Behavioral/Cognitive

Neurocomputational Consequences of Evolutionary Connectivity Changes in Perisylvian Language Cortex

Malte R. Schomers,[1,2] Max Garagnani,[1,3,4] and Friedemann Pulvermüller[1,2]

[1]Brain Language Laboratory, Freie Universität Berlin, 14195 Berlin, Germany, [2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, 10099 Berlin, Germany, [3]Centre for Robotics and Neural Systems, University of Plymouth, Plymouth PL4 8AA, United Kingdom, and [4]Department of Computing, Goldsmiths, University of London, London SE14 6NW, United Kingdom

The human brain sets itself apart from that of its primate relatives by specific neuroanatomical features, especially the strong linkage of left perisylvian language areas (frontal and temporal cortex) by way of the arcuate fasciculus (AF). AF connectivity has been shown to correlate with verbal working memory—a specifically human trait providing the foundation for language abilities— but a mechanistic explanation of any related causal link between anatomical structure and cognitive function is still missing. Here, we provide a possible explanation and link, by using neurocomputational simulations in neuroanatomically structured models of the perisylvian language cortex. We compare networks mimicking key features of cortical connectivity in monkeys and humans, specifically the presence of relatively stronger higher-order "jumping links" between nonadjacent perisylvian cortical areas in the latter, and demonstrate that the emergence of working memory for syllables and word forms is a functional consequence of this structural evolutionary change. We also show that a mere increase of learning time is not sufficient, but that this specific structural feature, which entails higher connectivity degree of relevant areas and shorter sensorimotor path length, is crucial. These results offer a better understanding of specifically human anatomical features underlying the language faculty and their evolutionary selection advantage.

Key words: action–perception cycle; arcuate fasciculus; cortical connectivity; neurocomputational modeling; perisylvian cortex; verbal working memory

---

Significance Statement

Why do humans have superior language abilities compared to primates? Recently, a uniquely human neuroanatomical feature has been demonstrated in the strength of the arcuate fasciculus (AF), a fiber pathway interlinking the left-hemispheric language areas. Although AF anatomy has been related to linguistic skills, an explanation of how this fiber bundle may support language abilities is still missing. We use neuroanatomically structured computational models to investigate the consequences of evolutionary changes in language area connectivity and demonstrate that the human-specific higher connectivity degree and comparatively shorter sensorimotor path length implicated by the AF entail emergence of verbal working memory, a prerequisite for language learning. These results offer a better understanding of specifically human anatomical features for language and their evolutionary selection advantage.

Introduction

One of the key questions about human nature addresses the brain mechanisms underlying the language faculty. In sharp contrast to their closest relatives, humans learn novel words effortlessly and extremely rapidly (Shtyrov et al., 2010; Kimppa et al., 2015) and build vocabularies of tens of thousands of words (Pinker, 1994; Brysbaert et al., 2016), which can also be stored in verbal working memory (VWM). We here ask which neural mechanisms and features of brain–structural connectivity might enable these uniquely human abilities.

Correspondence should be addressed to Malte R. Schomers, Brain Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: malte2011@cantab.net.

Comparative neuroanatomical investigations using diffusion tensor imaging (DTI) and diffusion weighted imaging (DWI) along with invasive tracer studies in nonhuman primates have greatly advanced the search for the specific structural features of the human brain. Lesion evidence shows that inferior frontal [including Brodmann areas (BAs) 44/45] and superior temporal areas (BAs 42/22) of the left perisylvian cortex are most crucial for language, as lesions therein lead to aphasias involving both language production and comprehension (Bates et al., 2003). These core language areas are connected by a dorsal fiber bundle, the arcuate fasciculus (AF; Schüz and Braitenberg, 2002), providing a bidirectional link (Matsumoto et al., 2004). Whereas the ventral connections between these areas do not seem to have changed massively in primate evolution, this dorsal bundle via the AF is rich and strong in humans (Rilling et al., 2008), available already shortly after birth, and strongly lateralized to the left hemisphere (Dubois et al., 2009, 2014)—the language-dominant hemisphere in most people. Invasive tracing studies of macaque brains revealed a similar dorsal link between temporal parabelt and prefrontal areas (Petrides and Pandya, 2009), but parallel DTI/DWI and tractography in humans and macaques indicate relatively richer direct connections between inferior prefrontal and temporal parabelt areas in humans (Rilling et al., 2012). In addition to this quantitative statement, specific qualitative differences appear to be present within the AF, where some area-specific long-distance connections seem to have strengthened massively or may even have newly emerged in the evolution from macaque and chimpanzee to human. Whereas comparative neuroanatomical DTI studies show connections between prefrontal cortex and temporal areas in the auditory parabelt in both humans and monkeys (Thiebaut de Schotten et al., 2012, their Fig. 3), the additional links between prefrontal and auditory belt and between premotor and auditory parabelt areas are well documented with DTI/DWI in humans but not so in macaques or chimpanzees (Fig. 1A; Rilling et al., 2012; Thiebaut de Schotten et al., 2012); as these connections introduce shortcuts to what can be described as a 5-step next-neighbor architecture (Fig. 1C,D), we call them "jumping links." Although not implying a complete absence of jumping links in nonhuman primates (Romanski et al., 1999; Smiley et al., 2007; Scott et al., 2015), the DTI-documented evolutionary change in dorsal connectivity leads to a shorter path length (defined as minimal number of synaptic steps) of strong links between auditory and articulatory motor areas. The AF appears crucially important for language, not only because of this evolutionary change, but also because its strength correlates with numerous human language abilities (Yeatman et al., 2011; López-Barroso et al., 2013; Saygin et al., 2013). However, a neuromechanistic explanation for why, among other factors, the quantitative topological differences in connectivity may be vital for the emergence of human-like language is still missing.

We here address this question using a novel approach of neurocomputational modeling, which has key advantages over both comparative studies and correlational evidence linking AF strength to language abilities. In those studies, a range of alternative features also distinguishing between monkey and human brains (including cortical area size and fiber diameters) could partly explain the observed performance differences. In contrast, models can be specifically designed to differ only in their connectivity structure, so that any functional change between them allows for definitive causal conclusions. We asked whether word learning or VWM abilities of humans could be causally linked to the presence of relatively stronger

jumping links in human perisylvian cortex, as suggested by DTI/DWI data.

Materials and Methods

Network structure and function

We used a neurocomputational model of the perisylvian language cortex. These networks were composed of graded response cells thought to represent the average activity of a local pool of neurons (Eggert and Van Hemmen, 2000). Networks were subdivided into model areas of $25 \times 25 = 625$ excitatory and the same number of inhibitory neurons each (Fig. 1E). One model area was established for each of the following perisylvian areas (Fig. 1B; Garagnani et al., 2008): superior-temporal primary auditory cortex (A1), auditory belt (AB), and parabelt (PB) areas and inferior-frontal articulatory motor (M1), premotor (PM) and prefrontal (PF) cortex. Adjacent areas in all models were connected, based on reciprocal links documented between the corresponding brain areas [Fig. 1B, green-colored arrows (e.g., A1 to AB, AB to PB); Pandya and Yeterian, 1985; Braitenberg and Pulvermüller, 1992; Pandya, 1994; Young et al., 1994, 1995; Kaas and Hackett, 2000; Rauschecker and Tian, 2000].

As outlined in the Introduction, the rationale for this study was to investigate the functional consequence of qualitative and quantitative differences in connectivity between temporal and frontal regions along the dorsal stream. We therefore implemented two model architectures, a monkey architecture (MA) and human architecture (HA). In creating these architectures, we focused on major differences in the connectivity structure between monkey and human perisylvian regions that have been suggested by DTI/DWI-based tractography. This method currently offers the only prospect for comparative neuroanatomy of cortical long-distance connectivity, as invasive tracer studies are not possible in humans. We did not aim at modeling the full complexity of the connectivity structures in each species, because even tractography data of exceptionally high quality are not as accurate as neuroanatomical tracing data (Thomas et al., 2014) and therefore may not allow one to uncover all functionally relevant links in a given species. However, DTI/DWI tractography studies converge on showing stronger left frontotemporal connections in humans compared with nonhuman primates and more specifically the unique presence of strong jumping links (see Introduction). Therefore, we focus on modeling these differences between species, rather than complete architectures. Whereas the MA included only next-neighbor connections between adjacent areas, the HA included additional jumping links (Fig. 1B–D, purple). The strengths of all links were identical.

In addition to the between-area connectivity, which differed between the network architectures, both architectures were designed so as to mimic a range of biologically realistic properties and therefore included the following features: (1) within-area connectivity, which was random, sparse (thus realizing only a small fraction of all possible connections), patchy, and topographic (Gilbert and Wiesel, 1983; Amir et al., 1993; Braitenberg and Schüz, 1998), and such that local connection probability fell off with distance (Braitenberg and Schüz, 1998; Perin et al., 2011); (2) local and area-specific inhibition mechanisms (Fig. 1E, caption; Palm, 1982; Bibbig et al., 1995; Wennekers et al., 2006), which act as a means to regulate and control activity in the network (Braitenberg, 1978; Palm, 1982; Garagnani et al., 2008); (3) synaptic modification by way of Hebb-type learning including both long-term potentiation (LTP) and long-term depression (LTD; Artola et al., 1990; Artola and Singer, 1993); and (4) constant presence of uniform, uncorrelated white noise during all phases of learning and retrieval in all parts of the network (Rolls and Deco, 2010).

The implementation of the computational model follows that used in previous publications (Garagnani et al., 2008, 2009; Garagnani and Pulvermüller, 2013; Pulvermüller and Garagnani, 2014). Details about the underlying computations are also given in the section on Full model specification.

Simulation procedures

Simulations consisted of the following two phases: the learning phase and the testing phase. Twelve pairs of network instances were built, with each
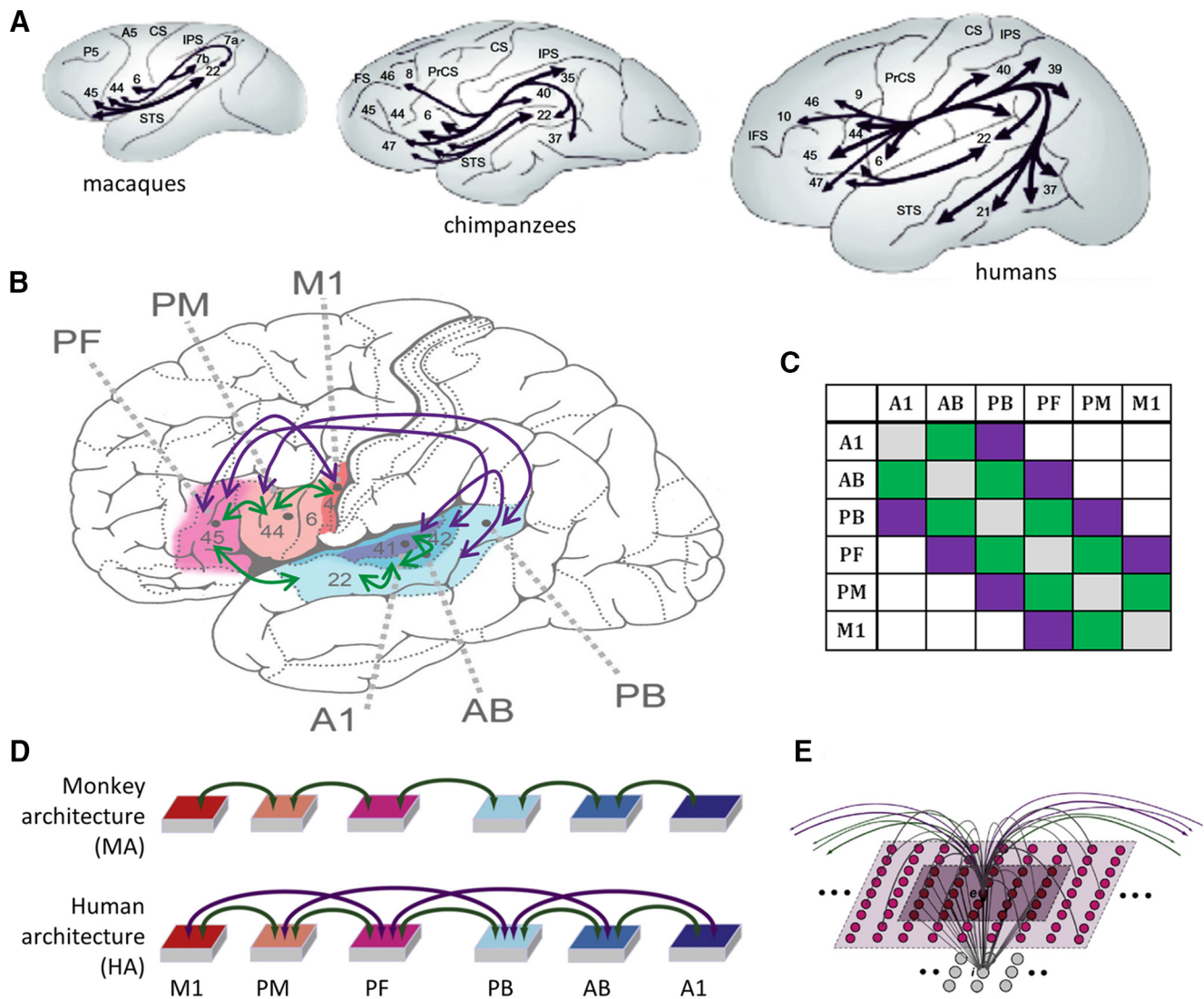
Figure 1. *A*, Illustration of perisylvian connectivity structure in macaques, chimpanzees, and humans as revealed by tractography studies [adapted by permission from Macmillan Publishers Ltd: Nature Neuroscience (Rilling et al., 2008), copyright 2008]. Note the strong frontotemporal connectivity of the latter, especially through the dorsal AF curving around the sylvian fissure, and the presence of ventral connections in both. *B*, A human brain schematic is used to illustrate the area subdivision of the primate frontotemporal perisylvian cortex into M1, PM, and PF, and A1, AB, and PB areas (Garagnani et al., 2008). Green arrows give the connections available in both the human and monkey architecture (HA, MA); purple arrows give connections unique to the human architecture. The purple arrows present only in the HA are meant to reflect the additional connectivity strength available only in humans, as shown by comparative DTI/DWI studies (see main text for detailed discussion). *C*, Connectivity matrix schematizing the connections according to next-neighbor (green) and indirect, jumping links (purple) skipping one intermediate area. *D*, Schematic depiction of the neural network architectures, equivalent to *B*. *E*, Microstructure of the connectivity of one single excitatory cell (labeled "e"). Local (lateral) inhibition is implemented by an underlying cell "i" (representing a cluster of inhibitory interneurons situated within the same cortical column), which receives excitatory input from all cells situated within a local (5 × 5) neighborhood (dark-colored area) and projects back to e, inhibiting it. Within-area sparse excitatory links (in gray) to and from e are limited to a (19 × 19) neighborhood (light-colored area); between-area excitatory projections (green and purple arcs) are topographic and target 19 × 19 neighborhoods in other areas (not depicted). Panel *B* has been adapted from Garagnani and Pulvermüller (2013); panels *D* and *E* have been adapted from Cortex, 57, Pulvermüller, F. and Garagnani, M., "From sensorimotor learning to memory cells in prefrontal and temporal association cortex: A neurocomputational study of disembodiment", pp. 1–21, copyright 2014, with permission from Elsevier.

pair consisting of one MA and one HA network (i.e., 24 networks in total). In each instance, we first initialized an HA network, which entailed (1) randomizing all synaptic links (and corresponding weights) between cells in neighboring areas (and within areas) and (2) randomly generating 14 sensorimotor patterns ("words") to be used during training. Following this HA initialization, the network was copied, preserving the initial random links and the set of to-be-learned patterns, and the additional jumping links (Fig. 1C, purple connections) were removed, resulting in an initialized MA network. Both network architectures were then trained separately but in exactly the same way (see below).

While each of the 12 different pairs of network instances had its own initial randomization of synaptic links and its own set of to-be-learned patterns, these features were identical to both pair members. Thus, there

was some degree of 'between-subject' variability among the 12 network instances (because of randomly generated patterns and weight initializations for each pair), but there was parallelism with respect to these features between the two instances of each MA–HA pair. This ensured that the only difference between each MA–HA pair was in their long-distance connectivity—our variable of interest—while keeping all other factors identical. One may see this as the simulation of the same brain, once with human and once with monkey architecture, and thus as a "within-subject" manipulation.

Training phase

The 14 different acoustic–articulatory patterns were generated for each pair of network instances including 17 specific cells in A1 and another 17

in M1, equaling 2.72% of the neurons in each respective 25 × 25 area. These neurons were thought to represent abstract articulatory and acoustic phonological features (including articulatory and acoustic phonological distinctive features and coarticulatory information) about spoken word forms. The selection of neurons was random and (again) identical between HA–MA pairs. When producing a spoken word form, specific articulatory movements yield acoustic signals, which, in turn, stimulate, with only minimal delay, the auditory system. To model this undeniable correlation of sensorimotor neuronal activity related to speech, which also receives support from recent electrocorticography studies (Cheung et al., 2016; Leonard et al., 2016), stimulus patterns were presented to the sensory and motor areas A1 and M1 networks. By "presenting a stimulus pattern," we mean that its 2 × 17 cells were activated together for 16 time steps. We wanted to avoid any possibly contaminating activity related to the previously presented stimulus pattern, and hence an interstimulus interval (ISI) followed each stimulus presentation. This ISI lasted for at least 30 time steps, until network activity had returned to the baseline value. During these ISIs the only input to the network was uniform white noise, simulating the spontaneous baseline neuronal firing observed in real neurons. Note that all parts of the network were subjected to the same amount of noise. The trial-to-trial presentation sequence of the different patterns was random. Hebbian learning was effective throughout learning trials, both during stimulus presentation and ISIs.

After stimulation to M1 and A1, activation spread throughout the model areas. As a consequence of activation spreading and the resultant coactivation of neurons across the network, associative learning led to the formation of circuits interlinking the articulatory and auditory patterns, as documented in several previous studies (Garagnani et al., 2008; Garagnani and Pulvermüller, 2013, 2016; Pulvermüller and Garagnani, 2014; Tomasello et al., 2016). Due to sensorimotor activation, neural activity was present in specific neurons in A1 and M1, which partly activated further neural elements connected to these stimulated ones. Correlated activity and Hebbian learning mechanisms led to synaptic strengthening so that, eventually, sensorimotor stimulation led to increasingly stronger activation spreading to specific neuron sets throughout the network, which finally led to the formation of a distributed long-term memory (LTM) trace, or cell assembly (CA; for a detailed description and analyses of cell assembly formation in this type of network, see Garagnani et al., 2008, 2009; Pulvermüller and Garagnani, 2014).

Testing phase

The functionality of the circuits developing in the HA and MA was then compared in the testing phase, where all previously learned auditory patterns were presented once again, in random order. Auditory stimulation (without articulatory pattern stimulation) was chosen to simulate speech perception. Stimulation was for two time steps; network responses were recorded during stimulation and the 30 subsequent time steps (i.e., 32 time steps in total).

Data analysis

Structural network properties: cell assembly sizes. To assess whether articulatory–acoustic learning led to cell assembly formation, the presence and sizes of these circuits were assessed in each network instance. To identify the neurons forming cell assemblies across the different network areas, the activity of all 3750 excitatory network cells was monitored in response to one specific stimulation pattern. For each area, we calculated the maximum firing rate occurring across all 625 excitatory cells of a given area at any time during the 30 time steps following sensory stimulation. A cell was considered a member of a given cell assembly if and only if at any time step its firing rate reached at least 50% of the firing rate of the maximally responsive cell in the given area at that time step (provided that such maximum firing rate was at least 0.2). These procedures and thresholds were chosen on the basis of simulation results obtained with the present and previous networks (Garagnani et al., 2008, 2009).

Dynamics of network activation. We also analyzed neural dynamics within each area separately in response to learned patterns. To quantify

Table 1. Parameter values used for the simulations

Equations	Parameters
1	Excitatory cells: $\tau = 2.5$ (in simulation time steps)
	Inhibitory cells: $\tau = 5$ (in simulation time steps)
	Scaling factor: $k_1 = 0.01$
	Noise scaling factor (training phase): $k_2 = 15\sqrt{48}$
	Noise scaling factor (testing phase): $k_2 = 5\sqrt{48}$
	Global inhibition strength (training phase): $k_s = 95$
	Global inhibition strength (testing phase): $k_s = 60$
3	Adaptation: $\alpha = 0.026$
4.1	Time constant for computing gliding average of cell activity:
	$\tau_A = 15$ (in simulation time steps)
4.2	$\tau_S = 8$
5	Postsynaptic potential thresholds for LTP: $\theta_+ = 0.15$
	Postsynaptic potential thresholds for LTD: $\theta_- = 0.15$
	Presynaptic output activity required for any synaptic change:
	$\theta_{pre} = 0.05$
	Learning rate: $\Delta w = 0.0007$

differences in activation dynamics, we first calculated, for each area, the time point at which the firing rate was highest (T_{max}). This value was then used to quantify the area-specific duration of sustained activity (which we interpret as a measure of verbal working memory; Fuster and Bressler, 2012), defined as the length of the interval (in simulation time steps) during which activity in an area remained significantly above the pre-stimulation average (≥ 2 SDs of the average firing rate in the 10 time steps immediately before stimulation). We refer to this quantity as the (area-specific) "sustained memory period" (SMP). For both T_{max} and SMP data, we conducted repeated-measures ANOVAs with factors model architecture (MA/HA) and area (six areas), both as within-subjects factor (see section Simulation procedures).

Full model specification [1]

Each model area consists of two layers of 625 excitatory and 625 inhibitory cells (Fig. 1E). Each excitatory cell represents a cluster of cortical neurons (pyramidal cells), and the underlying inhibitory cell models the cluster of inhibitory interneurons situated within the same cortical column (Wilson and Cowan, 1972; Eggert and Van Hemmen, 2000). The state of each cell x is uniquely defined by its membrane potential $V(x, t)$, representing the average of the sum of all (excitatory and inhibitory) postsynaptic potentials acting upon neural pool (cluster) x at time t, and governed by the following equation (see Table 1 for the parameter values used):

$$\tau \cdot \frac{dV(x, t)}{dt} = -V(x, t) + k_1(V_{In}(x, t) + k_2\eta(x, t)), \quad (1)$$

where $V_{In}(x, t)$ is the net input to cell x at time t (sum of all IPSPs and EPSPs; inhibitory synapses are given a negative sign), τ is the time constant of the membrane, k_1, k_2 are scaling constants, and $\eta(x, t)$ is a white noise process with uniform distribution over $[-0.5, 0.5]$. Time is in arbitrary units. Cells produce a graded response that represents the average firing rate of the neuronal cluster; in particular, the output (transformation function) of an excitatory cell x at time t is as follows:

$$O(x, t) = \begin{cases} 0 & \text{if } V(x, t) \leq \varphi \\ (V(x, t) - \varphi) & \text{if } 0 < (V(x, t) - \varphi) \leq 1, \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

where $O(x, t)$ represents the average (graded) firing rate (number of action potentials per time unit) of cluster x at time t; it is a piecewise linear sigmoid function of the cell membrane potential $V(x, t)$, clipped into the range $[0, 1]$ and with slope 1 between the lower and upper thresholds φ and $\varphi + 1$. The output $O(x, t)$ of an inhibitory cell is 0 if $V(x, t) < 0$, and $V(x, t)$ otherwise. In excitatory cells, the value of the threshold φ in Equation 2 varies in time, tracking the recent mean activity of the cell so

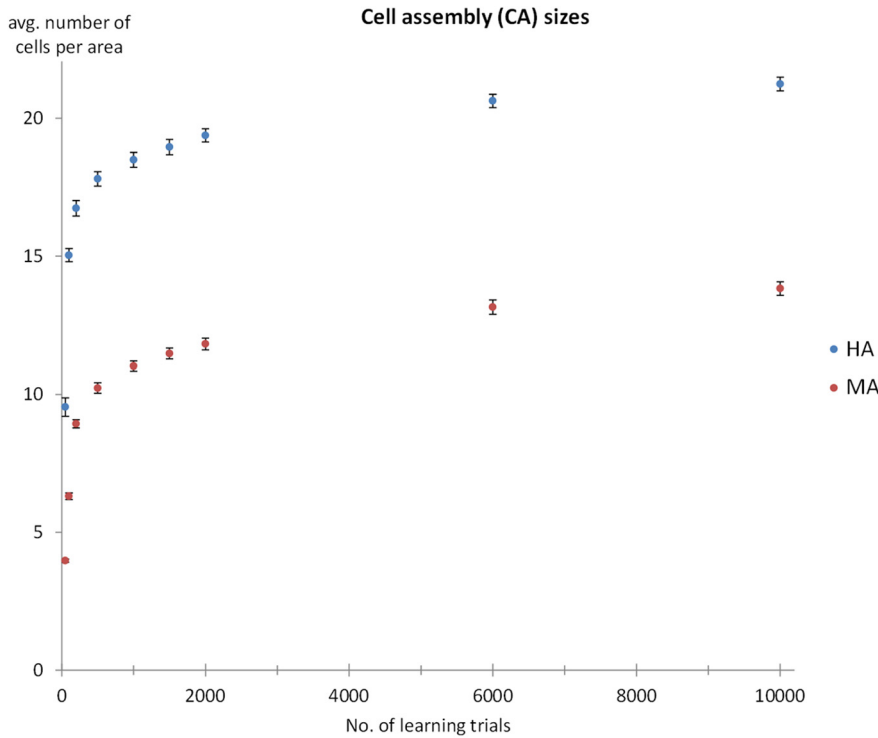[1]This has been adapted from Garagnani and Pulvermüller (2013).

Figure 2. Cell assembly sizes (number of cells in CA) as a function of the number of learning trials. Data are presented separately for the MA (in red) and the HA (in blue). Each data point represents the average of 12 network instances with 14 patterns per network (N = 168). Error bars show SEM after removing between-network variance (Morey, 2008). Note the asymptotic behavior of both architectures with an increasing number of learning trials.

as to implement a simple version of neuronal adaptation (Kandel et al., 2000; higher activity leads to a higher threshold). More precisely, it is written as follows:

$$\varphi(x, t) = \alpha \cdot \omega(x, t), \qquad (3)$$

where $\omega(x, t)$ is the time average of the recent output of the cell, and α is the "adaptation strength."

For an excitatory cell x, the approximate time average $\omega(x, t)$ of its output $O(x, t)$ is estimated by integrating the linear differential Equation 4.1 below with time constant τ_A, assuming an initial average $\omega(x, 0) = 0$, as follows:

$$\tau_A \cdot \frac{d\omega(x, t)}{dt} = -\omega(x, t) + O(x, t). \qquad (4.1)$$

Local (lateral) inhibitory connections (Fig. 1E) and area-specific inhibition are also implemented, realizing, respectively, local and global competition mechanisms (Duncan, 2006) and preventing activation from falling into nonphysiological states (Braitenberg and Schüz, 1998). More formally, in Equation 1 the input $V_{In}(x, t)$ to each excitatory cell of the same area includes an area-specific ("global") inhibition term $k_S \cdot \omega_S(x, t)$, which is subtracted from the total sum of the IPSP and EPSP postsynaptic potentials V_{In} in input to the cell, with $\omega_S(x, t)$ defined as follows:

$$\tau_s \cdot \frac{d\omega_s(x, t)}{dt} = -\omega_s(x, t) + \sum_{x \in \text{area}} O(x, t). \qquad (4.2)$$

The low-pass dynamics of the cells (Eqs. 1, 2, 4.1, 4.2) are integrated using the Euler scheme with step size Δt, where $\Delta t = 0.5$ (in arbitrary time units).

Excitatory links within and between (possibly nonadjacent) model areas are random and limited to a local (topographic) neighborhood;

weights are initialized at random, in the range [0, 0.1]. The probability of a synapse to be created between any two cells falls off with their distance (Braitenberg and Schüz, 1998) according to a Gaussian function clipped to 0 outside the chosen neighborhood (a square of size $n = 19$ for excitatory cell projections and $n = 5$ for inhibitory cell projections). This produces a sparse, patchy, and topographic connectivity, as typically found in the mammalian cortex (Amir et al., 1993; Kaas, 1997; Braitenberg and Schüz, 1998; Douglas and Martin, 2004).

The Hebbian learning mechanism implemented simulates well documented synaptic plasticity phenomena of LTP and LTD, which are believed to play a key role in experience-dependent plasticity, memory, and learning (Rioult-Pedotti et al., 2000; Malenka and Bear, 2004). In particular, the learning rule is an implementation of the Artola–Bröcher–Singer model of LTP/LTD (Artola et al., 1990; Artola and Singer, 1993). In the model, we discretized the continuous range of possible synaptic efficacy changes into two possible levels, $+\Delta w$ and $-\Delta w$ (with $\Delta w \ll 1$ and fixed). We defined as "active" any link from an excitatory cell x such that the output $O(x, t)$ of cell x at time t is larger than θ_{pre}, where $\theta_{pre} \in \,]0, 1]$ is an arbitrary threshold representing the minimum level of presynaptic activity required for LTP (or LTD) to occur. Thus, given any two cells x and y connected by a synaptic link with weight $w_t(x, y)$, the new weight $w_{t+1}(x, y)$ is calculated as follows:

$$w_{t+1}(x, y)$$

$$= \begin{cases} w_t(x, y) + \Delta w & (LTP) & \text{if } O(x, t) \geq \theta_{pre} \text{ and } V(y, t) \geq \theta_+ \\ w_t(x, y) - \Delta w & (LTD) & \text{if } O(x, t) \geq \theta_{pre} \text{ and } \theta_- \leq V(y, t) < \theta_+ \\ w_t(x, y) - \Delta w & (LTD) & \text{if } O(x, t) < \theta_{pre} \text{ and } V(y, t) \geq \theta_+ \\ w_t(x, y) & (no\ change) & \text{otherwise} \end{cases}$$

$$(A5)$$

Results

Cell assembly sizes

Before analyzing the dynamics of network activation, we computed the resulting cell assembly sizes obtained after representing the sensory part of a previously learned pattern to the model again (for cell assembly definition, see Materials and Methods). We computed CA sizes for 50, 100, 200, 500, 1000, 1500, 2000, 6000, and 10,000 learning trials. The resulting CA sizes are shown in Figure 2. CA sizes were always larger for the HA than the MA (all $p < 0.001$), regardless of the number of learning trials (presentations per pattern). During the first 1000 presentations, the number of CA cells grew at a very fast rate (relative ratios of CA sizes at 1000 versus 50 presentations were 1.82 for HA and 2.87 for MA). Growth rate fell off after 1000 presentations for both types of architectures (relative ratios of CA size ratios at 2000 versus 1000 presentations were 1.04 for HA networks and 1.06 for MA networks). These observations were supported by Tukey's HSD tests, which confirmed that CA sizes differed between 50 and 1000 presentations for both architectures (both $p < 0.001$), but did not significantly change between 1000 and 2000 presentations for either HA or MA networks (HA: $p = 0.17$; MA: $p = 0.14$). In addition, we approximated the derivative of the CA size changes at 500 and 1000 learning trials and found that at 500 time

steps the size increase per additional learning trial (i.e., CA growth rate) was larger for MA (0.3 cells/100 learning trials) than HA (0.2 cells/100 learning trials). However, at 1000 time steps, this growth rate was 0.1 cells/100 learning trials for both architectures. Hence, we assume that at 1000 learning trials both networks had become relatively saturated with respect to learning, such that additional learning trials produced very small increases in CA sizes. We therefore focused further analyses on networks trained to 1000 learning trials.

Dynamics of network activation

Figure 3 shows network dynamics (sum of firing rates as a function of simulation time step) induced by presentation of the sensory part of a previously learned word pattern to area A1 in the MA (top) and HA (bottom). Inspection of these plots reveals three qualitative differences in the dynamics of activation: (1) overall sum of firing rates are higher for the HA than for the MA (in part reflecting larger CA sizes; see Fig. 2); (2) activation is parallel for the HA, with areas AB/PB and PF/PM activating nearly simultaneously; in contrast, in the MA, activation spreads in a serial manner throughout the six areas; and (3) activation persists for a much longer time in the HA than in the MA.

Note that the modeling results of serial versus parallel activation seem to match recent experimental results. Whereas in the auditory system of macaques, "latencies [of auditory-evoked activity sometimes] increase with increasing hierarchical region" (Camalier et al., 2012), a feature that Camalier et al. see as partly "consistent with [serial] anatomical predictions," recordings from humans have been found to be "not supportive of a strict serial model envisioning principal flow of information" along the A1 to parabelt pathway (Nourski et al., 2014) but were supportive of largely parallel auditory area activation instead. This contrast, although coming from methodologically very different studies and only reflecting some aspects of extremely rich datasets, seems consistent with the tendencies toward serial versus parallel processing implicated by our MA and HA models, respectively (Figs. 3, 4A).

To investigate aspects 2 and 3 mentioned above (i.e., the seriality and persistence of activation quantitatively), we used the following measures (separately for each area and network type): the time step at which the maximum firing rate in a given area was reached, T_{max} (Fig. 4A) and the SMP (see Materials and Methods; Fig. 4B).

We also conducted separate ANOVAs on these two measures. For T_{max} data, the ANOVA revealed significant main effects of type ($F_{(1,11)} = 446$, $p < 0.001$) and area [$F_{(5,55)} = 1878$; Greenhouse-Geisser epsilon (GGe) = 0.52; p [GGe] < 0.001] and
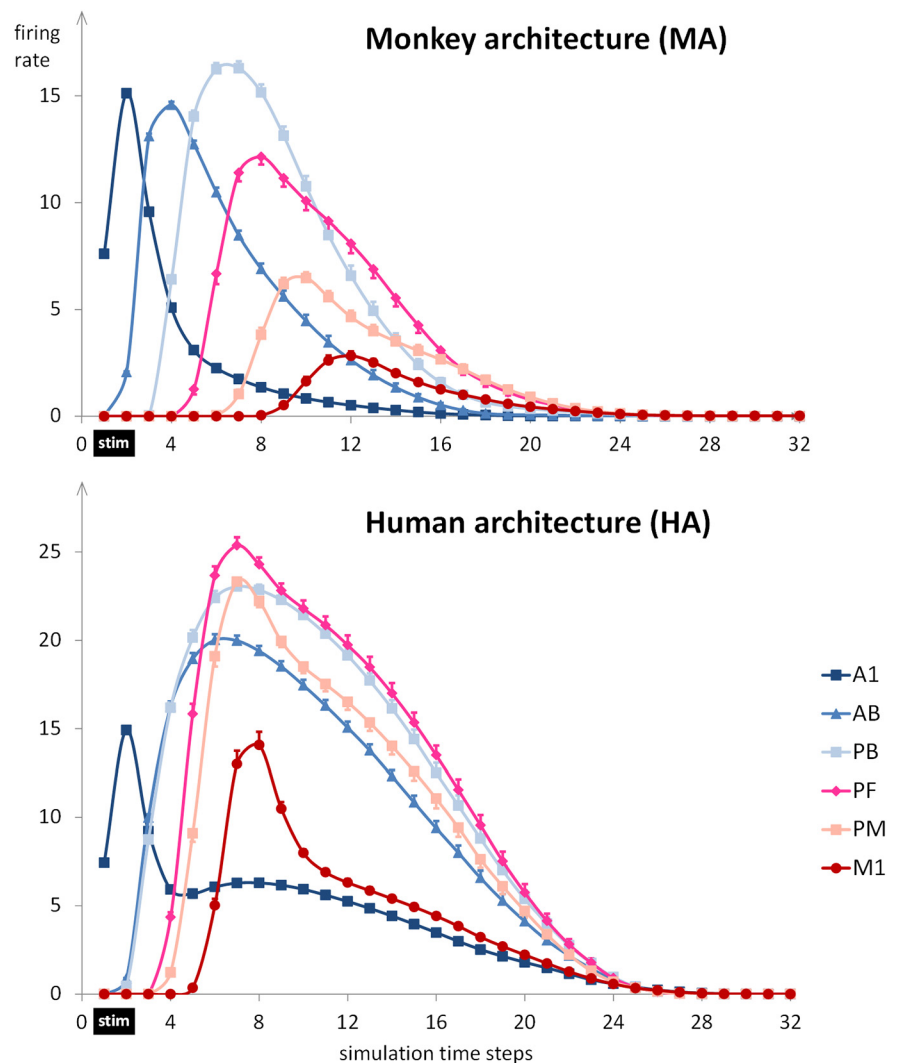

Figure 3. Dynamics of network activation after sensory stimulation. The panels show the sum of firing rates after presenting the sensory components of previously learned patterns to A1. Stimulation was for the first two time steps (marked by a black bar, "stim"), and, following this, firing rates were recorded for 30 time steps. As the sum of firing rates is shown, this measure reflects the total amount of activity in an area rather than average firing rate per cell. Each data point represents the average of 12 network instances with 14 patterns per network ($N = 168$). Error bars show SEM after removing between-network variance (Morey, 2008).

a significant interaction of type × area ($F_{(5,55)} = 466$; GGe = 0.49; p [GGe] < 0.001).

We conducted *post hoc* Tukey's HSD tests comparing, separately for HA and MA, T_{max} for adjacent areas. For the MA, all pairwise comparisons between adjacent areas were significant ($p < 0.001$), confirming the seriality of activation of adjacent areas. In contrast, for the HA, comparisons were not significant between adjacent areas PB and PF (mean difference = 0.19 time steps; $p = 0.98$), between PF and PM (mean difference = 0.18 time steps; $p = 0.99$), between PM and M1 (mean difference = 0.37; $p = 0.37$), or even between the nonadjacent areas PB and PM (mean difference = 0.37; $p = 0.37$). All other comparisons between adjacent areas (A1 and AB, AB and PB) for the human architecture were significant (all $p < 0.001$). This indicates that in the HA, activation initially spread serially from A1 via AB to PB, at which point the remaining areas of PF, PM, and M1 activated nearly simultaneously (Fig. 4A).

For SMP data, the ANOVA revealed significant main effects of type ($F_{(1,11)} = 1388$, $p < 0.001$) and area ($F_{(5,55)} = 186$; GGe =
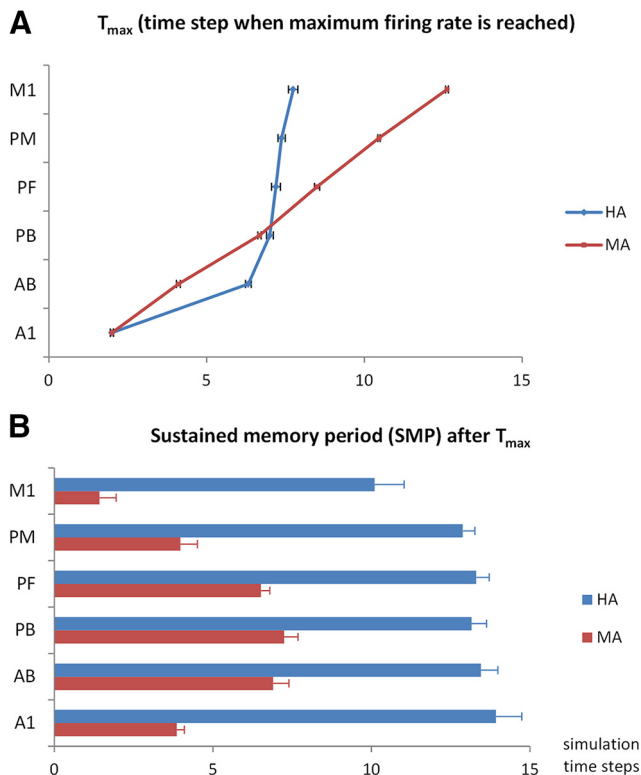
A

T$_{max}$ (time step when maximum firing rate is reached)

B

Sustained memory period (SMP) after T$_{max}$

Figure 4. Quantitative analyses of the dynamics of network activation in the MA (in red) and the HA (in blue). **A**, Time step when the maximum firing rate is reached, T$_{max}$ (within the 30 poststimulation time steps only). Note the serial activation of the MA and the nearly simultaneous "ignition" effect of all areas except A1/AB in the HA. **B**, SMP, defined as the duration (in time steps) starting from T$_{max}$ during which the firing rate remained at ≥2 SDs of the average firing rate of the prestimulation phase. Note the significantly larger SMP values for the HA in all areas. Each data point represents the average of 12 network instances with 14 patterns per network ($N = 168$). Error bars show SEM after removing between-network variance (Morey, 2008). Both **A** and **B** are calculated based on the same data depicted in Figure 3 (prestimulation baseline period not depicted).

0.63; p [GGe] < 0.001) and a significant interaction of type × area ($F_{(5,55)} = 62$; GGe = 0.54; p [GGe] < 0.001).

We conducted *post hoc* Tukey's HSD tests on the SMP data, comparing HA and MA at for each area separately. These comparisons showed that the SMP was significantly larger for the HA than for the MA in all six areas (all p < 0.001).

Repetition of analyses comparing HA 1000 to MA 10,000

Although, as described in the Results section for cell assembly sizes, after 1000 learning trials, both models were at comparable learning stages, we wanted to rule out the possibility that the MA is simply slower in learning but still able to achieve qualitatively similar results. Therefore, we tested whether all statistical analyses obtained by comparing the HA and MA still obtained even when comparing the HA after 1000 learning trials to the MA after 10,000 learning trials. The ANOVAs for T$_{max}$ and SMP data provided the same significance level for all effects. Thus, even when giving the weaker architecture the benefit of a 10-fold increase in learning trials, fundamental differences remained between the dynamics of network activation, and, hence, we can rule this out as a confound.

Discussion

We used a neural network model mimicking fronto-temporal perisylvian language areas, including primary sensorimotor, sec-

ondary, and multimodal brain areas, to simulate word learning and examine network responses to the sensory ("auditory") component of a previously learned pattern, akin to perceiving a familiar spoken word. Crucially, we compared the performances of two types of architectures, MA (monkey architecture) and HA (human architecture), implementing differences in the connectivity of perisylvian areas suggested by DTI/DWI tractography in monkeys/apes and humans. Our results showed the following advantages of the HA over the MA: (1) larger overall size of cell assemblies or action–perception circuits (APCs; Fig. 2), and thus stronger and more robust circuit activation (Fig. 3); (2) parallel rather than serial activation reflecting cell assembly ignition (Figs. 3, 4A); and (3) long-lasting activity in the network, reflecting cell assembly reverberation, and hence, emergence of VWM only in the HA (Figs. 3 and 4B).

Crucially, the disadvantages of the MA could not be compensated by longer training (up to 10,000 learning trials; see Results).

What are the linguistic implications of the observed functional changes?

These results suggest that a change in neuroanatomical connectivity structure emerging in primate evolution underlies the build-up of a functionally robust lexicon of neuronal memory traces for multimodal articulatory-auditory patterns. Although the present simulations did not implement semantics, the emergent neuronal assemblies with long-lasting reverberating activity can be seen as a prerequisite for building a cortical lexicon of meaningful words. As complementary simulation studies show, such semantic learning is possible based on the same mechanisms of correlation mapping as those functional in the current model (Garagnani and Pulvermüller, 2016; Tomasello et al., 2016). In contrast to the large and fast-activating cell assemblies in the HA, the smaller and functionally sluggish circuits in the MA activated in a serial fashion, area by area, and, despite this prolonged activation process, there was little-to-no reverberatory activity (SMP; Fig. 4B). In contrast, the HA yielded longer-persisting activity in its APCs, which we interpret as signifying verbal (phonological) working memory (Fuster and Bressler, 2012; Zipser et al., 1993).

The evolutionary change in neuroanatomical structure may provide a partial explanation for why nonhuman primates have extremely weak auditory working memory, not only compared to humans but also compared to primates' working memory abilities in other sensory modalities (Fritz et al., 2005; Scott et al., 2012, 2014), and why, even after extensive training, nonhuman primates achieve vocabularies of only a fraction of those seen in humans (Savage-Rumbaugh et al., 1993; Call and Tomasello, 2007).

The functional relevance of the motor system for VWM

It is widely agreed that a main function of the arcuate fasciculus is to map acoustic to articulatory representations. In our models, when presenting learned auditory patterns to A1, sustained activation in motor areas (M1 and PM)—those areas most distant from the sensory input—was observed only in the HA, and it occurred earlier than in the MA (Figs. 3, 4B). This motor activity in our model can be viewed as reflecting (subvocal) articulation or rehearsal processes in a "phonological loop" (Baddeley, 2003). Our results thus support the idea that VWM is not subserved by any dedicated module but, rather, consists in reverberating activity between frontal and temporoparietal areas, in line with patient, neuroimaging, and transcranial magnetic stimulation (TMS) evidence demonstrating the importance of speech percep-

tion and production areas in VWM (Belleville et al., 1992; Paulesu et al., 1993; Wilson, 2001; Buchsbaum et al., 2005; Jacquemot and Scott, 2006; Romero et al., 2006; Buchsbaum and D'Esposito, 2008; Acheson et al., 2011; Liao et al., 2014). Hence, an obvious explanation of why AF strength is important for VWM is that the AF enables quick and efficient sensory-to-motor coupling along the dorsal stream for retrieving word form representations when listening and thereby activates bidirectional auditory-to-motor and motor-to-auditory loops for activity maintenance in reverberating working memory circuits (Pulvermüller and Garagnani, 2014).

Individual differences in the degree of motor systems recruitment during speech perception could also contribute to differential working memory abilities. Correlations between verbal working memory performance and speech motor system activations during speech perception have been demonstrated, both in fMRI (Szenkovits et al., 2012) and using motor-evoked potentials (Murakami et al., 2015). Hence, one can speculate that higher verbal working memory abilities are driven by stronger motor systems recruitment, although the existing studies do not allow definite conclusions about the causality of this relationship.

We note that one prediction emerging from the present account is that the producibility of incoming auditory stimuli should influence working memory. Producibility of speech sounds influences the activation of motor areas (Wilson and Iacoboni, 2006) and auditory-to-motor connectivity (Londei et al., 2010) during perception. If this motor activation is also functionally relevant for verbal working memory, then producibility should similarly influence the learning of novel word forms. Indeed, producibility has been shown to influence recognition accuracy in word learning (Schulze et al., 2012). Furthermore, neurophysiological memory traces for newly learned word forms differ depending on whether they exhibit native-like—and hence pronounceable—phonology (Kimppa et al., 2015) and also depending on whether they are actually articulated during learning (Pulvermüller et al., 2012).

The relation between working memory and language learning

Just like large vocabularies, VWM is a unique feature of humans, and even across human individuals there seem to be intrinsic relationships between VWM and language-learning abilities (Baddeley et al., 1988; Baddeley, 1993; Papagno and Vallar, 1995). Furthermore, speech production deficits can lead to reduced vocabulary size, presumably due to impairments in overt or covert repetition of novel pseudowords (Bishop et al., 1990). Finally, the perisylvian areas implicated in articulatory rehearsal have been shown to also be important for word recognition memory by fMRI (Wagner et al., 1998; Davachi et al., 2001; Clark and Wagner, 2003; Paulesu et al., 2009) and non-invasive brain stimulation experiments (Karabanov et al., 2015; Savill et al., 2015).

In essence, current theory and data strongly support that word learning in humans requires and relies on VWM. Human (anterior and posterior) perisylvian cortex provides the substrate for VWM, and the perisylvian dorsal connection by way of the AF plays a crucial role in word learning (López-Barroso et al., 2013), likely in concert with the extreme capsule (López-Barroso et al., 2011). This is not to say that perisylvian connectivity is the only relevant factor, as other structures, notably the hippocampus (Breitenstein et al., 2005; Sederberg et al., 2007) and the amygdala (Ripollés et al., 2014), play important complementary roles in word learning.

What are the critical variables and benefits of the evolutionary network topological change?

Although there is agreement that the human AF is important for language (Wernicke, 1874; Rilling et al., 2008) and experimental evidence demonstrates its importance for verbal working memory (Benson et al., 1973; Damasio and Damasio, 1980; Catani et al., 2007; Rauschecker et al., 2009; Buchsbaum et al., 2011), the precise reason and underlying cortical mechanisms for these structure–function relationships had long remained unclear. Carefully controlled comparison of neural architectures may help here, as these can be exactly parallelized so that any functional difference between architectural "twins" can be uniquely attributed to the one and only manipulated structural feature. In our present case, this specific feature was the implementation of strong corticocortical "jumping link" connections, which, as suggested by comparative neuroanatomical studies using DTI/DWI tractography (see Introduction), may constitute an important structural difference between human and nonhuman primate brains. These connections provide "shortcuts" in the auditory-articulatory pathway in left perisylvian cortex, leading to shorter sensorimotor path length. In general, path length is an important feature of cortical neuroanatomy, which can be used to characterize functionally relevant differences (Kaiser and Hilgetag, 2006; van den Heuvel and Sporns, 2013). Furthermore, as more connections were present in the HA, multiple parallel links became available for projecting acoustic and articulatory phonological information onto each other. Rapid activation flow between articulatory and auditory regions appears necessary for building "actively" reverberating loops, providing the rehearsal mechanism in human verbal working memory, and it is precisely this active memory component that nonhuman primates lack (Scott and Mishkin, 2016). Hence, we propose that these two features taken together—the more numerous connections and their shorter path lengths—are the crucial variables determining the more robust "word representations" and the emergence of verbal (phonological) working memory in humans. Short sensorimotor path length may offer a mechanism not only for verbal working memory, but also for the coupling of auditory and motor information related to speech (for a related computational model, see Westermann and Miranda, 2004). This coupling could also explain why auditory-articulatory interactions are pervasive in speech perception and comprehension (Schomers and Pulvermüller, 2016; Skipper et al., 2017).

Conclusions

Our results suggest that the AF plays a critical role in word learning because its rich connectivity in humans allows for efficient binding of auditory and articulatory information about speech into persistently active neuronal circuits carrying VWM functions. As VWM is necessary for acquiring a vast repertoire of meaningful words, such strongly reverberating circuits may be essential for explaining human language. We believe that the present comparative-neurocomputational research approach may open new and exciting pathways for explanatory evolutionary neuroscience.

References

Acheson DJ, Hamidi M, Binder JR, Postle BR (2011) A common neural substrate for language production and verbal working memory. J Cogn Neurosci 23:1358–1367. CrossRef Medline

Amir Y, Harel M, Malach R (1993) Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. J Comp Neurol 334:19–46. CrossRef Medline

Artola A, Singer W (1993) Long-term depression of excitatory synaptic

transmission and its relationship to long-term potentiation. Trends Neurosci 16:480–487. CrossRef Medline

Artola A, Bröcher S, Singer W (1990) Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. Nature 347:69–72. CrossRef Medline

Baddeley A (1993) Short-term phonological memory and long-term learning: a single case study. Eur J Cogn Psychol 5:129–148. CrossRef

Baddeley A (2003) Working memory: looking back and looking forward. Nat Rev Neurosci 4:829–839. CrossRef Medline

Baddeley A, Papagno C, Vallar G (1988) When long-term learning depends on short-term storage. J Mem Lang 27:586–595. CrossRef

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF (2003) Voxel-based lesion–symptom mapping. Nat Neurosci 6:448–450. CrossRef Medline

Belleville S, Peretz I, Arguin M (1992) Contribution of articulatory rehearsal to short-term memory: evidence from a case of selective disruption. Brain Lang 43:713–746. CrossRef Medline

Benson DF, Sheremata WA, Bouchard R, Segarra JM, Price D, Geschwind N (1973) Conduction aphasia: a clinicopathological study. Arch Neurol 28:339–346. CrossRef Medline

Bibbig A, Wennekers T, Palm G (1995) A neural network model of the cortico-hippocampal interplay and the representation of contexts. Behav Brain Res 66:169–175. CrossRef Medline

Bishop DVM, Brown BB, Robson J (1990) The relationship between phoneme discrimination, speech production, and language comprehension in cerebral-palsied individuals. J Speech Lang Hear Res 33:210–219. CrossRef

Braitenberg V (1978) Cell assemblies in the cerebral cortex. In: Theoretical approaches to complex systems: proceedings, Tübingen, June 11–12, 1977 (Heim R, Palm G), pp 171–188. New York: Springer-Verlag.

Braitenberg V, Pulvermüller F (1992) Entwurf einer neurologischen Theorie der Sprache. Naturwissenschaften 79:103–117. CrossRef Medline

Braitenberg V, Schüz A (1998) Cortex: statistics and geometry of neuronal connectivity. Berlin: Springer.

Breitenstein C, Jansen A, Deppe M, Foerster AF, Sommer J, Wolbers T, Knecht S (2005) Hippocampus activity differentiates good from poor learners of a novel lexicon. Neuroimage 25:958–968. CrossRef Medline

Brysbaert M, Stevens M, Mandera P, Keuleers E (2016) How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. Front Psychol 7:1116. CrossRef Medline

Buchsbaum BR, D'Esposito M (2008) The search for the phonological store: from loop to convolution. J Cogn Neurosci 20:762–778. CrossRef Medline

Buchsbaum BR, Olsen RK, Koch P, Berman KF (2005) Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron 48:687–697. CrossRef Medline

Buchsbaum BR, Baldo J, Okada K, Berman KF, Dronkers N, D'Esposito M, Hickok G (2011) Conduction aphasia, sensory-motor integration, and phonological short-term memory—an aggregate analysis of lesion and fMRI data. Brain Lang 119:119–128. CrossRef Medline

Call JE, Tomasello ME (2007) The gestural communication of apes and monkeys. Mahwah, NY: Erlbaum.

Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. Proc Natl Acad Sci U S A 109:18168–18173. CrossRef Medline

Catani M, Allin MP, Husain M, Pugliese L, Mesulam MM, Murray RM, Jones DK (2007) Symmetries in human brain language pathways correlate with verbal recall. Proc Natl Acad Sci U S A 104:17163–17168. CrossRef Medline

Cheung C, Hamiton LS, Johnson K, Chang EF (2016) The auditory representation of speech sounds in human motor cortex. eLife 5:e12577. CrossRef Medline

Clark D, Wagner AD (2003) Assembling and encoding word representations: fMRI subsequent memory effects implicate a role for phonological control. Neuropsychologia 41:304–317. CrossRef Medline

Damasio H, Damasio AR (1980) The anatomical basis of conduction aphasia. Brain 103:337–350. CrossRef Medline

Davachi L, Maril A, Wagner AD (2001) When keeping in mind supports later bringing to mind: neural markers of phonological rehearsal predict subsequent remembering. J Cogn Neurosci 13:1059–1070. CrossRef Medline

Douglas RJ, Martin KA (2004) Neuronal circuits of the neocortex. Annu Rev Neurosci 27:419–451. CrossRef Medline

Dubois J, Hertz-Pannier L, Cachia A, Mangin JF, Le Bihan D, Dehaene-Lambertz G (2009) Structural asymmetries in the infant language and sensori-motor networks. Cereb Cortex 19:414–423. CrossRef Medline

Dubois J, Dehaene-Lambertz G, Kulikova S, Poupon C, Hüppi PS, Hertz-Pannier L (2014) The early development of brain white matter: a review of imaging studies in fetuses, newborns and infants. Neuroscience 276:48–71. CrossRef Medline

Duncan J (2006) EPS Mid-Career Award 2004: brain mechanisms of attention. Q J Exp Psychol 59:2–27. CrossRef

Eggert J, Van Hemmen JL (2000) Unifying framework for neuronal assembly dynamics. Phys Rev E 61:1855. CrossRef

Fritz J, Mishkin M, Saunders RC (2005) In search of an auditory engram. Proc Natl Acad Sci U S A 102:9359–9364. CrossRef Medline

Fuster JM, Bressler SL (2012) Cognit activation: a mechanism enabling temporal integration in working memory. Trends Cogn Sci 16:207–218. CrossRef Medline

Garagnani M, Pulvermüller F (2013) Neuronal correlates of decisions to speak and act: spontaneous emergence and dynamic topographies in a computational model of frontal and temporal areas. Brain Lang 127:75–85. CrossRef Medline

Garagnani M, Pulvermüller F (2016) Conceptual grounding of language in action and perception: a neurocomputational model of the emergence of category specificity and semantic hubs. Eur J Neurosci 43:721–737. CrossRef Medline

Garagnani M, Wennekers T, Pulvermüller F (2008) A neuroanatomically grounded Hebbian-learning model of attention—language interactions in the human brain. Eur J Neurosci 27:492–513. CrossRef Medline

Garagnani M, Wennekers T, Pulvermüller F (2009) Recruitment and consolidation of cell assemblies for words by way of Hebbian learning and competition in a multi-layer neural network. Cogn Comput 1:160–176. CrossRef

Gilbert CD, Wiesel TN (1983) Clustered intrinsic connections in cat visual cortex. J Neurosci 3:1116–1133. Medline

Jacquemot C, Scott SK (2006) What is the relationship between phonological short-term memory and speech processing? Trends Cogn Sci 10:480–486. CrossRef Medline

Kaas JH (1997) Topographic maps are fundamental to sensory processing. Brain Res Bull 44:107–112. CrossRef Medline

Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A 97:11793–11799. CrossRef Medline

Kaiser M, Hilgetag CC (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. PLoS Comput Biol 2:e95. CrossRef Medline

Kandel ER, Schwartz, JH, Jessel TM (2000) Principles of neural sciences. New York: McGraw-Hill.

Karabanov AN, Paine R, Chao CC, Schulze K, Scott B, Hallett M, Mishkin M (2015) Participation of the classical speech areas in auditory long-term memory. PLoS One 10:e0119472. CrossRef Medline

Kimppa L, Kujala T, Leminen A, Vainio M, Shtyrov Y (2015) Rapid and automatic speech-specific learning mechanism in human neocortex. Neuroimage 118:282–291. CrossRef Medline

Leonard MK, Cai R, Babiak MC, Ren A, Chang EF (2016) The peri-Sylvian cortical network underlying single word repetition revealed by electrocortical stimulation and direct neural recordings. Brain Lang. Advance online publication. Retrieved February 12, 2017. doi:10.1016/j.bandl.2016.06.001. CrossRef Medline

Liao DA, Kronemer SI, Yau JM, Desmond JE, Marvel CL (2014) Motor system contributions to verbal and non-verbal working memory. Front Hum Neurosci 8:753. CrossRef Medline

Londei A, D'Ausilio A, Basso D, Sestieri C, Gratta CD, Romani GL, Belardinelli MO (2010) Sensory-motor brain network connectivity for speech comprehension. Hum Brain Mapp 31:567–580. CrossRef Medline

López-Barroso D, de Diego-Balaguer R, Cunillera T, Camara E, Münte TF, Rodríguez-Fornells A (2011) Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. Cereb Cortex 21:2742–2750. CrossRef Medline

López-Barroso D, Catani M, Ripollés P, Dell'Acqua F, Rodríguez-Fornells A, de Diego-Balaguer R (2013) Word learning is mediated by the left arcu-

ate fasciculus. Proc Natl Acad Sci U S A 110:13168–13173. CrossRef Medline

Malenka RC, Bear MF (2004) LTP and LTD: an embarrassment of riches. Neuron 44:5–21. CrossRef Medline

Matsumoto R, Nair DR, LaPresto E, Najm I, Bingaman W, Shibasaki H, Lüders HO (2004) Functional connectivity in the human language system: a cortico-cortical evoked potential study. Brain 127:2316–2330. CrossRef Medline

Morey RD (2008) Confidence intervals from normalized data: a correction to Cousineau (2005). Tutor Quant Methods Psychol 4:61–64. CrossRef

Murakami T, Kell CA, Restle J, Ugawa Y, Ziemann U (2015) Left dorsal speech stream components and their contribution to phonological processing. J Neurosci 35:1411–1422. CrossRef Medline

Nourski KV, Steinschneider M, McMurray B, Kovach CK, Oya H, Kawasaki H, Howard MA 3rd (2014) Functional organization of human auditory cortex: investigation of response latencies through direct recordings. Neuroimage 101:598–609. CrossRef Medline

Palm G (1982) Neural assemblies. Berlin: Springer.

Pandya DN (1994) Anatomy of the auditory cortex. Rev Neurol (Paris) 151:486–494. Medline

Pandya DN, Yeterian EH (1985) Architecture and connections of cortical association areas. In: Association and auditory cortices (Peters A, Jones EG, eds), pp 3–61. New York: Springer.

Papagno C, Vallar G (1995) Verbal short-term memory and vocabulary learning in polyglots. Q J Exp Psychol 48:98–107. CrossRef

Paulesu E, Frith CD, Frackowiak RS (1993) The neural correlates of the verbal component of working memory. Nature 362:342–345. CrossRef Medline

Paulesu E, Vallar G, Berlingeri M, Signorini M, Vitali P, Burani C, Perani D, Fazio F (2009) Supercalifragilisticexpialidocious: how the brain learns words never heard before. Neuroimage 45:1368–1377. CrossRef Medline

Perin R, Berger TK, Markram H (2011) A synaptic organizing principle for cortical neuronal groups. Proc Natl Acad Sci U S A 108:5419–5424. CrossRef Medline

Petrides M, Pandya DN (2009) Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. PLoS Biol 7:e1000170. CrossRef Medline

Pinker S (1994) The language instinct: the new science of language and mind. London: Penguin UK.

Pulvermüller F, Garagnani M (2014) From sensorimotor learning to memory cells in prefrontal and temporal association cortex: a neurocomputational study of disembodiment. Cortex 57:1–21. CrossRef Medline

Pulvermüller F, Kiff J, Shtyrov Y (2012) Can language-action links explain language laterality? An ERP study of perceptual and articulatory learning of novel pseudowords. Cortex 48:871–881. CrossRef Medline

Rauschecker AM, Deutsch GK, Ben-Shachar M, Schwartzman A, Perry LM, Dougherty RF (2009) Reading impairment in a patient with missing arcuate fasciculus. Neuropsychologia 47:180–194. CrossRef Medline

Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci U S A 97:11800–11806. CrossRef Medline

Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TE (2008) The evolution of the arcuate fasciculus revealed with comparative DTI. Nat Neurosci 11:426–428. CrossRef Medline

Rilling JK, Glasser MF, Jbabdi S, Andersson J, Preuss TM (2012) Continuity, divergence, and the evolution of brain language pathways. Front Evol Neurosci 3:11. CrossRef Medline

Rioult-Pedotti MS, Friedman D, Donoghue JP (2000) Learning-induced LTP in neocortex. Science 290:533–536. CrossRef Medline

Ripollés P, Marco-Pallarés J, Hielscher U, Mestres-Missé A, Tempelmann C, Heinze HJ, Rodríguez-Fornells A, Noesselt T (2014) The role of reward in word learning and its implications for language acquisition. Curr Biol 24:2606–2611. CrossRef Medline

Rolls ET, Deco G (2010) The noisy brain: stochastic dynamics as a principle of brain function. Oxford: Oxford UP.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–1136. CrossRef Medline

Romero L, Walsh V, Papagno C (2006) The neural correlates of phonological short-term memory: a repetitive transcranial magnetic stimulation study. J Cogn Neurosci 18:1147–1155. CrossRef Medline

Savage-Rumbaugh ES, Murphy J, Sevcik RA, Brakke KE, Williams SL, Rumbaugh DM, Bates E (1993) Language comprehension in ape and child. Monogr Soc Res Child Dev 58:1–222. Medline

Savill N, Ashton J, Gugliuzza J, Poole C, Sim Z, Ellis AW, Jefferies E (2015) tDCS to temporoparietal cortex during familiarisation enhances the subsequent phonological coherence of nonwords in immediate serial recall. Cortex 63:132–144. CrossRef Medline

Saygin ZM, Norton ES, Osher DE, Beach SD, Cyr AB, Ozernov-Palchik O, Yendiki A, Fischl B, Gaab N, Gabrieli JDE (2013) Tracking the roots of reading ability: white matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children. J Neurosci 33:13251–13258. CrossRef Medline

Schomers MR, Pulvermüller F (2016) Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. Front Hum Neurosci 10:435. CrossRef Medline

Schulze K, Vargha-Khadem F, Mishkin M (2012) Test of a motor theory of long-term auditory memory. Proc Natl Acad Sci U S A 109:7121–7125. CrossRef Medline

Schüz A, Braitenberg V (2002) The human cortical white matter: quantitative aspects of cortico–cortical long-range connectivity. In: Cortical areas: unity and diversity (Schüz A, Miller R, eds), pp 377–385. London: Taylor & Francis.

Scott BH, Mishkin M (2016) Auditory short-term memory in the primate auditory cortex. Brain Res 1640:264–277. CrossRef Medline

Scott BH, Mishkin M, Yin P (2012) Monkeys have a limited form of short-term memory in audition. Proc Natl Acad Sci U S A 109:12237–12241. CrossRef Medline

Scott BH, Mishkin M, Yin P (2014) Neural correlates of auditory short-term memory in rostral superior temporal cortex. Curr Biol 24:2767–2775. CrossRef Medline

Scott BH, Leccese PA, Saleem KS, Kikuchi Y, Mullarkey MP, Fukushima M, Mishkin M, Saunders RC (2015) Intrinsic connections of the core auditory cortical regions and rostral supratemporal plane in the macaque monkey. Cereb Cortex. Advance online publication. Retrieved February 12, 2017. doi:10.1093/cercor/bhv277. CrossRef Medline

Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, McCarthy DC, Brandt A, Tully MS, Kahana MJ (2007) Hippocampal and neocortical gamma oscillations predict memory formation in humans. Cereb Cortex 17:1190–1196. CrossRef Medline

Shtyrov Y, Nikulin VV, Pulvermüller F (2010) Rapid cortical plasticity underlying novel word learning. J Neurosci 30:16864–16867. CrossRef Medline

Skipper JI, Devlin JT, Lametti DR (2017) The hearing ear is always found close to the speaking tongue: review of the role of the motor system in speech perception. Brain Lang 164:77–105. CrossRef Medline

Smiley JF, Hackett TA, Ulbert I, Karmas G, Lakatos P, Javitt DC, Schroeder CE (2007) Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys. J Comp Neurol 502:894–923. CrossRef Medline

Szenkovits G, Peelle JE, Norris D, Davis MH (2012) Individual differences in premotor and motor recruitment during speech perception. Neuropsychologia 50:1380–1392. CrossRef Medline

Thiebaut de Schotten M, Dell'Acqua F, Valabregue R, Catani M (2012) Monkey to human comparative anatomy of the frontal lobe association tracts. Cortex 48:82–96. CrossRef Medline

Thomas C, Ye FQ, Irfanoglu MO, Modi P, Saleem KS, Leopold DA, Pierpaoli C (2014) Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proc Natl Acad Sci U S A 111:16574–16579. CrossRef Medline

Tomasello R, Garagnani M, Wennekers T, Pulvermüller F (2016) Brain connections of words, perceptions and actions: a neurobiological model of spatio-temporal semantic activation in the human cortex. Neuropsychologia. Advance online publication. Retrieved February 12, 2017. doi:10.1016/j.neuropsychologia.2016.07.004. CrossRef Medline

van den Heuvel MP, Sporns O (2013) Network hubs in the human brain. Trends Cogn Sci 17:683–696. CrossRef Medline

Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998) Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. Science 281:1188–1191. CrossRef Medline

Wennekers T, Garagnani M, Pulvermüller F (2006) Language models

based on Hebbian cell assemblies. J Physiol-Paris 100:16–30. CrossRef Medline

Wernicke C (1874) Der aphasische symptomenkomplex. Eine psychologische studie auf anatomischer basis. Breslau, Germany: Kohn und Weigert.

Westermann G, Reck Miranda E (2004) A new model of sensorimotor coupling in the development of speech. Brain Lang 89:393–400. CrossRef Medline

Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J 12:1–24. CrossRef Medline

Wilson M (2001) The case for sensorimotor coding in working memory. Psychon Bull Rev 8:44–57. CrossRef Medline

Wilson SM, Iacoboni M (2006) Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. Neuroimage 33:316–325. CrossRef Medline

Yeatman JD, Dougherty RF, Rykhlevskaia E, Sherbondy AJ, Deutsch GK, Wandell BA, Ben-Shachar M (2011) Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. J Cogn Neurosci 23:3304–3317. CrossRef Medline

Young MP, Scannell JW, Burns GA, Blakemore C (1994) Analysis of connectivity: neural systems in the cerebral cortex. Rev Neurosci 5:227–250. Medline

Young MP, Scannell JW, Burns G (1995) Analysis of cortical connectivity. Heidelberg: Springer.

Zipser D, Kehoe B, Littlewort G, Fuster J (1993) A spiking network model of short-term active memory. J Neurosci 13:3406–3420. Medline