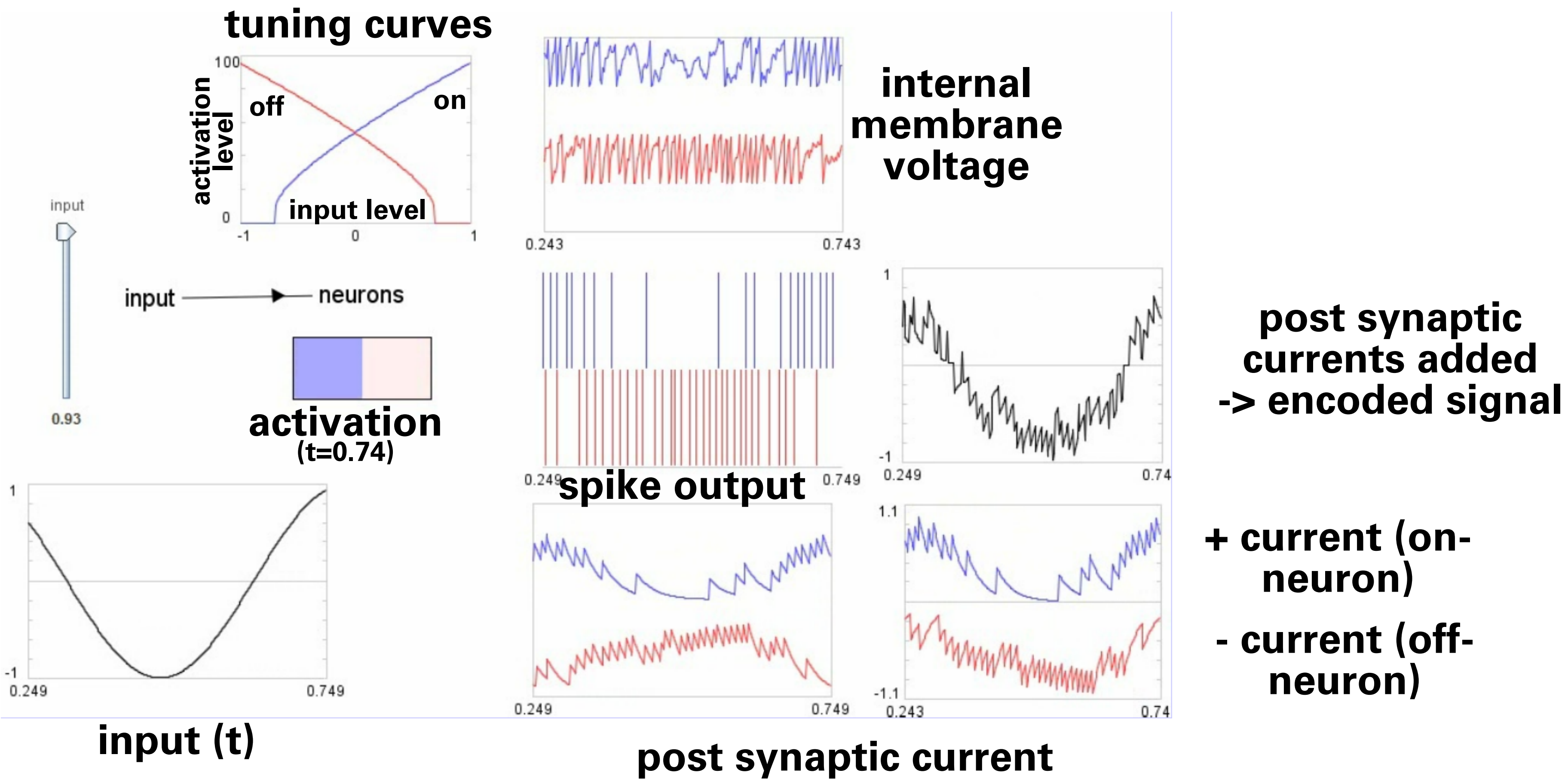
- Neuron = Information processing unit:
- Presynaptic activity (potential) triggers postsynaptic activity (current, based on number of spikes per time interval)
- ... triggers activity level of each neuron cell (soma)

- First goal: representing an input value range by neural activity
- -> we need more than one neuron!

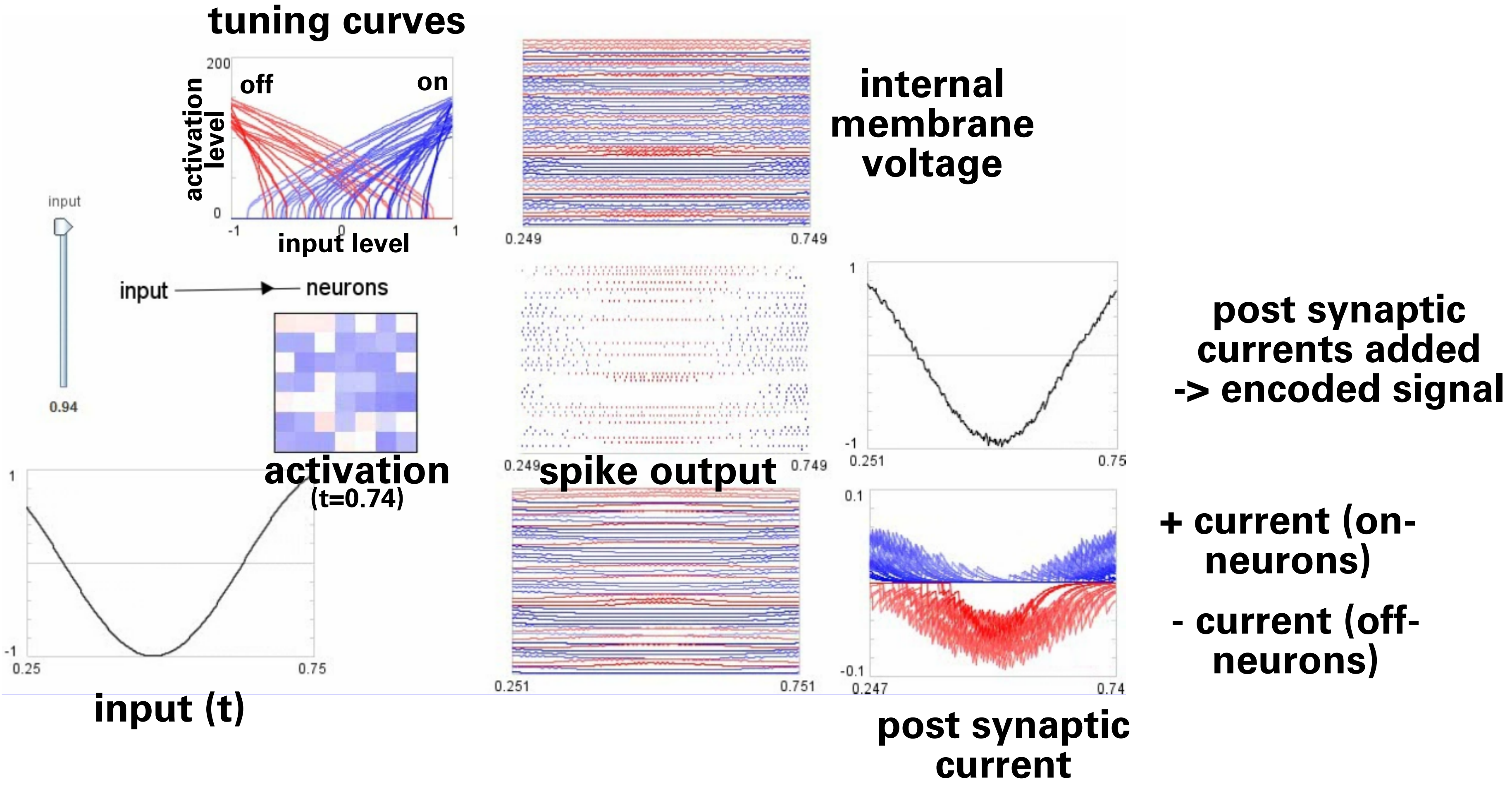
Example: one neuron, sensory input



Example: two neurons

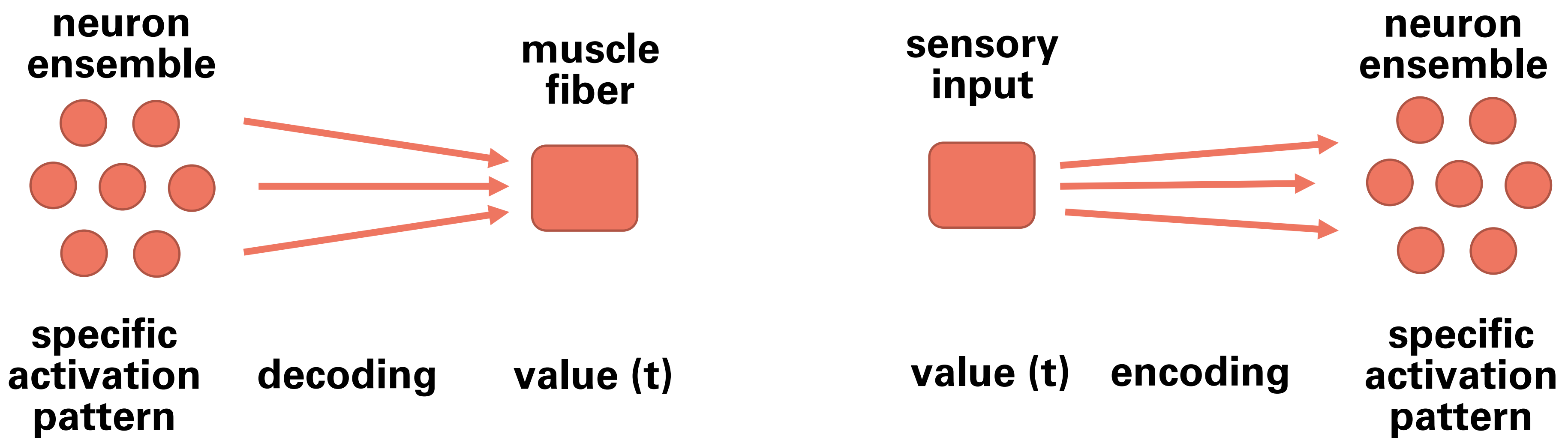


Example: 50 neurons



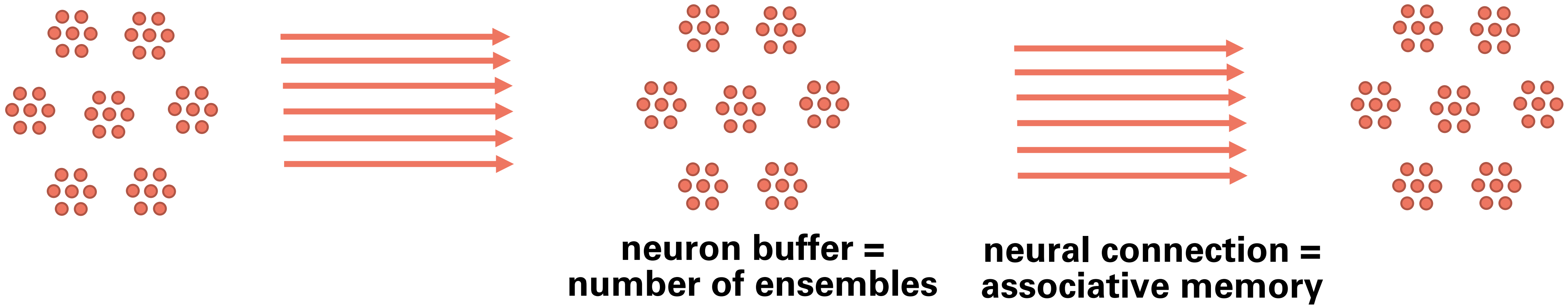
50-100 neurons -> neuron ensemble

- represent “values” (e.g., a muscular activation level; a specific sensory input intensity level over time)
- each value (winthin the range) is
  - represented by specific neural activity within the ensemble
  - en-/decoded by synaptic weights (connections between nenurons)



50-100 ensembles -> neuron buffer

- SPA (Semantic Pointer Architecture as part of NENGO)
- buffer represents complex information (items)
- -> coded by a number of “values”; mathematically represented by “vectors” (S-pointers)
- S-pointers represent:
  - phonetic or phonological items (sounds, syllables words)
  - lexemes (language-specific)
  - concepts (meanings) ((-> “thoughts”))

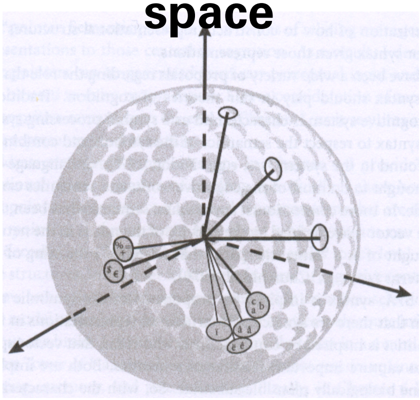


NENGO: Decoding complex info (S-pointers)

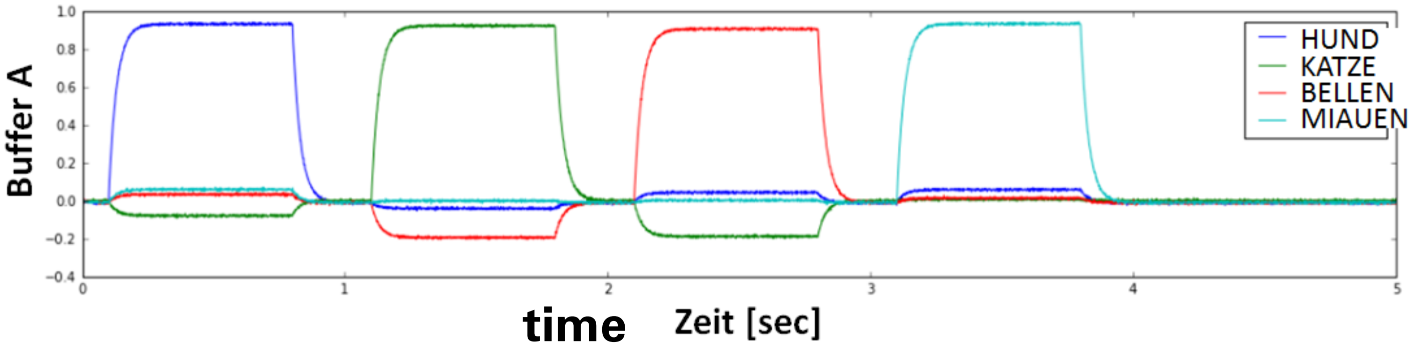
- by: S-pointer activity plots "similarity plots" for each neuron buffer
- only the most activated items within a buffer are shown (similar items are co-activated)

neuron buffer

D-dimensional space



S-pointer activity plot over time

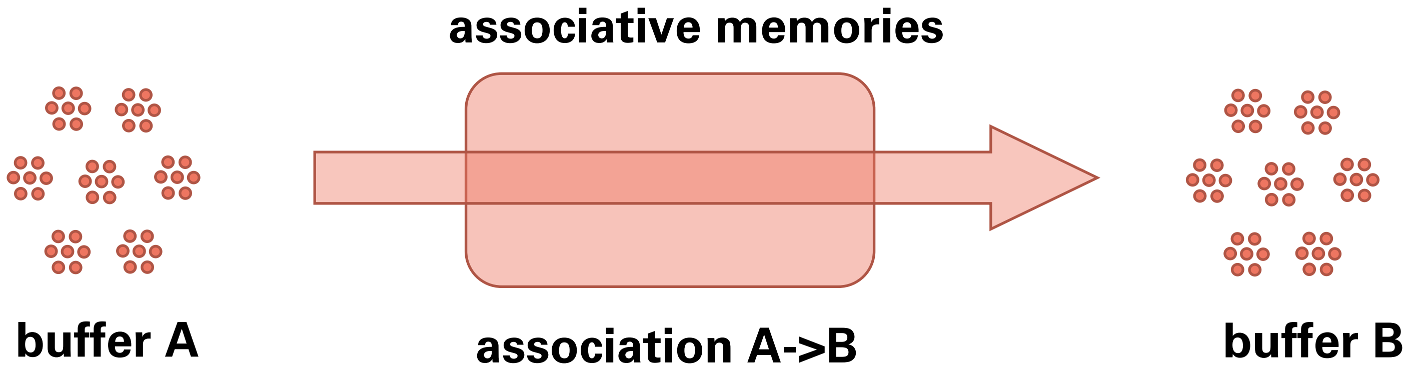


time

dog
cat
to bark
to meow

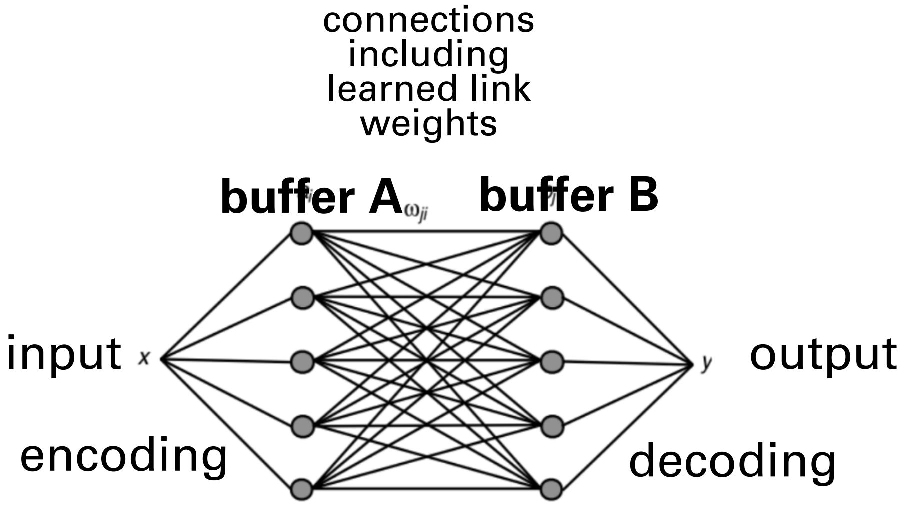
NENGO: associative memories

- define the associations of items activated in different buffers: e.g., from phonological form /dOg/ -> lexeme "dog"-> concept <dog/chien/Hund>
- neural connections between different buffers
- develop by learning -> adjustment of synaptic link weights



Neural associations

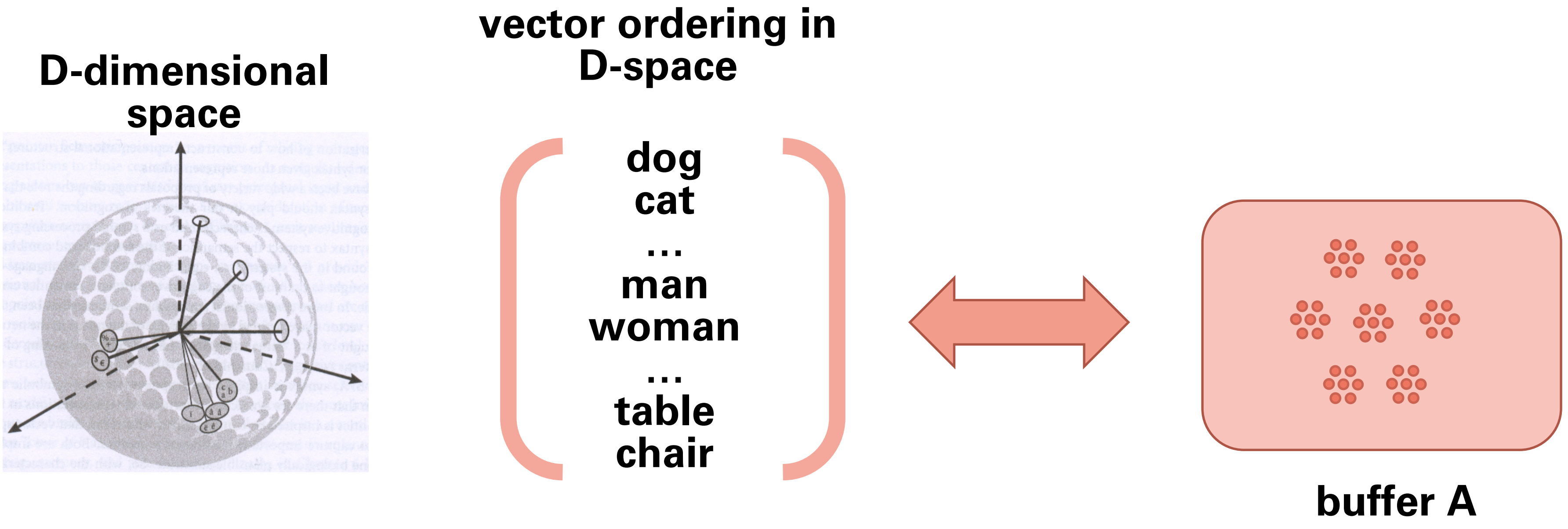
- connect each neuron with each other neuron



transformation of states by neural associations

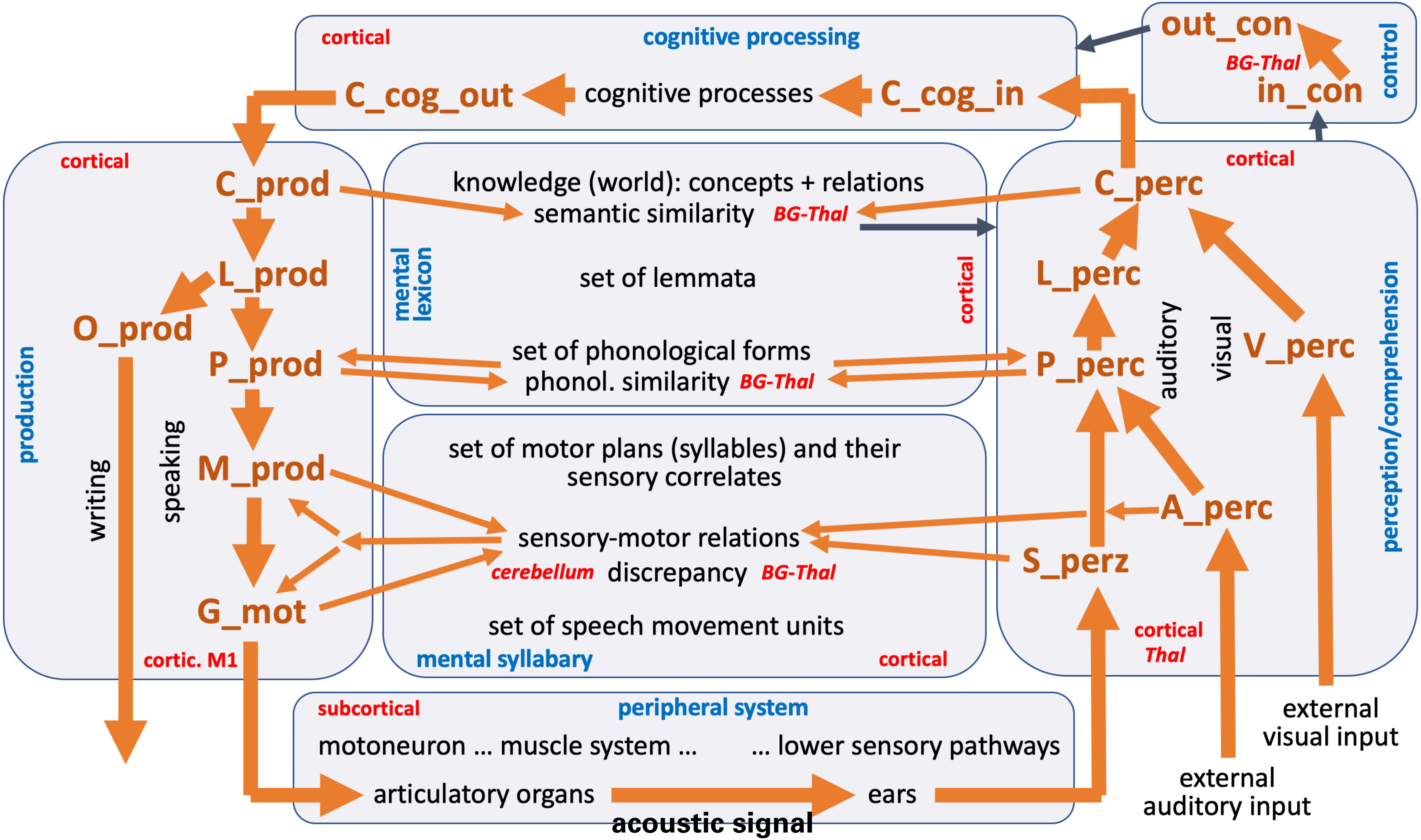
NENGO: S-pointer-networks

- define the associations of items within a buffer: e.g.,:
  - at concept level: <dog> - <cat> -> animals; <chair> - <table> -> furniture
  - at phonological level: /car/ - /cat/ -> begin with /k/; /far/ - /fat/ -> with /f/
- similar S-pointers point in a similar direction:



Speech processing model: Architecture

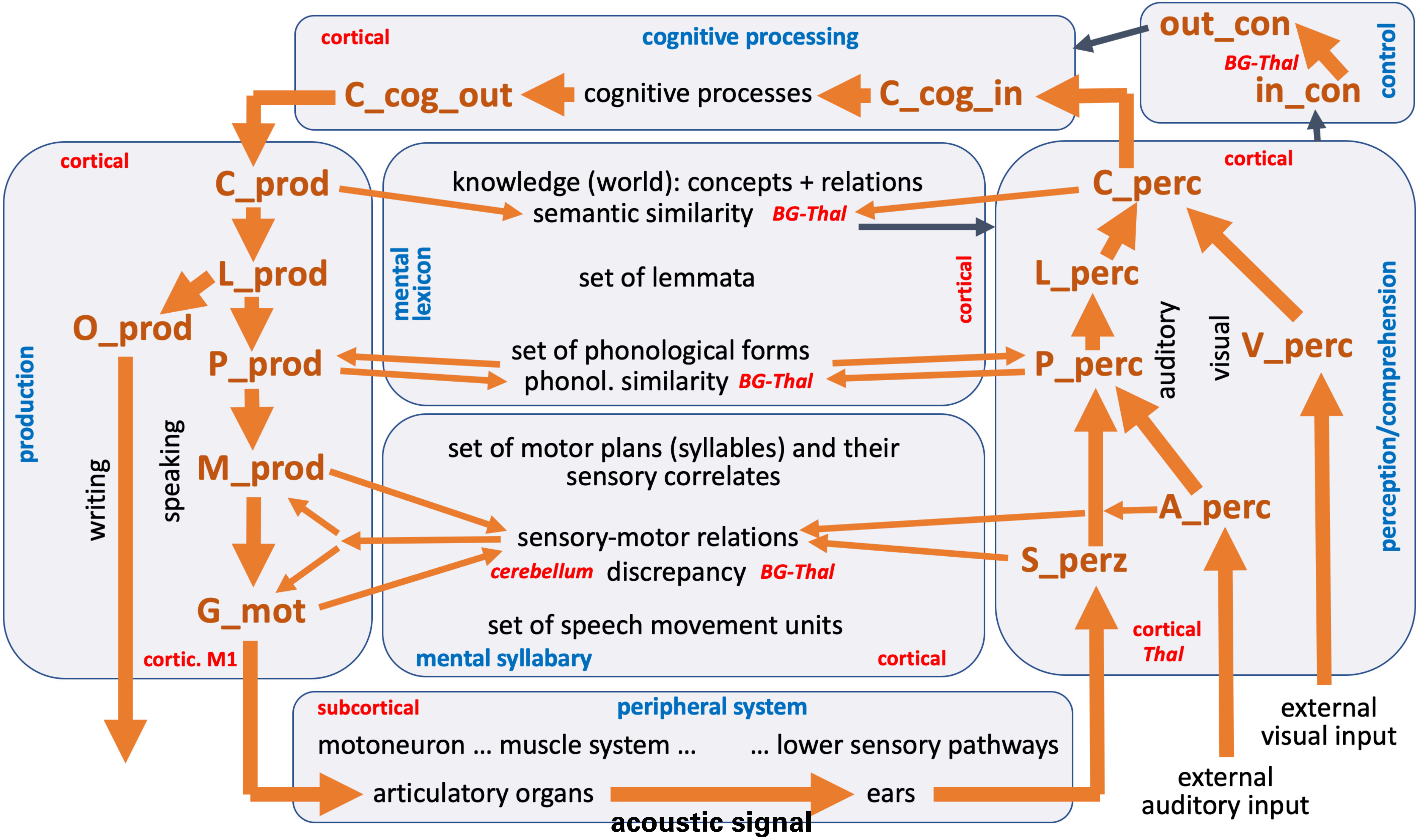
- The complete large-scale model
- Perception pathway: Auditory input (values, processed in neuron ensembles) -> cognition (understanding; comprehension) ((buf -> assoc buf -> buf -> assoc buf …))
- Production pathway: message, word (S-pointer, processed in neuron buffers) -> articulatory-acoustic output ((same))
- Including: mental lexicon, mental syllabary as skill or knowledge repositories ((learned S-pointer-networks!))



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity BG-Thal
set of lemmata
set of phonological forms
phonol. similarity BG-Thal
mental lexicon
cortical
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum discrepancy BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal
external visual input
external auditory input

Speech processing model: Architecture

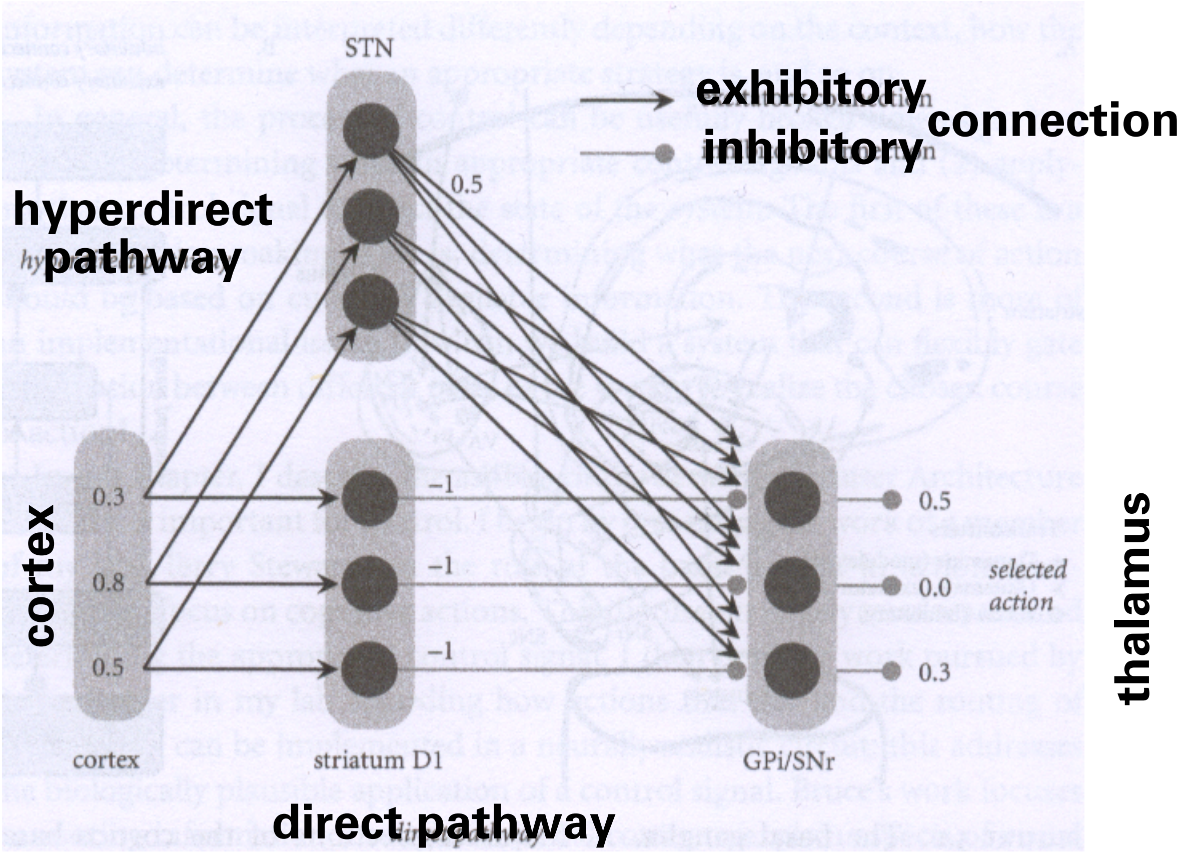
- In addition:
- Internal feedback loops: semantic level, phonological level, sensorimotor level
- External feedback loop: motor-articulatory-acoustic-auditory



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal
external visual input
external auditory input

NENGO: Basal ganglia and Thalamus

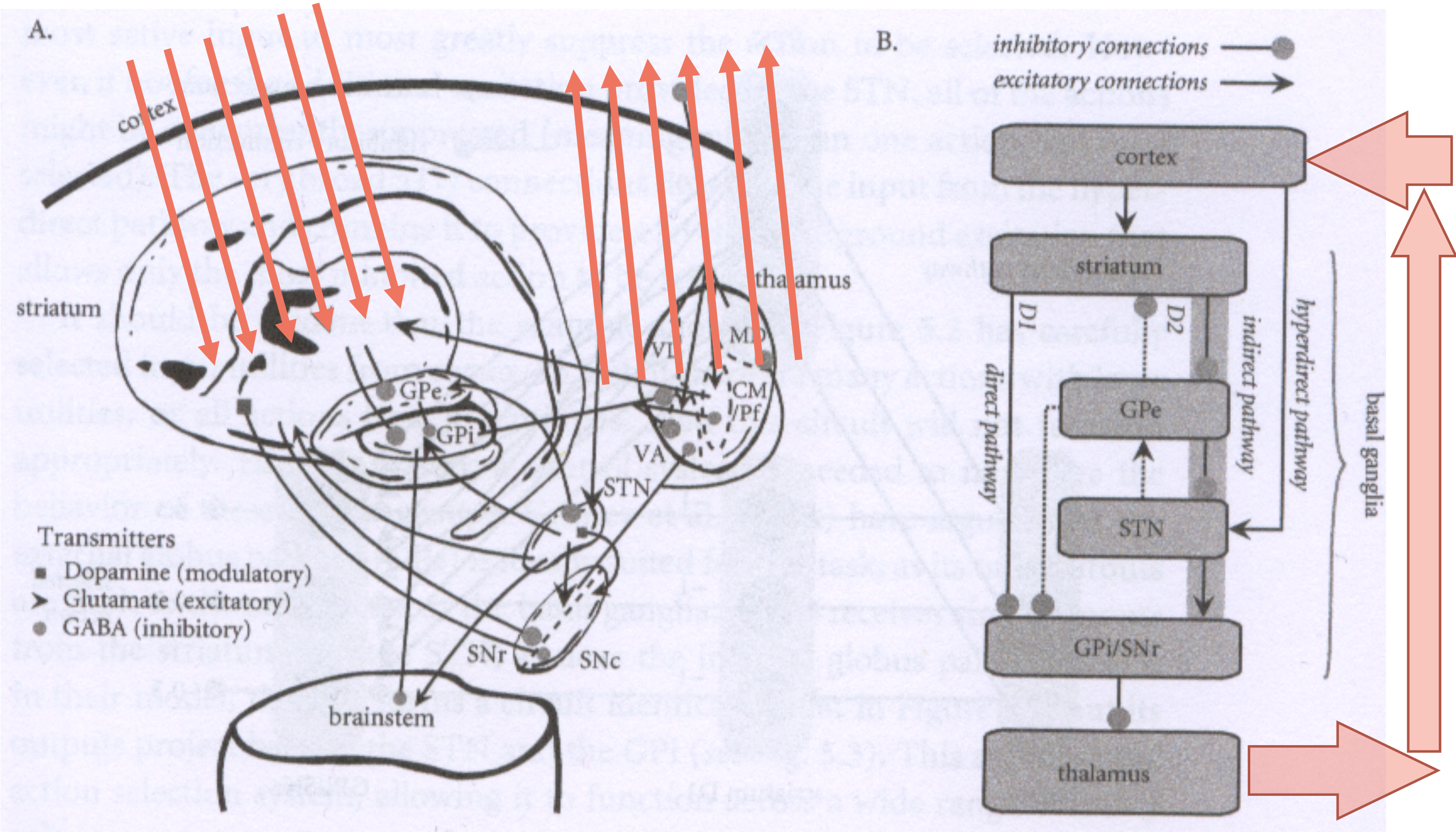
- the cortico-cortical control loop
- ... controls the temporal sequence of (neural) actions:
  - calculate utility values for each available action (cortex -> striat)
  - choosing the most appropriate action as next action based of the set of current utility values (thalamus)
  - Modules (neuron clusters) of BG:
    - striatum,
    - STN: subthalamic nucleus
    - GPi: globus pallidus, ...



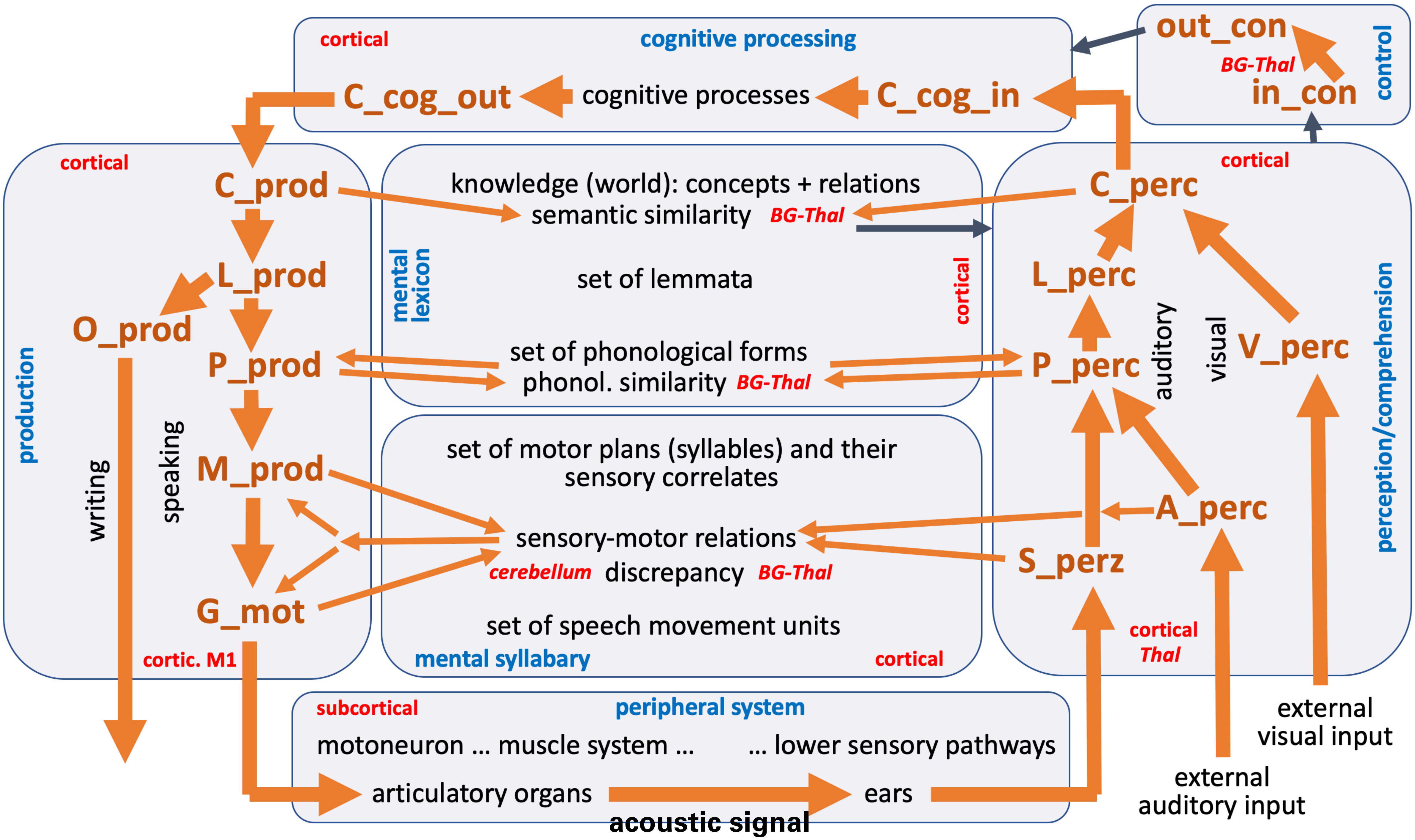
Eliasmith (2013)

BG-Thal: the cortico-cortical loop

- direct / hyperdirect pathway
- different types of neurons (of synaptic connections) in different neuron clusters
- inhibitory and exhibitory pathways

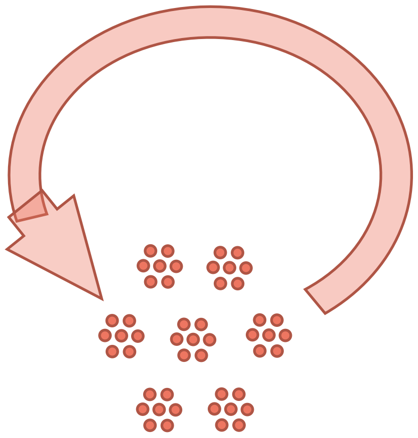


Eliasmith (2013)


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
external visual input
external auditory input
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal

The cognitive processing level

- Short-term memories, e.g., for memorizing word lists (-> serial recall task)
- realized as recursive buffers (neural within-associations)



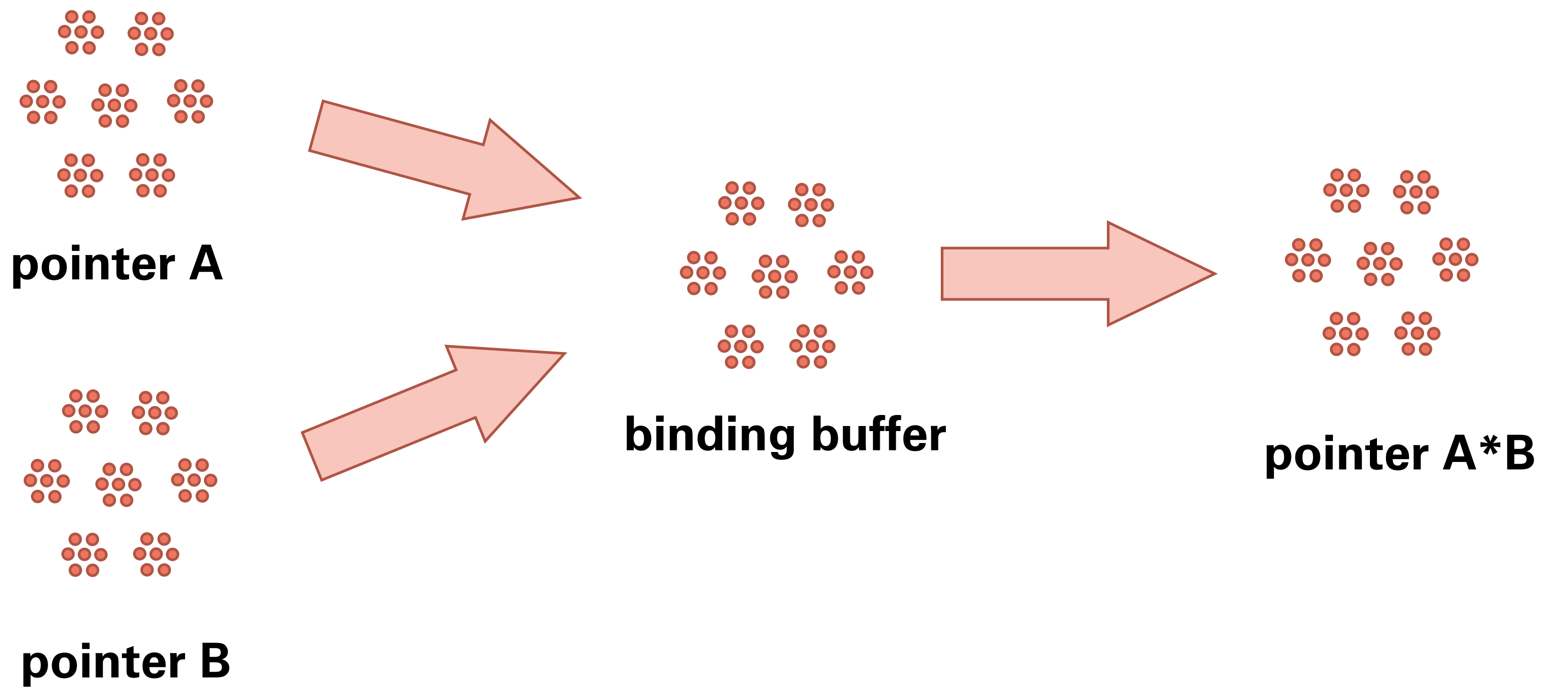
recursive buffer

The cognitive processing level

- Cognitive processing needs:
- Binding and inverse binding (unbinding), e.g., for building sentences (-> the dog barks and the cat meows)
  - binding: <dog>*<to bark> ; <cat>*<to meow>
  - Sentence: (<dog>*<actor> + <bark>*<acting>
  - unbinding: what can the dog / cat do? -> <acting>$^{-1}$
- … needed for memorizing orderings of list items (-> serial recall vs. free recall)

The cognitive processing level

- Binding / unbinding is realized by specific neural structures -> binding buffers vs. normal S-pointer buffers



The cognitive processing level

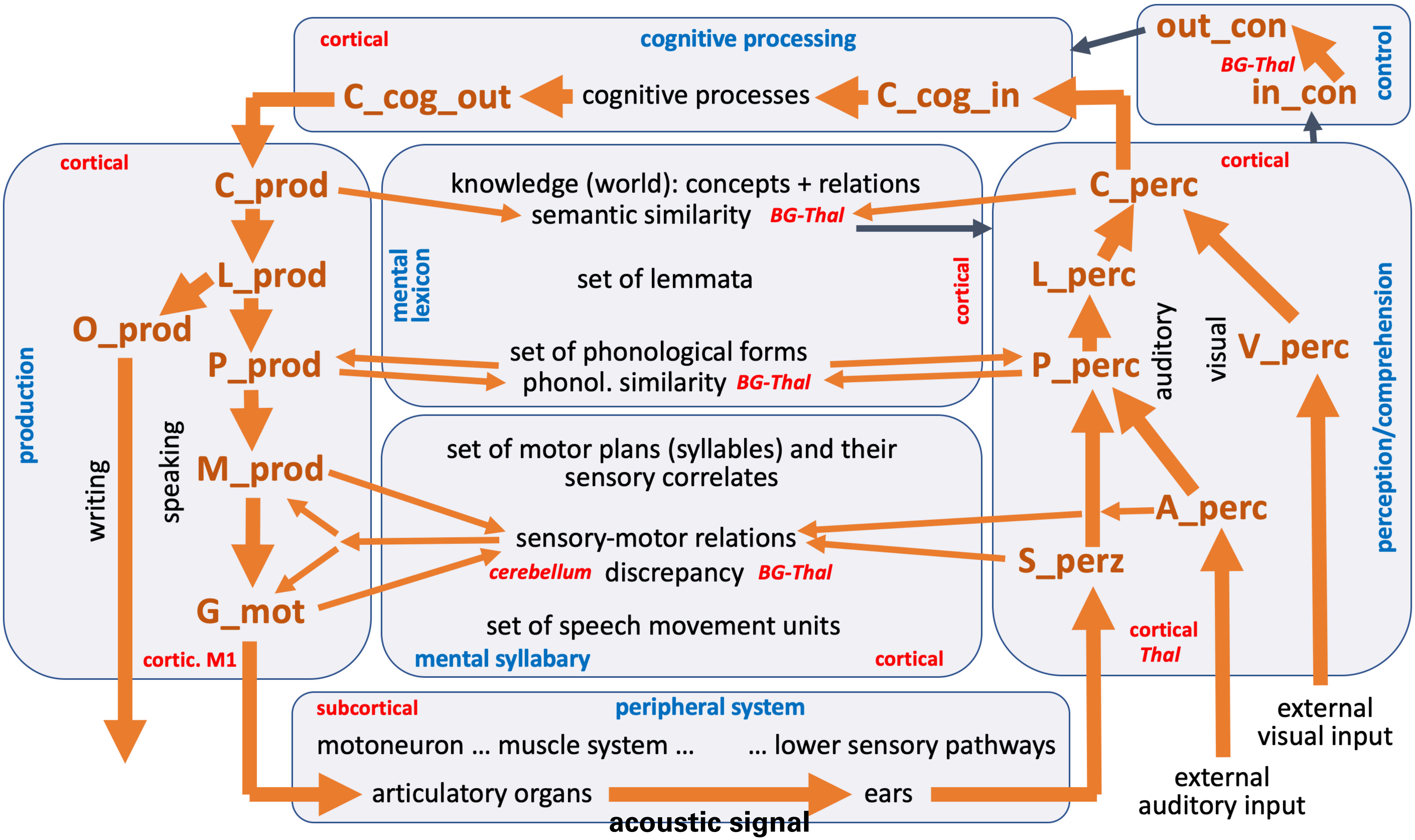
- Activating superordinate terms (e.g., a dog is an animal; a chair is a furniture, is an object)
- binding processes in S-pointer networks, using relations like “is a”: <chair>*<is_a> => <object>

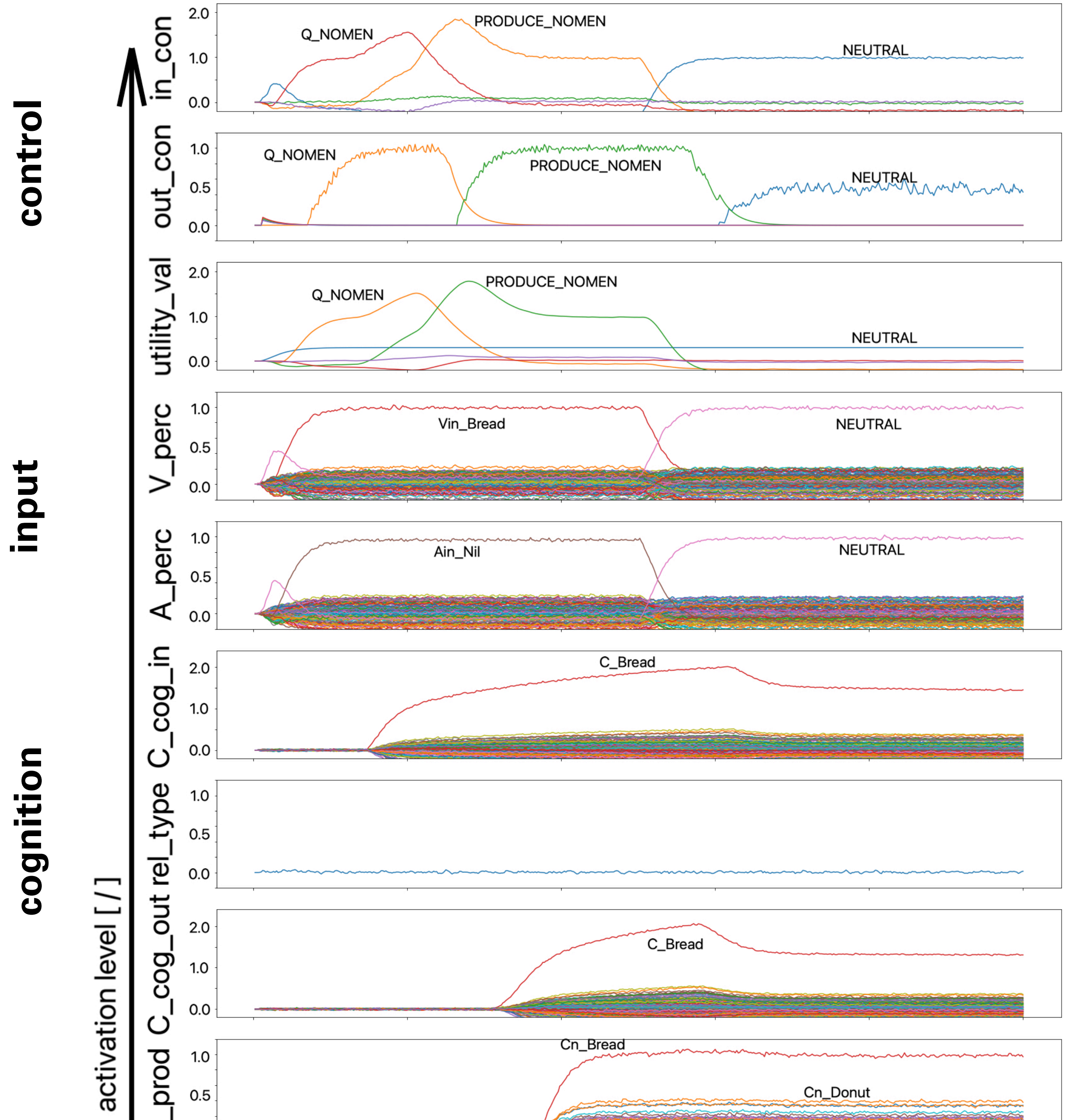
Download Video3:
https://www.youtube.com/watch?v=fDNNuM0H3j8

The definition of scenarios

- Our large-scale model is more than just an architecture:
- In addition: we can define the "world" in which the speaker has to "act" (define scenarios)
- Scenario + model (avatar) -> simulated behavior
- Scenarios: speech screening tasks: (medical screenings)
  - Picture naming (visual input -> production)
  - Word comprehension (auditory input -> naming of superordinate term)
  - Nonsense word (logatome / syllable) repetition

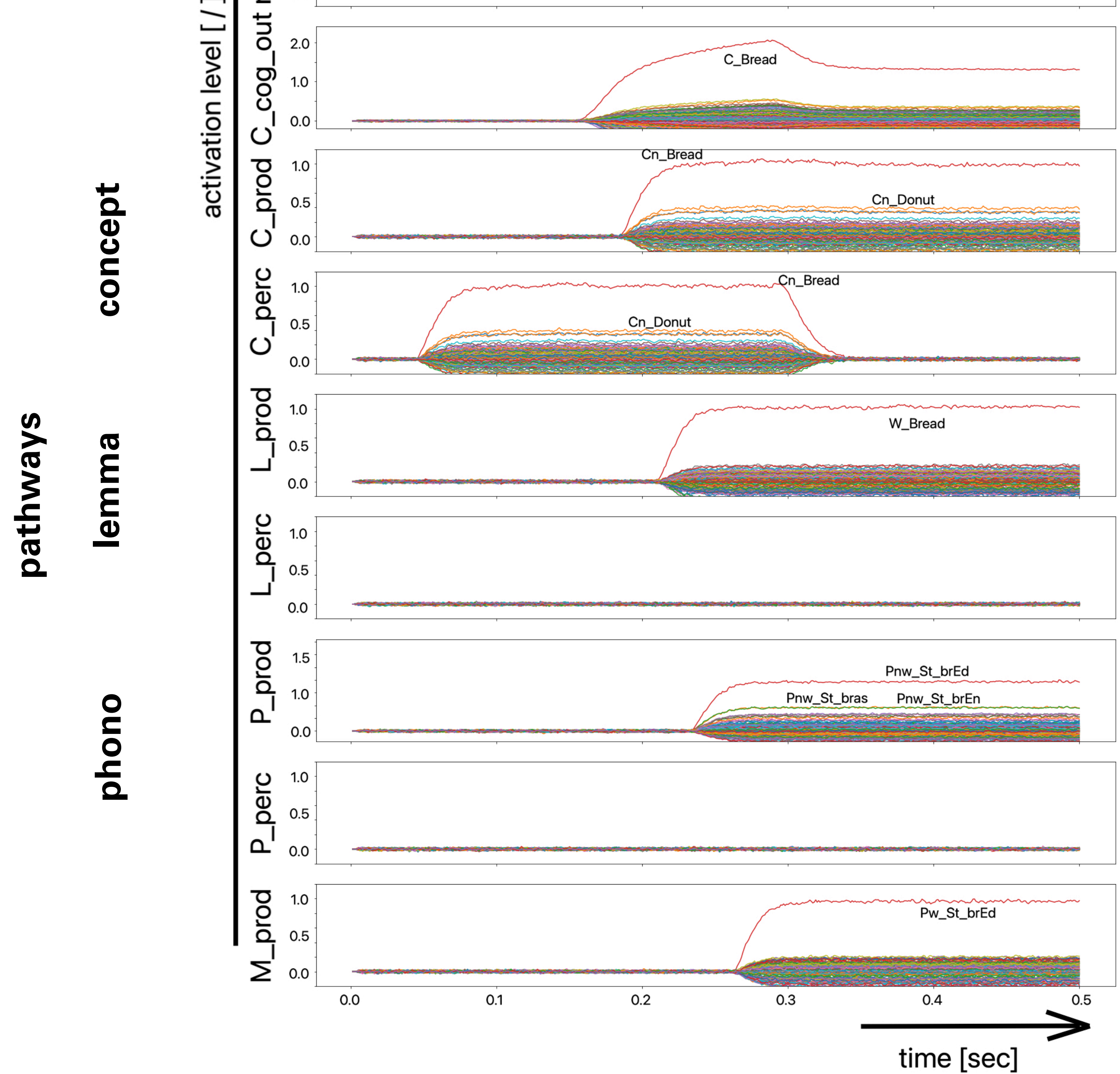
Kroeger BJ, Stille C, Blouw P, Bekolay T, Stewart TC (2020) Hierarchical sequencing and feedforward and feedback control mechanisms in speech production: A preliminary approach for modeling normal and disordered speech. Frontiers in Computational Neuroscience, 14:99. www.speechtrainer.eu **-> publications**


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their
sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal
external
visual input
external
auditory input

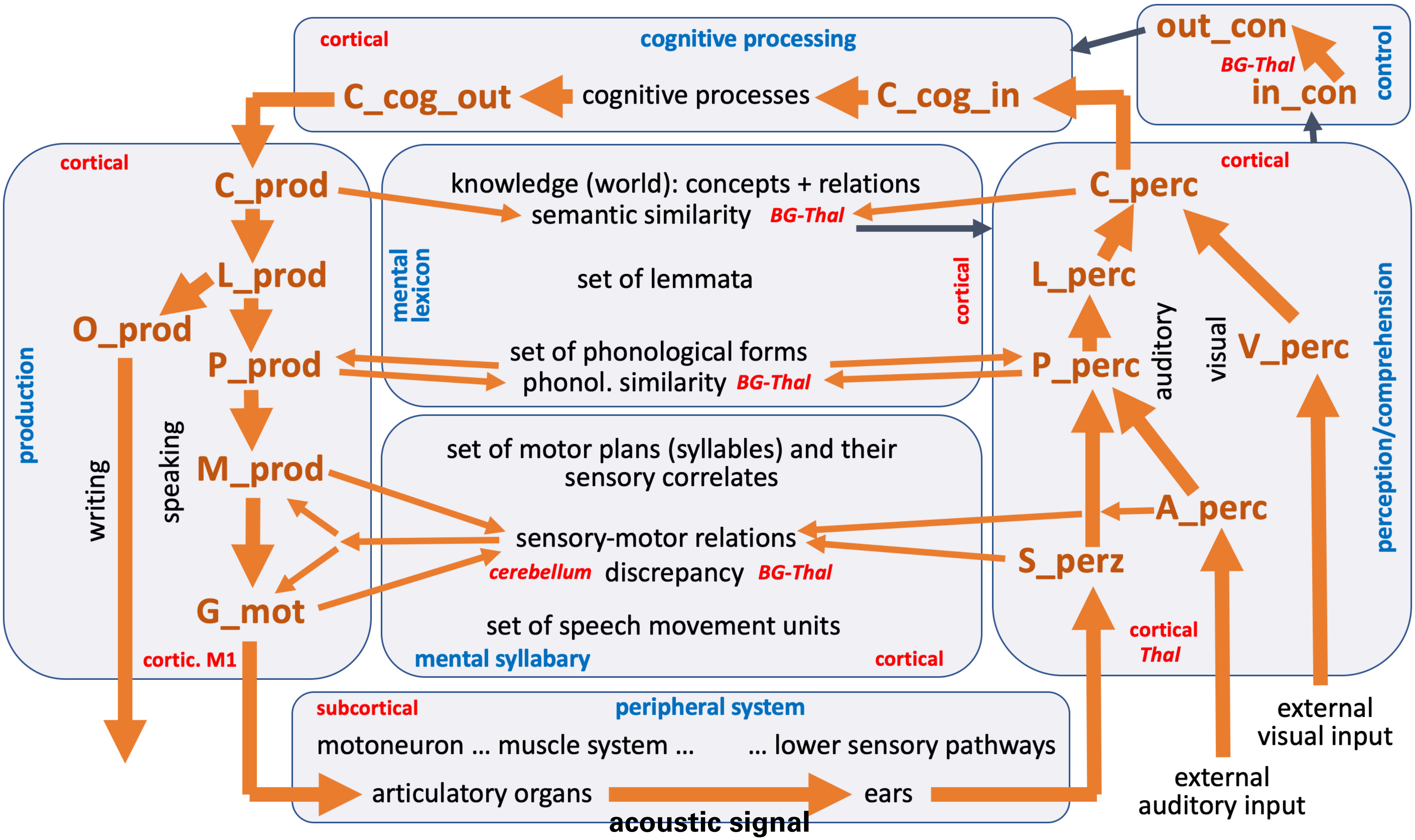


← **priming: 100 msec: "look@pic and activate noun"**
+150 msec : "utter noun"

Kröger et al. (2020)

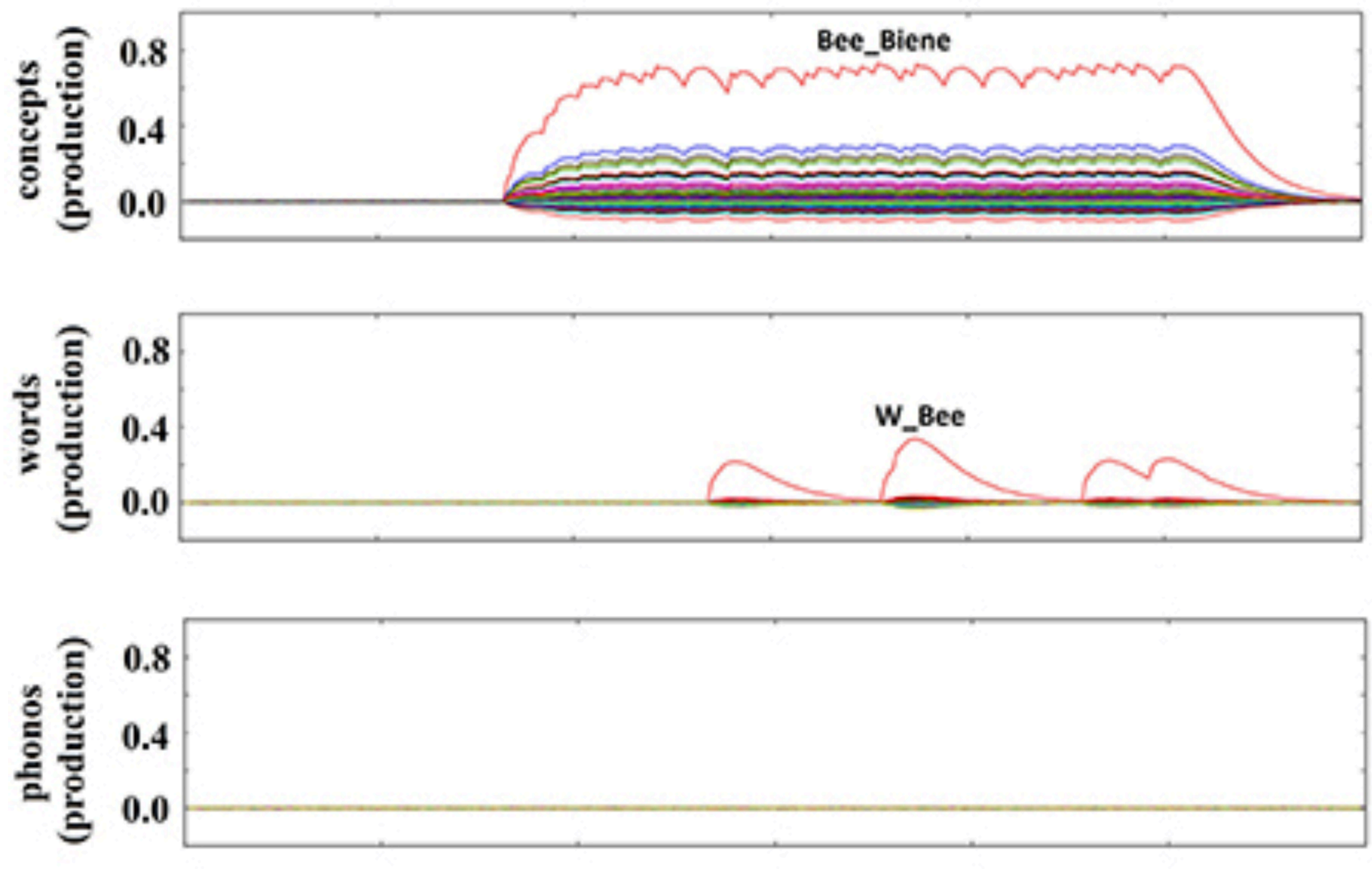


Kröger et al. (2020)


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
external visual input
external auditory input
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal

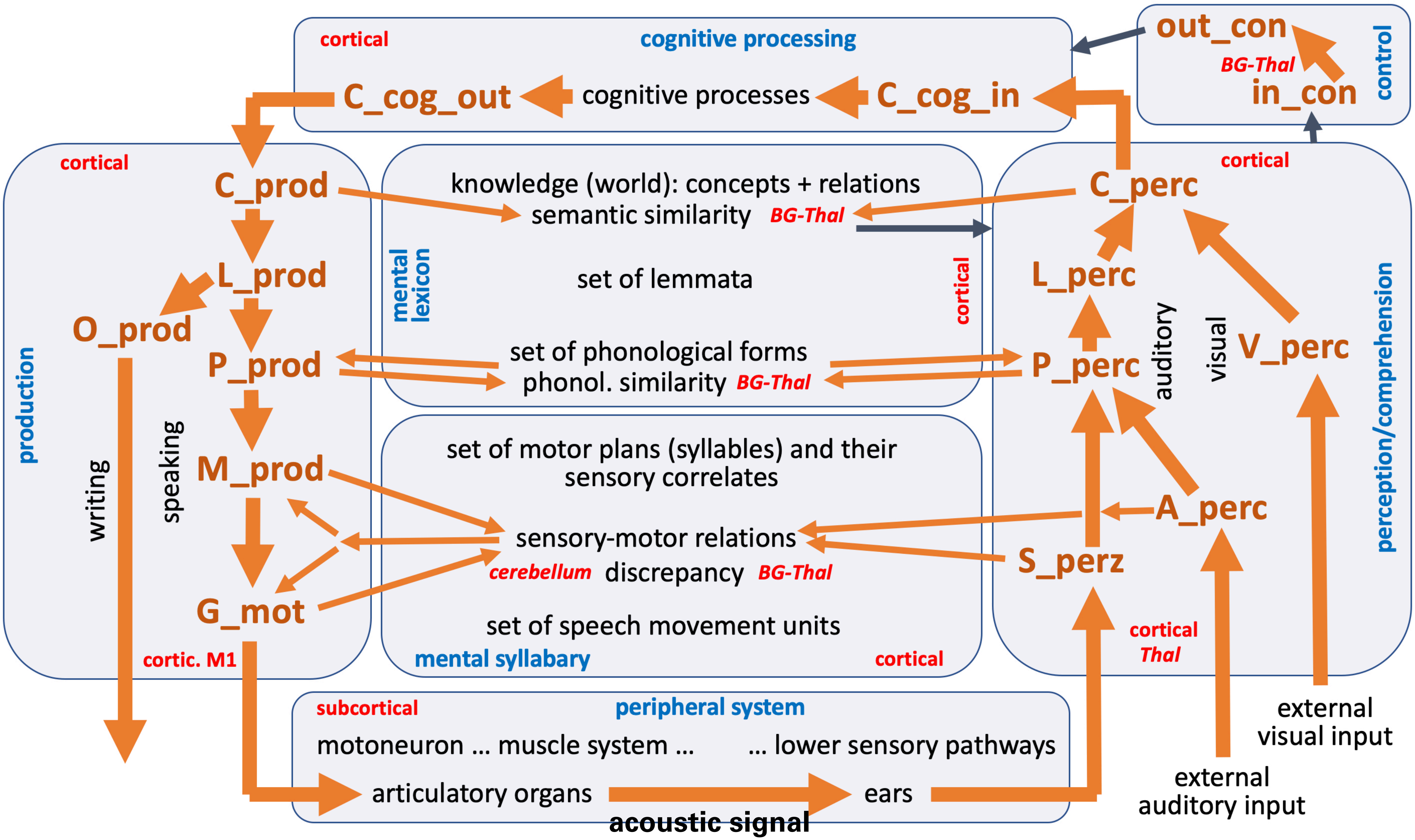
The definition of scenarios

- Speech screening tasks:
  - Picture naming (visual input -> production) rare cases (<1% of simulations)
  - -> underline the neural character of the model (neural spikes are statistic events! -> neural signals are noisy!)



normal model;
no neural
dysfunctions
implemented!

Kröger BJ, Crawford E, Bekolay T, Eliasmith C (2016) Modeling interactions between speech production and perception: speech error detection at semantic and phonological levels and the inner speech loop. Frontiers in Computational Neuroscience 10:51 www.speechtrainer.eu -> publications

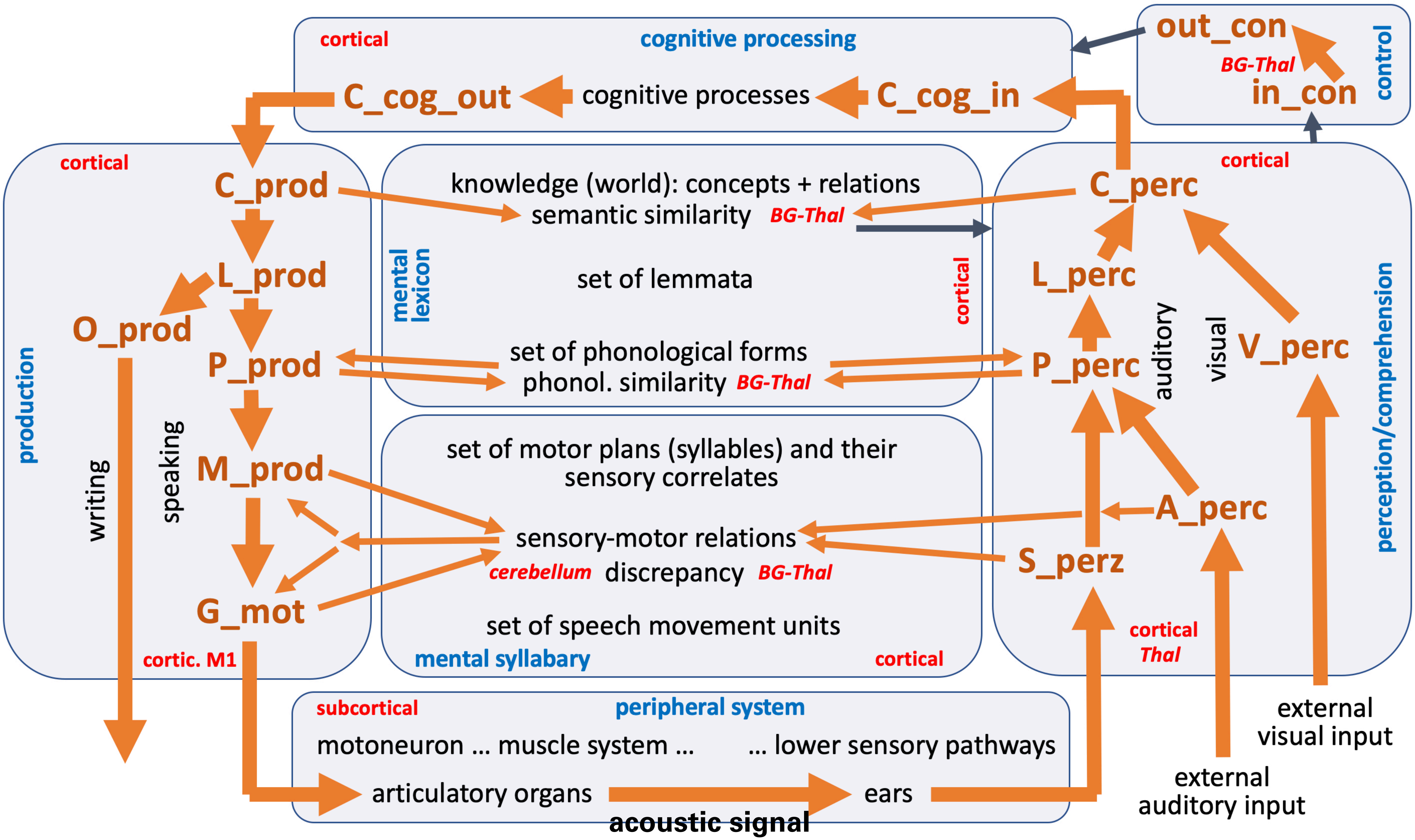


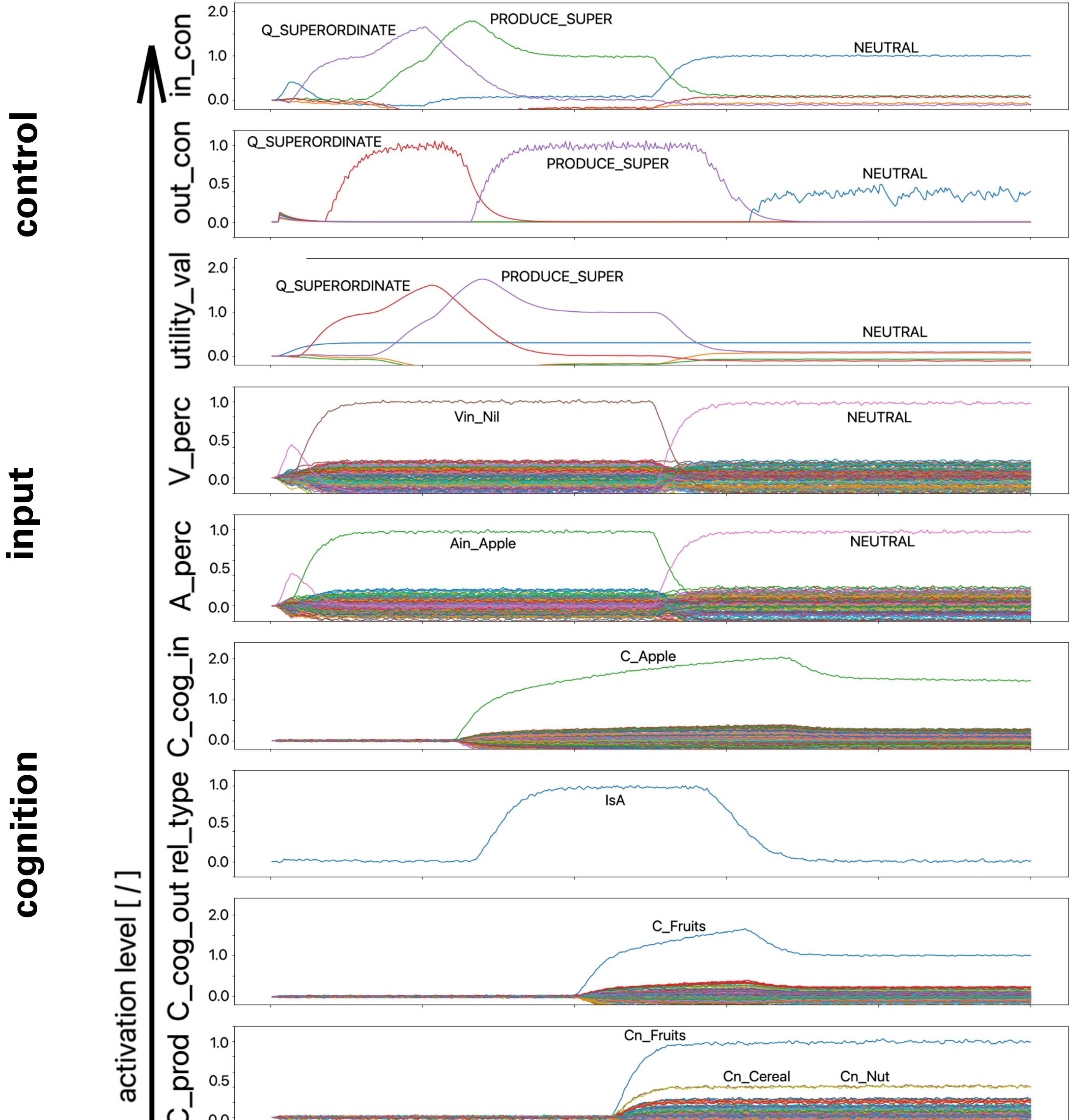
cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their
sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron … muscle system …
… lower sensory pathways
articulatory organs
ears
acoustic signal
external
visual input
external
auditory input

The definition of scenarios

- defines the “world” in which the speaker has to “act”
- -> behavior
- Speech screening tasks:
  - Picture naming (visual input -> production)
  - Word comprehension (auditory input -> naming of superordinate term)
    - Is a cognitive task including binding
  - Nonsense word (logatome) repetition

Kroeger BJ, Stille C, Blouw P, Bekolay T, Stewart TC (2020) Hierarchical sequencing and feedforward and feedback control mechanisms in speech production: A preliminary approach for modeling normal and disordered speech. Frontiers in Computational Neuroscience, 14:99


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal
external visual input
external auditory input



← priming: 100 msec: ”listen and think” +150 msec : “utter noun”

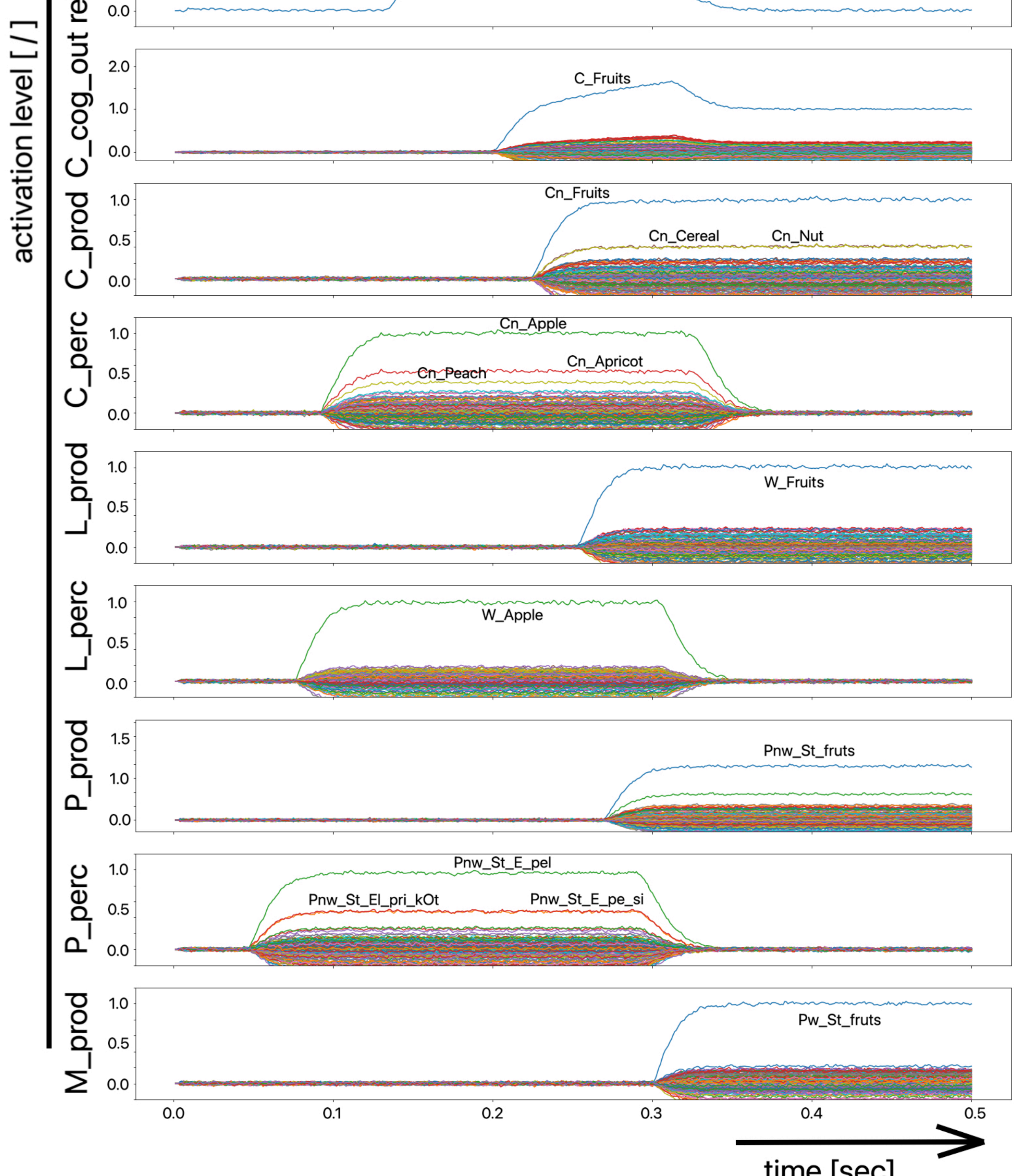
Kröger et al. (2020)

pathways

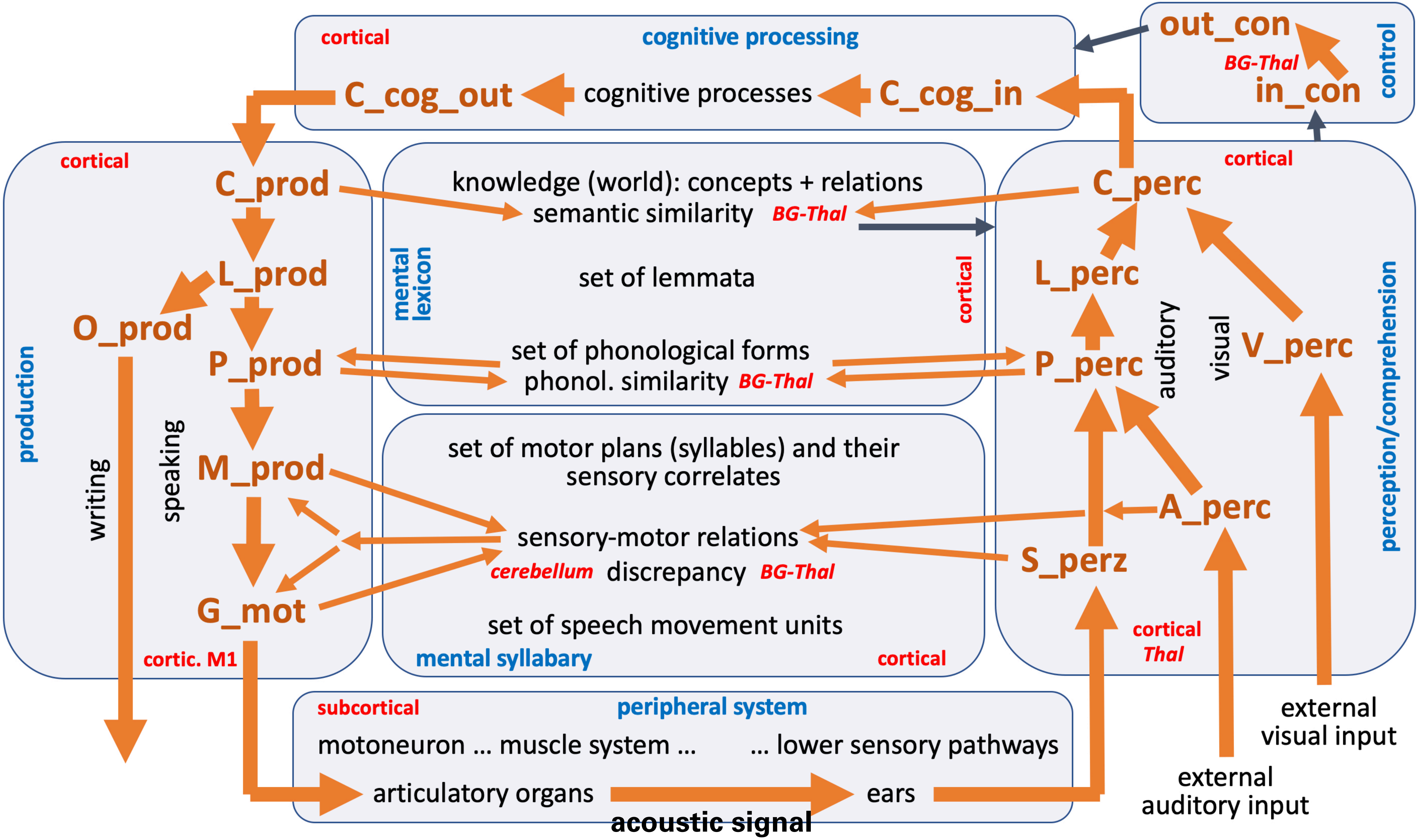
concept

lemma

phono

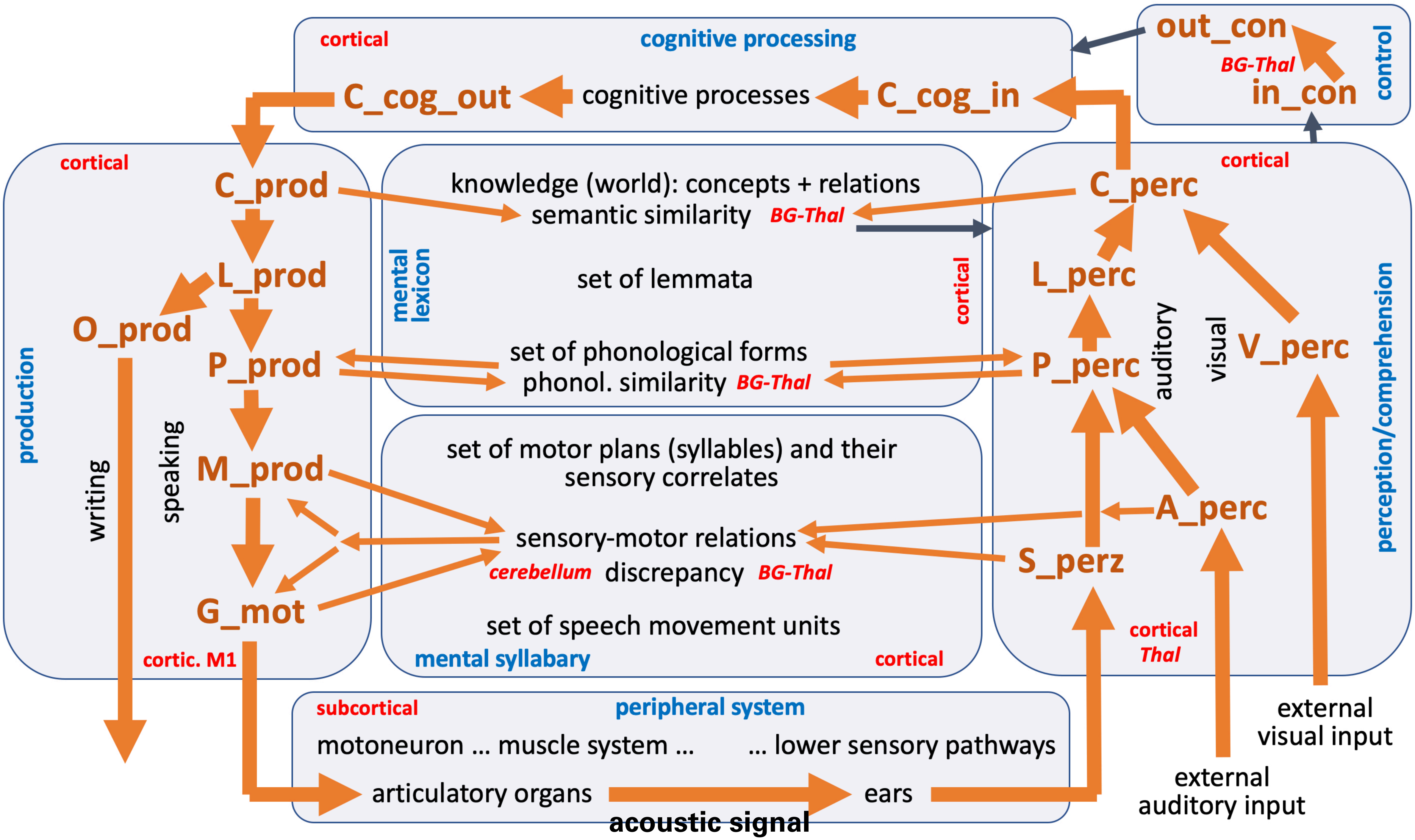


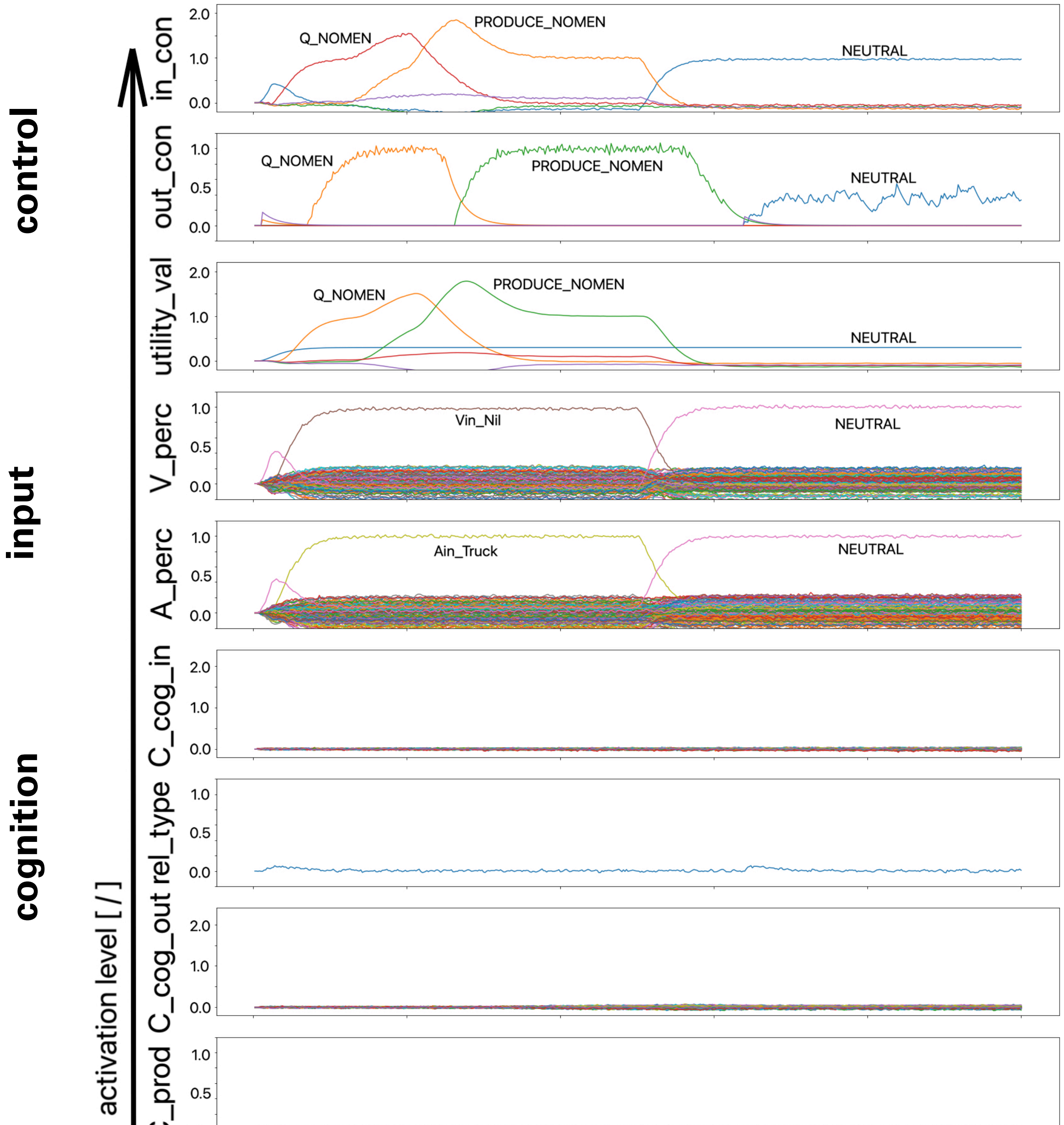
Kröger et al. (2020)


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal
external visual input
external auditory input

The definition of scenarios

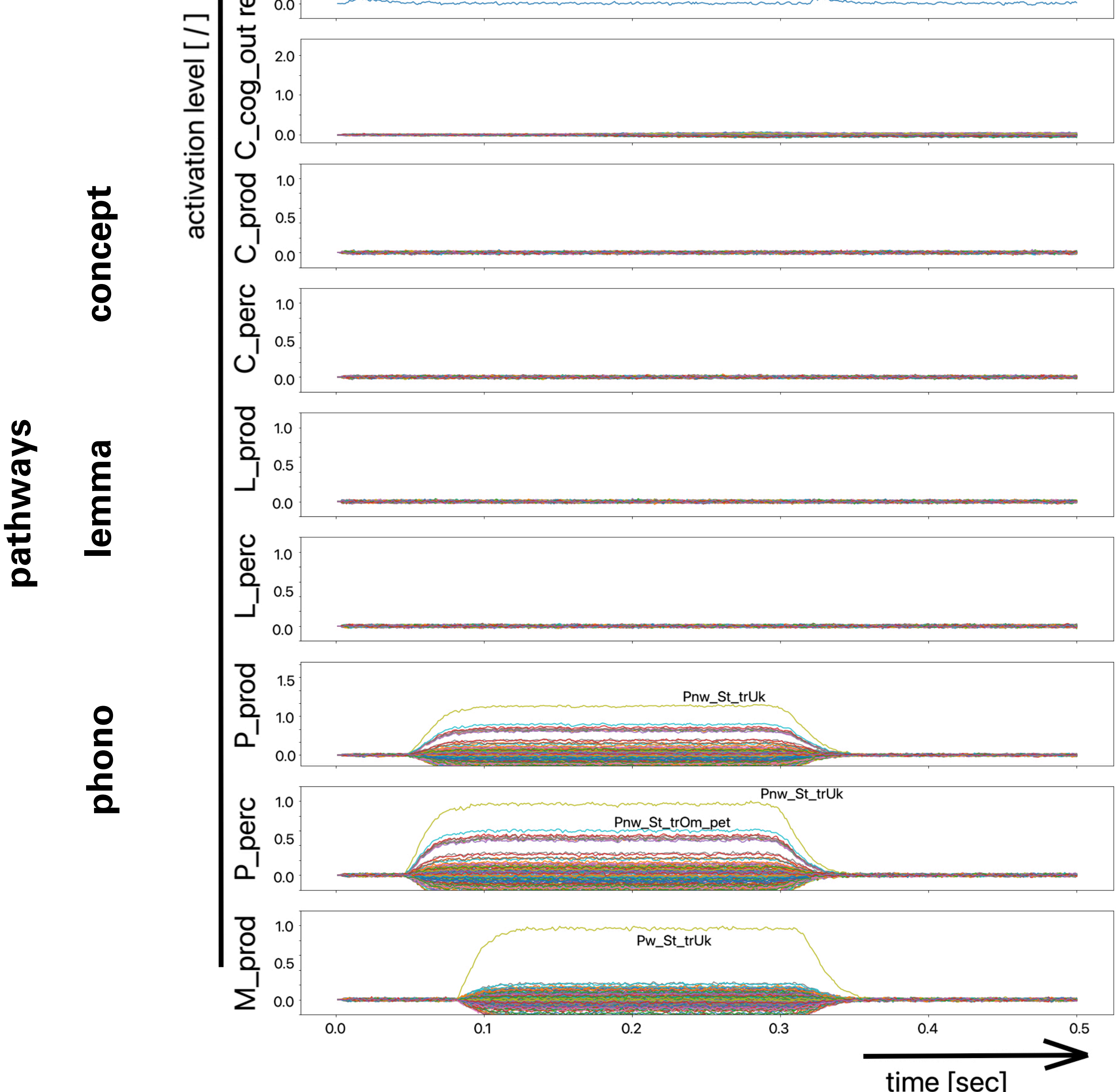
- defines the “world” in which the speaker has to “act”
- -> behavior
- Speech screening tasks:
  - Picture naming (visual input -> production)
  - Word comprehension (auditory input -> cognitive activation and naming of superordinate term)
  - Nonsense word (logatome) repetition
    - Do not need higher level processing; just shortcut at phonological level
    - Phono-shortcut means: repetition without comprehension; without any lexical access


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
perception/comprehension
A_perc
S_perz
cortical
Thal
external visual input
external auditory input
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal

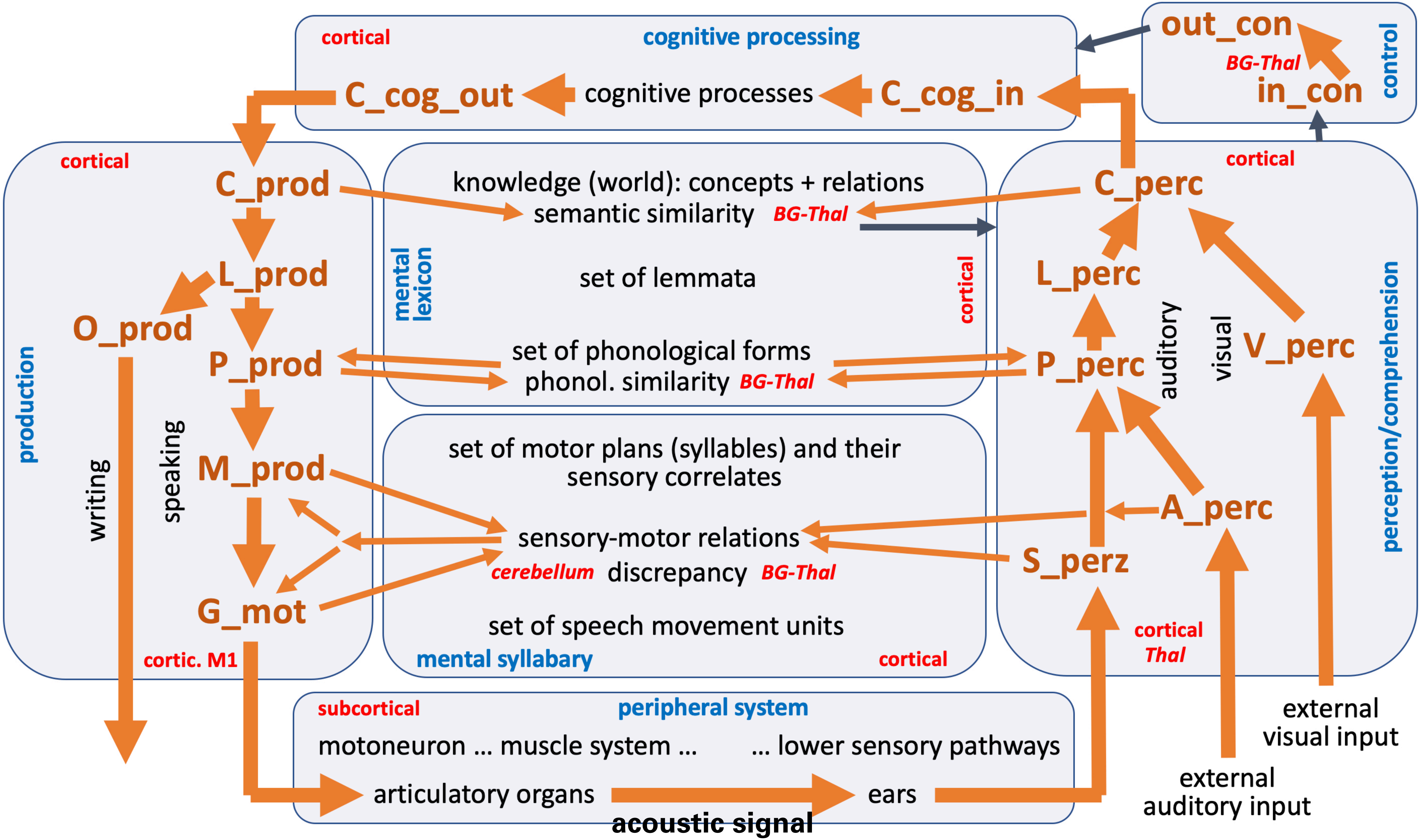


← **priming: 100 msec: "listen" + "reproduce"**
150 msec : "further reproduce"

Kröger et al. (2020)



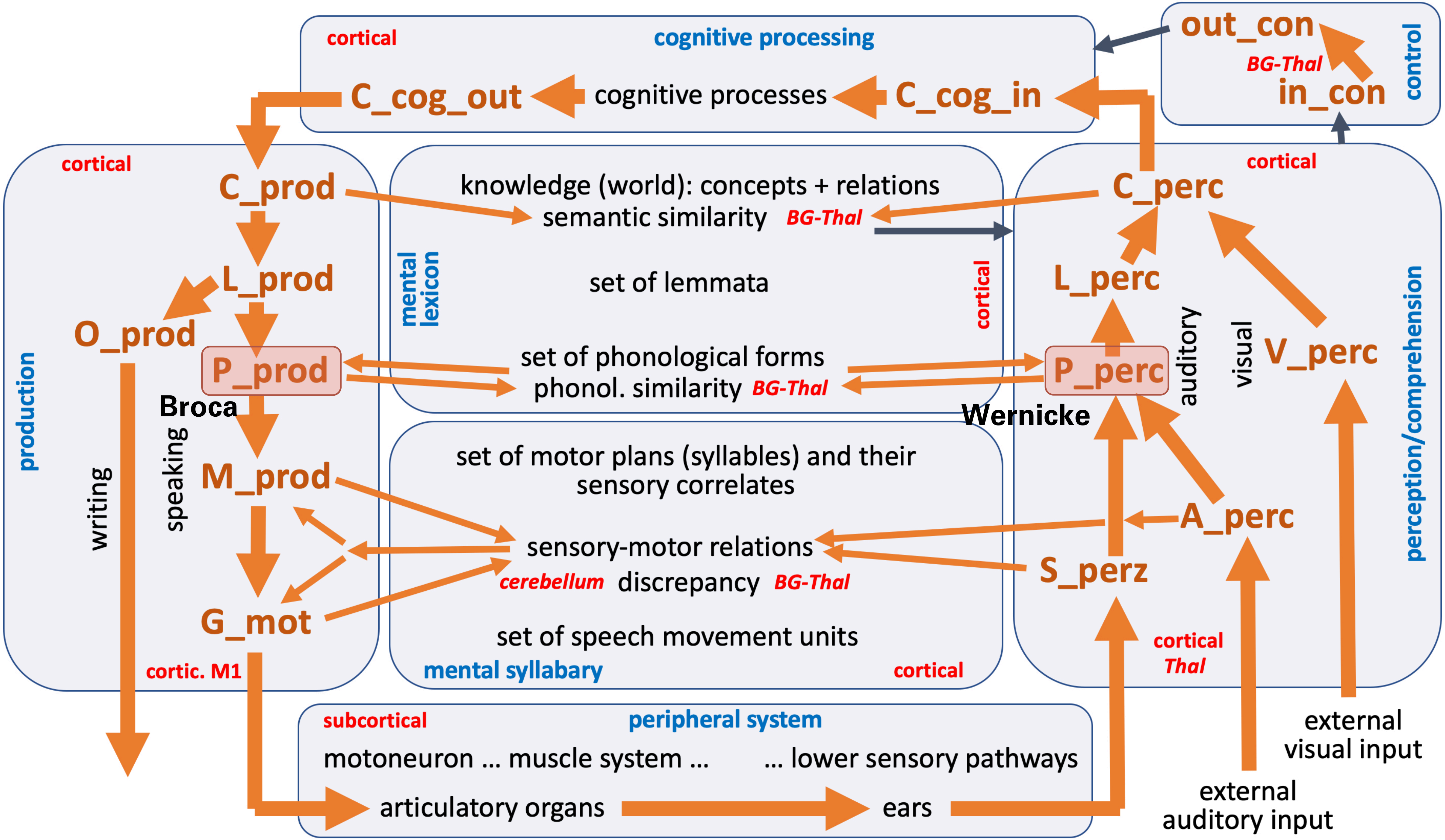
Kröger et al. (2020)



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
acoustic signal
external visual input
external auditory input

Integration of neural dysfunctions

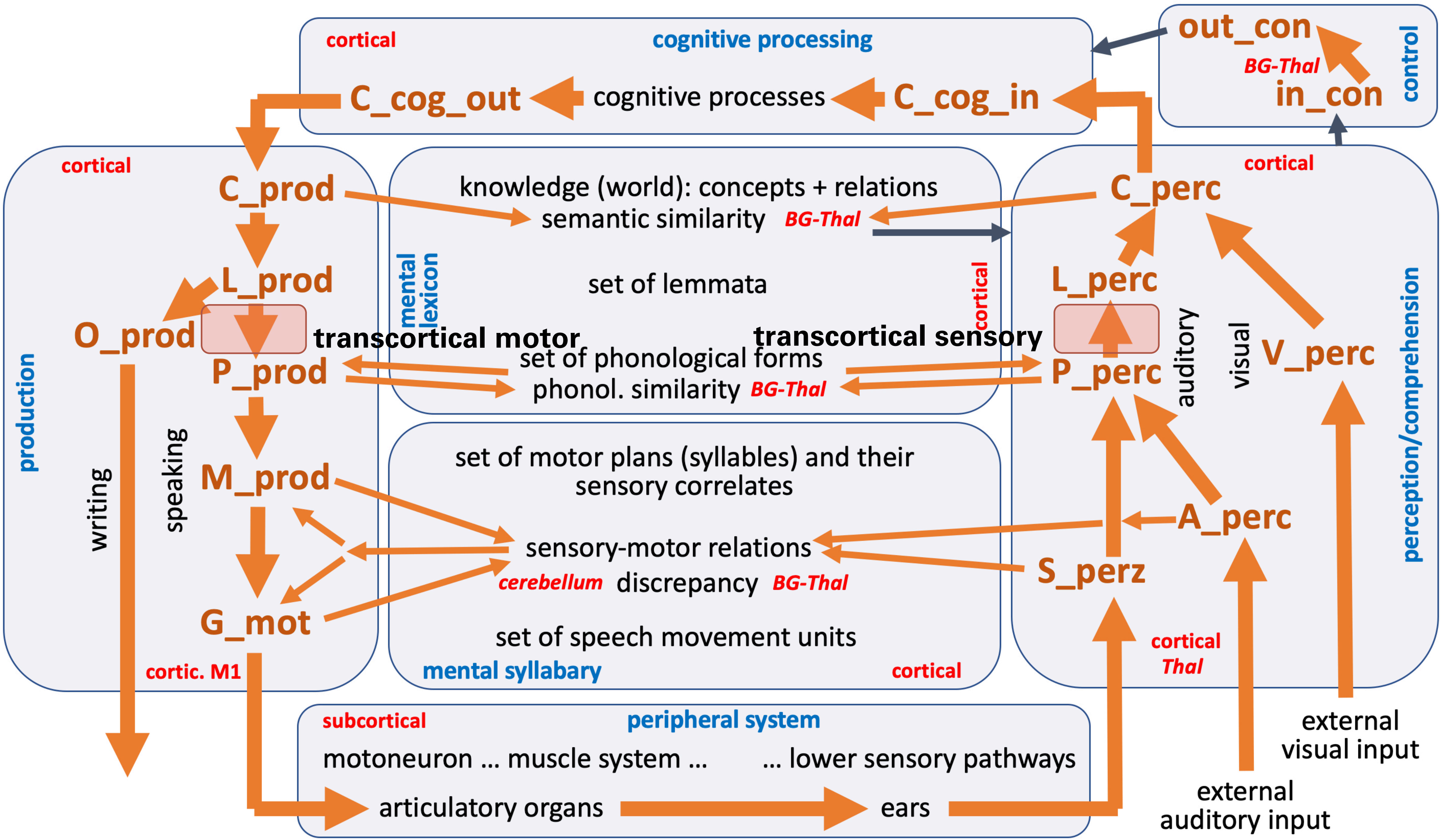
- We will perform the same tasks, but now including:
- Lesions at different levels of the model:
  - Phonological state buffer -> Broca / Wernicke aphasia
  - Associative memories between phono-lemma buffer -> transcortical motor / sensory aphasia
  - Associative memories between lemma-concept buffers -> mixed aphasia
  - Neural association between phono-phono buffers -> conduction aphasia
- Lesions -> decrease in task performance



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
Broca
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
Wernicke
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron … muscle system …
… lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

Integration of neural dysfunctions

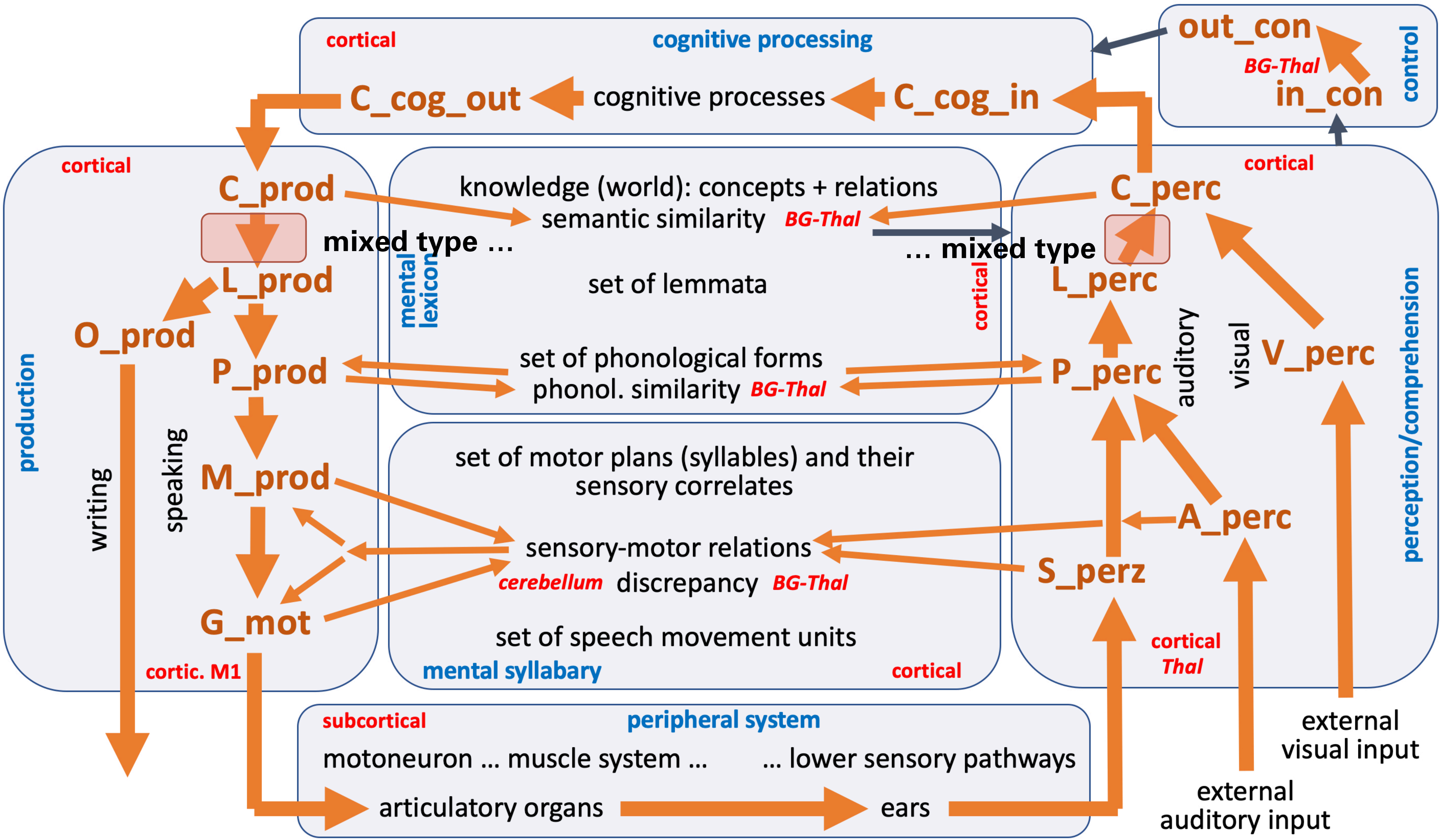
- Lesions at different levels of the model:
  - Phonological state buffer -> Broca / Wernicke aphasia
  - Associative memories between phono-lemma buffer -> transcortical motor / sensory aphasia
  - Associative memories between lemma-concept buffers -> mixed aphasia
  - Neural association between phono-phono buffers -> conduction aphasia
- Lesions -> decrease in task performance



out_con
BG-Thal
in_con
control
cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
transcortical motor
transcortical sensory
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their
sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
perception/comprehension
A_perc
S_perz
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external
visual input
external
auditory input

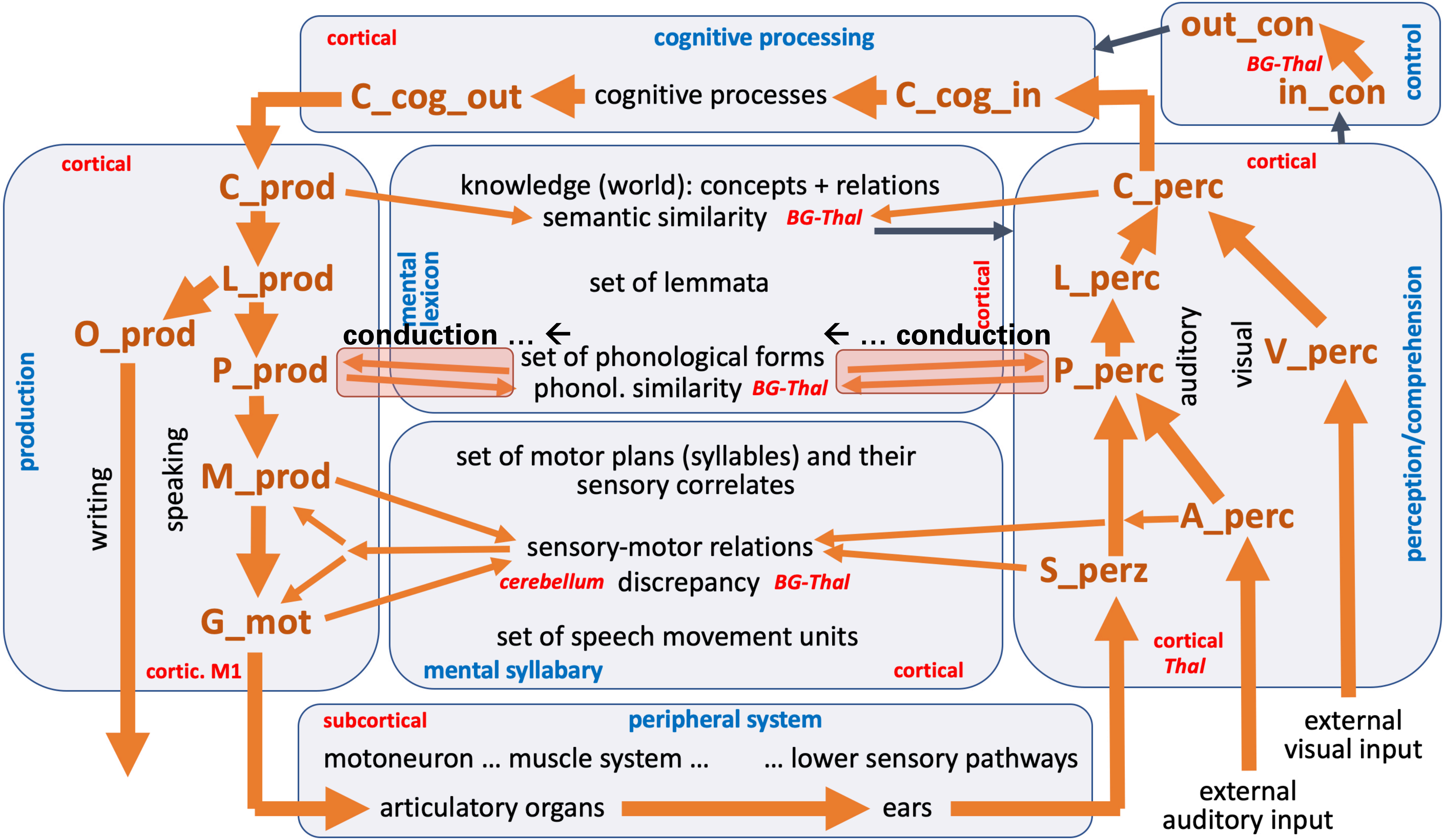
Integration of neural dysfunctions

- Lesions at different levels of the model:
  - Phonological state buffer -> Broca / Wernicke aphasia
  - Associative memories between phono-lemma buffer -> transcortical motor / sensory aphasia
  - Associative memories between lemma-concept buffers -> mixed aphasia
  - Neural association between phono-phono buffers -> conduction aphasia
- Lesions -> decrease in task performance


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
mixed type …
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
… mixed type
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron … muscle system …
… lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

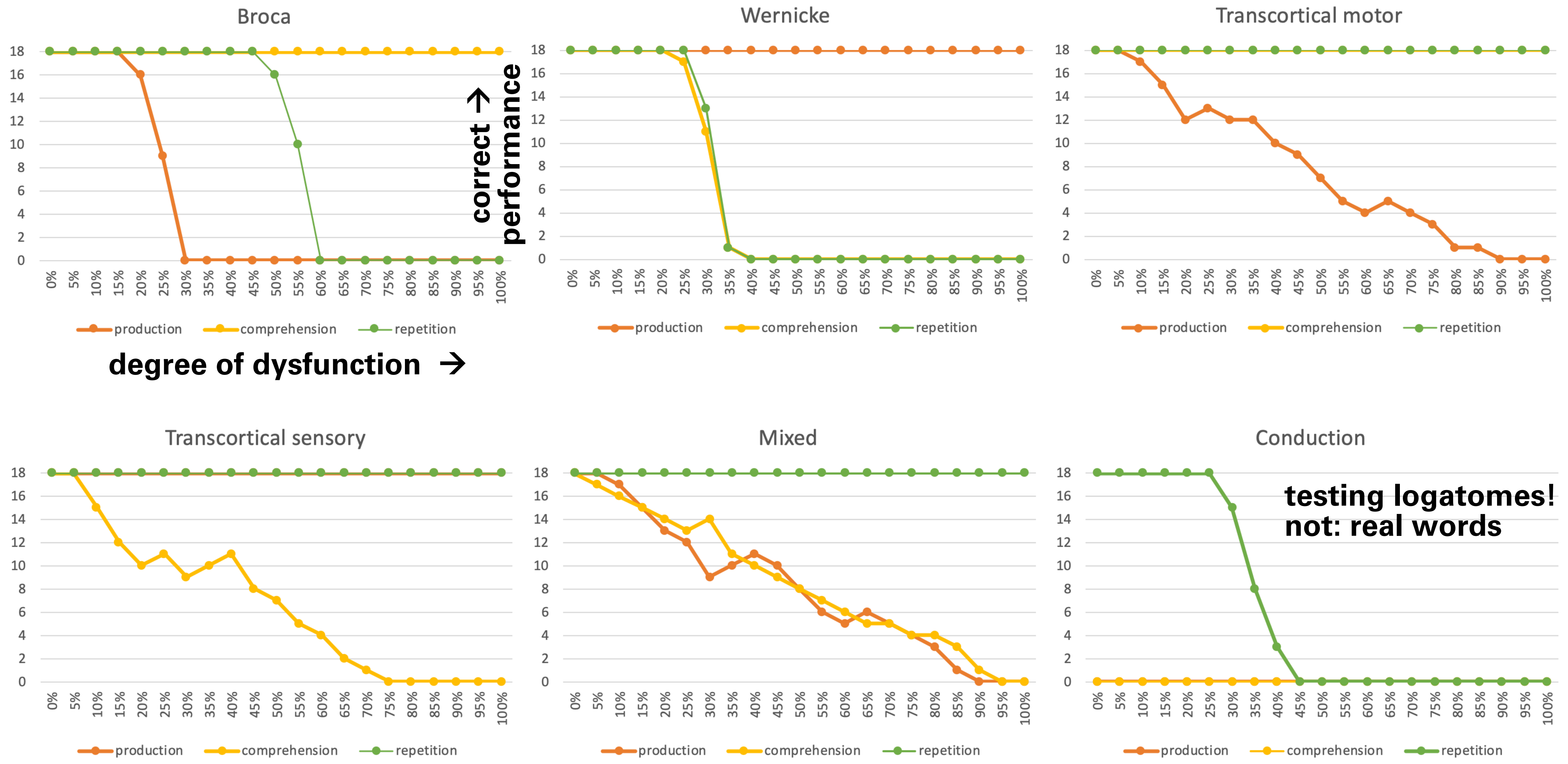
Integration of neural dysfunctions

- Lesions at different levels of the model:
  - Phonological state buffer -> Broca / Wernicke aphasia
  - Associative memories between phono-lemma buffer -> transcortical motor / sensory aphasia
  - Associative memories between lemma-concept buffers -> mixed aphasia
  - Neural association between phono-phono buffers -> conduction aphasia
- Lesions -> decrease in task performance

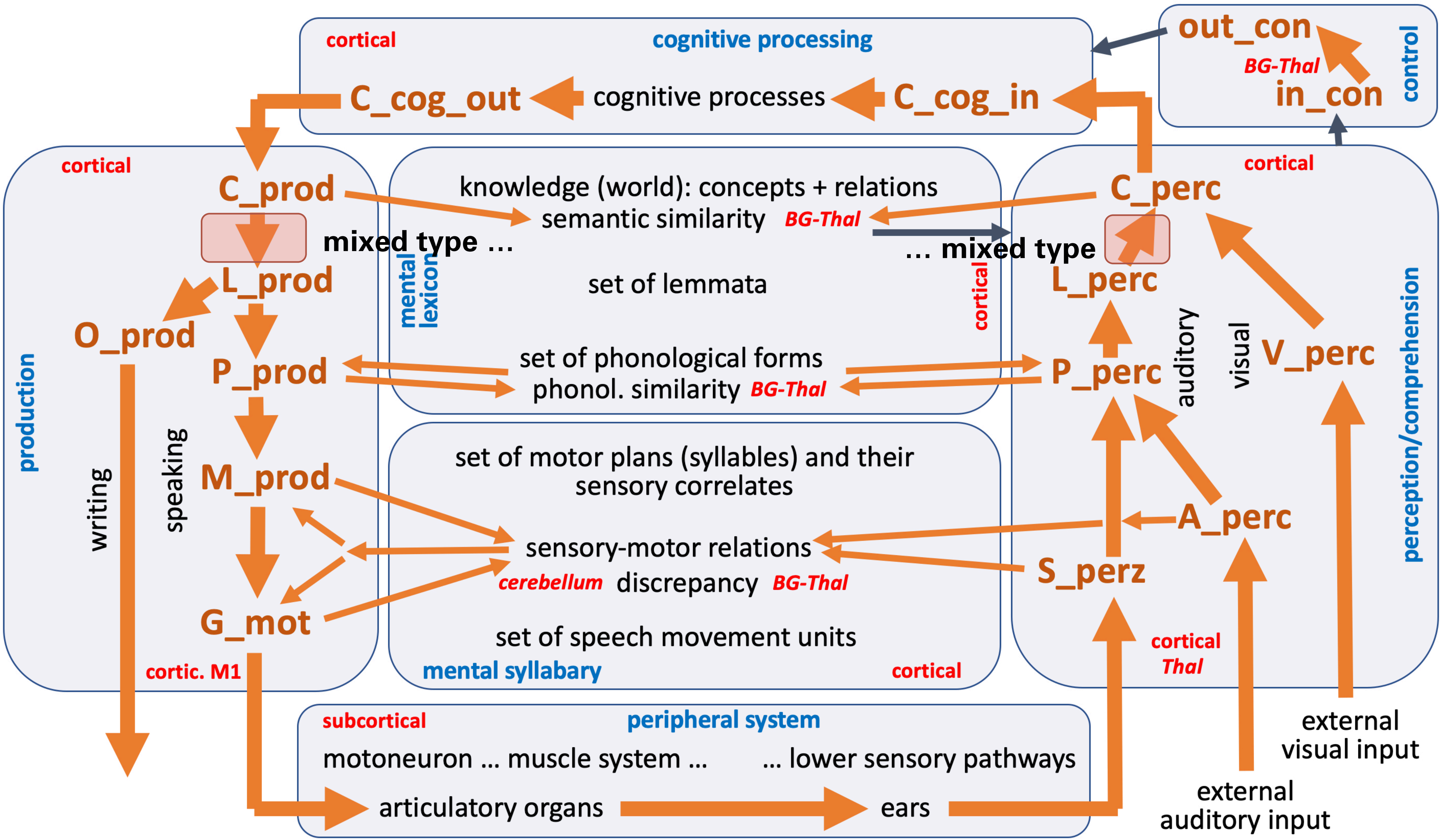

cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
conduction ... ←
← ... conduction
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their
sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external
visual input
external
auditory input

Integration of neural dysfunctions

- Lesions at different levels of the model:
  - Phonological state buffer -> Broca / Wernicke aphasia
  - Associative memories between phono-lemma buffer -> transcortical motor / sensory aphasia
  - Associative memories between lemma-concept buffers -> mixed aphasia
  - Neural association between phono-phono buffers -> conduction aphasia
- Lesions -> decrease in task performance (symptoms!)



Kröger et al. (2020)



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
mixed type …
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
set of lemmata
mental lexicon
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
… mixed type
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron … muscle system …
… lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

Download Video4:
https://www.youtube.com/watch?v=xzAgOnna9Mw

The definition of scenarios

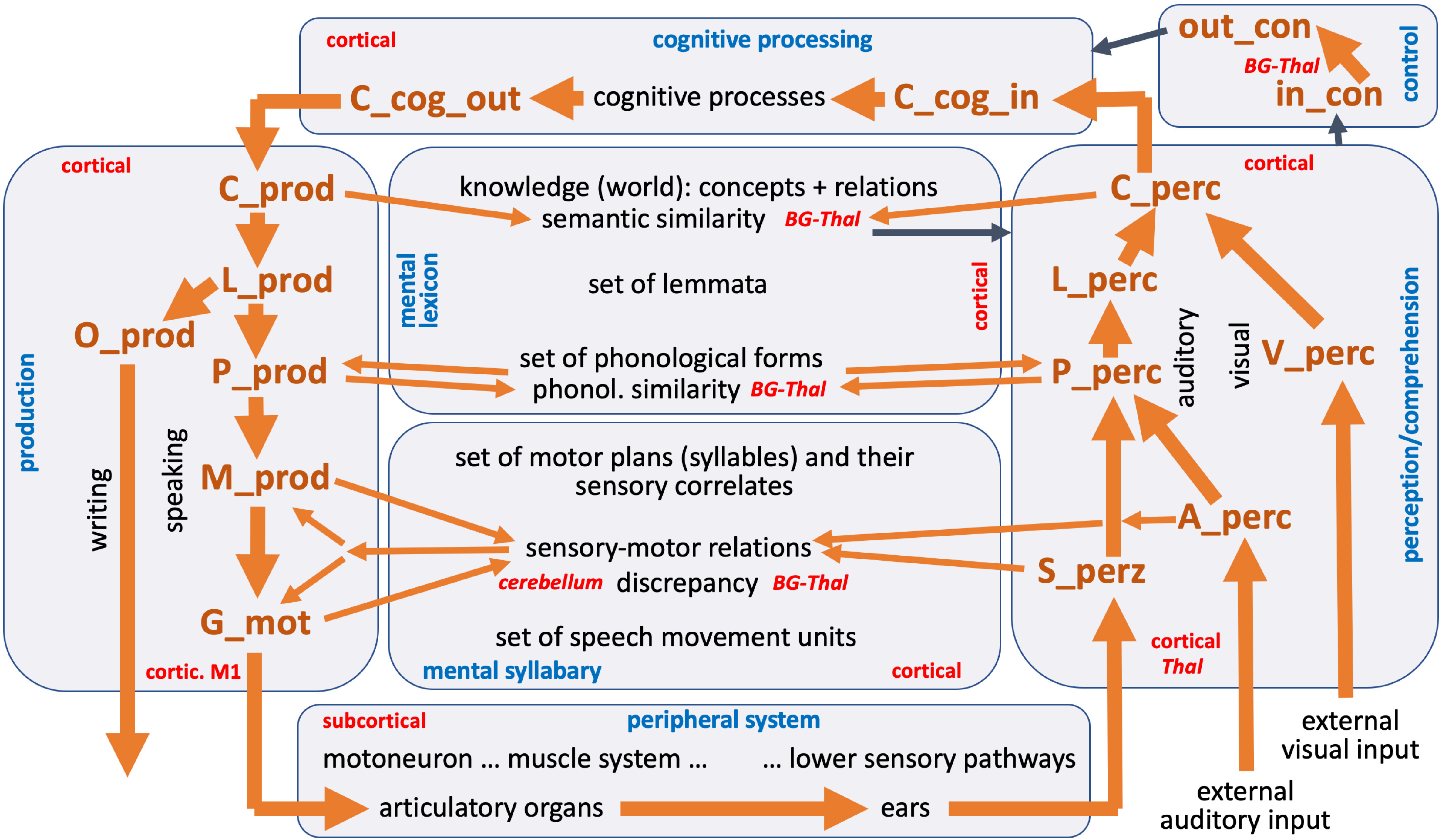
- So far:
  - simple production task (picture naming)
  - comprehension tasks
  - simple repetition task
- more complex tasks:
- Further speech screenings: (-> demonstrating realism of model)
  - Serial recall (10/15 words; Choo 2010, master thesis, Waterloo Univ.)
  - Picture naming with phonological or semantic distractors words
  - Picture naming in case of lexical access/retrieval problems: “tip of the tongue” case: -> help by introducing phonological and semantic cues

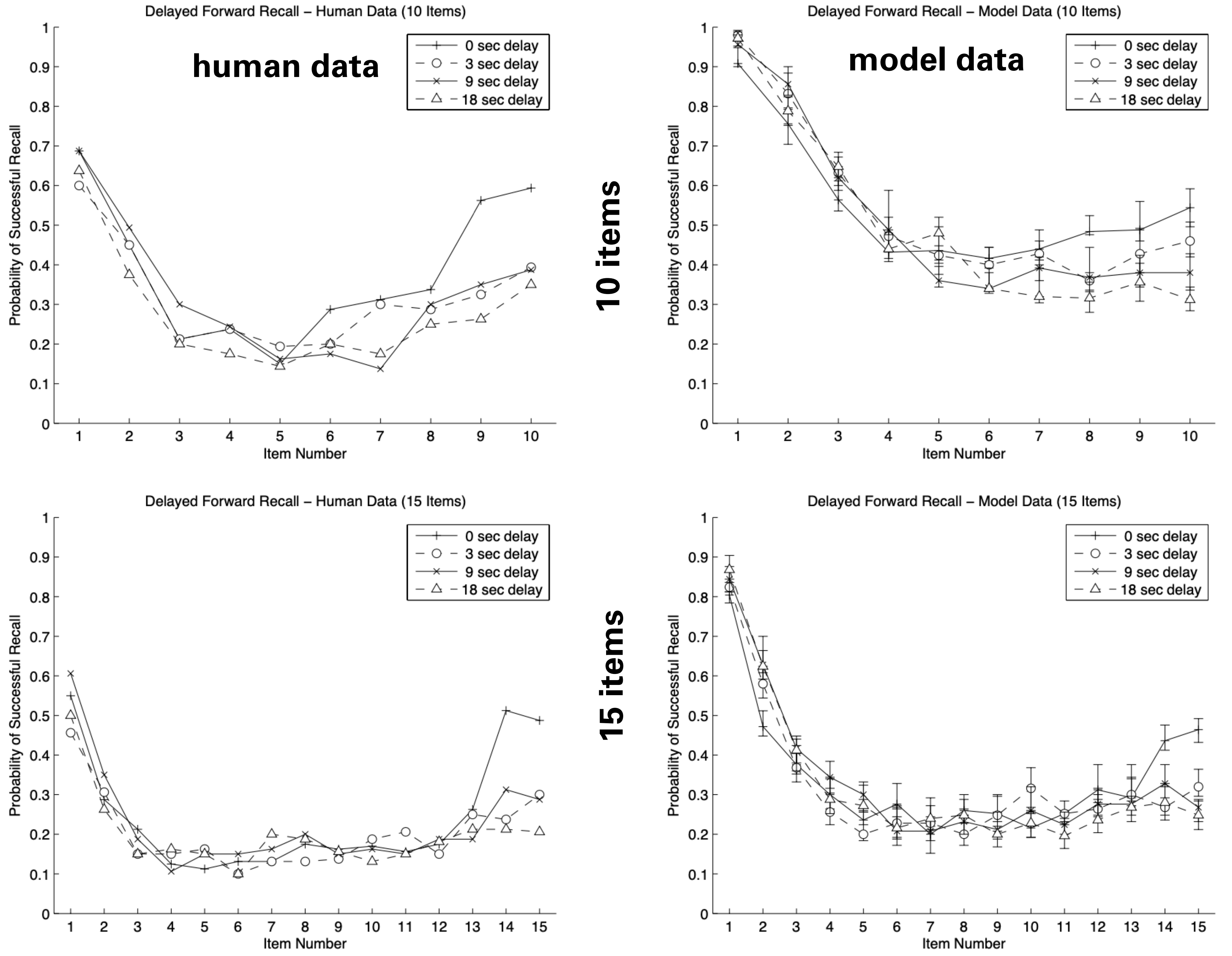
Why more simulations?

- Demonstrating the realism of neural models
- -> we already demonstrated:
  - normal behavior (naming, comprehension, repetition tasks)
  - rare cases (fail in naming in < 1% of all simulations)
  - introduction of neural dysfunctions -> simulation of symptoms in case of different speech disorders (diff. types of aphasias)
- Further simulations:
  - simulating “overburdening” (-> serial recall of 15 words)
  - simulating ”difficult task”: -> get more faults / mistakes
    - e.g.,: simulation of “speech errors” by using distractors;
    - e.g.,: simulating “tip of the tongue” cases;

The definition of scenarios

- more complex tasks:
  - Serial recall (10/15 words; Choo 2010, master thesis, Waterloo Univ.)
    - -> lowest repetition rate for words in the middle (concave result graph)
    - Short term memory and binding of positions and concepts
  - Picture naming with phonological or semantic distractors (acoustically introduced) -> decrease in performance
  - Picture naming in case of lexical access problems ("tip of the tongue" case): help by introducing phonological and semantic cues -> increase in performance


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their
sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron … muscle system …
… lower sensory pathways
articulatory organs
ears
external
visual input
external
auditory input



Choo FX (2010), master thesis, Univ. Waterloo

The definition of scenarios

- Further speech screenings: (-> demonstrating realism of model)
  - Picture naming with phonological or semantic distractors (acoustic)
    - specific word list: including phonological and semantic dis-/similar words
    - -> generation of direct HALTS / no or late HALTS in different cases by including a evaluation of similarity of items at phonological and at semantic level as part of the internal feedback loops

Kroeger BJ, Stille C, Blouw P, Bekolay T, Stewart TC (2020) Hierarchical sequencing and feedforward and feedback control mechanisms in speech production: A preliminary approach for modeling normal and disordered speech. Frontiers in Computational Neuroscience, 14:99. **www.speechtrainer.eu -> publications**

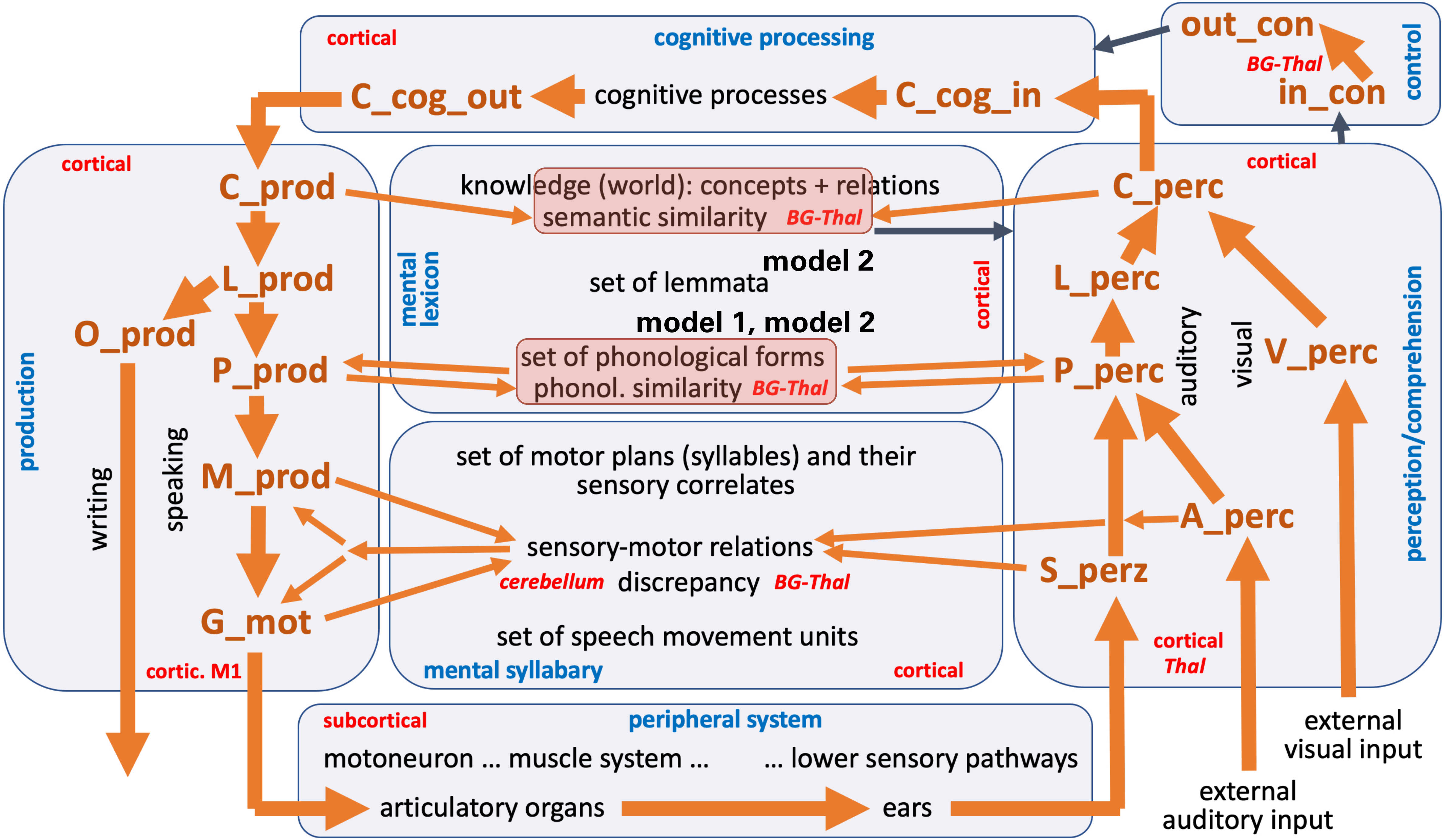
Word list -> mental lexicon: 18x6 =104 words

Target word	Semantically similar word	Phonologically similar word	Phonologically and semantically similar word	Dissimilar word	Semantically superordinate item (semantic cues)	Similar sound segments (phonological cues)
Apple	Peach	Apathy	Apricot	Couch	Fruits	/Ep/
Basket	Crib	Ban	Bag	Thirst	Bin	/bE/
Bee	Spider	Beacon	Beetle	Flag	Crawler	/bi/
Bread	Donut	Brick	Bran	Nail	Cereal	/br/
Camel	Pig	Cash	Calf	Bucket	clovenHooved	/kE/
Carrot	Spinach	Cast	Cabbage	Evening	Veg	/kE/
Duck	Raven	Sub	Dove	Brass	Bird	/da/
Elephant	Moose	Elm	Elk	Stripe	HornAnimal	/El/
Fly	Moth	Flu	Flea	Rake	Bluebottle	/fl/
Lamp	Candle	Landing	Lantern	Package	LightSource	/lE/
Peanut	Almond	Piano	Pecan	Dress	Nut	/pi/
Rabbit	Beaver	Raft	Rat	Coffee	Rodent	/rE/
Snake	Eel	Snack	Snail	Fire	Invertebrate	/snE/
Spoon	Ladle	Sparkle	Spatula	Cable	Lifter	/sp/
Squirrel	Mole	Skate	Skunk	Chain	HairySkin	/sk/
Train	Bus	Trophy	Trolley	Fox	PublicTrans	/tr/
Truck	Jeep	Trap	Tractor	Celery	UtilityVehicle	/tr/
Trumpet	Horn	Traffic	Trombone	Corner	BrassWind	/tr/

The definition of scenarios

- Further speech screenings: (-> demonstrating realism of model)
  - Picture naming with phonological or semantic distractors (acoustic)
    - specific word list: including phonological and semantic dis-/similar words
    - -> generation of direct HALTS / no or late HALTS in different cases by including a evaluation of similarity of items at phonological and at semantic level as part of the internal feedback loops

Kroeger BJ, Stille C, Blouw P, Bekolay T, Stewart TC (2020) Hierarchical sequencing and feedforward and feedback control mechanisms in speech production: A preliminary approach for modeling normal and disordered speech. Frontiers in Computational Neuroscience, 14:99. **www.speechtrainer.eu -> publications**


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity BG-Thal
model 2
set of lemmata
model 1, model 2
set of phonological forms
phonol. similarity BG-Thal
mental lexicon
cortical
set of motor plans (syllables) and their
sensory correlates
sensory-motor relations
cerebellum discrepancy BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external
visual input
external
auditory input

Results: picture naming with distractors

- 18 items, 3 runs -> 54 simulations per sub-experiment
  - 4 different distractor words, 2 different models -> 432 simulations
- -> HALT events in case of dissimilar distractor words
  - No HALT in case of similar distractor words (suppress early HALT)

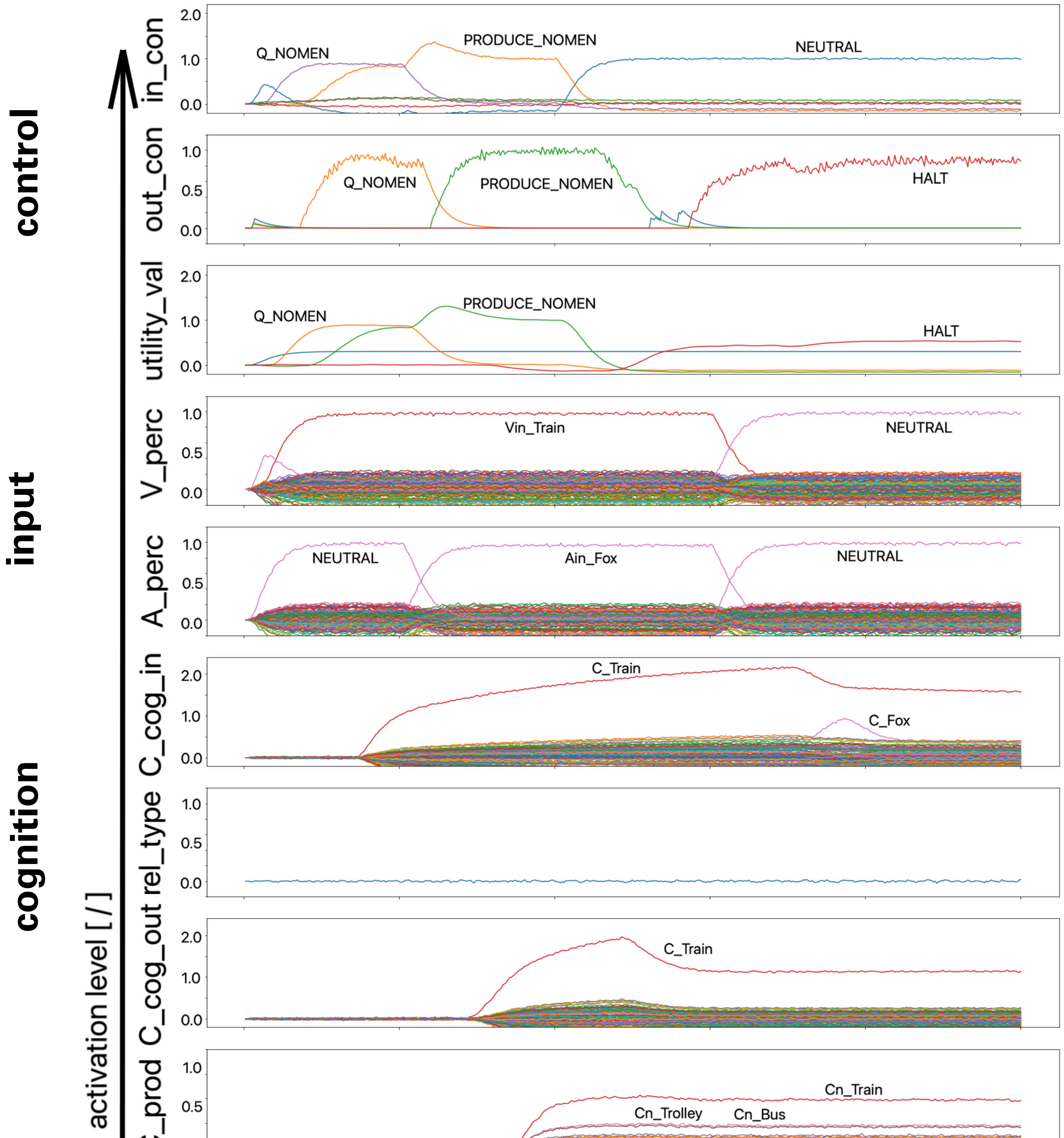
Type of model	Type of distractor word	Number of stops (test series 1)	Number of stops (test series 2)	Number of stops (test series 3)	Number of stops (sum)	Number of simul. Without stop (sum)
1	Semantic similar	1	2	0	3	51
1	Phonological similar	10	10	9	no strong suppression effect here --> 29	25
1	Sem + phono similar	1	1	0	--> 2	52
1	Dissimilar	17	18	15	--> 50	4
2	Semantic similar	15	15	16	46	8
2	Phonolgical similar	2	2	2	6	48
2	Sem + phono similar	0	1	1	2	52
2	Dissimilar	17	14	13	44	10

diff-eval @ semantic + phono level (model 1)

diff-eval @ phono level only (model 2)

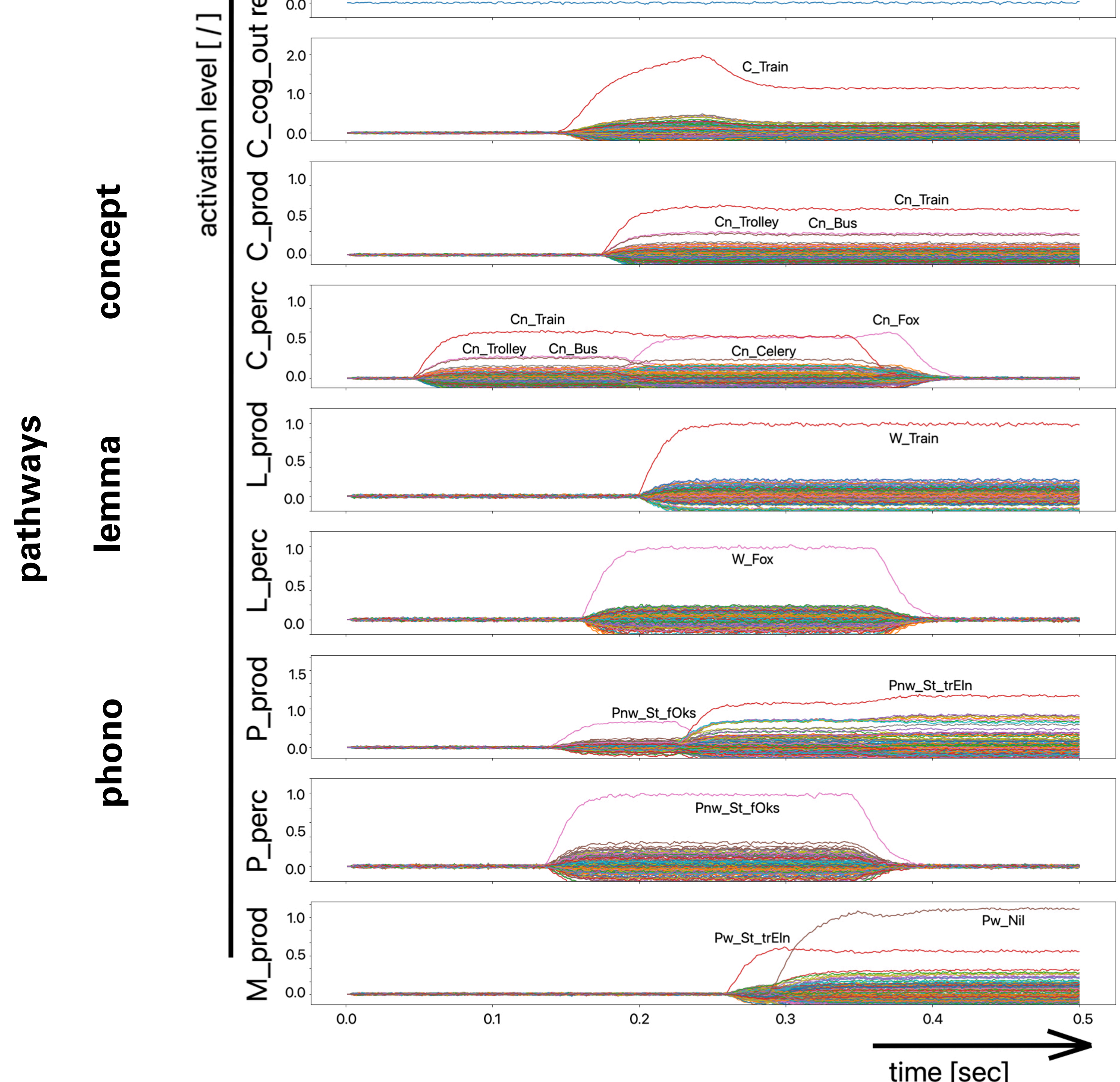
The definition of scenarios

- Speech screenings: (-> demonstrating realism of model)
  - Picture naming with phonological or semantic distractors (acoustic)
    - Case: phonological and semantic dissimilar word
    - -> generation of HALT signal (-> word is NOT produced)



← priming: 100 msec:
"look and think" +150 msec :
"utter noun" + later:
… HALT may be generated!

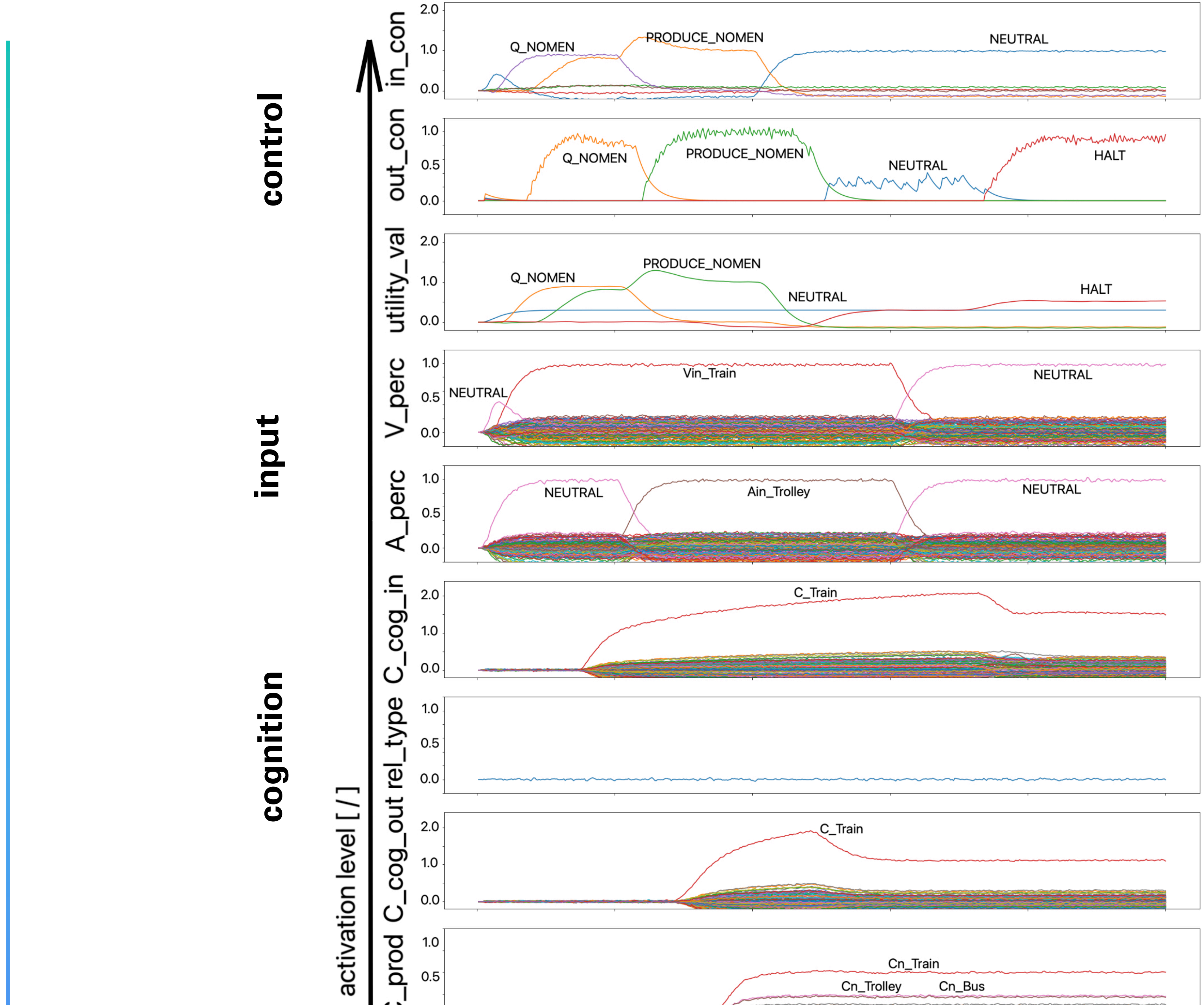
Kröger et al. (2020)



Kröger et al. (2020)

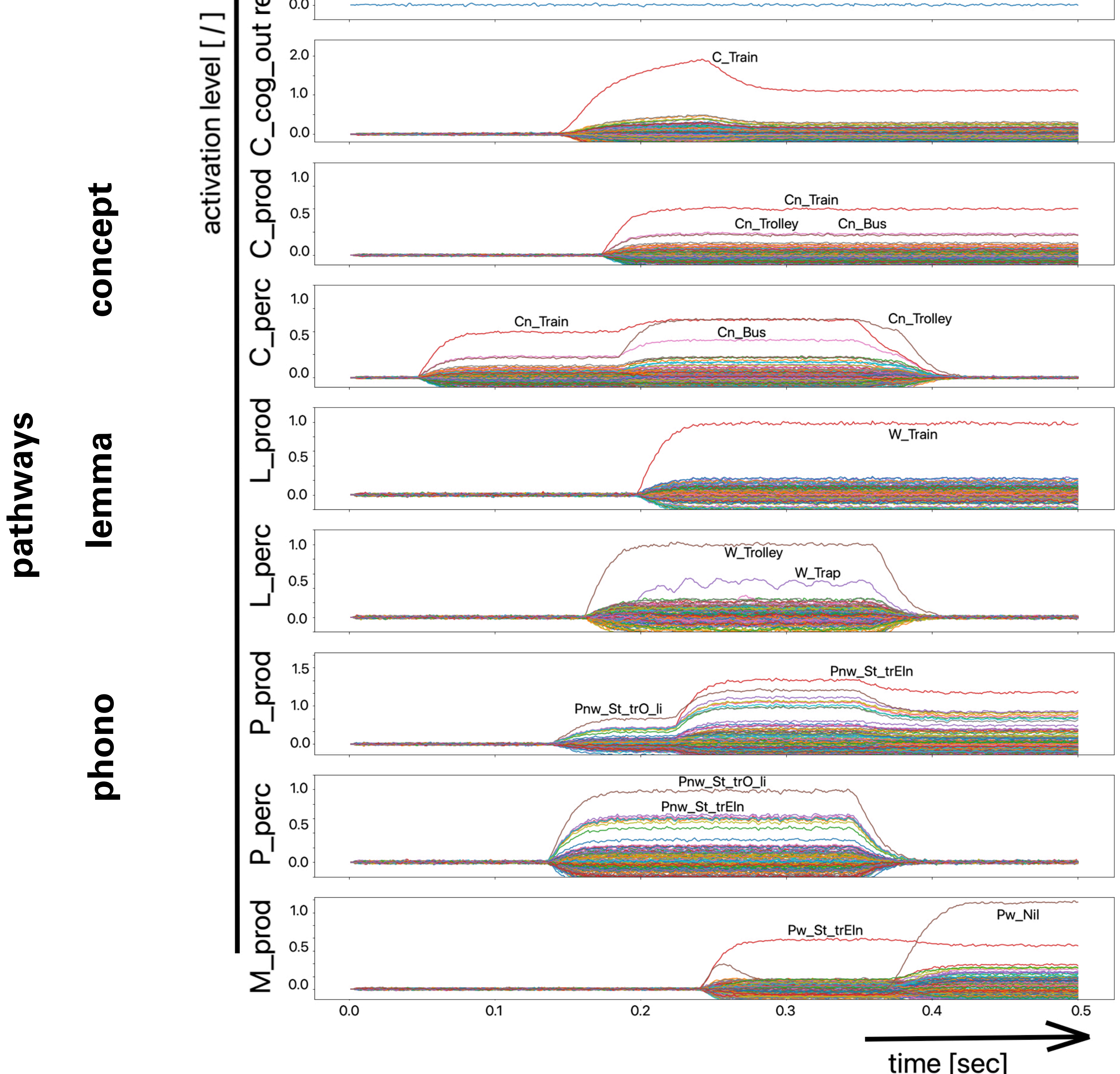
The definition of scenarios

- Speech screenings: (-> demonstrating realism of model)
  - Picture naming with phonological or semantic distractors (acoustic)
    - Case: phonological and semantic similar word
    - -> suppress early HALT signal (-> word is produced)



← priming: 100 msec:
"look and think" +150 msec :
"utter noun" + later:
… HALT may be generated!

Kröger et al. (2020)



Kröger et al. (2020)

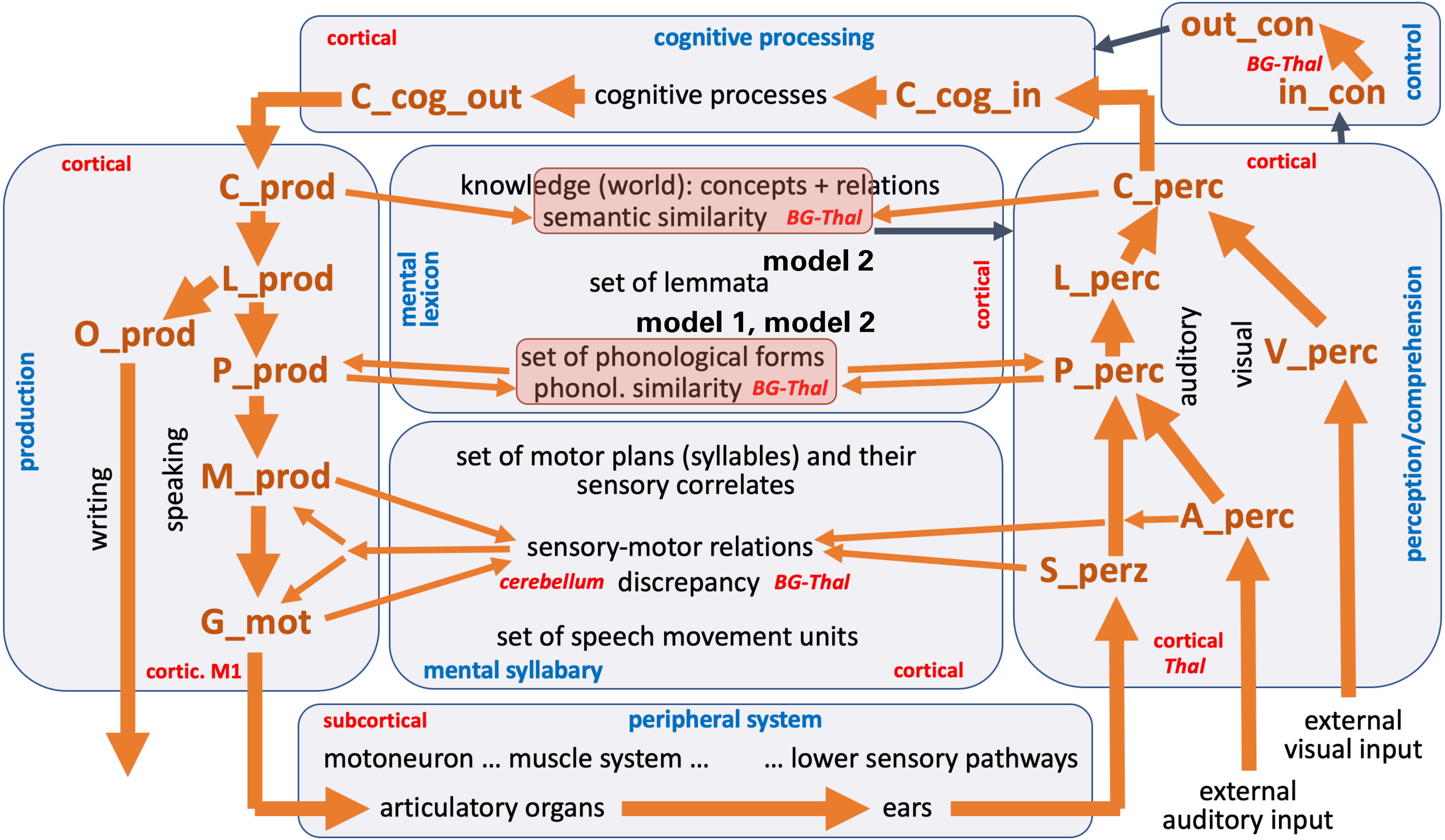
Results: picture naming with distractors

- -> HALT events in case of dissimilar distractor words
  - No HALT in case of similar distractor words (suppress early HALT)
- Which model is realistic? Neither!
  - Need: diff-eval of 0.5 * semantic + phono level (??)
  - But: Basic law in modeling: do not adapt the model parameters (← exp. results)

Type of model	Type of distractor word	Number of stops (test series 1)	Number of stops (test series 2)	Number of stops (test series 3)	Number of stops (sum)	Number of simul. Without stop (sum)
1	Semantic similar	1	2	0	too strong sup -> 3	51
1 diff-eval @ semantic + phono level	Phonological similar	10	10	9	too weak sup -> 29	25
1	Sem + phono similar	1	1	0	too strong sup -> 2	52
1	Dissimilar	17	18	15	50	4
2	Semantic similar	15	15	16	46	8
2 diff-eval @ phono level only	Phonolgical similar	2	2	2	too strong sup -> 6	48
2	Sem + phono similar	0	1	1	too strong sup -> 2	52
2	Dissimilar	17	14	13	44	10

Models need to be straight forward

- define the model qualitatively AND quantitatively on the basis of rules for neural functioning
- Generate (qualitative and quantitative) simulation resultson this basis! without changing / adjusting the model
- See the model:
    - in all cases (neurons, neural connections) we use the default parametrization given by NENGO
    - we suggest a post-adaptation of the model only in case of one black arrow:


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity BG-Thal
model 2
set of lemmata
model 1, model 2
set of phonological forms
phonol. similarity BG-Thal
mental lexicon
cortical
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum discrepancy BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

Download Video5:
https://www.youtube.com/watch?v=KdJ6KV0sA-8

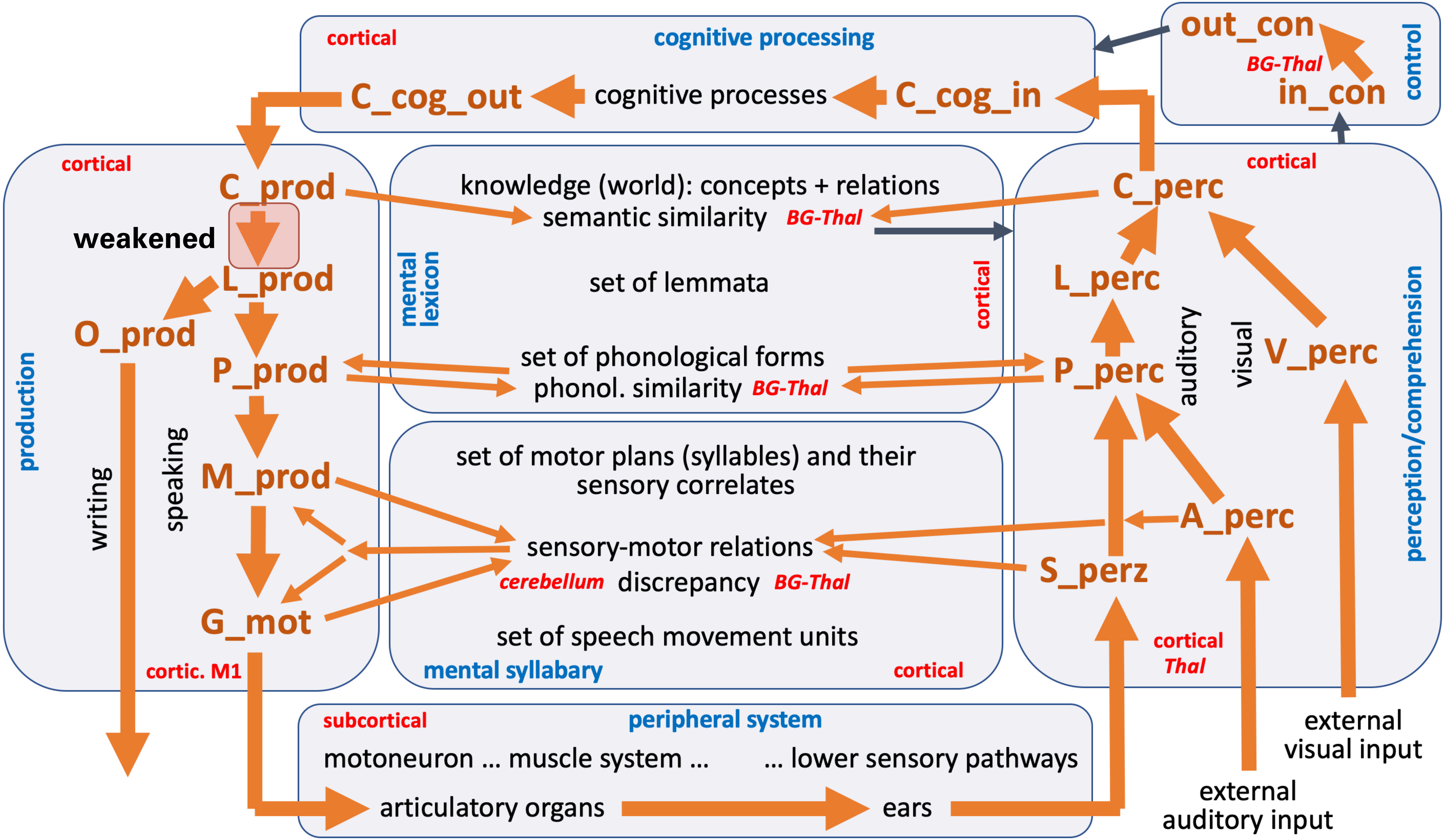
The definition of scenarios

- more complex tasks:
- Further speech screenings: (-> demonstrating realism of model)
  - Serial recall (10/15 words; Choo 2010, master thesis, Waterloo Univ.)
    - -> lowest repetition rate for middle words (concave result graph)
  - Picture naming with phonological or semantic distractors (acoustically introduced) -> decrease in performance; induces HALTs
  - Picture naming in case of lexical access problems ("tip of the tongue" case): help by introducing phonological and semantic cues -> increase in performance

The definition of scenarios

- Further speech screenings: (-> logopedic research)
  - Picture naming with phonological or semantic cues (acoustic)
    - Naming in case of “word is on tip of my tongue” (patient with problems in lexical access / lexical retrieval); ->
    - only 20-45% of correct productions occur without any cue
  - What if we use cues, if a word is not produced or not produced correctly?
    - Semantic cue: target word “car”; cue: “you can drive it, has four wheels ”
    - Phonological cue: target word “car”: cue: “it starts with /k/”

Kroeger BJ, Stille C, Blouw P, Bekolay T, Stewart TC (2020) Hierarchical sequencing and feedforward and feedback control mechanisms in speech production: A preliminary approach for modeling normal and disordered speech. Frontiers in Computational Neuroscience, 14:99. **www.speechtrainer.eu -> publications**

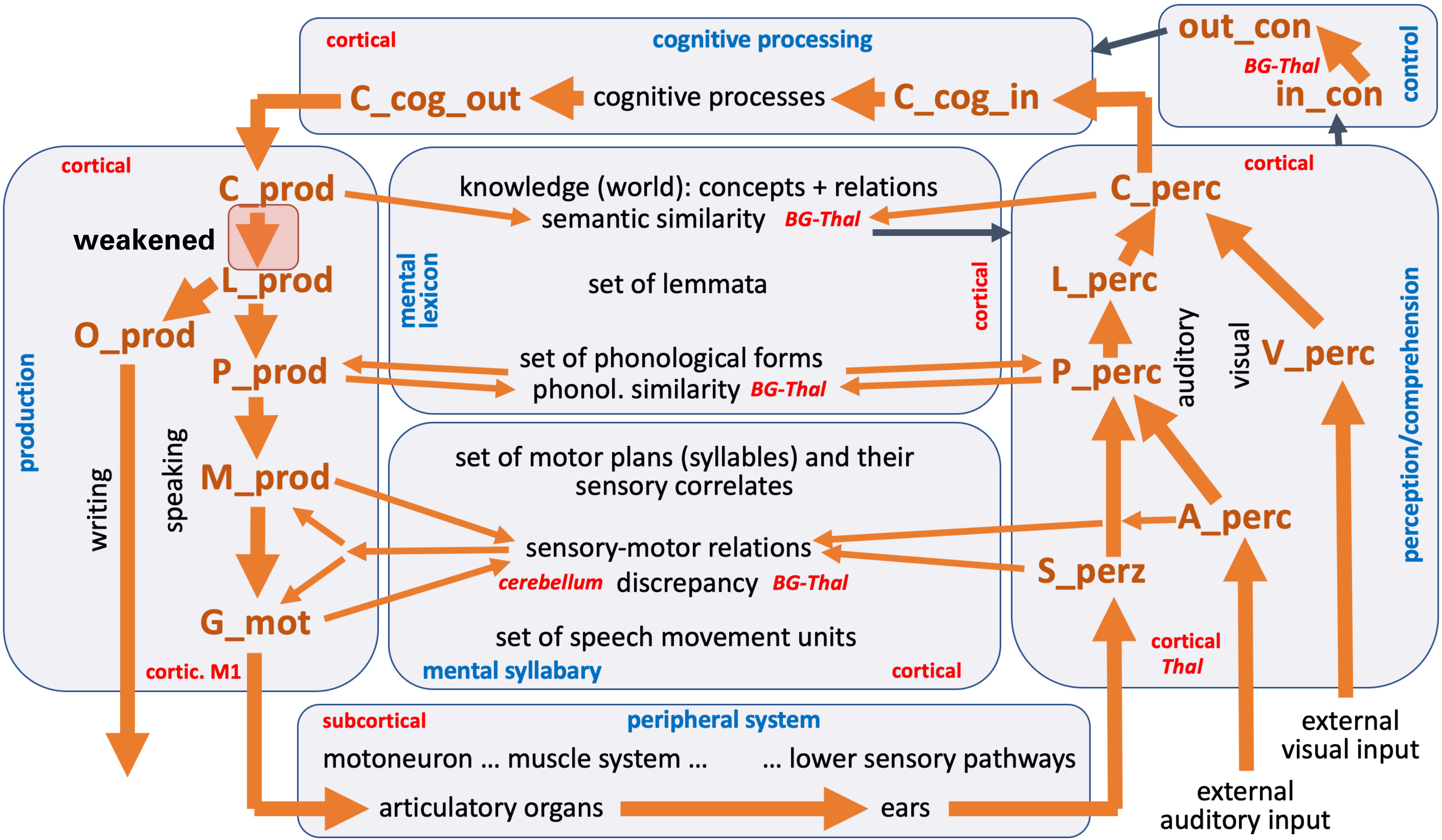


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
weakened
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
perception/comprehension
A_perc
S_perz
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

The definition of scenarios

- Further speech screenings: (-> logopedic research)
  - Picture naming with phonological or semantic cues (acoustic)
    - Naming in case of “word is on tip of my tongue” (patient with problems in lexical access / lexical retrieval); ->
    - only 20% of correct productions occur without any cue
  - How do cues work in the model, if a word is not produced or not produced correctly?
    - Semantic cue: increases neural activity C_perz -> lexical loop
    - Phonological cue: increases neural activity at P_perc -> lexical loop or:
    - phonological shortcut may be more effective!

Kroeger BJ, Stille C, Blouw P, Bekolay T, Stewart TC (2020) Hierarchical sequencing and feedforward and feedback control mechanisms in speech production: A preliminary approach for modeling normal and disordered speech. Frontiers in Computational Neuroscience, 14:99. **www.speechtrainer.eu** **-> publications**



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
weakened
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
set of lemmata
mental lexicon
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

Cues are helpful:

- 2 different cues; 2 different models: with/without shortcut perc->prod at phonological level
- Starting with 20-45% of correct word naming; increase to about 55-70%

Number of test series: type of model and type of cue	Number of target word activation before/after cue (test-series 1)	Number of target word activation before/after cue (test-series 2)	Number of target word activation before/after cue (test-series 3)	Number of correct target word activation before cue (sum)	further Number of correct target word activation after cue (sum)	Number of activation of incorrect target words (sum)
1: L-route, phono	8/5	6/6	10/3	24	14 -> 38	16
2: shortcut, phono	4/6	2/10	8/4	14	20 -> 34	20
3: L-route, semantic	6/7	4/5	4/4	14	16 -> 30	24

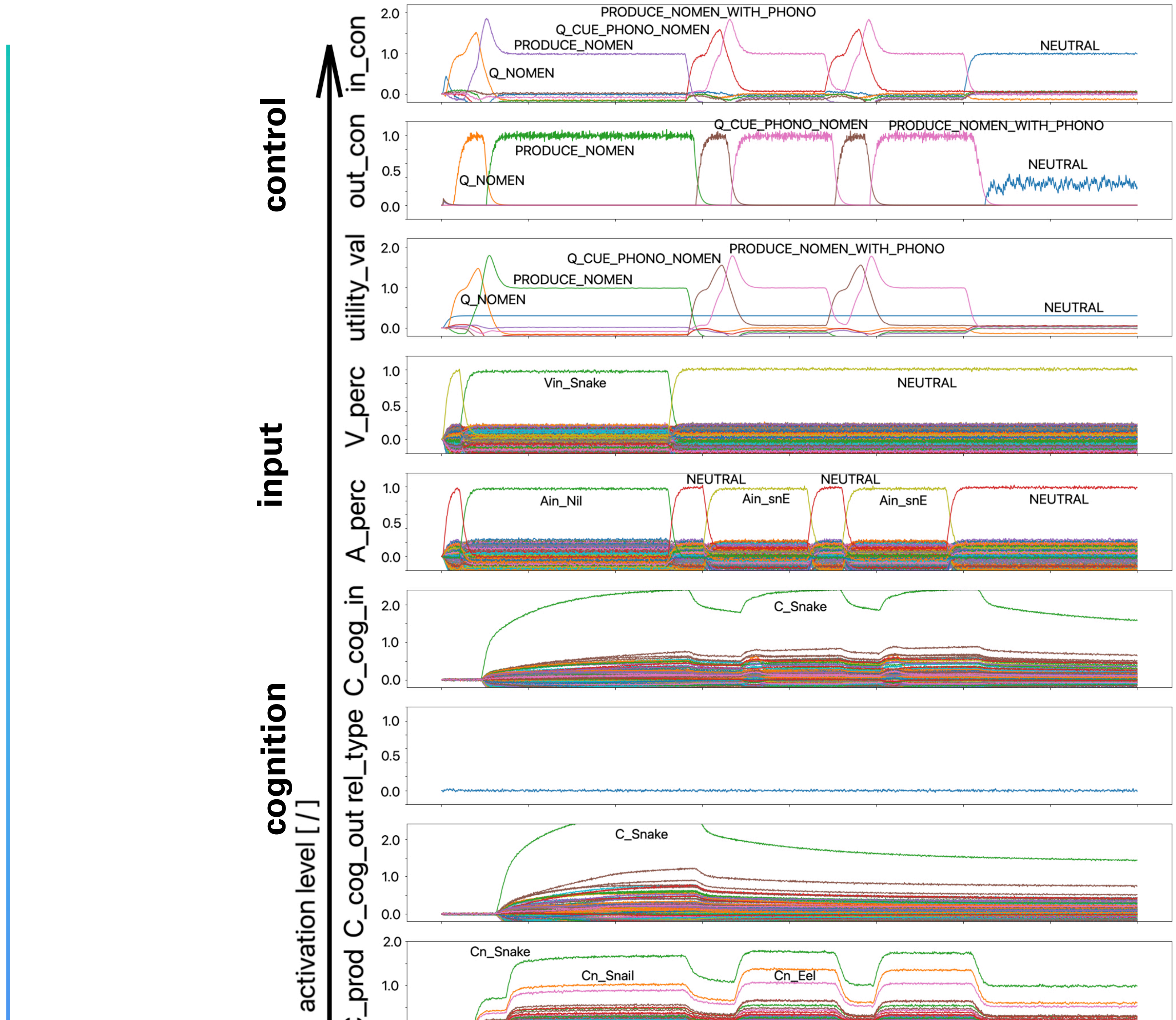
increase in correct productions

The maximum number of simulations per test series is 18. With a number of correct target word activation of 6, correct target word activation was not achieved in 12 out of 18 simulation cases. The total number of simulations per table row is 54. A distinction is made in this table regarding whether a target word is already correctly and completely activated in the motor plan state buffer before cues are given, i.e., within the first 750 ms of the simulation time, or whether the target word was only correctly activated after occurrence of cues. "L route" is model type "lexical route" without short-circuit at the phonological level; "Shortcut": model type with decoupled mental lexicon, but with short-circuit of the state buffers at the phonological level P_perc -> P_prod.

Kröger et al. (2020)

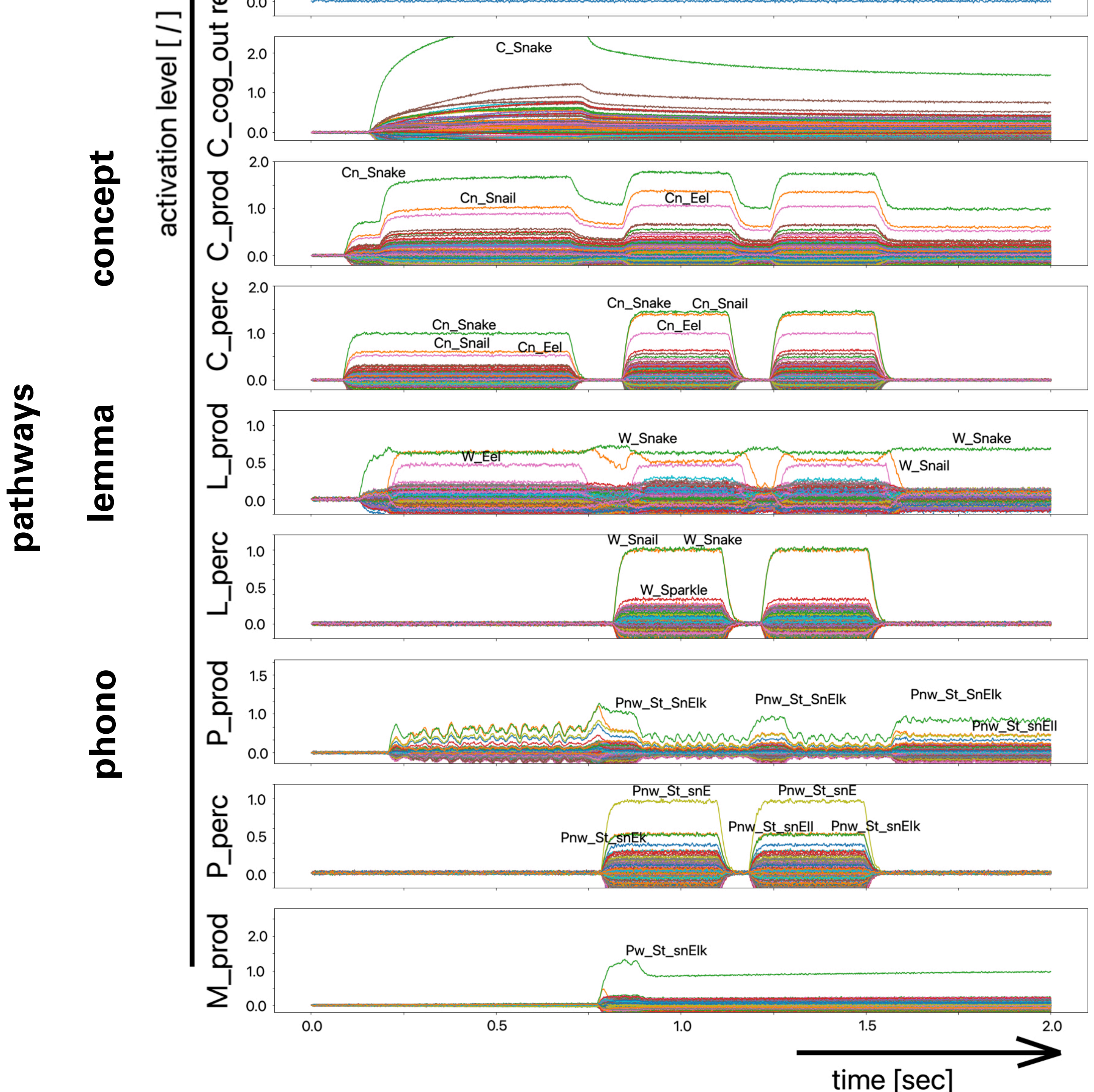
The definition of scenarios

- Further speech screenings: (-> logopedic research)
  - Picture naming with phonological or semantic cues (acoustic)
    - Naming without cues in case of “word is on tip of my tongue” -> no or incorrect word activation at P_prod after picture presentation
    - 4 cases:
      - Word activation starts with phonological cue (not earlier); 2 cases: “snake” and “apple”
      - … starts with semantic cue (not earlier) 2 cases: “duck” and “fly”
      - fourth case: -> rare case …



← priming: 100 msec:
"look and think" + 600 msec :
"utter noun"
+ later: "listen to cues" +
"utter noun"

Kröger et al. (2020)



C_perc -> C_prod shortcut!

No P_perc -> P_perc shortcut!

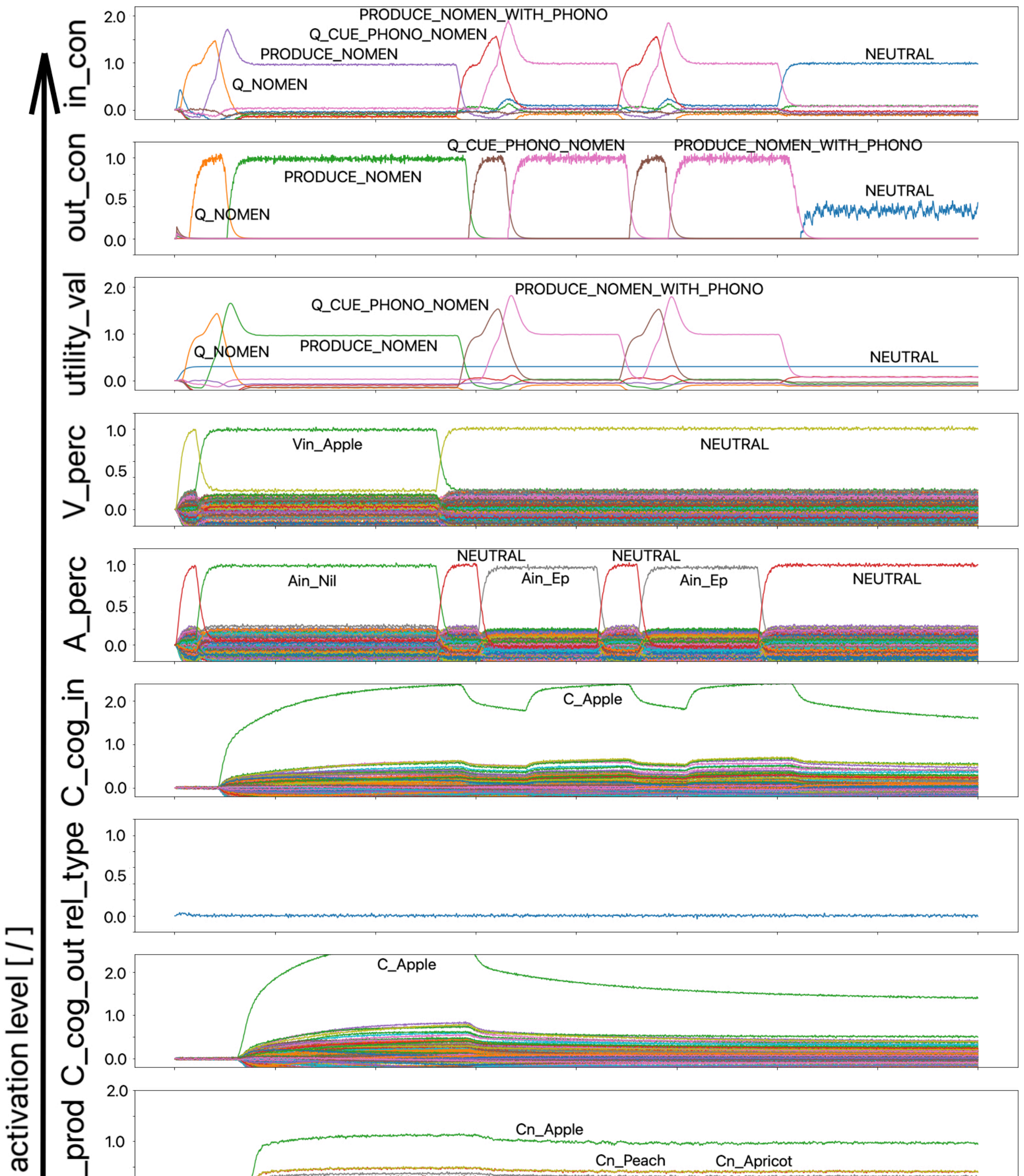
Increase in activation of all similar syllables because of S-pointer-networks!

Kröger et al. (2020)

control

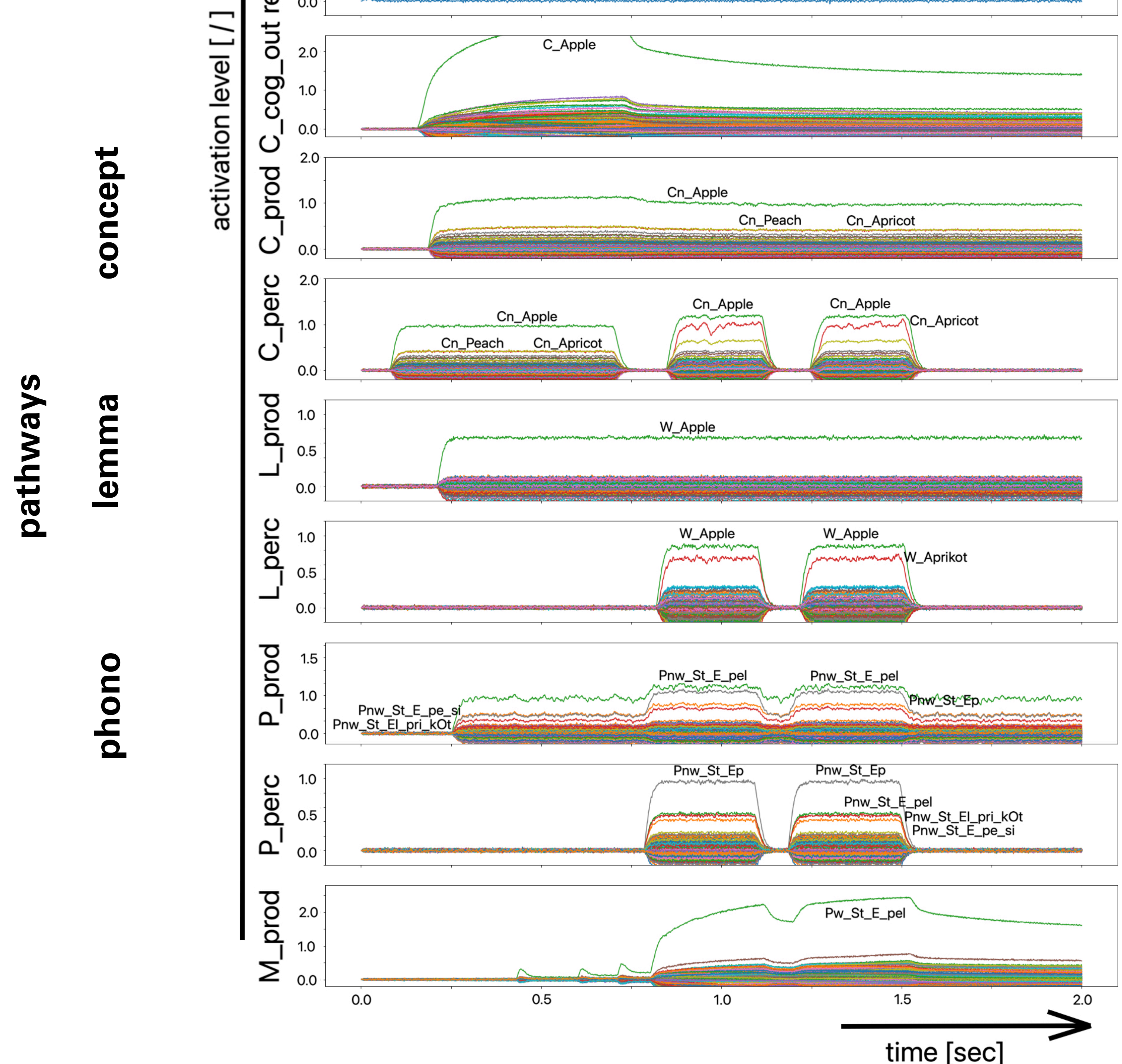
input

cognition



← priming: 100 msec:
"look and think" + 600 msec :
"utter noun"
+ later: "listen to cues" +
"utter noun"

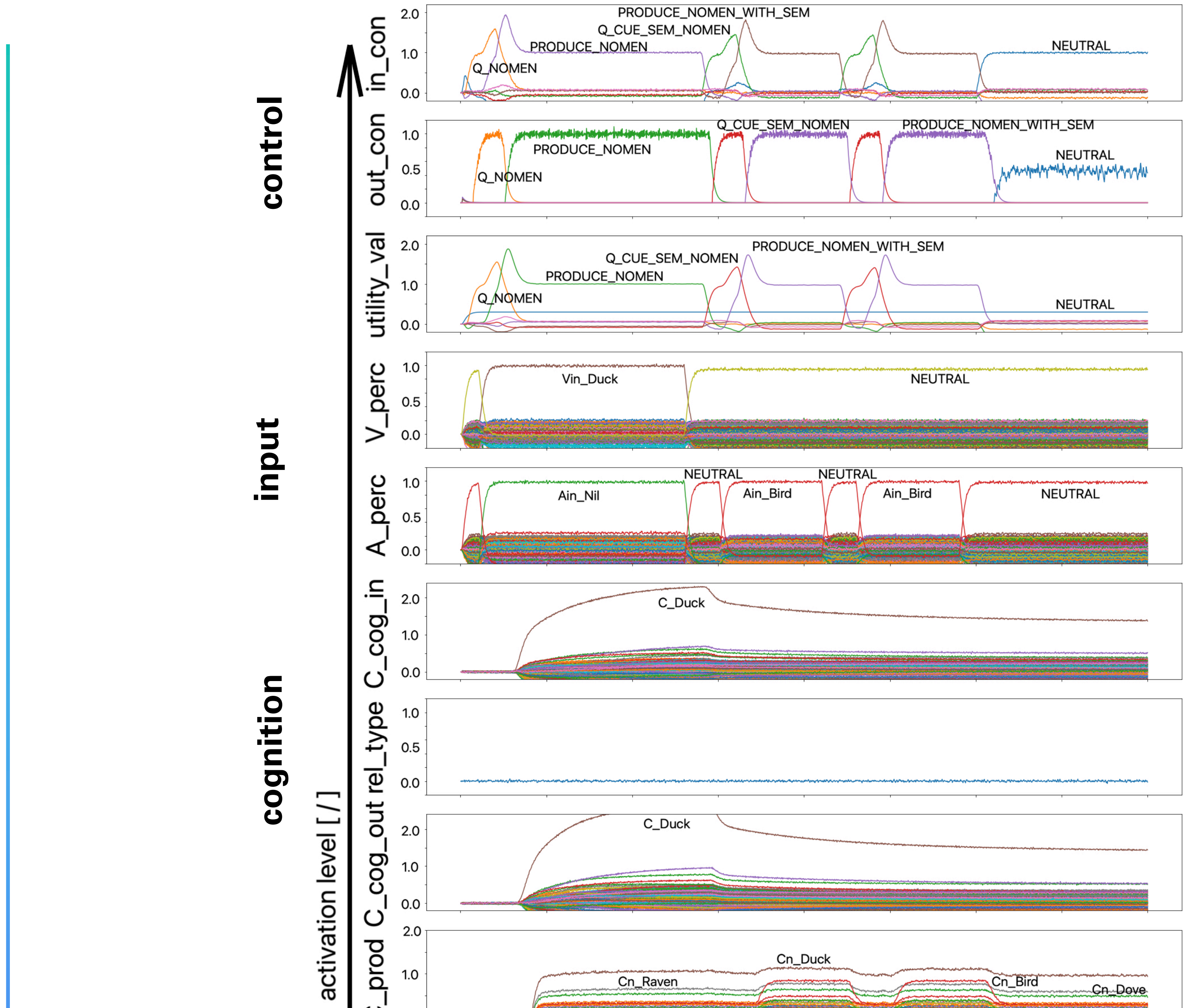
Kröger et al. (2020)



No C_perc -> C_prod shortcut!

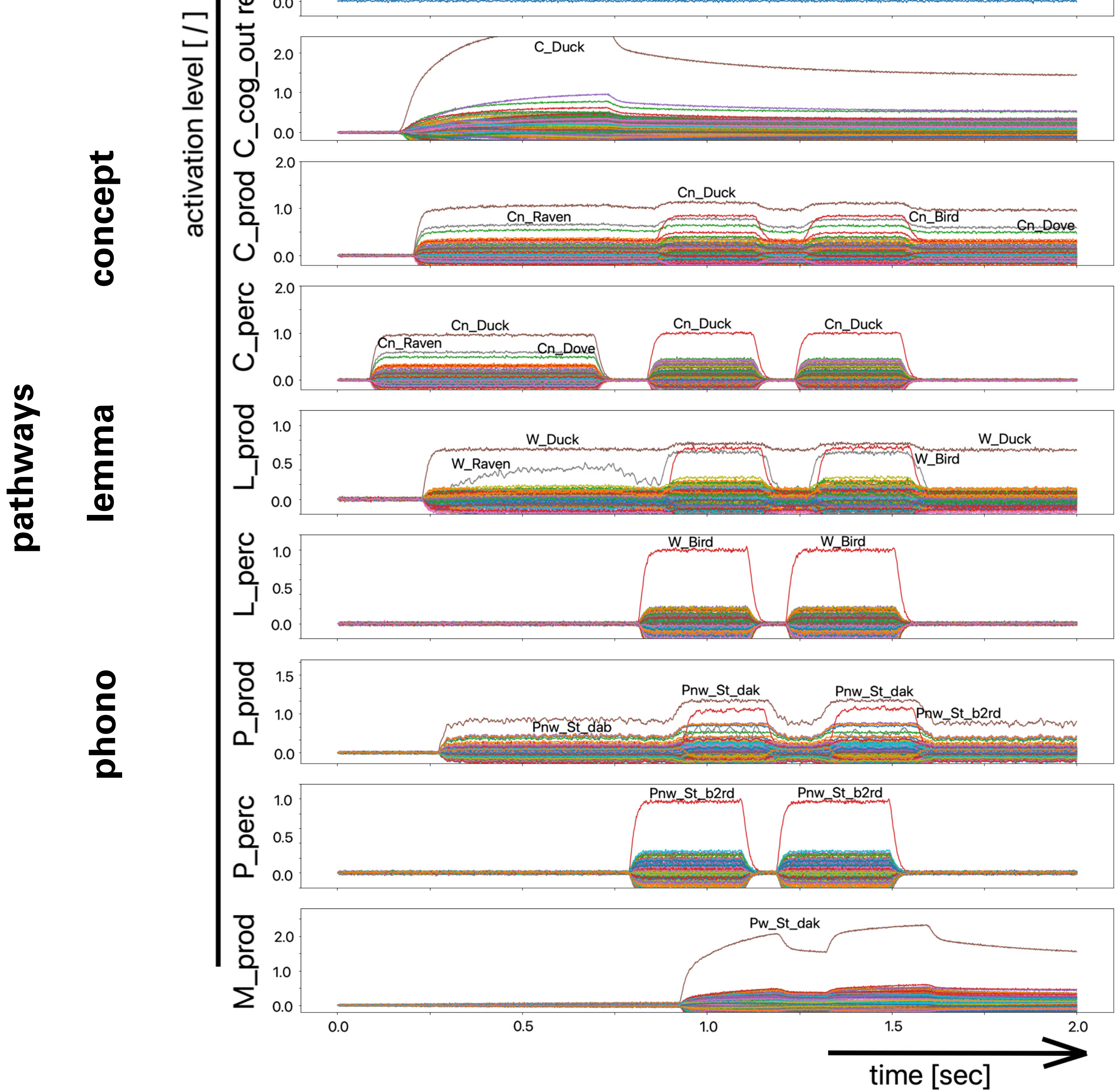
P_perc -> P_perc shortcut!

Kröger et al. (2020)



← priming: 100 msec:
"look and think" + 600 msec :
"utter noun"
+ later: "listen to cues" +
"utter noun"

Kröger et al. (2020)



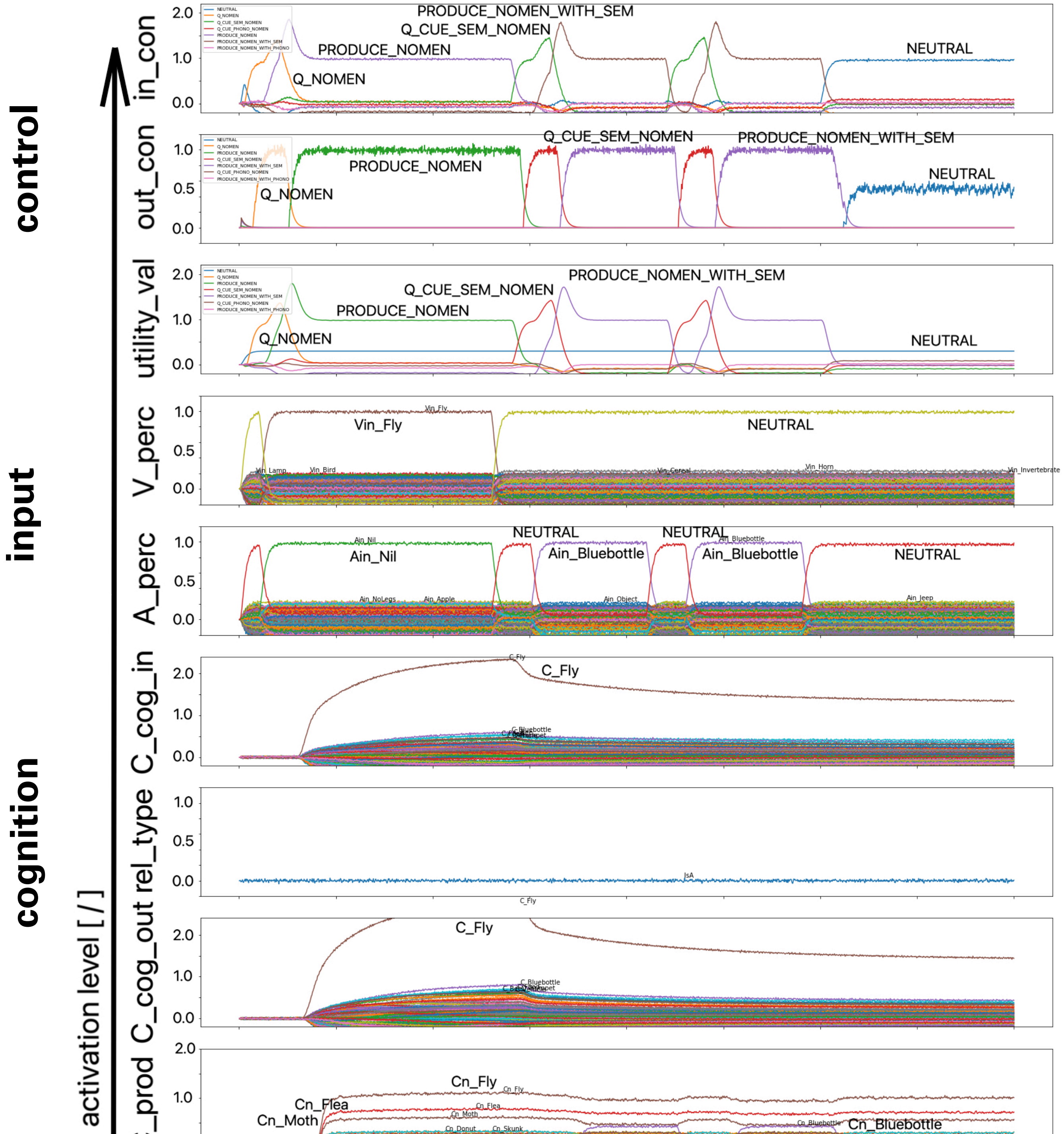
C_perc -> C_prod shortcut!

No P_perc -> P_perc shortcut!

Kröger et al. (2020)

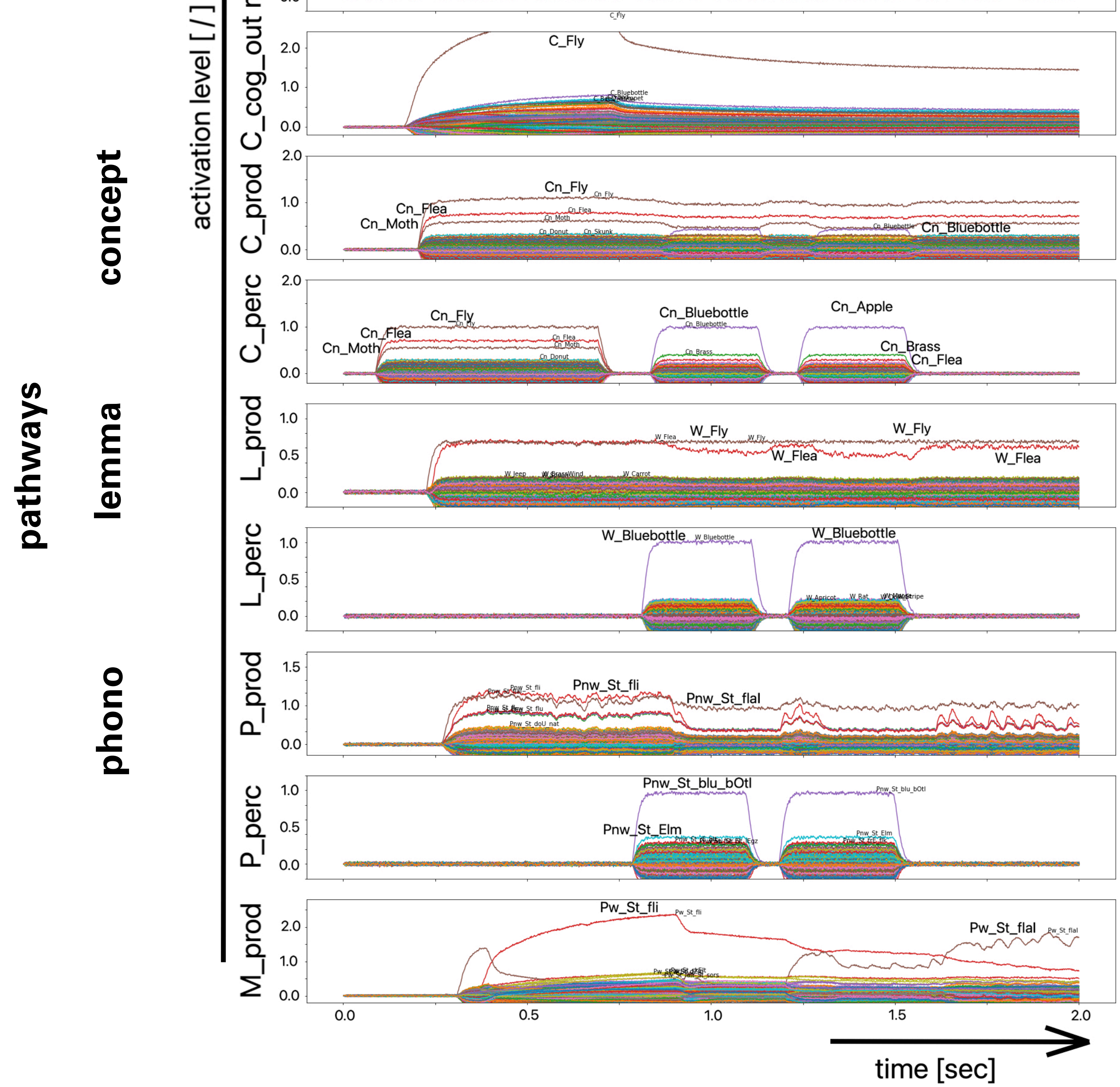
The definition of scenarios

- Further speech screenings: (-> logopedic research)
  - Picture naming with phonological or semantic cues (acoustic)
    - Naming without cues in case of “word is on tip of my tongue” with cues:
    - 4 cases: ->
      - Word activation starts with phonological cue (not earlier); 2 cases: “snake” and “apple”
      - … starts with semantic cue (not earlier) 2 cases: “duck” and “fly”
    - Fourth case: -> rare case
      - fourth case: A (semantic and) phonological similar word (“flea” (Floh)) will be activated first and corrected later by a semantic cue ((“bluebottle” (Schmeißfliege), also: “meat fly”))
      - -> rare case: correction processes may occur in neural models! (realism!!!)



← **priming: 100 msec:**
”look and think” + 600 msec :
“utter noun”
+ later: “listen to cues” +
“utter noun”

Kröger et al. (2020)



C_perc -> C_prod shortcut!

No P_perc -> P_perc shortcut!

Kröger et al. (2020)

Realism of the model:

- Is the model (modeling of single neurons, of connections, of buffers, etc.) biologically “realistic”?
- Most biological realism for
  - the single neuron models (leaky integrate and fire) -> spiking
  - the cortico-cortical BG-Thalamus loop for process control
- But: using the NENGO-concepts of neuron buffers, memories, associative buffers, binding buffers S-pointer-networks etc. leads to a good modelling of all types of behavior occurring in human data

Realism of the model:

- the large-scale model approximates human behavior surprisingly well in many cases: e.g.,
  - the serial recall task –> concave form: middle words have lower production probability (Choo 2010)
  - the rare cases; even for normal picture naming -> < 1% production errors or production fails (Kröger et al. 2016)
  - cues are helpful in case of mild mixed aphasia -> increase of correct word naming from 20-45% (no cues) to 55-70% (-> increase in performance)

Realism of the model:

- the large-scale model approximates human behavior surprisingly well in many cases: e.g.,
  - Modelling symptoms (task performance) for different types of aphasia (-> decrease in performance with increasing strength of disorder)
  - Modelling picture naming with distractor words (-> decrease in performance (fails or errors) in case of occurrence of distractor words)
- so: in some cases: approximation of human data on quantitative level is moderate;
  - But all "human behaviors" , all "effects" are modeled correctly (qualitative level)

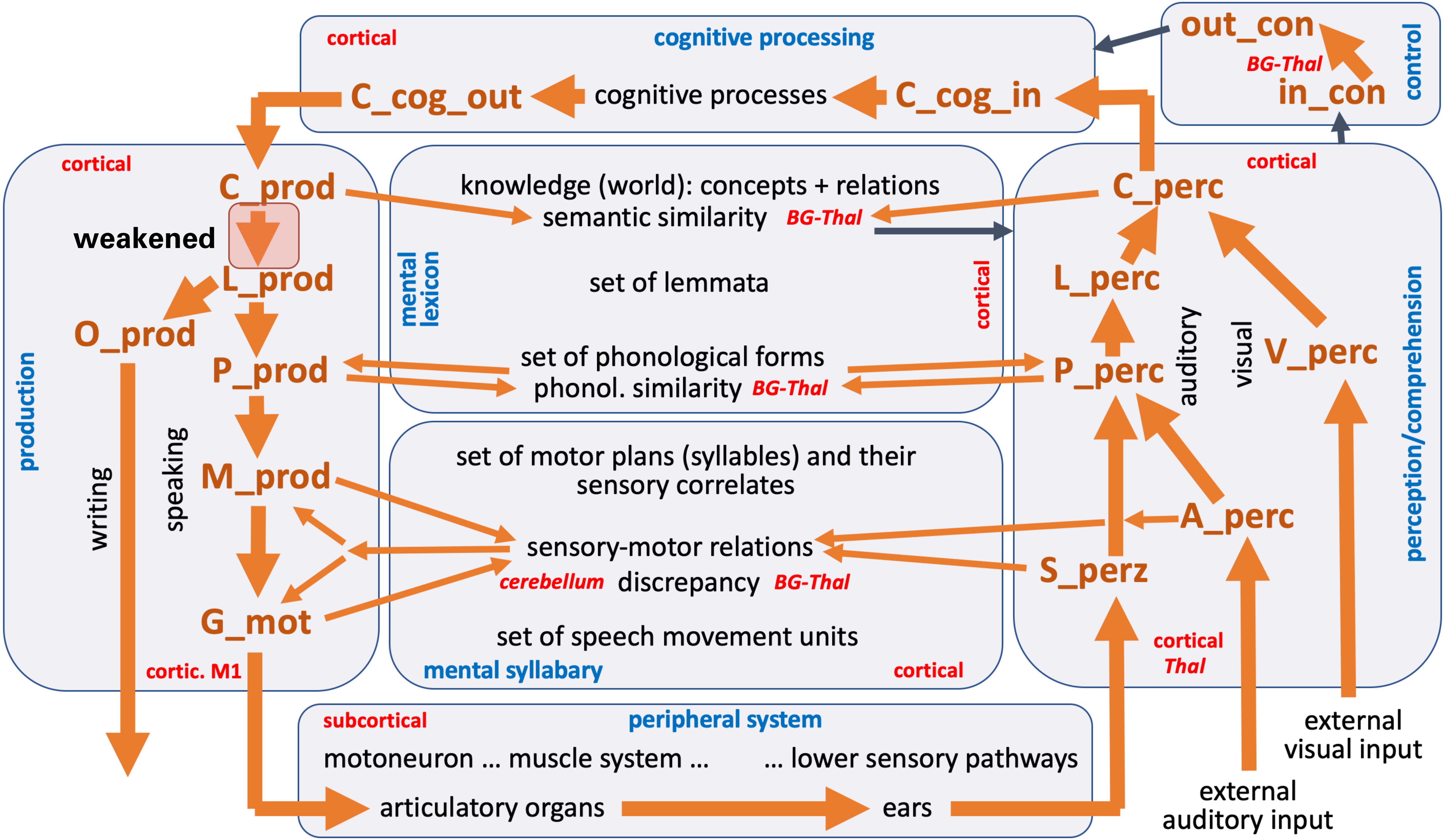
Benefits from the model:

- Medicine / speech therapy research:
- Simulation of behavior of patients in speech screenings is quite close to reality (model data <-> human data)
- Creation of "simulated patients" with specific neural dysfunctions is already possible!
  - -> that allows checking of different versions of a screening with respect to its increase in sensitivity for detecting a specific neural dysfunction (a specific type of speech disorder)

Download Video6:
https://www.youtube.com/watch?v=wz_KWL9ugn8

Addendum: Future work: Integration of articulatory-acoustic model

- Results shown so far -> cortical part of the model
- Further work: temporal control for syllable and speech gesture coordination -> Motor plans and their execution: time series of actions / gestures:
  - Vocalic gestures / actions: vocal tract form actions: /a/ , /i/, /u/,
  - Consonantal gestures / actions: oral full/near closing actions, labial, apical, dorsal … -> /b/, /p/, /m/, … /s/, /z/, …
  - Glottal gestures / actions: phonation, opening actions (voiceless sounds), closing actions (glottal stop)
  - Velum: opening / closing actions of velopharyngeal port

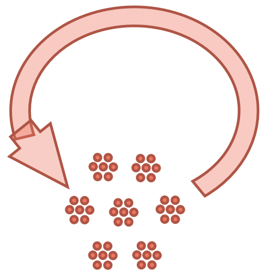


cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
production
C_prod
weakened
L_prod
O_prod
P_prod
M_prod
G_mot
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
cortical
Thal
perception/comprehension
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

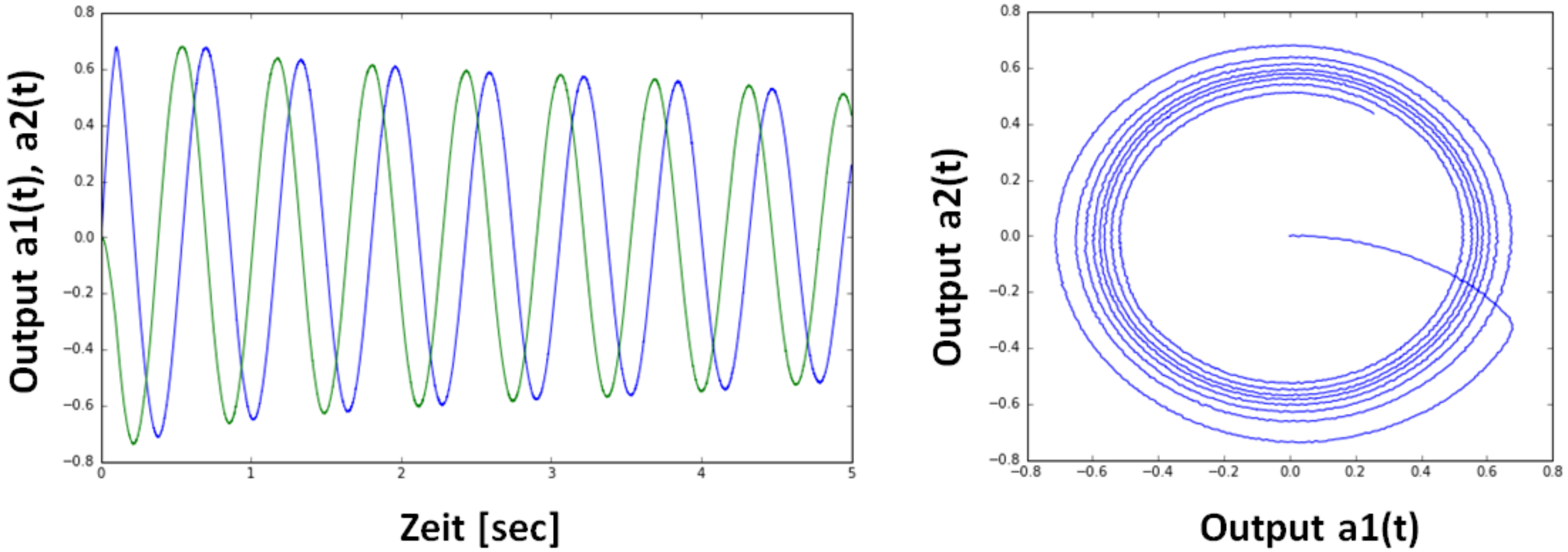
Addendum: Future work: articulatory-acoustic model integration

- -> neuro-biological control concept for speech movements including generation of articulatory movements and acoustic speech signal
- Modeling of articulatory gestures, phasing concept for timing of gestures, … -> part of quantitative "Articulatory Phonology" (Goldstein et al. USC, Haskins Labs)
- but our control concept is implemented as a part of a neuro-biologically based brain model: neuron ensembles -> oscillators; based on spiking LIF-neurons, … (differs from quantitative AP)

Buffers as oscillators

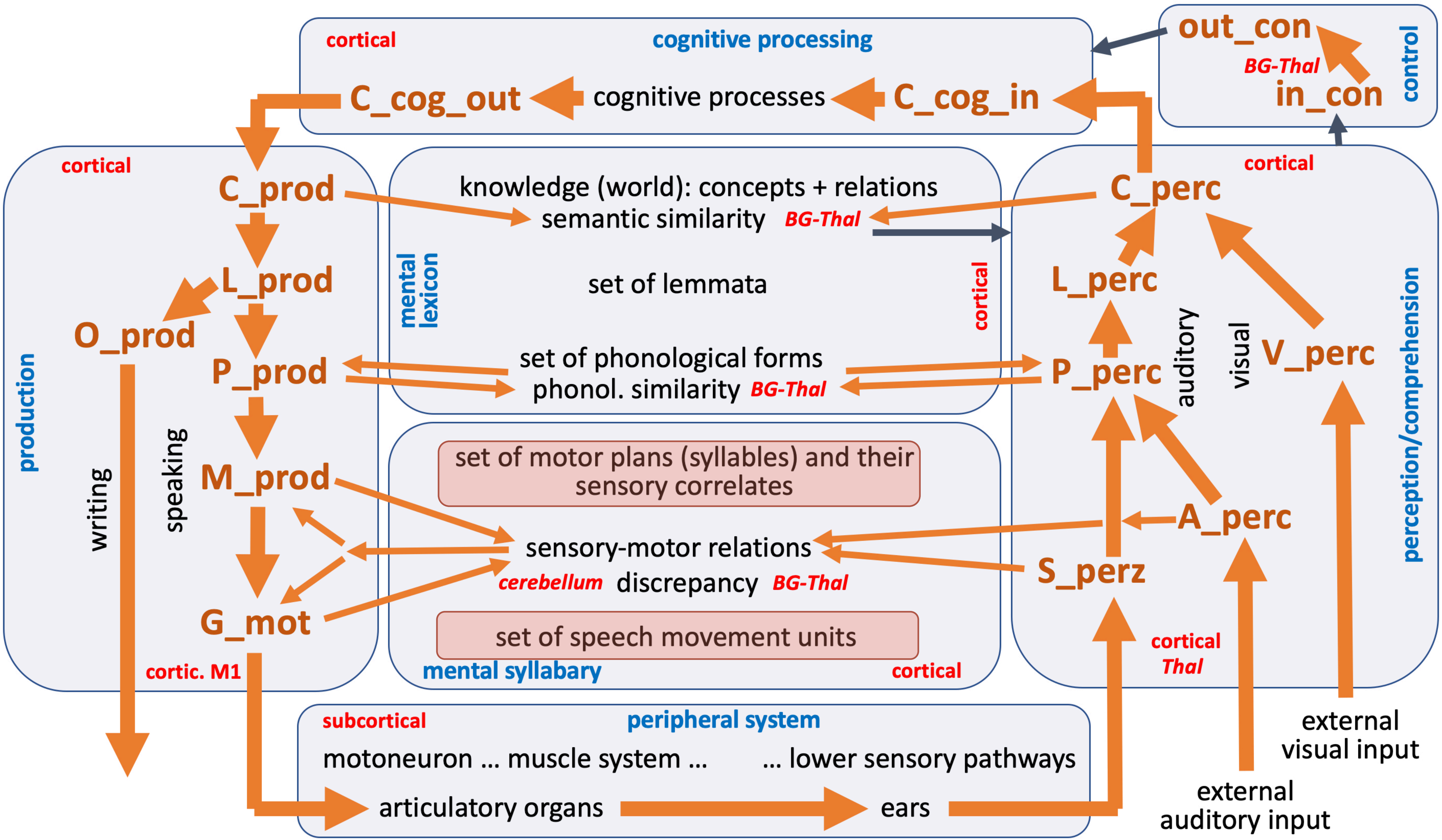


- Buffers with recursive neural connections -> short-term memories or oscillators
- activation transfer from 50% of the neurons to the other 50% of the neurons of the buffer and back …

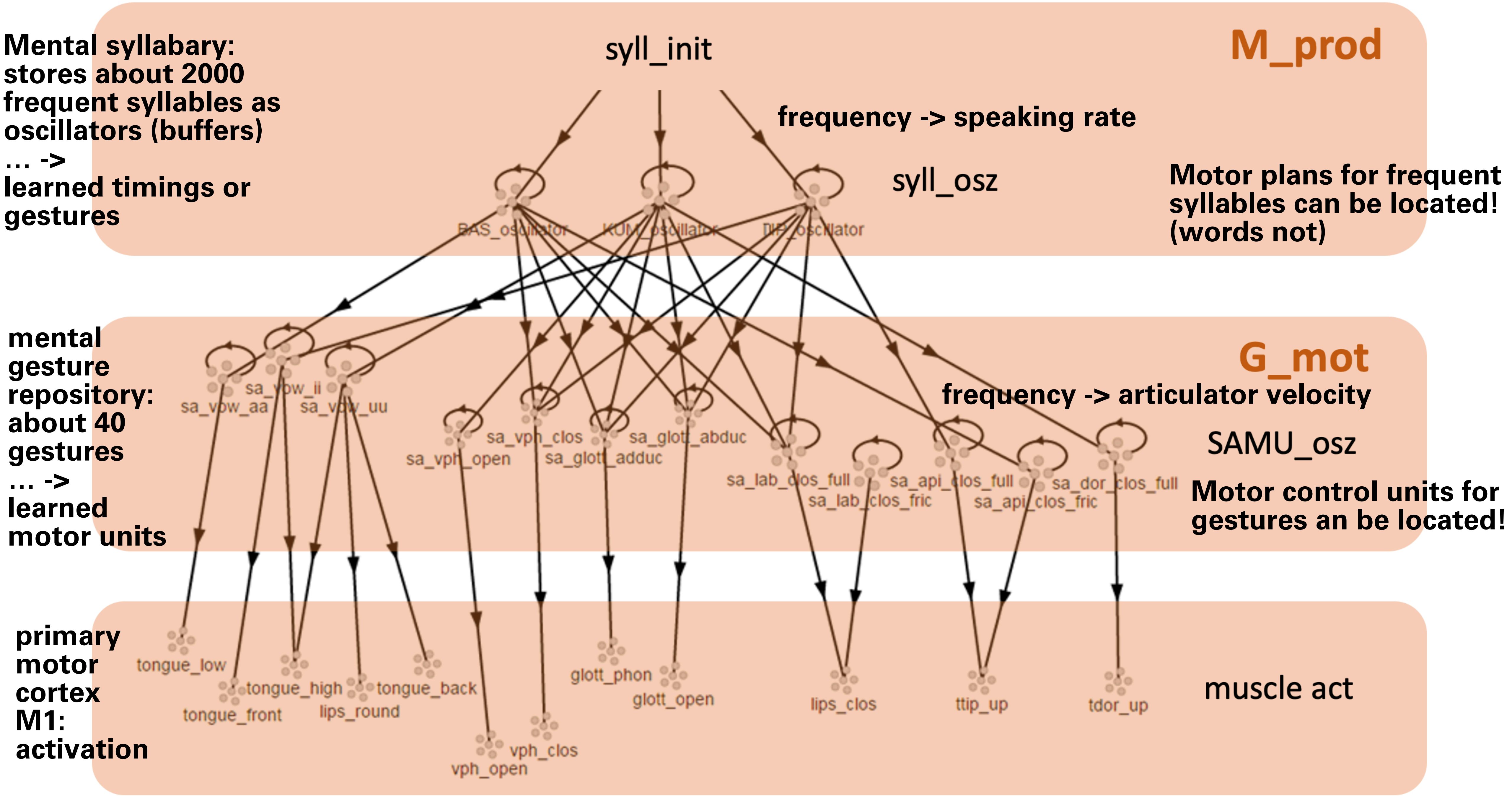


The motor part of the neural model

- from phonological form to motor plans to a temporal series of gestures / speech actions :
- Syllable as basic unit:
- motor plans are stored as a whole "mental syllabary"
- Motor plans store the timing (phasing) of smaller units: gestures or vocal tract action units (= basic motor control units)
- Learning of motor plans? -> somatosensory and auditory feedback is important



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external visual input
external auditory input

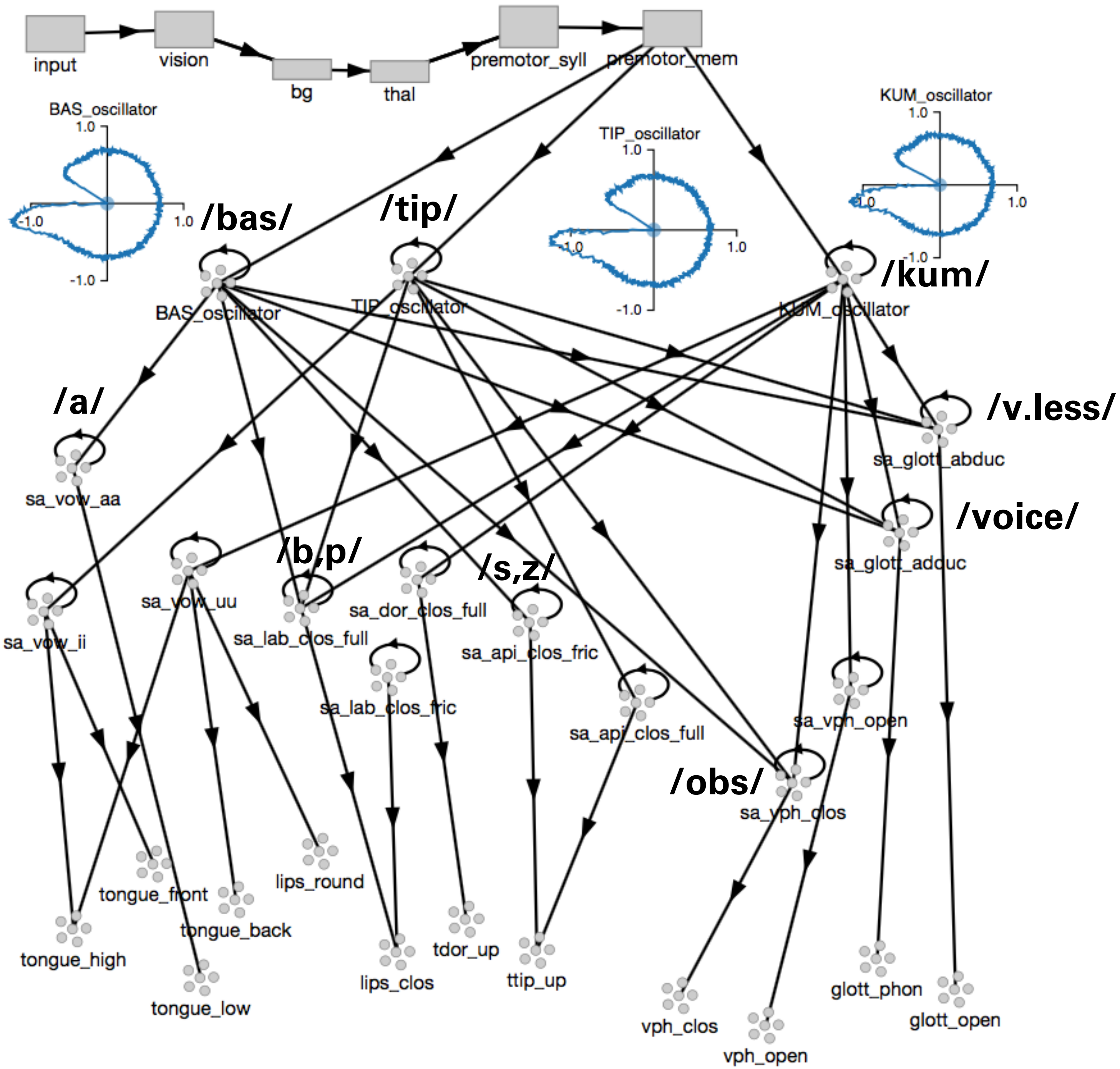


Mental syllabary: stores about 2000 frequent syllables as oscillators (buffers) … -> learned timings or gestures
syll_init
M_prod
frequency -> speaking rate
syll_osz
Motor plans for frequent syllables can be located! (words not)
BAS_oscillator
KUM_oscillator
TIP_oscillator
mental gesture repository: about 40 gestures … -> learned motor units
G_mot
frequency -> articulator velocity
SAMU_osz
sa_vow_aa
sa_vow_ii
sa_vow_uu
sa_vph_open
sa_vph_clos
sa_glott_adduc
sa_glott_abduc
sa_lab_clos_full
sa_lab_clos_fric
sa_api_clos_full
sa_api_clos_fric
sa_dor_clos_full
Motor control units for gestures an be located!
primary motor cortex M1: activation
muscle act
tongue_low
tongue_front
tongue_high
lips_round
tongue_back
vph_open
vph_clos
glott_phon
glott_open
lips_clos
ttip_up
tdor_up

Example word: /bas kum tip/

run Nengo_gui →

((or: run Nengo in python interpreter))

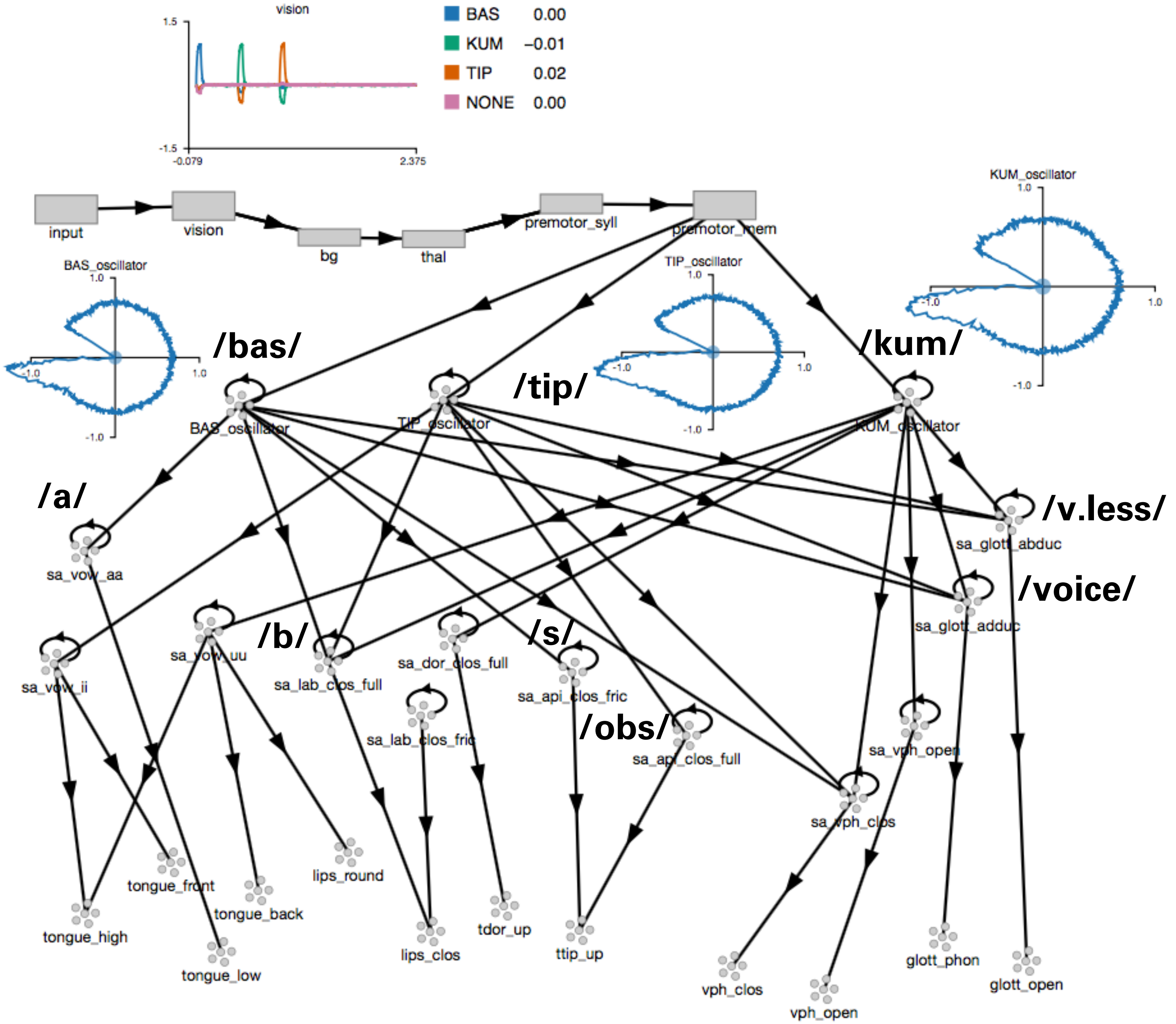


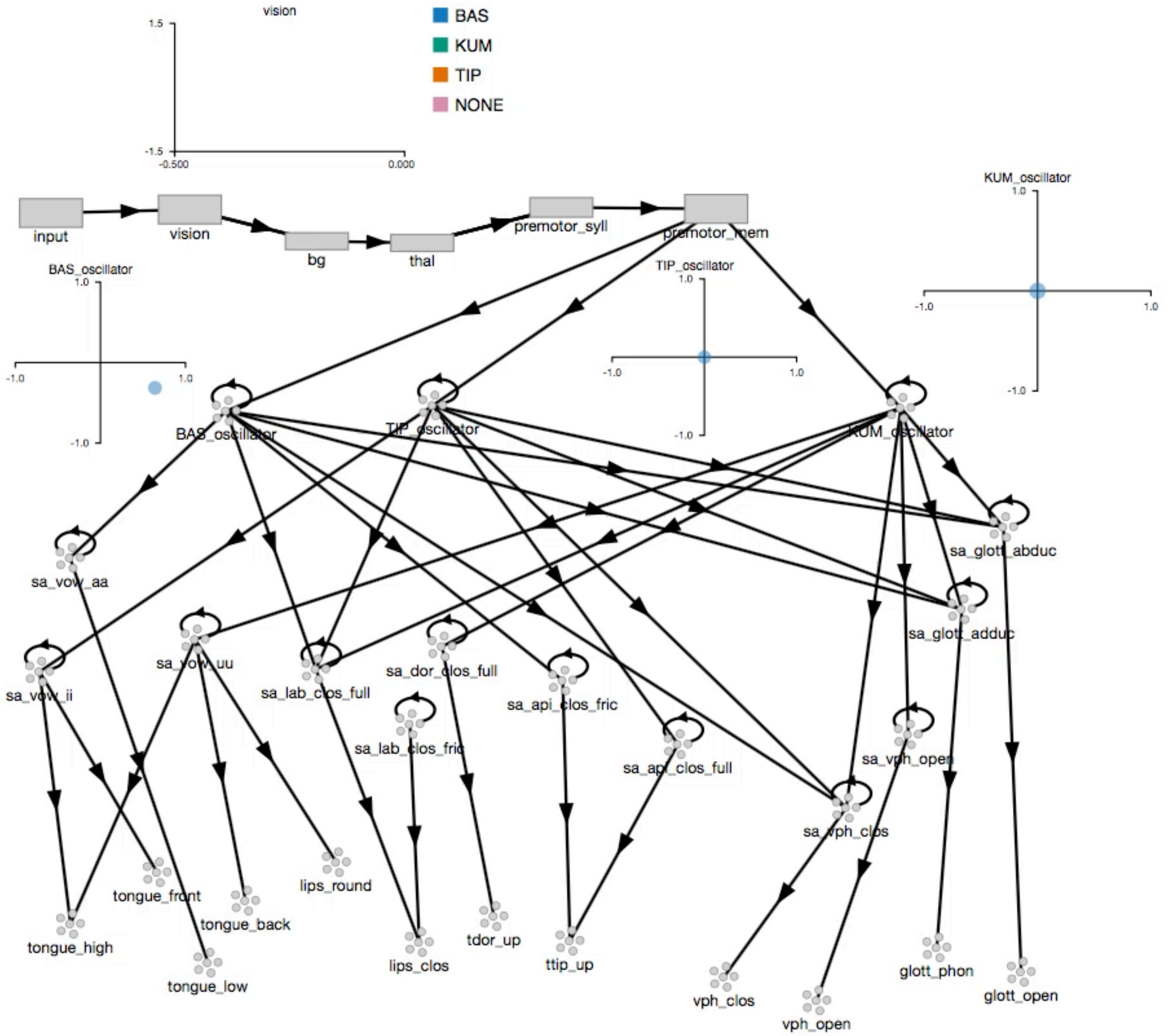
Why oscillators?

- Oscillators define time intervals (<-> duration of one oscillation cycle)
- Oscillators define timing of actions:
  - actions may start at specific points in time within an oscillation cycle = phase value
  - Timing of articulation remains phase-stable even for different frequencies of oscillation (for different duration intervals for each syllable; for different speaking rates)

Example word: /bas kum tip/

see the timing of the syllable oscillator activation: Video1_start100: simulation of three syllables



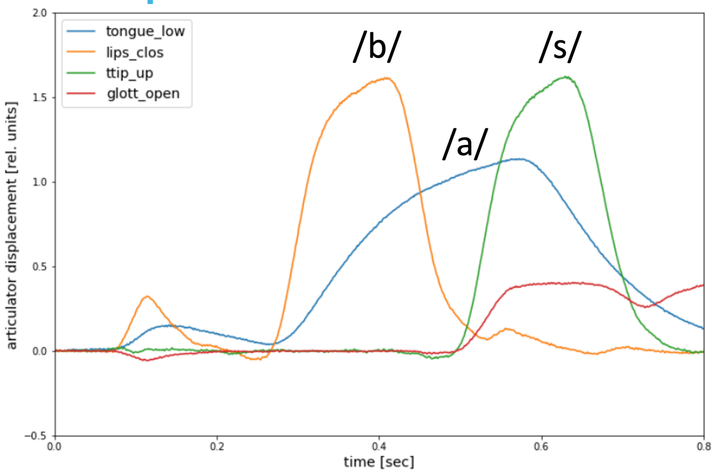


vision
1.5
-1.5
-0.500
0.000
BAS
KUM
TIP
NONE
input
vision
bg
thal
premotor_syll
premotor_mem
KUM_oscillator
BAS_oscillator
TIP_oscillator
BAS_oscillator
TIP_oscillator
KUM_oscillator
sa_glott_abduc
sa_glott_adduc
sa_vow_aa
sa_vow_uu
sa_vow_ii
sa_lab_clos_full
sa_dor_clos_full
sa_api_clos_fric
sa_lab_clos_fric
sa_api_clos_full
sa_vph_open
sa_vph_clos
tongue_high
tongue_front
tongue_back
tongue_low
lips_round
lips_clos
tdor_up
ttip_up
vph_clos
vph_open
glott_phon
glott_open

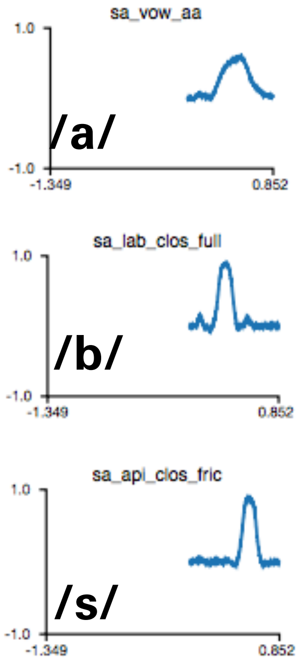
Syllable /bas/

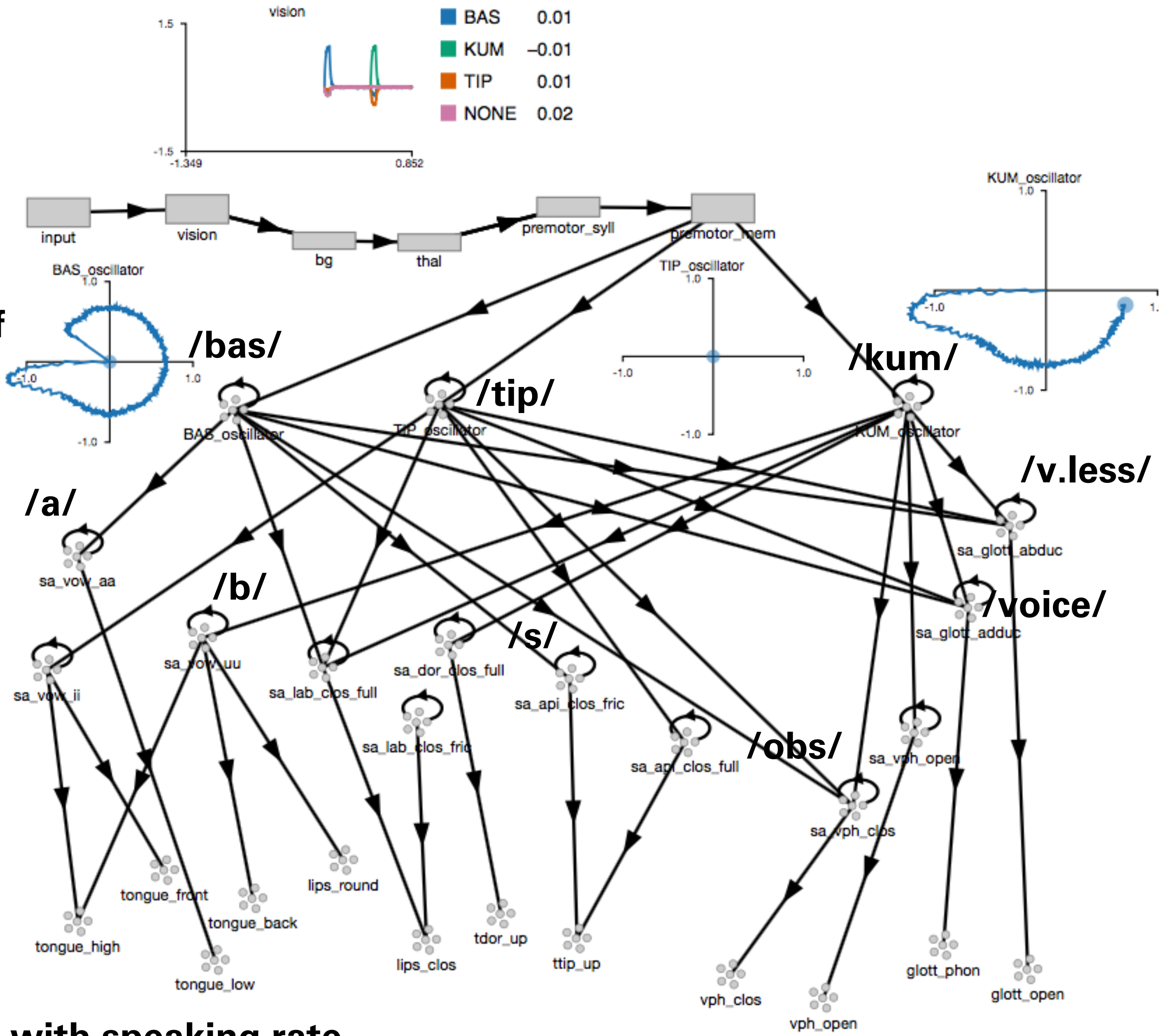
See the timing for speech gesture (action) activation:
Video2_start045: simulation of first syllable (neural model)

Introduction to syllable articulation:
Video3_repeat020: simulation of first syllable (articulation)



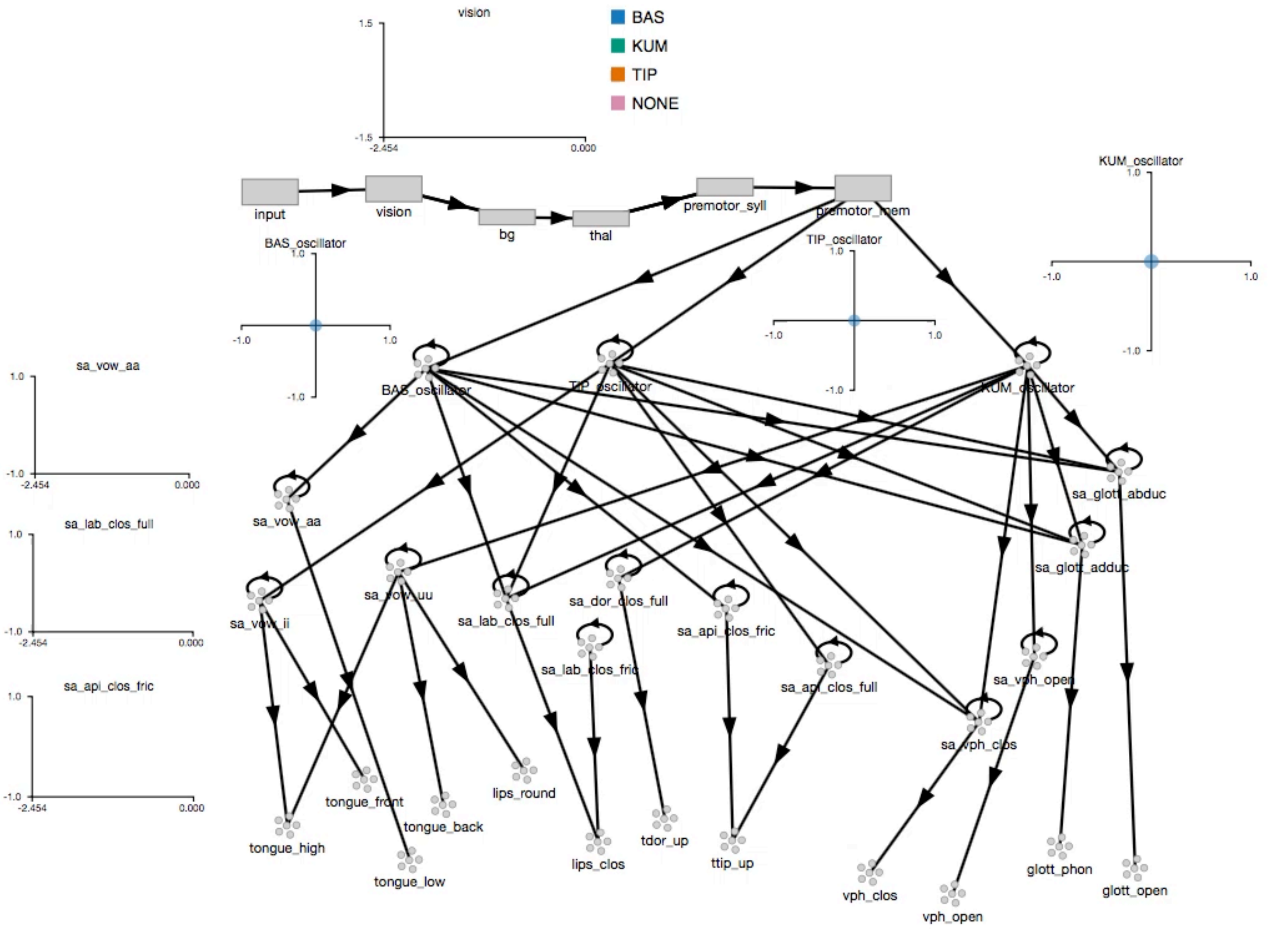
Gesture oscillators: degree of neural activation -> degree of realization of a gesture



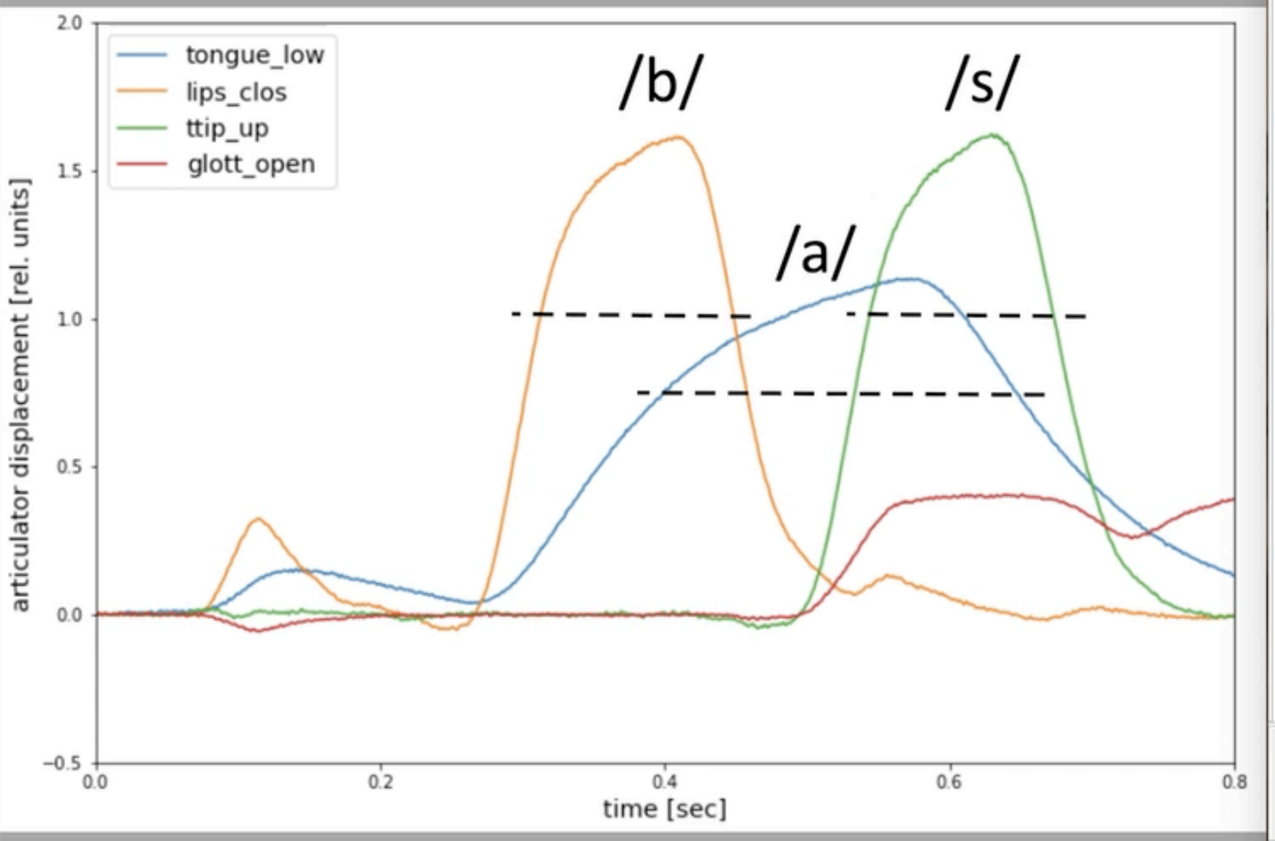


- **Artic.velocity = constant**
- **phasing: temporal overlap decreases with speaking rate**

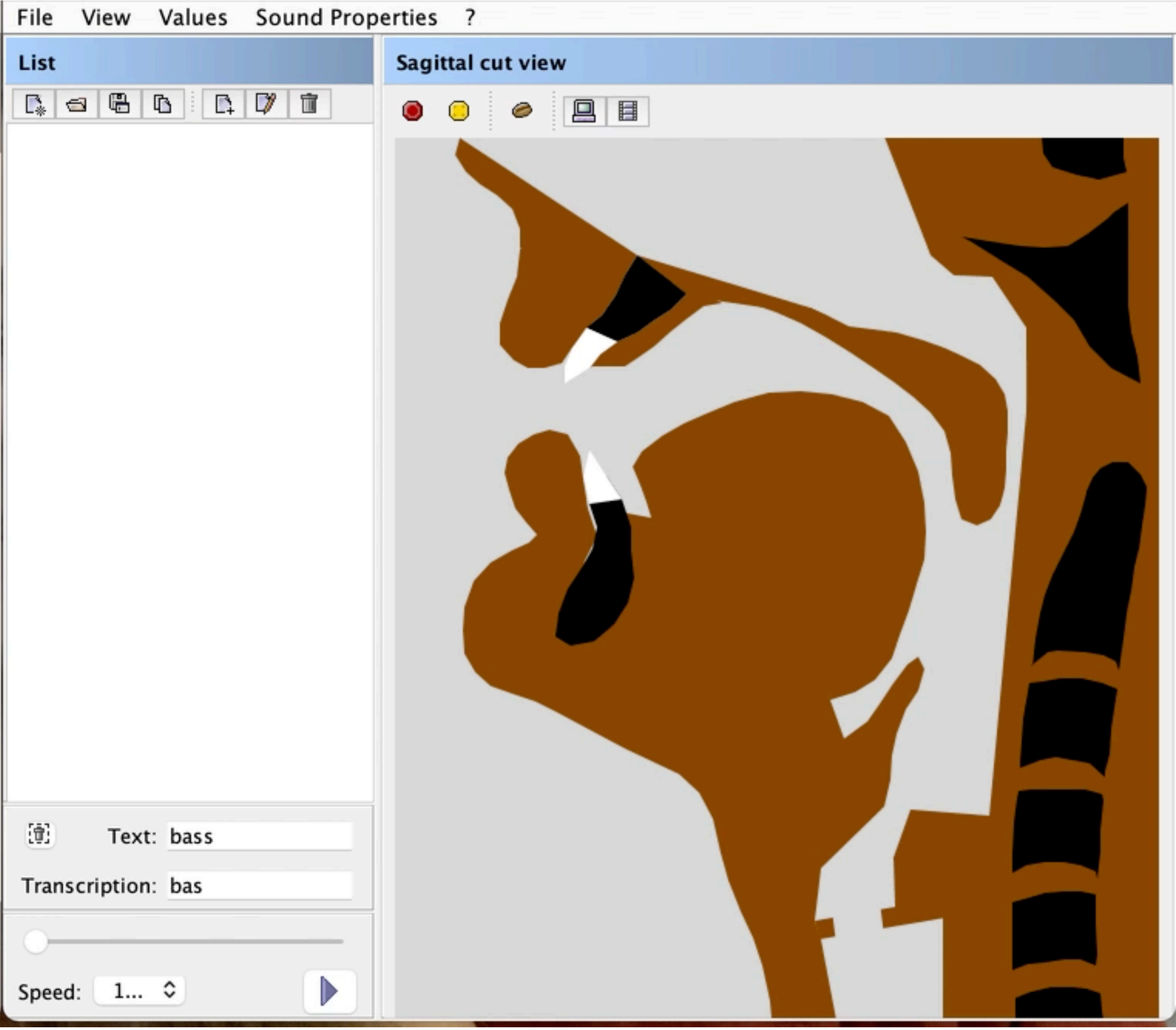
Gesture oscillators: degree of neural activation
-> degree of realization of a gesture
-> "abstract" movement trajectory



Gesture oscillators: degree of neural activation -> degree of realization of a gesture -> “abstract” movement trajectory



2D-articulatory model



Application of our approach: Modeling different speaking rates

- phase values (motor plan / syllable level) stay constant
- oscillator frequency for gestures stays constant (differentiates vocalic and consonantal gestures)
- oscillator frequency of syllable oscillators (motor plan level) changes

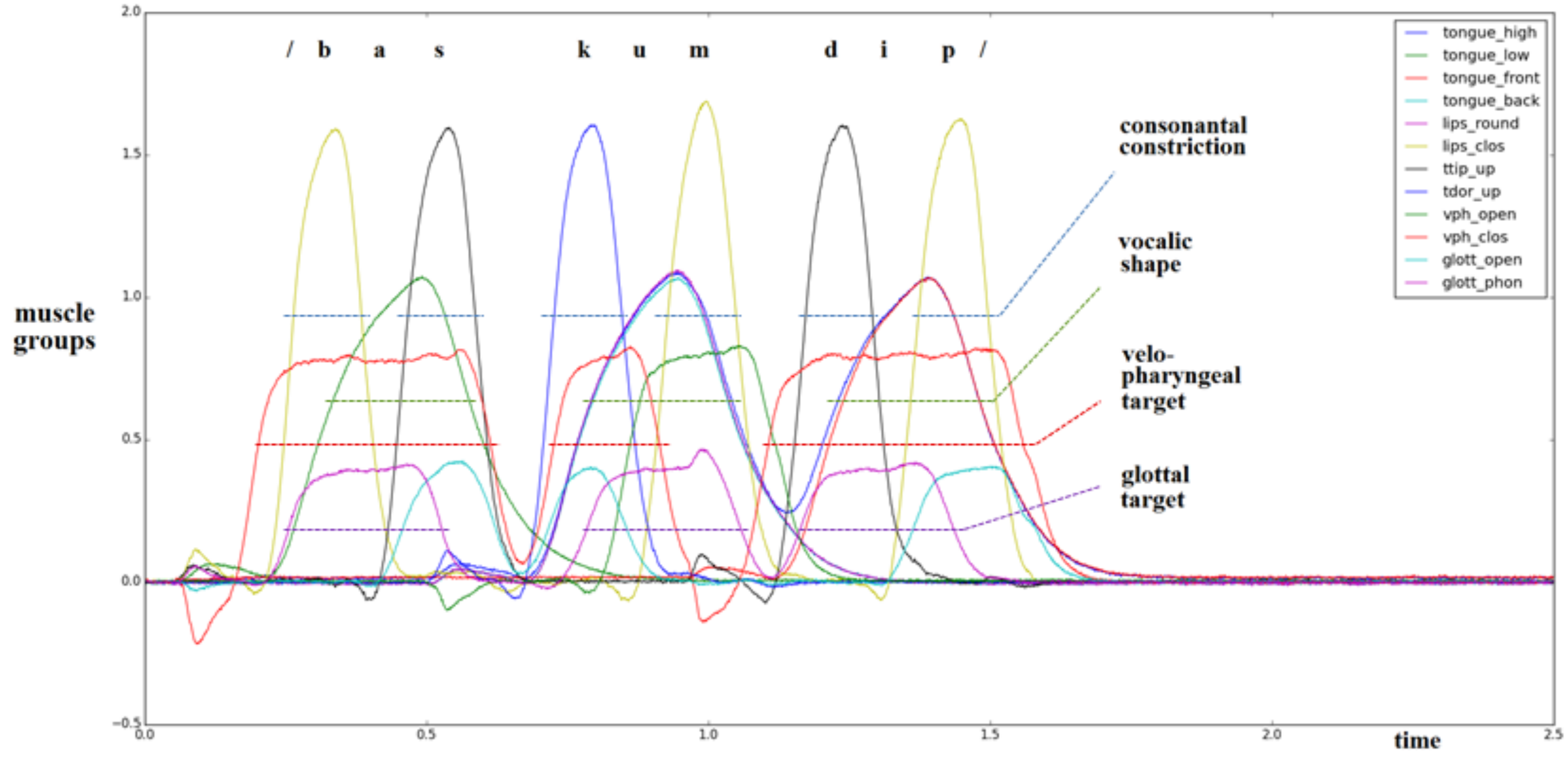
Kroeger, BJ, Bekolay T, Blouw P, Stewart TC (2020) Developing a model of speech production using the Neural Engineering Framework (NEF) and the Semantic Pointer Architecture (SPA). Proceedings of the International Seminar on Speech Production ISSP2020. Yale University, New Haven, CT. www.spechtrainer.eu -> publications

Results

- constant phasing of actions leads to correct production of speech sounds at each speaking rate:
- slow speaking rate:

no increase in articulatory effort per speech action

a

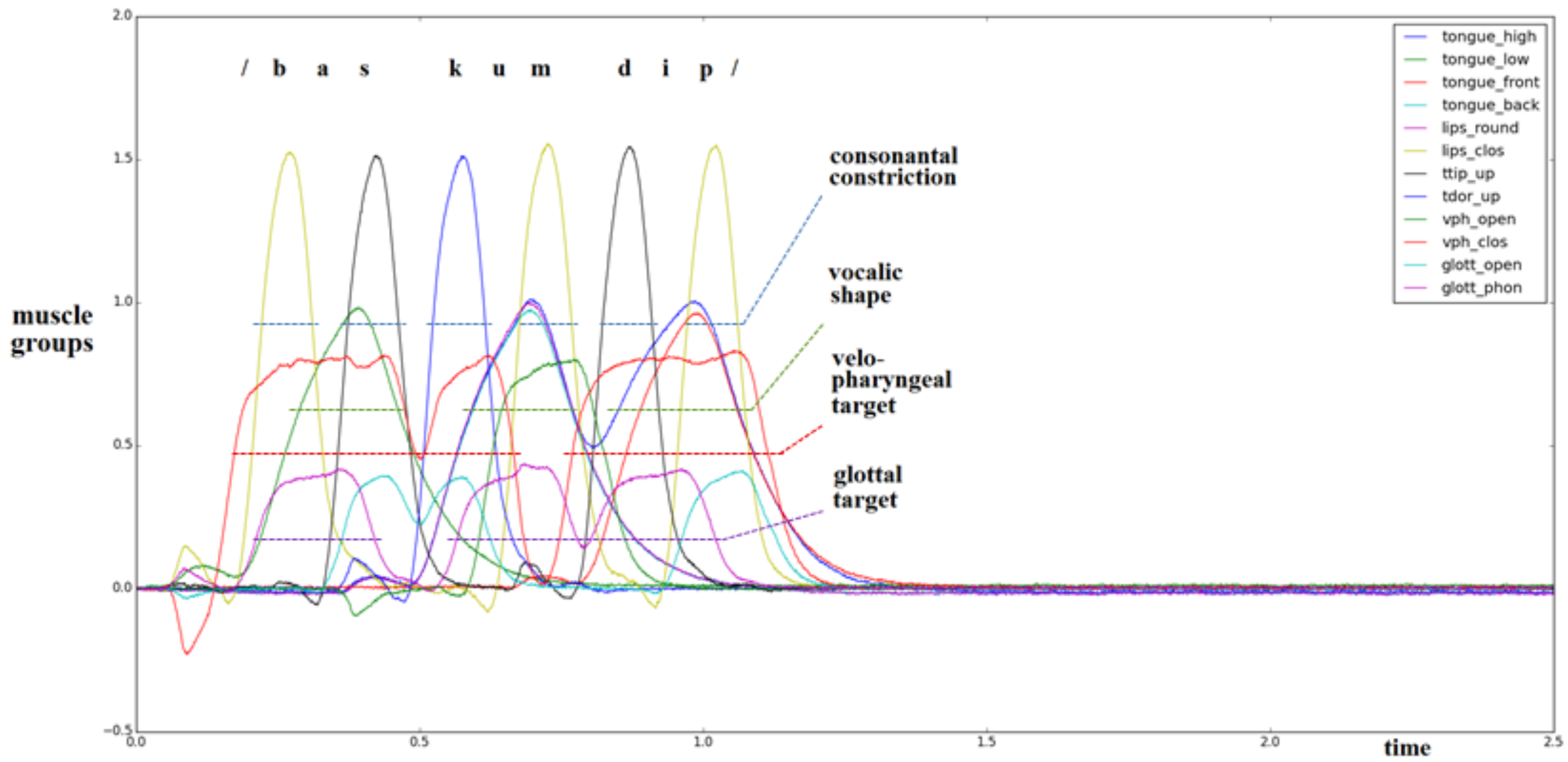


Results

- constant phasing of actions leads to correct production of speech sounds at each speaking rate:
- normal speaking rate:

no increase in articulatory effort per speech action

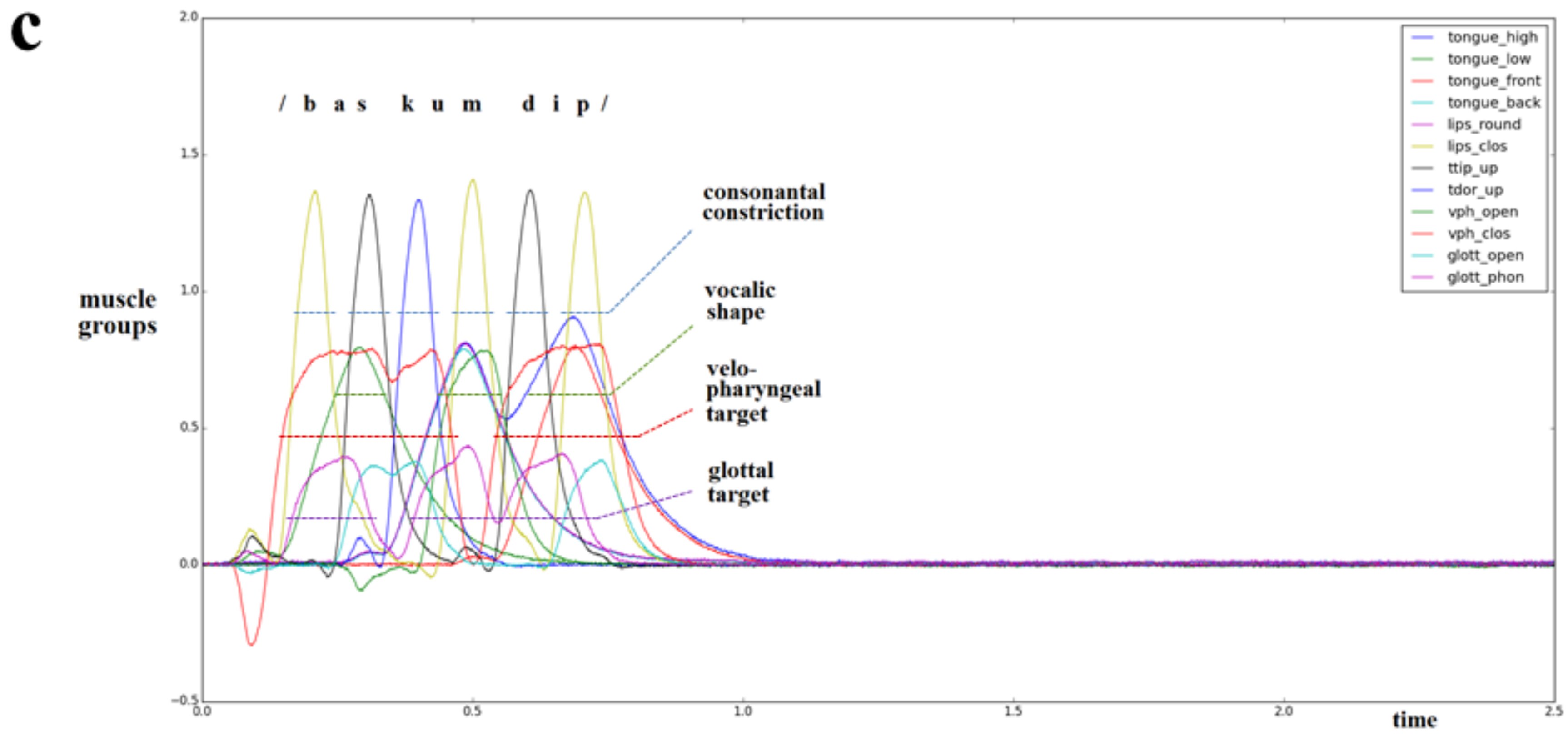
b



Results

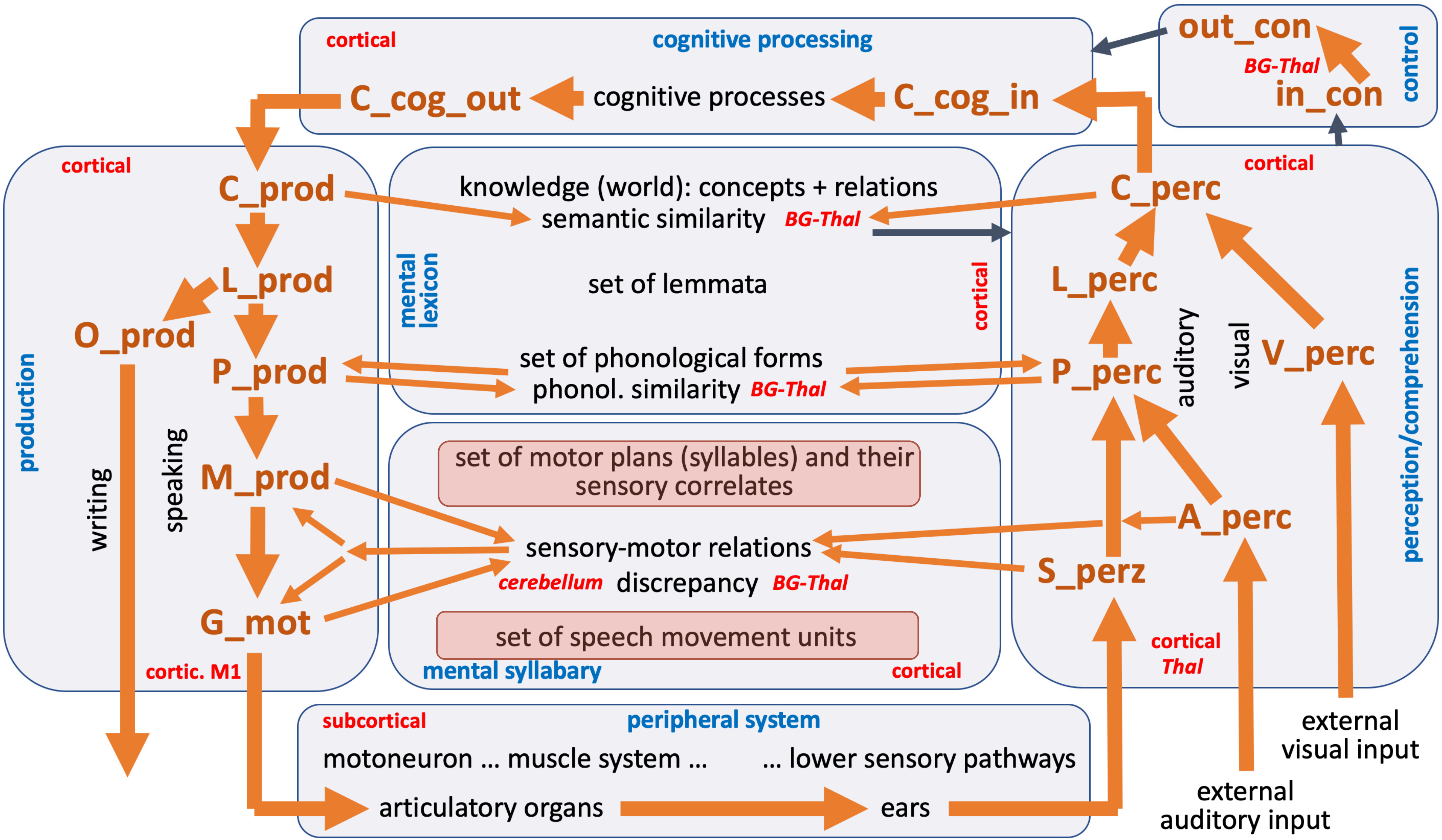
- constant phasing of actions leads to correct production of speech sounds at each speaking rate:
- fast speaking rate:

no increase in articulatory effort per speech action



Further work

- Further work: real-time integration of the articulatory-acoustic model into the neural brain model
- Integrating acoustic and somatosensory feedback allows motor (plan) learning
  - Timing of begin and release of constrictions / closures
  - Learning acoustic-motor relations
- -> implementation example for the articulatory model: see:
- SpeechArticulationTrainer (AppStore, GooglePlay)
- https://www.youtube.com/watch?v=C09EYber_T4



cortical
cognitive processing
C_cog_out
cognitive processes
C_cog_in
out_con
BG-Thal
in_con
control
cortical
C_prod
L_prod
O_prod
P_prod
M_prod
G_mot
production
writing
speaking
cortic. M1
knowledge (world): concepts + relations
semantic similarity
BG-Thal
mental lexicon
set of lemmata
cortical
set of phonological forms
phonol. similarity
BG-Thal
set of motor plans (syllables) and their sensory correlates
sensory-motor relations
cerebellum
discrepancy
BG-Thal
set of speech movement units
mental syllabary
cortical
cortical
C_perc
L_perc
P_perc
auditory
visual
V_perc
A_perc
S_perz
perception/comprehension
cortical
Thal
subcortical
peripheral system
motoneuron ... muscle system ...
... lower sensory pathways
articulatory organs
ears
external visual input
external auditory input