THE LEXICON-SYLLABARY MODEL: EVIDENCE FROM NEUROIMAGING STUDIES

Cornelia Eckers, Stefan Heim, and Bernd Kröger

RWTH Aachen University
cornelia.eckers@rwth-aachen.de

Abstract: A neurocomputational lexicon-syllabary model of speech acquisition, production and comprehension has been suggested. It contains modules representing the mental lexicon, the mental syllabary, the articulation, and the lower level sensory processing. In order to prove that the neurocomputational lexicon-syllabary model is realistic, it has been compared to functional imaging studies dealing with speech production and speech perception. The results of most of these neuroimaging studies are in agreement with and provide evidence for our suggested model. Only the neuronal localization of the mental syllabary remains unclear. A neuroimaging experiment is proposed in order to get more evidence of the neuronal correlates of the mental syllabary, i.e. a brain region where syllable states are represented in a supramodal manner.

1 Introduction

Neurobiologically based computational models of speech acquisition, production and comprehension are rare [1, 9]. The suggested model consists of two central modules, i.e. the mental lexicon and the mental syllabary. These modules comprise neural maps which are capable of storing semantic, phonological, sensory and motor representations of speech items like syllables or words [21]. But are this model and its maps realistic? That means: Are the maps and their connections represented in the central nervous system? The answers will provide more evidence for the current model and in general for the theory of speech processing. For that reason the lexicon-syllabary model is compared to the results of functional magnetic resonance imaging (fMRI) studies.

2 The lexicon-syllabary model and its functions

2.1 The organization of the model

Our proposed neurocompuational lexicon-syllabary model consists of different modules and maps which represent different functions (Fig. 1). There are processing modules, i.e. the semantic processing and phonological processing. Further processing modules are the mental lexicon and the mental syllabary. In this paper we will concentrate especially on these two last-named modules. Moreover there are two lower level processing modules, i.e. the articulatory processing, comprising the neuromuscular programming and execution and the articulatory-acoustic model, and the lower level sensory processing ending up in sensory short term representations (i.e. external auditory and somatosensory states, Fig. 1). Within the mental lexicon and mental syllabary two different types of maps occur: self-organizing maps (SOM) and state maps. Self-organizing maps are the higher level semantic map (S-Map) as central layer of the mental lexicon and the lower level phonetic map (P-Map) as central layer of the mental syllabary. These SOMs are interconnected in bidirectional ways by synaptic link weights to different domain-specific state maps (Fig. 2). The SOMs and their link weights are part of the long-term-memory. They are capable of representing specific states, i.e. words within the S-Map and syllables within the P-Map by local punctual neuron activations.

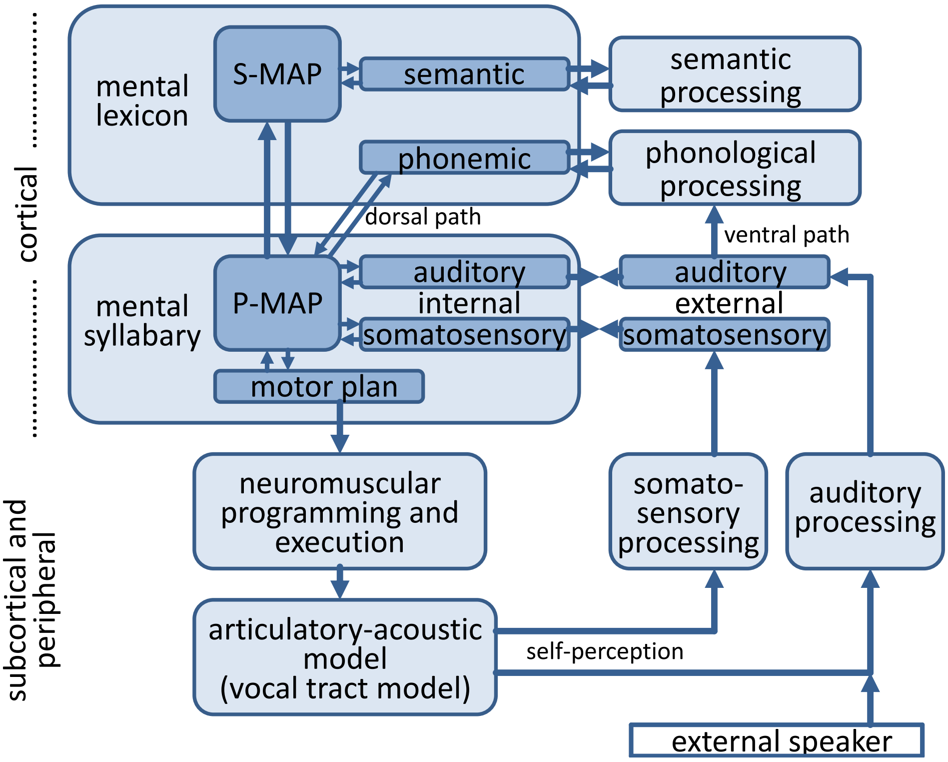

Figure 1 - Structure of our neurocomputational model of speech processing. Light blue boxes indicate processing modules; dark blue boxes indicate self-organizing maps (S-MAP and P-MAP) and neural state maps (i.e. semantic, phonemic, auditory, somatosensory and motor plan state map) (reproduced from Kröger et al., 2011, p. 2414, [20]).

The domain-specific state maps are part of the short-term memory. In terms of the mental lexicon these are the semantic state map and phonemic state map. Relevant to the mental syllabary these are the auditory state map, somatosensory state map and motor plan state map. The domain-specific state maps comprise ensembles of spatially closely connected model neurons. E.g. the neural representation of an auditory state within the auditory state map can be assumed to be a neural representation of a bark-scaled acoustic spectrogram [20]. By activating one neuron within the P-Map or S-Map (local activation) link weights are more or less activated and lead to an activation pattern of all neurons within the domain-specific state map which represents the word's or syllable's, e.g. auditory, state.

2.2 Speech production

A specific semantic activation pattern of a lexical item, occurring in the semantic state map (Fig. 1), is the starting point of word production. It represents semantic features, reached from different areas, from sensory, somatosensory, olfactory, auditory and motor areas to define a non-abstract object (e.g. a visible object like a cat). This activation pattern leads next to a local

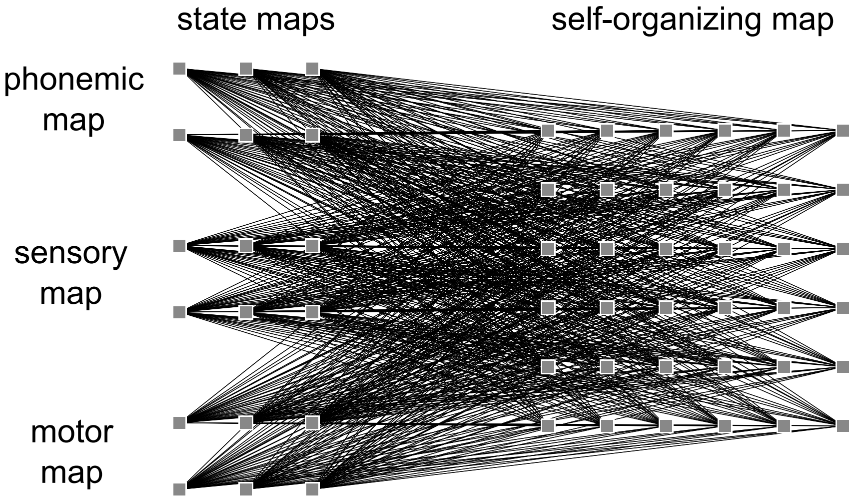

Figure 2 - Example of a self-organizing network for the mental syllabary (cf. [18]). Heaps of grey squares indicate neuron collectives (i.e. neural maps). Black lines (i.e. link weights) indicate neural connections between domain-specific state maps (i.e. phonemic, sensory, motor) and a self-organizing map (SOM) (here: phonetic map). The SOM and all neural connections represent the self-organizing network (reproduced from Kröger et al., 2009, p. 797, [24]).

punctual co-activation within the S-Map. The S-Map is a crucial part of the mental lexicon, capable of organizing lexical items with respect to their semantic features [28]. Subsequently there will be a co-activation within the P-Map and its connected state maps. The P-Map is the central layer of the mental syllabary. It is capable of activating phonological, sensory, and motor representations of frequent syllables [21] by bidirectional interconnections with the sub-symbolic domain specific state maps, i.e. the phonemic state map, auditory state map, somatosensory state map and motor plan state map. The phonemic state map offers symbolic phonological information concerning the syllables and is still part of the mental lexicon. The auditory state map is capable of activating the auditory state of a speech item, i.e. how it sounds like. Within the somatosensory state map the tactile and proprioceptive information of the articulation process, i.e. how it feels like, can be activated. The motor plan state map is capable of activating motor plan states of a syllable or word. Articulation starts from the motor plan state map towards a subsequent neuromuscular programming and execution of a succession of temporarily overlapping vocal tract actions as defined by the motor plan (also called gestural score or vocal tract action score) [19]. The vocal tract model generates the acoustic realization of syllables, words, and short utterances [2].

2.3 Feedback control

In order to monitor self-articulated speech items, there is a feedback control pathway, comprising the articulation module and the lower level sensory processing module (Fig. 1). The acoustic as well as the somatosensory result of articulation will be perceived. Feedback control starts by using peripheral sensory organs, e.g. auditory, proprioceptive and tactile sensors. These input signals are processed ending up in the sensory short term storage, i.e. external auditory and somatosensory states. Somatosensory information will activate a pattern within the (external) somatosensory state map, which can be compared with an already learned somatosensory state (internal). The same happened simultaneously concerning the acoustic information. The external auditory state will be compared to the internal auditory state. This leads to a comparison of what it sounds/feels like (external state map) and how it *has to* sound/feel like (internal state map). If the difference between the activation patterns within the external state maps and internal state maps accedes a certain limit, corrections can be initiated [10]. This feedback control is especially important for learning and training sensorymotor patterns by repeating speech items by monitoring them until they are learned (early phases of speech acquisition).

2.4 Speech perception

Apart from speech production and feedback control, routes for perception and comprehension of external speakers are also available (see figure 2). These routes are, except for monitoring, especially important for perceiving and comprehending speech items articulated by external speakers. An acoustic signal from an external speaker will pass the auditory processing, initiate a neural activation pattern within the auditory state map, and leads to a local co-activation within the P-Map. This local activation leads to a co-activation of a state within the phonemic state map, i.e. the phonemic representations of this lexical item. This is called the dorsal pathway in terms of Hickok and Poeppel [14]. Additional the activation within the P-Map causes a co-activation of the semantic representation within the S-Map and leads to a comprehension of the lexical item. Moreover the motor plan state could be co-activated via the P-Map. Except the dorsal pathway, [14] postulated a ventral pathway, which connects neural auditory representations of an external speech signal directly with phonemic and lexical representations. This can be included in our model via the connection of the auditory state map to the phonological processing (Fig. 1). This is not modeled in detail in the current version of this approach.0

3 The lexicon-syllabary model in context to fMRI

To get more evidence for the neuronal existence of the modules and maps of this neurocomputational lexicon-syllabary model of acquisition, production and comprehension, it has been compared to results of different functional imaging data from various experiments of speech processing. The model will be distinguished into four main modules to discuss their neuronal correlates in a more structured way: semantic modules (i.e. the mental lexicon, semantic and phonological processing), mental syllabary, articulatory processing and sensory processing.

3.1 Semantic modules

The mental lexicon in general is located mainly in the left frontal and temporal lobe [16]. The S-Map, as the central layer of the mental lexicon, can be compared to an amodal semantic hub [26], which is assumed to be located in the anterior temporal lobes [26]. The semantic states as well as the semantic processing module are hypothezised to represent a network which is widely distributed over the whole cerebral cortex including the anterior temporal lobe and anterior and posterior portions of the inferior frontal gyrus for controlled retrieval and selection of lexical items [25]. The phonological processing together with the phonemic state map, also representing a network, occur in particular in the left posterior inferior frontal gyrus, i.e. Broca's Area (BA 44) and in the posterior portion of the left posterior superior temporal gyrus (pSTG) [5, 6, 30]. Heim and Friederici (2003) figure out, that during comprehension first the pSTG is activated and than BA 44. During production tasks this will be the other way around [13]. The pSTG is mentioned to be the phonological word form store, whereas the left BA 44 mentioned to be involved in the feature manipulation of phonemes and in the process of syllabification of phonemes [16].

3.2 Mental Syllabary

The left BA 44 seems to be the interface between phonological processing and phonetic and motor processing. After the higher level processing there will be accessed phonetic and motor processes, i.e. the articulation, from BA 44 [12]. In a fMRI-study of Riecker et al. (2005) [27] a preparative loop as a neural network is mentioned, which prepared the articulation. It is connecting neuronal functions in the left supplementary motor area (SMA), left anterior insula, left dorsolateral prefrontal cortex, and right superior cerebellum (Fig. 3). A study by Eickhoff et al. (2009) [7] dissociated a subnetwork, which is relevant for motor planning, too. Similar to [27] they found the insula playing a major role in speech motor preparation, which transferred information further to the cerebellum and additional to the basal ganglia (Fig.4). The described activation patterns could be compared to speech motor planning including the P-Map and motor plan state map of the mental syllabary as well as the programming of the articulatory processing (see below).

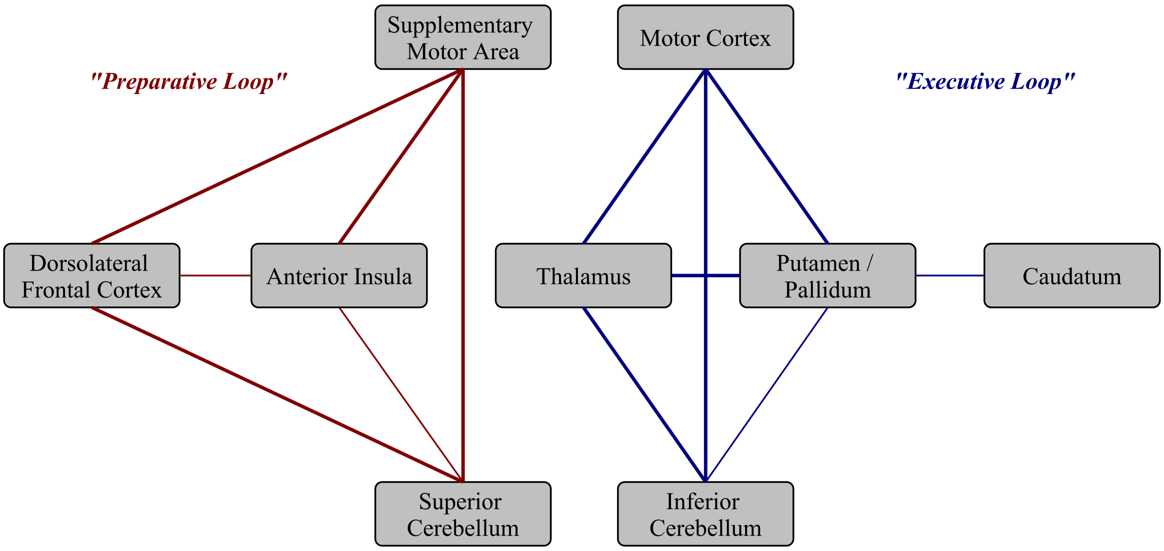

Figure 3 - Preparative and executive loop in terms of speech motor control (reproduced from Riecker et al., 2005, p. 704, [27])

Bohland and Guenther (2006) [3] mentioned the left perisylvian cortex as the central place of the mental syllabary resp. the P-Map. More precisely Gosh et al. (2008) [8] conclude that activation of the mental syllabary takes place in BA 44 as well as in the ventral part of the premotor cortex (PMC). Further fMRI-experiments lead to the assumption that the left anterior insula could be compared to the mental syllabary [4, 22, 27]. Further authors found neuronal correlates, which can be related to the mental syllabary in the posterior middle and inferior portions of both temporal lobes [14].

The auditory state map as short-term library of the sound-to-movement patterns seems to be located in the unimodal auditory and temporal multimodal sensory cortex (STG or BA 22). Further results of neuroimaging studies show an activation of this map in the posterior part of the STG [10]. Particularly Guenther et al. (2004) [11] and Ghosh et al. (2008) [8] found the auditory state map to be located in the transverse temporal gyri and planum temporale. In [7] they found activation for coordination of speech-articulation for sound-to-movement mappings in the anterior insula. Activation of processing corresponding to the somatosensory state map were found unimodal in the somatosensory cortex (BA 1, 2, 5 and anterior BA 7) and parietal multimodal sensory cortex (Gyrus angularis, gyrus supramarginalis, i.e. posterior BA 7, BA

39, BA 40) [10]. The results of [8] and Tourville et al. (2008) [29] reveal primarily the ventral somatosensory cortex as neuronal correlate for the somatosensory state map. Activations concerning the motor plan state map are hypothesized to happen especially in the PMC and/or in the SMA [7, 27].

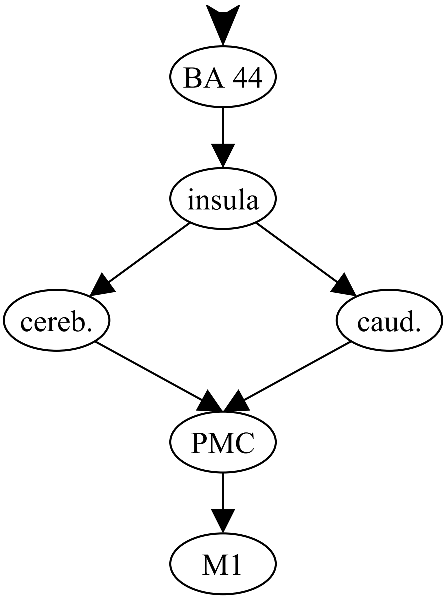

Figure 4 - Neural system model of overt speech production. Abbreviations: insula = anterior insula; cereb. = cerebellum; caud. = caudate nucleus; PMC = ventral premotor cortex; M1 = primary motor cortex (reproduced from Eickhoff et al., 2009, p. 2414, [7])

3.3 Articulatory processing

According to [7] programming and execution takes place in a network which involves information flowing from the above mentioned cerebellum and basal ganglia to the precentral motor cortex and further to the primary motor cortex (BA 4) (Fig. 4). The information proceeds further via subcortical structures towards the peripheral neuromuscular units, which directly controls the movements of the vocal tract articulator [7, 15]. In [3] as well as in [8] the primary motor cortex is mentioned, too, concerning the execution of motor plans. In [27] the execution of motor plans is also mentioned as a network, i.e. executive loop. It involves similar to [7] the left motor cortex, left thalamus, left basal ganglia, and right inferior cerebellum (Fig. 3).

3.4 Sensory processing

The auditory pathway starts at the ears and passes through peripheral and subcortical regions and finally reaches the primary auditory cortex (BA 41, 42, external auditory state map) [10]. The auditory processing is hypothesized to lie bilateral in the temporal lobes, i.e. the dorsal superior temporal gyrus [17]. The neuronal localization of somatosensory processing is hypothesized to lie bilateral in the parietal lobes. Somatosensory feedback signals reach the speech organ regions of the primary somatosensory cortex (BA 3, external somatosensory state map), i.e. the anterior inferior parietal lobe [17].

4 Conclusion and Discussion

The main goal of this paper is to describe the neurocomputational lexicon-syllabary model and to relate neuronal correlates to the modules of the current model in order to support that this model is realistic. There can be found many neuroimaging studies in literature, which provide evidence for the existence and neuronal localization of the different modules as well as maps and interconnections. Each map can be mapped to specific cortical or sub-cortical regions. But a rest of doubt remains, because the implemented neuroimaging experiments were not defined closely to the functions of the current model. But this will be done in further work concerning the mental syllabary resp. the P-Map and its connections to the state maps. Especially about the neuronal localization of the mental syllabary there is also further disagreement among different studies. It is hypothesized to lie in the perysilvian region [3] or more precisely in the left posterior inferior frontal gyrus as well as in the ventral part of the PMC [8] or in the left anterior insula [22] or even in the posterior middle and inferior portions of both temporal lobes [14]. It should be mentioned that most of these cortical regions are close to each other. But it can also be hypothesized, that the mental syllabary is not located in *one* specific area. This will be proved in a new neuroimaging experiment (i) in order to identify the P-Map, i.e. a brain region where syllable states are represented in a supramodal way and (ii) in order to underpin the hypothesis of the existence of a phonetotopic order (concerning phonetic-phonological and structure) of syllable states within the P-Map (not mentioned in detail in this paper) [23].

References

[1] BAILLY, G.: *Learning to speak: sensory-motor control of speech movements.* Speech Communication, 22:251–267, 1997.

[2] BIRKHOLZ, P.: *3D-Artikulatorische Sprachsynthese.* Logos Verlag, Berlin, 2005.

[3] BOHLAND, J. W. and F. H. GUENTHER: *An fMRI investigation of syllable sequence production.* NeuroImage, 32:821–841, 2006.

[4] BRENDEL, B., I. HERTRICH, M. ERB, A. LINDNER, A. RIECKER, W. GRODD and H. ACKERMANN: *The contribution of mesiofrontal cortex to the preparation and execution of repetitive syllable productions: An fMRI study.* NeuroImage, 50:1219–1230, 2010.

[5] BURTON, M. W., S. L. SMALL and S. E. BLUMSTEIN: *The role of segmentation in phonological processing: An fMRI investigation.* Journal of Cognitive Neuroscience, 12:679–690, 2000.

[6] DÉMONET, J. F., F. CHOLLET, D. RAMSAY, S.AND CARDEBAT, J. L. NESPOULOUS, R. WISE, A. RASCOL and R. FRACKOWIAK: *The anatomy of phonological and semantic processing in normal subjects.* Brain, 115:1753–1768, 1992.

[7] EICKHOFF, S. B., S. HEIM, K. ZILLES and K. AMUNTS: *A systems perspective on the effective connectivity of overt speech production.* pyiosophical Transactions of the Royal Society, 376:2399–2421, 2009.

[8] GHOSH, S. S.: *A Neuroimaging Study of Premotor Lateralization and Cerebellar Involvement in the Production of Phonemes and Syllables.* Journal of Speech, Language, and Hearing Research, 51:1183–1202, 2008.

[9] GUENTHER, F.: *A neural network model of speech acquisition and motor equivalent speech production.* Biological Cybernetics, 72:43–53, 1994.

[10] GUENTHER, F. H.: *Cortical interactions underlying the production of speech sounds.* Communication Disorders, 39:350–365, 2006.

[11] GUENTHER, F. H., A. NIETO-CASTANON and S. S. GHOSH: *Representation of Sound Categories in Auditory Cortical Map.* Journal of Speech, Language and Hearing Research, 47:46–57, 2004.

[12] HEIM, S., EICKHOFF, S. B. and K. AMUNTS: *Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modeling.* NeuroImage, 48:616–624, 2009.

[13] HEIM, S. and A. FRIEDERICI: *Phonological processing in language production: time course of brain activity.* Neuroreport, 14:2031–2033, 2003.

[14] HICKOK, G. and D. POEPPEL: *Towards a functional neuroanatomy of speech perception.* Trends in Cognitive Sciences, 4:131–138, 2007.

[15] HILLIS, A. E., M. WORK, P. B. BARKER, M. A. JACOBS, E. L. BREESE and K. MAURER: *Re-examining the brain regions crucial for orchestrating speech articulation.* Brain, 127:1479–1487, 2004.

[16] INDEFREY, P. and W. J. M. LEVELT: *The spatial and temporal signatures of word production components.* Cognition, 15:101–144, 2004.

[17] KANDEL, E. R., J. H. SCHWARTZ and T. M. JESSELL: *Principles of Neural Science. 4th edition.* McGraw, New York, 2000.

[18] KOHONEN, T.: *Self-organizing maps.* Springer, Berlin, 2001.

[19] KRÖGER, B. J. and P. BIRKHOLZ: *A gestural-based concept for speech movement control in articulatory speech synthesis.* In ESPOSITO, A., M. FAUNDEZ-ZANUY, E. KELLER and M. MARINO (eds.): *Verbal and Nonverbal Communication Behaviours.*, Berlin, 2007. Springer.

[20] KRÖGER, B. J., P. BIRKHOLZ, J. KANNAMPUZHA and C. NEUSCHAEFER-RUBE: *Towards the acquisition of a sensorimotor vocal tract action repository within a neural model of speech processing.* In ESPOSITO, A., A. VINCIARELLI, K. VICSI, C. PELACHAUD and A. NIJHOLT (eds.): *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issues*, Berlin, in press. Springer.

[21] KRÖGER, B. J., P. BIRKHOLZ and C. NEUSCHAEFER-RUBE: *Towards an Articulation-Based Developmental Robotics Approach for Word Processing in Face-to-Face Communication.* PALADYN Journal of Behavioral Robotics, in press.

[22] KRÖGER, B. J. and S. HEIM: *Mapping of functions to brain regions: a neuro-phonetic model of speech production, perception, and acquisition.* Faits de Langues, 37:165–174, 2011.

[23] KRÖGER, B. J., J. KANNAMPUZHA, A. LOWIT and C. NEUSCHAEFER-RUBE: *Phonetotopy within a neurocomputational model of speech production and speech acquisition.* In FUCHS, S., H. LOEVENBRUCK, D. PAPE and P. PERRIER (eds.): *Some aspects of speech and the brain*, Berlin, 2009. Peter Lang.

[24] KRÖGER, B. J., J. KANNAMPUZHA and C. NEUSCHAEFER-RUBE: *Towards a neurocomputational model of speech production and perception.* Speech Communication, 51:793–809, 2009.

[25] LAU, E. F., P. C. and D. POEPPEL: *A cortical network for semantics: (de)constructing the N400.* Nature Reviews Neuroscience, 9:920–933, 2008.

[26] PATTERSON, K., P. J. NESTOR and T. T. ROGERS: *Where do you know what you know? The representation of semantic knowledge in the human brain.* Nature, 8:976–987, 2007.

[27] RIECKER, A., J. KASSUBEK, K. GRÖSCHEL, W. GRODD and H. ACKERMANN: *fMRI reveals two distinct cerebral networks subserving speech motor control.* Neurology, 64:700–706, 2005.

[28] RITTER, H. and T. KOHONEN: *Self-Organizing Semantic Maps.* Biological cybernetics, 61:241–254, 1989.

[29] TOURVILLE, J. A., K. J. REILLY and F. H. GUENTHER: *Neural mechanisms underlying auditory feedback control of speech.* NeuroImage, 39:1429–1443, 2008.

[30] ZATORRE, R. J., E. MEYER, A. GJEDDE and A. C. EVANS: *PET studies of phonetic processing in speech: Review, replication, and reanalysis.* Cerebral Cortex, 6:21–30, 1996.