

# Emergent Jaw Predominance in Vocal Development through Stochastic Optimization

Clément Moulin-Frier\* ‡ \*\*, Jules Brochard\* § \*\*, Freek Stulp\* † ¶, Pierre-Yves Oudeyer\* ||

\* Inria and Ensta ParisTech, France

† German Aerospace Center (DLR), Institute of Robotics and Mechatronics, Wessling, Germany

‡ clement.moulinfrier@gmail.com, now at the SPECS laboratory at the Universitat Pompeu Fabra, Barcelona, Spain

§ brochard.jules@gmail.com, now at the Institut du Cerveau et de la Moelle épinière, Paris, France

¶ mail@freekstulp.net

|| pierre-yves.oudeyer@inria.fr

\*\* The two first authors contributed equally to the paper.

**Abstract**—Infant vocal babbling strongly relies on jaw oscillations, especially at the stage of canonical babbling, which underlies the syllabic structure of world languages. In this paper, we propose, model and analyze an hypothesis to explain this predominance of the jaw in early babbling.

This hypothesis states that general stochastic optimization principles, when applied to learning sensorimotor control, automatically generate ordered babbling stages with a predominant exploration of jaw movements in early stages. The reason is that those movements impact the auditory effects more than other articulators.

In previous computational models, such general principles were shown to selectively freeze and free degrees of freedom in a model reproducing the proximo-distal development observed in infant arm reaching. The contribution of this paper is to show how, using the same methods, we are able to explain such patterns in vocal development.

We present three experiments. The two first ones show that the recruitment order of articulators emerging from stochastic optimization depends on the target sound to be achieved but that on average the jaw is largely chosen as the first recruited articulator. The third experiment analyses in more detail how the emerging recruitment order is shaped by the dynamics of the optimization process.

## I. INTRODUCTION

In the course of early infant vocal development the vocal tract articulators are not used equally [12], [13]. A recruitment structure is displayed where the jaw seems to play a special role [22], paving the way towards an important developmental change around the age of 7 months called *canonical babbling* [26]. Before this period, human infants produce non-speech vocalisations which do not display the syllabic structure specific to human speech. This particular structure appears with canonical babbling, where the infant suddenly and robustly alternate vowels and consonants in its first proto-syllables (typically “babababa”). Although controversial [26], this is often considered as the onset of speech learning. Some vocal articulators are more relevant than others for producing such proto-syllables, where the jaw and the lips seem to play a predominant role [12]. Experimental data shows that in the earliest stages of speech development the articulatory control of the jaw precedes the one of the lips, suggesting that this

developmental pattern influences the pattern of speech sound acquisition [13]. Moreover, the predominance of jaw movements has been considered as an important factor influencing both speech evolution [21] and development [22].

A number of hypotheses, which we discuss in more detail in Section II-B, have been proposed to explain the particular role of the jaw in speech development and in particular in canonical babbling; Warlaumont provides a recent review [44]. Note however that both issues (on the origins of jaw predominance and of canonical babbling) are two distinct ones in speech science, although they are obviously linked (see above). This paper focuses on modeling the emergence of jaw predominance from learning mechanisms. One set of hypotheses poses that the predominance of the jaw in vocalisation arises from other precursor behaviors, such as mastication and ingestion behaviors in the so-called *Frame/Content* theory [21], or communicative orofacial gestures, such as lip-smacking in non-human primates [10]. Other works consider that such a rhythmic behavior is not specific to speech because arm babbling also appears around 6 months [18]. This suggests that these rhythmic patterns could be due to a general brain dynamics reorganization during the first year of life. Another line of work considers that rhythmic jaw movement can be the result of sensorimotor and social learning processes. Warlaumont [41], [43], [42] proposes computational models of syllabic structure emergence based on social or intrinsic reinforcement. Curiosity-driven self-exploration and imitation have also been considered as causal mechanisms to model the emergence of canonical babbling in speech acquisition [25].

Such developmental organization observed in vocal learning has analogs in the developmental organization of the acquisition of other motor skills, such as infant arm reaching which is characterized by the proximo-distal law: when learning to reach visual targets with their hands, infants first explore movements of its proximal joints (shoulder) before exploring more distal joints [7]. In previous work, we have shown that basic stochastic optimization principles allow the reproduction of this proximodistal ordering [38], [37]. This was explained by the fact that stochastic optimization drives automatically

the learner to explore more the degrees of freedom which have maximal impact on the sensory effect at any given point of the learning process, and that these “maximal impact” degrees of freedom were changing as the learning process got the system closer to sensory goals.

The contribution of this paper is to show that the same principles are also able to explain the predominance of the jaw in vocal development. Thus, the hypothesis we propose is that the predominance of the jaw may be (partially) explained by fundamental optimization principles applied in the vocal sensorimotor space.

Our model uses stochastic optimization to reach a number of auditory goals (vocalising different vowels) by searching in the space of articulatory movements. The optimization and trajectory generation algorithms are the same as in [38], [37], except that trajectories now correspond to movements of a simulated vocal tract, rather than a simulated arm. As a side effect of this process, the available degrees of freedom are autonomously frozen or released according to which articulator is the most impactful with respect to the optimization process. For the development of reaching, we previously reproduced proximodistal ordering [38], [37], independently of the goal to be reached. Analogously, we now observe that the jaw is in average the most used articulator, independent of the auditory goals, i.e. the vowel to be vocalised.

The advantage of our hypothesis is that it does not rely on phylogenetic precursors as in the Frame/Content theory. It also does not require curiosity-driven intrinsic motivations or social reinforcement (but either or both these mechanisms could model the processes which generate the auditory goals that are set by the experimenter in this article). The fact that canonical babbling arises so robustly across different groups of infants [26], suggests the existence of several driving forces [44]. We do not suggest that our model is the *only* force. It should rather be considered as a general, domain-independent force, complementary to domain-specific forces.

## II. RELATED WORK

Understanding how sensorimotor learning is autonomously structured through developmental processes is a major goal of developmental sciences. Computational contributions in the field of developmental robotics have in particular focused on the role of exploration strategies in the structuring of learning processes. Indeed, complex sensorimotor behavior acquisition by developmental agents, whatever their biological or mechanical/computational nature, implies dynamical interactions in a complex embodiment with the environment. In such a context, pure random exploration does not provide adequate data to allow an efficient learning [27]. This is due to the high dimensionality of the involved sensorimotor spaces, the non-linearity and redundancy of the sensorimotor mappings and the significant cost in time and resources of performing informative sensorimotor interaction with the environment. These constraints force the agent to develop efficient exploration strategies, resulting in the formation of a structured learning schedule.

The freezing and freeing of degrees of freedoms is an important family of structuring mechanisms, organizing efficiently learning and exploration in animals, and infants in particular [5], [7]. In this section, we now discuss two domains, arm control and vocalisation development, in which these phenomena have been observed in infant development.

### A. Development of Arm Control

When learning to progressively control their arms, infants obey the so-called *proximodistal law*. The infant first starts to learn how to control his shoulders, and later focuses on its elbows, its wrists and finally its fingers. A number of experimental studies characterized this phenomenon. For example, Berthier et al. [7] showed that the development of early reaching in infants [6] follows a proximodistal structure, where infants first learn to reach by freezing the elbow and the hand, while varying shoulder and trunk movements, and then progressively used more distal joints of the elbow and hand. Studies in adult motor skill acquisition showed similar patterning of freezing and freeing of degrees of freedom, applied to the acquisition or racketing skills [35], soccer [16] or skiing [40].

Although neural myelinisation process may guide aspects of this developmental process by progressively impacting neural structures in a proximodistal manner, it has been shown that the proximodistal law can also be explained by general learning and optimization mechanisms [31], [38], [37]. Our previous work has shown that a quite simple stochastic optimization process, allowing a progressive learning of reaching arm movements by the minimization of a cost function, can display an emergent ordering of degree-of-freedom recruitment [38], [37]. Without being pre-coded, such an optimization process naturally recruits in priority proximal joints because they induce a wider range of effects by allowing the movement of the entire arm (e.g. a wider range of reached hand positions) when compared to distal joints.

Therefore, an organism exploring its own sensorimotor abilities has interest in recruiting proximal joints first because such a strategy allows him to rapidly make a reasonable approximation of the range of possible effects. It is actually not surprising that both biology, through the myelination process, and cognition, through exploration strategies, appear to converge to a similar solution (proximodistal release of degrees of freedom) for the same optimization problem: how to efficiently learn the control of a complex motor apparatus.

### B. Vocal Development

During vocal development, the involved articulators (the jaw, the tongue, the lips ...) are not recruited equally [12], [13]. In particular, at seven months canonical babbling appears in a very robust way [26] and the predominance of jaw movements in vocal development has been suggested as an important factor influencing this phenomena [22].

The Frame/Content theory [21], [20] suggests that jaw predominance is due to the role of feeding movement in speech evolution by providing a powerful sound modulation

ability when coupled with phonation. Although controversial (see e.g. [1]), this hypothesis is supported by infant data displaying statistical vowel-consonant associations in line with the theory predictions. A relatively similar hypothesis has been proposed where the considered phylogenetic precursor would be stereotyped communicative orofacial actions like lip-smacking in macaque monkeys [10], [9].

Sensorimotor and social learning processes have also been used to explain rhythmic jaw movement. Warlaumont [41], [43], [42] proposes computational models of syllabic structure emergence based on social or intrinsic reinforcement. The model starts with random vocalisations produced using an articulatory synthesizer, i.e a computer model of the human vocal tract able to synthesize sound from articulatory movements. These random vocalisations mainly result in non-speech sounds. In the first case (social reinforcement), a human subject then listens to these vocalisations and is asked to reinforce them or not according to his own judgment of speech-likeness. In the second case (intrinsic reinforcement), vocalisations are reinforced using an objective salience measure (more salient sounds are reinforced more). When reinforced, a learning rule drives the system towards producing the corresponding articulation more often. In both cases (social and intrinsic), the model converges towards the production of syllabic vocalisations.

Still based on sensorimotor and social learning, some of the authors of the present paper have proposed a model of speech acquisition based on curiosity-driven self-exploration and imitation [25]. Here the agent drives its own vocalisations according to a learning-progress maximization principle: it self-generates auditory goals according to the progress it observes in learning how to achieve them. Such a mechanism has been shown to be highly efficient in learning inverse models in high-dimensional and redundant robotics setups [30], [3], [28], [11] and concretely formalizes concepts of intrinsic motivation described in the psychology literature into algorithmic architectures that can be experimented in computers and robots [32], [4], [29], [2]. A side effect of this exploration strategy is to self-organize developmental pathways where the agent autonomously focuses on tasks of increasing complexity. In particular, this leads to the emergence of canonical babbling where proto-syllabic structures are observed.

In the following sections, we adapt the aforementioned model explaining the proximodistal law of arm development from stochastic optimization mechanisms [38], [37] to the vocal domain. For this aim, we use a computer-simulated vocal tract model, and show that a similar effect can occur in particular conditions, recruiting the jaw in priority due to the wider range of auditory effects it produces, and displaying a coherent ordering of the other motor parameters.

### III. METHODS

This section first describes how vocal articulation and auditory perception are implemented in the proposed model. Then we define the cost function that expresses the aim of learning: to produce a vocalisation acoustically close to a target vowel.

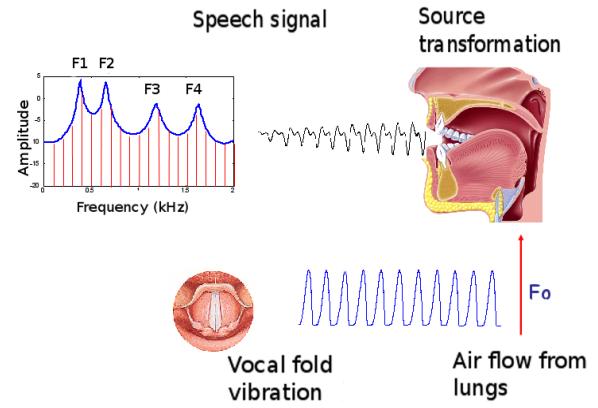


Fig. 1. An articulatory synthesizer is a model of human vocal production. The lung air flow triggers vocal fold vibration, providing a source signal with fundamental frequency  $F_0$ . According to the vocal tract shape, acting as a resonator, the harmonics of the source fundamental frequency are selectively amplified or faded, resulting in a sound wave originating at the lips. Represented in a frequency-amplitude space, this sound wave typically displays a number of amplitude local maxima, called the formants  $F_i$  and ordered from the lowest to the highest frequency. The formants are known to be important auditory features in human speech perception.

Finally we explain the stochastic optimization algorithm used to optimize this cost function.

#### A. Simulating Vocal Production and Perception

We now describe the articulatory synthesizer used, as well as how the motor trajectories that determine the vocal tract's shape over time are generated. We also describe how the output of the synthesizer – peak frequency trajectories in Hertz – is converted into a perceptual scale.

*1) Articulatory synthesis:* Our computational model involves the articulatory synthesizer of the DIVA model described in [14]<sup>1</sup>, based on Maeda's model [23]. It is a computational approximation of the general speech production principles illustrated in Fig. 1.

The model receives 13 articulatory parameters as input. The first 10 are extracted from a principal component analysis (PCA) performed on sagittal contours of images of the vocal tract of a human speaker, allowing the reconstruction of those contours from a 10-dimensional vector. In this study, we only retain the 7 first parameters which have a major impact on the shape of the vocal tract (labeled P1 to P7 on Fig. 2). Since these parameters are derived from a statistical analysis of the vocal tract contours, they do not directly correspond to the biological speech articulators (the jaw, the lips, the tongue etc...). However, the principal components extracted from the PCA relatively match the effect of vocal articulators on the vocal tract shape as it can be seen on Fig. 2: P1 mainly correspond to jaw movements, P2 to tongue vertical

<sup>1</sup> DIVA is available online at <http://www.bu.edu/speechlab/software/diva-source-code>. DIVA is a complete neurocomputational model of speech acquisition, in which we only use the synthesizer computing the articulatory-to-auditory function.

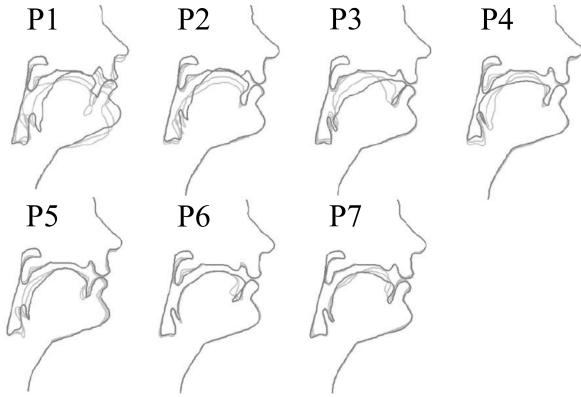


Fig. 2. We use 7 articulatory dimensions to control the vocal tract shape (figure adapted from the DIVA source code documentation). Each subplot shows a sagittal contour of the vocal tract, where we can identify the nose and the lips on the right side. Bold contours correspond to a positive value of the articulatory parameter, the two thin contours are for a null (neutral position) and negative values. These dimensions globally correspond to the degrees of freedom of the human vocal tract articulators. For example, *P*1 mainly controls the jaw height, whereas *P*3 rather controls the tongue front-back position.

movements, *P*3 to tongue horizontal movements, *P*4 to tongue flattening, *P*5 to low movements of the tongue at the level of the throat, *P*6 to horizontal movements of the tongue tip and *P*7 to small rotation movements of the tongue. The 3 last parameters of the 10-dimensional vector having a low influence on the vocal tract are fixed to their neutral position. The effect of these 7 retained articulatory parameters on the vocal tract shape are displayed Fig. 2. Through an area function, associating sections of the vocal tract with their respective areas for a given motor configuration, the model is able to compute the acoustic properties of the resulted signal if phonation occurs. This latter is controlled through 3 others parameters (glottal pressure, vocal cord tension and voicing), that we set at a value assuring normal phonation. We restrain our study to vowel sounds implying a sufficiently open configuration of the vocal tract. The synthesizer is then able to compute the 3 first formants (see Fig. 1) of the signal through the area function.

2) *Motor trajectory generation:* A complete vocalisation corresponds to the time evolution of the vocal tract's shape. We constrain to vocalisations lasting 500ms and control them through the acceleration of the 7 articulatory parameters. Taking inspiration from the Dynamic Movement Primitive framework [17], each acceleration profile is a linear combination of predefined basis functions. Starting from the neutral position of the vocal tract with a null velocity, these accelerations are integrated twice over the 500ms, thus resulting in position trajectories for each articulators, as illustrated on Fig. 3.

The acceleration  $\ddot{q}_{m,t}$  of the  $m$ -th articulator at time  $t$  is determined as a linear combination of basis functions (Equation 1, 2 and 3), where  $\theta_{m,b}$  is the weight of the  $b$ th

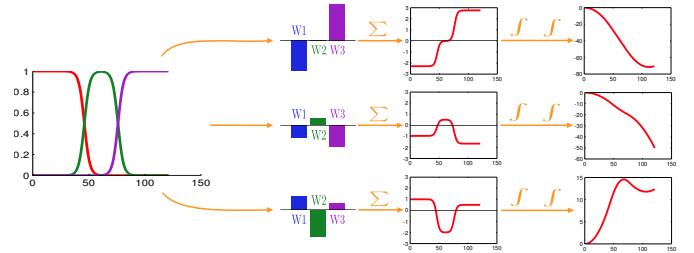


Fig. 3. Examples of basis functions' uses. Left: 3 basis functions (x-axis is time). Center: 3 triplets of weights and the weighted sum. Right: the 3 position trajectories computed from the weighted sum of the basis functions, integrated twice.

basis function of the  $m^{\text{th}}$  articulatory parameter:

$$\ddot{q}_m(t) = \sum_{b=1}^B g_b(t) \theta_{m,b} \quad \text{Acc. of } m^{\text{th}} \text{ articulatory} \quad (1)$$

$$g_b(t) = \frac{\Psi_b(t)}{\sum_{i=1}^B \Psi_i(t)} \quad \text{Normalized basis functions} \quad (2)$$

$$\Psi_b(t) = \exp(-(t - c_b)^2/w^2) \quad \text{Kernel} \quad (3)$$

The centers  $c_{b=1\dots B}$  of the kernels  $\Psi$  are spaced equidistantly in the 500ms duration of the movement, and all have a width of  $w = 50\text{ms}$ . In our experiments we will use  $B = 4$  basis functions for each articulator. Using 7 articulators as stated above, a full motor command is therefore a 28-dimensional vector.

3) *Auditory perception:* The two first formants of the speech signal are known to be the main features of vowel perception [19], [8]. The DIVA synthesizer provides formant trajectories as output, expressed in Hertz. We convert the values in Hertz into a perceptual scale, typically linear at low frequencies and logarithmic at high frequencies. We use the Bark scale proposed by [33] to reflect the psycho-acoustics of speech, defined as follow:

$$\text{Barks} = 7 \times \sinh^{-1} \left( \frac{\text{Hertz}}{650} \right) \quad (4)$$

However, giving the same importance to all formants does not take into account spectral masking phenomena, according to which low-frequency components decrease the perceptual role of higher-frequency components. This led to the proposal [34] that  $F1$  should have typically three times the weight of an “effective second formant”  $F'2$  grouping the roles of  $F2$  and  $F3$ . Following this psycho-acoustic proposition, we modified the sensori space and used only the  $(3F1, F2)$  plane. A similar modeling was used in [24] to successfully explain statistical tendencies in world-language phonological systems. Following this psychoacoustic proposition, we use  $F1$  and  $F2$  as auditory features and in the distance computation in this  $(F1, F2)$  plane we will weight the first dimension with

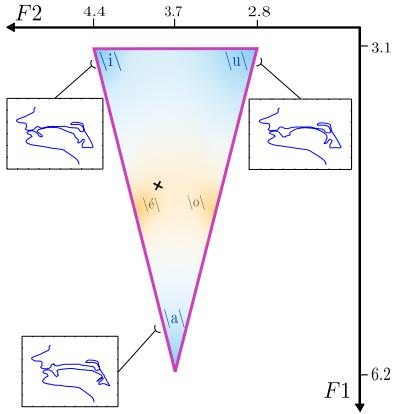


Fig. 4. Variations of  $F1-F2$  (in Barks scale): they take place in what is called the vocalic triangle, ideally represented here with the 5 canonical vowels  $\langle a, e, i, o, u \rangle$ . Vowels  $\langle i, u \rangle$  and  $\langle a \rangle$  are shown in blue together with a possible tract shape for each of them. Vowels  $\langle e \rangle$  and  $\langle o \rangle$  are shown in brown. The cross represents the neutral vocalisation of the synthesizer, i.e. when all the articulators have a null value. Note that the  $F1-F2$  axes are traditionally reversed in speech science to better reflect the relationship between formant values and articulatory configurations: high  $F1$  values correspond to open configurations of the vocal tract (*lowering* the jaw as in  $\langle a \rangle$ , hence at the bottom of the vocalic triangle), whereas high  $F2$  values correspond to a constriction toward the front of the vocal tract as in  $\langle i \rangle$ .

a factor 3. The perception that the vocal agent has of its own vocalisation is the value in Barks of  $F1$  and  $F2$  at the end of the vocalisation. Fig. 4 displays possible values and associated vocal tract shapes.

### B. Cost function

The main purpose of vocalizations in this model is to minimize the difference between a desired vocalization (set by the experimenter) and the actual vocalization (starting from an initial vocalization corresponding to the resting state of the articulators). We express this as a cost function that includes the distance to the goal in the formant-space:  $\|s_g - s\|^2$ , where

$s_g = \begin{pmatrix} 3F1_g \\ F2_g \end{pmatrix}$  is the goal in the sensory formant space and  $s = \begin{pmatrix} 3F1_{t_N} \\ F2_{t_N} \end{pmatrix}$  is formant values actually reached at the end of the vocalisation. Note the factor 3 introduced for  $F1$  values that models the higher weight of this formant as argued above.

Secondary costs that should also be minimized are the energy consumption, and awkward positions far from the rest position. The first is expressed as a penalty on accelerations:  $\sum_{m=1}^7 \sum_{t=1}^T \frac{a_{m,t}^2}{2}$ , where  $a_{m,t}$  is the acceleration of the  $m^{th}$  articulator at the time step  $t$  of the vocalisation. To avoid pathological positions which are far from the resting position we add the term  $\max_m (|P_{m,t_N}|)$ , where  $P_{m,t_N}$  is the position of the  $m^{th}$  articulator at the end of the vocalisation.

Weighting each part according to their impact on the cost, we get the function:

$$J = 10^4 \underbrace{\left\| \begin{pmatrix} 3F1_g \\ F2_g \end{pmatrix} - \begin{pmatrix} 3F1_{t_N} \\ F2_{t_N} \end{pmatrix} \right\|^2}_{\text{Distance to goal}} + \underbrace{\max_m (|P_{m,t_N}|)}_{\text{Avoid awkward positions}} + 10^{-1} \underbrace{\sum_{m=1}^7 \sum_{t=1}^T \frac{a_{m,t}^2}{2}}_{\text{Penalize accelerations}} \quad (5)$$

A low-cost vocalisation is therefore a configuration which approaches the goal with a simple configuration and minimal energy. Note that this cost function is very similar to the one used to learn arm reaching behaviors [37]. The factors  $10^4$  and  $10^{-1}$  have two purposes: 1) a scaling factor to compensate for different range of values the different cost components have, 2) a weighting factor enabling the prioritization of tasks. The order of priorities is: achieve a sound close to the target, achieve end-state comfort, minimize accelerations.

### C. Stochastic optimization

The 28-dimensional vector that determines the movements of the vocal tract (see Section III-A2) defines a search space, whose optimum with respect to the cost function (Section III-B) defines a correct, low-energy vocalization of a particular vowel. In this section, we describe the stochastic optimization algorithm used to find this optimum.

In principle, any stochastic optimization algorithm could be used, as long as it has the following properties: 1) It is model-free, because we assume the learner does not have access to the details of the model described in Section III-A. Such algorithms are known as black-box optimization algorithms, because they make no assumptions about the cost function, i.e. they treat it as a black box. 2) Is able to adapt its exploration along different directions in the search space, because we want to study which parts of the vocal tract are explored most at different phases during learning. 3) Does not contain any assumptions about vocal development.

Examples of algorithms that meet these requirements are natural gradient descent, cross-entropy methods or Covariance Matrix Adaptation – Evolutionary Strategy (CMA-ES) [15]. We use a special case of CMA-ES, which is considered to be state-of-the-art in stochastic optimization, called  $PI^{BB}$  [36]. We do so because  $PI^{BB}$  is even simpler, and sets certain open parameters of CMA-ES to values that are derived from stochastic optimal control [39].

Within the search space,  $PI^{BB}$  represents a Gaussian distribution  $\mathcal{N}(\theta, \Sigma)$ , where the aim is to find  $\theta \equiv \theta_g$ , where  $\theta_g$  is the optimum.  $PI^{BB}$  is an iterative algorithm which consists of three phases during each iteration, which are visualized in Fig. 5.

- **Exploration.** Sample  $K$  parameter vectors  $\theta_k$  from  $\mathcal{N}(\theta, \Sigma)$ , and determine the cost  $J_k$  of each sample. In the visualization of our illustratory example task  $K = 15$ , and the cost  $J(\theta)$  is the distance  $\|\theta\|$  to the center of the

```

input :
 $\theta^{\text{init}}$ ;  $J : \theta \mapsto R$ ;
 $\lambda^{\text{init}}, \lambda^{\min}, \lambda^{\max}$ ; initial parameter vector
 $K$ ;  $h$ ; cost function
 $\Sigma$ ; exploration levels
 $\langle \theta, \Sigma \rangle$ ; roll-outs per update
 $h$ ; eliteness parameter

 $\theta = \theta^{\text{init}}$ 
 $\Sigma = \lambda^{\text{init}} \mathbf{I}$ 
while cost not converged do
    Exploration: sample parameters
    foreach  $k$  in  $K$  do
         $\theta_k \sim \mathcal{N}(\theta, \Sigma)$ 
         $J_k = J(\theta_k)$  cost
    end
     $J^{\min} = \min(\{J_k\}_{k=1}^K)$ 
     $J^{\max} = \max(\{J_k\}_{k=1}^K)$ 
    Evaluation: compute weight for each sample
    foreach  $k$  in  $K$  do
         $P_k = \frac{\exp(-h \frac{J_k - J^{\min}}{J^{\max} - J^{\min}})}{\sum_{l=1}^K \exp(-h \frac{J_l - J^{\min}}{J^{\max} - J^{\min}})}$ 
    end
    Update: Weighted averaging over  $K$  samples
     $\Sigma \leftarrow \sum_{k=1}^K [P_k (\theta_k - \theta)(\theta_k - \theta)^T]$ 
     $\Sigma \leftarrow \text{boundcovar}(\Sigma, \lambda^{\min}, \lambda^{\max})$ 
     $\theta_{\text{new}} \leftarrow \sum_{k=1}^K [P_k \theta_k]$ 
end

```

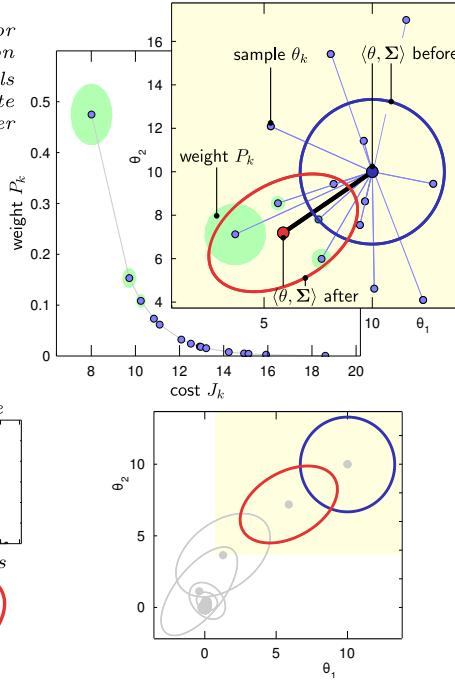


Fig. 5. Illustration of The  $PI^{BB}$  algorithm in pseudo-code (left), as well as a illustratory optimization (right). For simplicity, the search space is 2-dimensional and the cost of a sample  $\theta$  is simply the distance to the origin:  $J(\theta) = \|\theta\|$ . Top-right: Visualization of one parameter update with  $PI^{BB}$ , and the decreasing exponential used to weight the samples. Bottom-right: Evolution of the parameters over several updates, illustrating how the distribution converges towards the minimum  $\theta_g = [0, 0]$ . The algorithm is initialized by setting the mean and covariance parameters  $\langle \theta, \Sigma \rangle$  to  $\theta^{\text{init}}$  and  $\lambda^{\text{init}} \mathbf{I}$  respectively, visualized as a dark blue ellipse. These parameters are updated at each iteration of the main loop. The red ellipse illustrates the first update.

plane. In the Fig. 5 this distance between the blue dots and the center lies approximately between 8 and 19.

- **Evaluation.** Determine the weight  $P_k$  of each sample, given its cost. Essentially, low-cost samples have higher weights, and vice versa. The normalized exponentiation function that maps costs to weights is visualized in the top-right graph. Larger green circles correspond to higher weights.
- **Update.** Update the parameters  $\langle \theta, \Sigma \rangle$  with weighted averaging. In the visualization, the updated parameters are depicted in red. Because low-cost samples (e.g. a cost of 8-10) have higher weights, they contribute more to the update, and  $\theta$  therefore moves in the direction of the optimum  $\theta_g = [0, 0]$ .

The bottom-right inset of Fig. 5 highlights two important properties of  $PI^{BB}$ , which it shares with cross-entropy methods and CMA-ES. First, the mean of the Gaussian distribution iteratively comes closer and closer to the optimum (the origin, in the example). Second, the covariance matrix of the distribution is elongated towards the optimum.  $PI^{BB}$  thus adapts its exploration, such that the amplitude of exploration is greater in the direction of the origin (i.e. the algorithm finds automatically that it is more efficient to try variations of  $\theta$  along this direction at this stage of the optimization process). We define the *exploration magnitude* as the largest eigenvalue of the covariance matrix  $\Sigma$ . Intuitively, this means that the exploration magnitude is measured by the width of the gaussian in its largest direction. Below, this  $PI^{BB}$  optimization

process is applied to the 7 groups of 4 parameters, each corresponding to one of the 7 articulators. Then, for each group the largest eigenvalue of each subset allows to measure the exploration magnitude of each articulator at any given point in the optimization process.

1) *Application of  $PI^{BB}$  to the synthesizer:* Before describing the experiments, Fig. 6 first summarizes how the vocal synthesizer, motor trajectory generation, auditory perception and stochastic optimization work together to learn how to vocalize vowels.

The following steps 1-6 are separately performed  $K = 20$  times.

- 1) For each of the  $m = 7$  vocal parameters, a 4D-vector is sampled from the distribution  $\mathcal{N}(\theta_m, \Sigma_m)$ . Thus, one distribution is used per vocal parameter.
- 2) Each of these 7 4D-vector defines the weights of the basis functions of its corresponding vocal parameters.
- 3) The weights multiplied with the basis functions yield the accelerations for each of the 7 vocal parameters.
- 4) Integrating these accelerations yields reference trajectories for the 7 vocal parameters.
- 5) DIVA is then used to determine the vocalization (the formants in Hz), which are converted to auditory trajectories in the formant space.
- 6) The cost for each of the sample is determined using the cost function  $J$ , see Equation (5).

Finally a step 7 is performed: the distribution  $\mathcal{N}(\theta_m, \Sigma_m)$  is

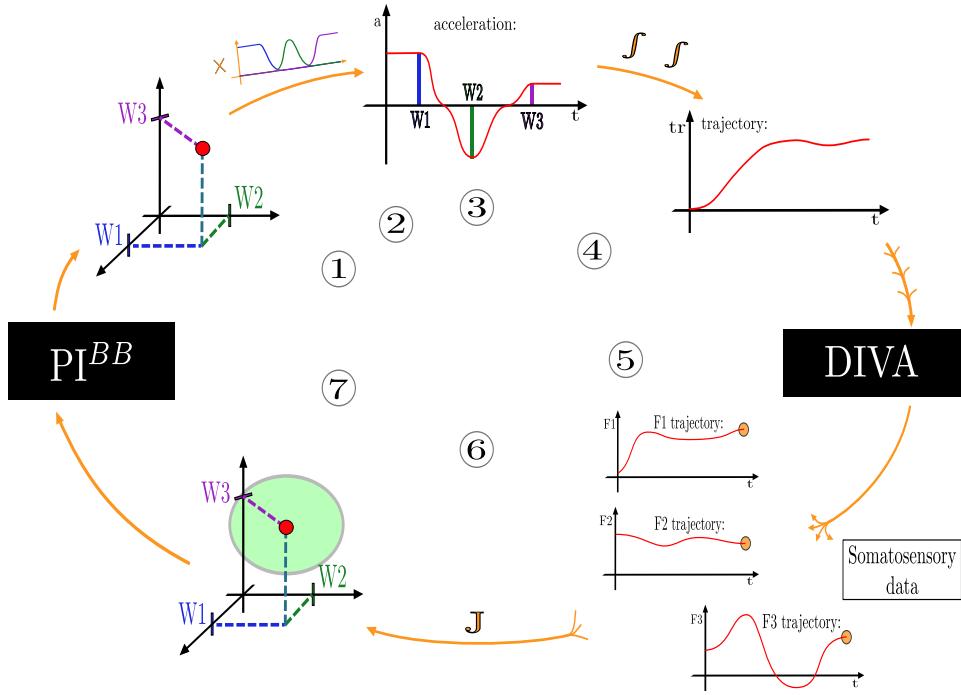


Fig. 6. Illustration of applying  $PI^{BB}$  to vocalization learning.

updated, by using the  $K$  costs  $J$  to weight the  $K$  4D-vectors. This is done separately for each of the 7 vocal parameters.

The steps above continue for 50 updates. After each update, we determine the exploration magnitude in vocal parameter  $m$  by determining the largest eigenvalue  $\lambda_m$  of the covariance matrix  $\Sigma_m$ . Large values indicate that a lot of exploration is occurring within this vocal parameter.

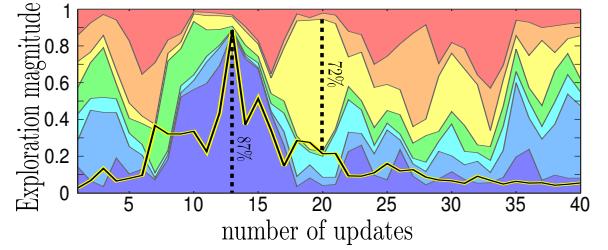
Fig. 7 illustrates one experiment, in which  $PI^{BB}$  learns to vocalize the vowel \a/ within 50 updates. This graph illustrates the relative exploration magnitude within each of the 7 vocal parameters as learning progresses, i.e.  $\lambda_m^{\text{rel.}} = \frac{\lambda_m}{\sum_{i=1}^7 \lambda_i}$ . Thus, we see that initially the relative exploration magnitudes are the same, but between updates 5-20, most of the exploration happens in the first vocal parameter  $P1$  (dark blue), which mostly corresponds to jaw movements.

We interpret the relative exploration magnitude to be an indication of the relative freezing and freeing of degrees of freedom. For instance, the first vocal parameter  $P1$  (dark blue, mostly controlling the jaw) is performing 89% of the relative exploration at update 12. Thus, this degree of freedom is freed, whereas the others are frozen. At update 26, the  $P5$  (back and forth tongue position) is freed, performing 83% of the exploration.

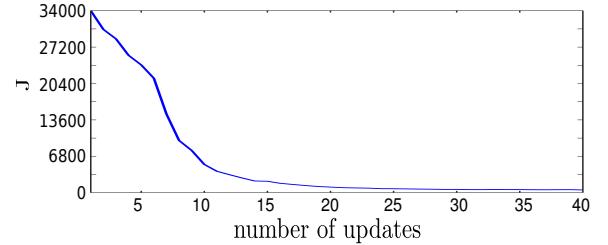
In Fig. 7, the total exploration magnitude  $\lambda^{\text{total}} = \sum_{m=1}^7 \lambda_m$  is visualized as a thick yellow/black line. It is normalized to 1 within an experiment.

#### IV. EXPERIMENTS

In this section, we first present experiments that investigate the freeing and freezing of degrees of freedom in the context



(a) The relative exploration magnitude of each vocal parameter exploration during the optimization process are displayed by the seven color patches, from  $P1$  (dark blue color patch at the bottom) to  $P7$  (pink color patch at the top). Two exploration peaks are shown: one at update 13 (87% relative exploration of  $P1$ ) and one at update 20 (72% of relative exploration of  $P5$ ). The total exploration magnitude is plotted as a thick yellow/black line. Note that the values are smoothed to filter out the exploration peaks that only last one update: each value of the articulators' exploration is averaged with the 4 last and next values.



(b) Learning curve of  $J$  over the updates.  
Fig. 7. Illustration of stochastic optimization to learn to vocalize the vowel \a/. (a) displays the balance between the exploration of each articulators. (b) shows the optimization of the cost function  $J$

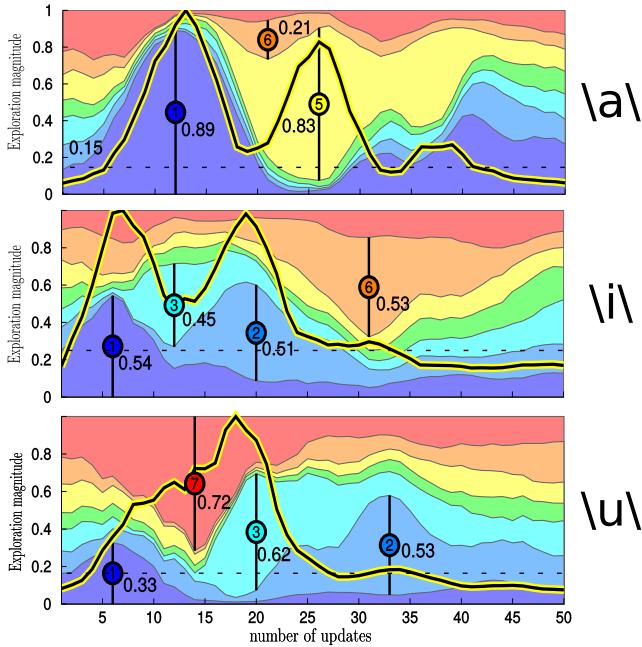


Fig. 8. Results of running stochastic optimization to learn to vocalize the vowels  $\backslash a \backslash$ ,  $\backslash i \backslash$ , and  $\backslash u \backslash$ . Same conventions as in Fig. 7.

of learning to vocalize the vowels  $\backslash a \backslash$ ,  $\backslash i \backslash$ , and  $\backslash u \backslash$ . Then, we do the same for arbitrary goals in the vocal space. Finally, we perform a sensitivity analysis providing further explanation of the obtained results.

#### A. Experiment 1: Recruitment Order for Vowels as Goals

This experiment analyses the recruitment order of parameter when learning how to achieve the three canonical vowels  $\backslash a \backslash$ ,  $\backslash i \backslash$  and  $\backslash u \backslash$ .

1) *Method and results:* We perform three optimization runs, one for each of the vowels  $\backslash a \backslash$ ,  $\backslash i \backslash$  and  $\backslash u \backslash$ . The results are shown in Fig. 8, highlighting the order in which degrees of freedom are freed (a degree of freedom is said to be freed when it has a large relative exploration magnitude). We do so by annotating the graph with the maximum relative exploration, e.g. 0.89 for  $P1$  in the first graph. This annotation is only added if the absolute normalized exploration (thick yellow/black line) is 5 percentage points above the baseline at the beginning; this threshold is indicated by a dashed line. Our motivation for this threshold is that, if there is no (absolute) exploration at all, none of the degrees are being freed, and it is therefore not relevant to study the relative exploration magnitudes.

By analyzing the peaks of the maximum relative exploration, we can determine a *recruitment order*, i.e. the order in which the degrees of freedom are freed. For the vowel  $\backslash a \backslash$ , this order is  $P1$ ,  $P6$ ,  $P5$ . For the vowel  $\backslash i \backslash$  it is  $P1$ ,  $P3$ ,  $P2$ ,  $P6$ , and for the vowel  $\backslash u \backslash$  it is  $P1$ ,  $P7$ ,  $P3$ ,  $P2$ .

2) *Discussion:* In these results we observe that the most recruited parameters are those controlling the distinctive phonetic features of the vowel to vocalize. Producing the vowel

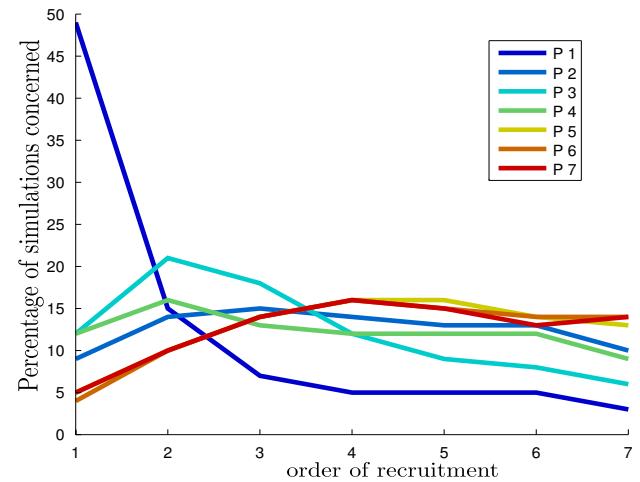


Fig. 9. Rank frequencies for vowels uniformly distributed in the vocalic triangle.  $P1$  is mostly related to jaw movements, see Fig. 2.

$\backslash a \backslash$  necessitates a vocal tract constriction at the level of the throat, what is typically obtained by opening the jaw and placing the tongue in a back position, hence the recruitment of  $P1$  (jaw opening) and  $P5$  (tongue movement at the throat level). Producing  $\backslash i \backslash$  necessitates a constriction of the tongue behind the teeth, what is typically obtained with a rather closed jaw and the tongue in the front position, hence the recruitment of  $P1$  (jaw),  $P3$  (tongue horizontal movement) and  $P2$  (tongue vertical movement). Finally, producing  $\backslash u \backslash$  necessitates constriction back in the palate, what is typically obtained by placing the tongue in a high-back position, which is here obtain through the conjoint action of  $P1$  (jaw),  $P7$  (tongue rotation) and  $P3$  (tongue horizontal movement).

It therefore seems that attempting to achieve different vowel goals results in different recruitment orders. Moreover the recruited articulators for a given vowel are mainly those involved in the production of the associated phonetic features.

Another observation is that  $P1$ , the parameter related to jaw movements, is freed first for all the vowels. To investigate whether this is due to the choice of these particular vowels, or rather a general feature of learning vocalizations in this sensorimotor space, we conduct a second experiment.

#### B. Experiment 2: Recruitment Order for Random Goals

This experiment analyses the global tendency of the recruitment order on a large number of vowel goals uniformly distributed in the formant space.

1) *Method and results:* We perform separate optimization runs for achieving 40 000 vowel goals drawn randomly in the vocalic triangle. For each run, we analyze the recruitment order using the same method as in the previous experiment.

The results are shown in Fig. 9. We observe a strong tendency to recruit  $P1$  first: it is the case in almost 50% of the simulations, whereas the other articulators are recruited first only between 4 and 13% of the simulations.

2) *Discussion:* Our interpretation of why  $P1$  (mainly corresponding to a jaw movement, see Fig. 2) is the most frequent degree of freedom to be freed first is that:

- The jaw position has a strong influence on the  $F1$  frequency, especially around the articulatory rest position from which the optimization process starts. This effect is shown on Fig. 10.
- $F1$  has an important influence on auditory perception [33], as we modeled using a scaling factor, and thus on the cost to be optimized (see the following section for a detailed analysis).

Therefore,  $P1$  is the most useful articulator to achieve a variety of vowel goals in our model, because it covers the wider range of  $F1$  values (see Fig. 10), which has a strong influence on the cost function because of its particular weighting (three times more than  $F2$  in our model). It is important to note that the result of Fig. 9 strongly depends on this particular weighting, which is based on psychoacoustic considerations [33], [34]. Therefore, our model shows that the strong effect attributed to  $F1$  in psychoacoustics can result in the predominance of the articulator principally controlling it (i.e.  $P1$  in our model, related to jaw movements) as a way to efficiently achieve a wide variety of auditory goals. This point will be analyzed in more detail in the next experiment.

These results provide an original interpretation for explaining the predominant role of the jaw in human speech evolution and acquisition. As we discussed in Section II-B, it has been proposed that this particular role could be due to evolutionary precursor behaviors involving jaw cycling such as mastication and ingestion [21], as well as non-human primates communicative gestures such as lipsmacks and tonguesmacks [10]. Here we show that pure learning mechanisms, driven by stochastic optimization within a space constrained by vocal morphological properties, can also be involved due to the particular role of the jaw and the first formant in vowel production and perception.

The next section provides a sensitivity analysis emphasizing how the stochastic optimization process allows a dynamical “freezing” and “freeing” of the different articulators according to their impact on the cost function minimization, and in particular analyzing in more detail the relative impact of jaw exploration during the course of vocal optimization.

### C. Experiment 3: Sensitivity Analysis

To better understand how the stochastic optimization process structures exploration by automatically freeing and freezing degrees of freedom, this experiment performs a sensitivity analysis quantifying the respective influence of articulatory parameters on the cost function during the realization of a particular vowel.

1) *Method:* We analyze the respective influence on the cost function of the articulatory parameters during the course of an optimization run to achieve a particular vowel. The aim is to show that stochastic optimization recruits in priority articulators that have a greater influence on the cost function,

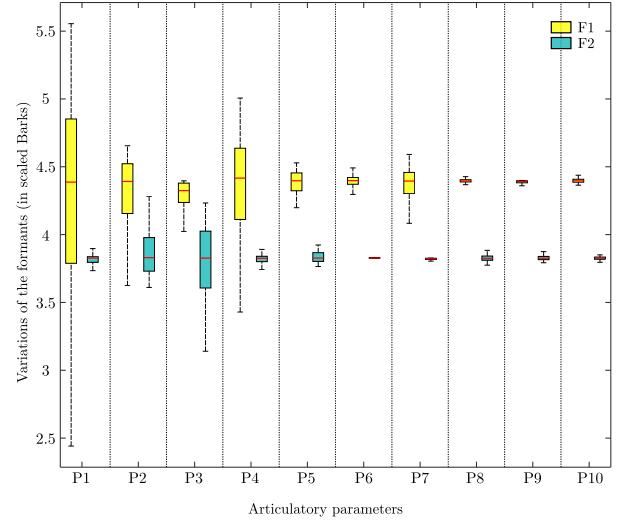


Fig. 10. Variations of the formant values (y-axis) implied by movements of each articulatory parameter (x-axis) uniformly sampled in the interval  $[-1, 1]$ , with 1000 samples per parameter. When a given articulator varies, the others are kept in their rest position.

providing a more detailed understanding of how such a mechanism can explain the predominance of jaw movements in early vocal development.

The optimization is run on the seven articulatory parameters with the goal of achieving the vowel \u. However, for the sake of simplicity, we restrict our analysis to the role of  $P1$  (mainly controlling the jaw position) and  $P3$  (mainly controlling the tongue front/back position) on the cost function. We chose these articulators because  $P1$  mainly influences  $F1$  values whereas  $P3$  mainly influences  $F2$  ones (as shown in Fig. 10), allowing to better distinguish their respective roles in achieving the vowel \u. The gradient of the cost function with respect to  $P1$  and  $P3$  is computed on regularly-sampled ( $P1, P3$ ) tuples. The partial derivatives are computed in an empirical manner using Equation (6), by looking at the cost function variations induced by small variations of the articulatory parameters around the sampled ( $P1, P3$ ) values. The values of the cost function  $J$  in Equation (6) are computed by fixing the five other articulators to values allowing the production of the target vowel \u, where we use the values obtained at the end of the optimization run.

$$\frac{dJ}{dP_i}(P_i) = \frac{J(P_i + dP_i) - J(P_i - dP_i)}{2 \times dP_i} \quad (6)$$

This allows the quantification of the influence of  $P1$  and  $P3$  on the cost function depending on their values. Then, we define the relative influence on the cost function of  $P1$  over  $P3$  by the ratio:

$$ratio(P1, P3) = log_{10} \left( \frac{| \frac{dJ}{dP_1}(P1) |}{| \frac{dJ}{dP_3}(P3) |} \right). \quad (7)$$

Therefore  $\text{ratio}(P1, P3)$  is the logarithm of the ratio between the absolute partial derivatives with respect to  $P1$  and  $P3$  (the partial derivatives being defined in Equation 6). It is positive when  $P1$  has a greater influence than  $P3$  on the cost function, negative otherwise.

2) *Results:* Fig.11a to 11f show the respective influence of  $P1$  and  $P3$  at regular time steps during a particular simulation, as computed by Equation (7). The cost function is computed with the vowel '\u' as the auditory goal. We observe that the starting neutral configuration is in a region where  $P1$  has more influence than  $P3$  on the cost function, i.e. where the ratio (Equation (7)) has a high value.

Fig.11g shows that exploration mainly appears on  $P1$  until the 9<sup>th</sup> update. This is explained by the greater influence of  $P1$  on the cost function at the beginning of the optimization process, i.e. around the rest position and in early vocal trials achieving small perturbations of this initial position, as observed on Fig.11a where the red color indicates that perturbations of  $P1$  have a higher influence on the cost function. As the learning progresses towards the goal (end position of the trajectories getting closer to '\u'), the optimization process reaches a blue ridge (Update 9): from this state, perturbations of  $P3$  therefore display a greater influence on the cost function than perturbations of  $P1$  and this explains the freeing of  $P3$  and the freezing of  $P1$  as observed on Fig.11g.

3) *Discussion:* This sensitivity analysis provides evidence that the stochastic optimization process will favor the recruitment of articulators having greater influence on the cost function minimization at a given update. In other words, exploration takes place on articulators which are the most useful to move closer to the auditory goal according to the current configuration of the vocal tract. Since  $P1$  (mainly controlling jaw movements) has a greater influence on  $F1$  (Fig.10) and  $F1$  has a higher weight in the cost function computation (Eq. 5), this analysis provides a computational understanding of the results obtained in Fig. 9, where we observed a strong tendency of recruiting in priority  $P1$  to achieve various vowel goals.

## V. CONCLUSION

It has been proposed that canonical babbling is so robust that a number of forces probably act on it [26]. Traditional propositions of these forces concern the derivation of prelinguistic behaviors in the course of human evolution, for example non-human primate orofacial communicative gestures [21]; the influence of the social environment [43]; or the role of curiosity-driven learning in exploring the articulatory-auditory space [25]. This paper proposes an original hypothesis regarding the predominant role of the jaw in infant vocal development, where we suggest that it could be a result of stochastic learning processes allowing the production of various auditory effects. This resulting predominance of jaw movements would then favor the emergence of canonical babbling in infant vocal development. This work takes inspiration of a previous model showing that such exploration strategies implied by a

stochastic optimization process provides a cognitive reason for the proximo-distal law of arm development [38], [37].

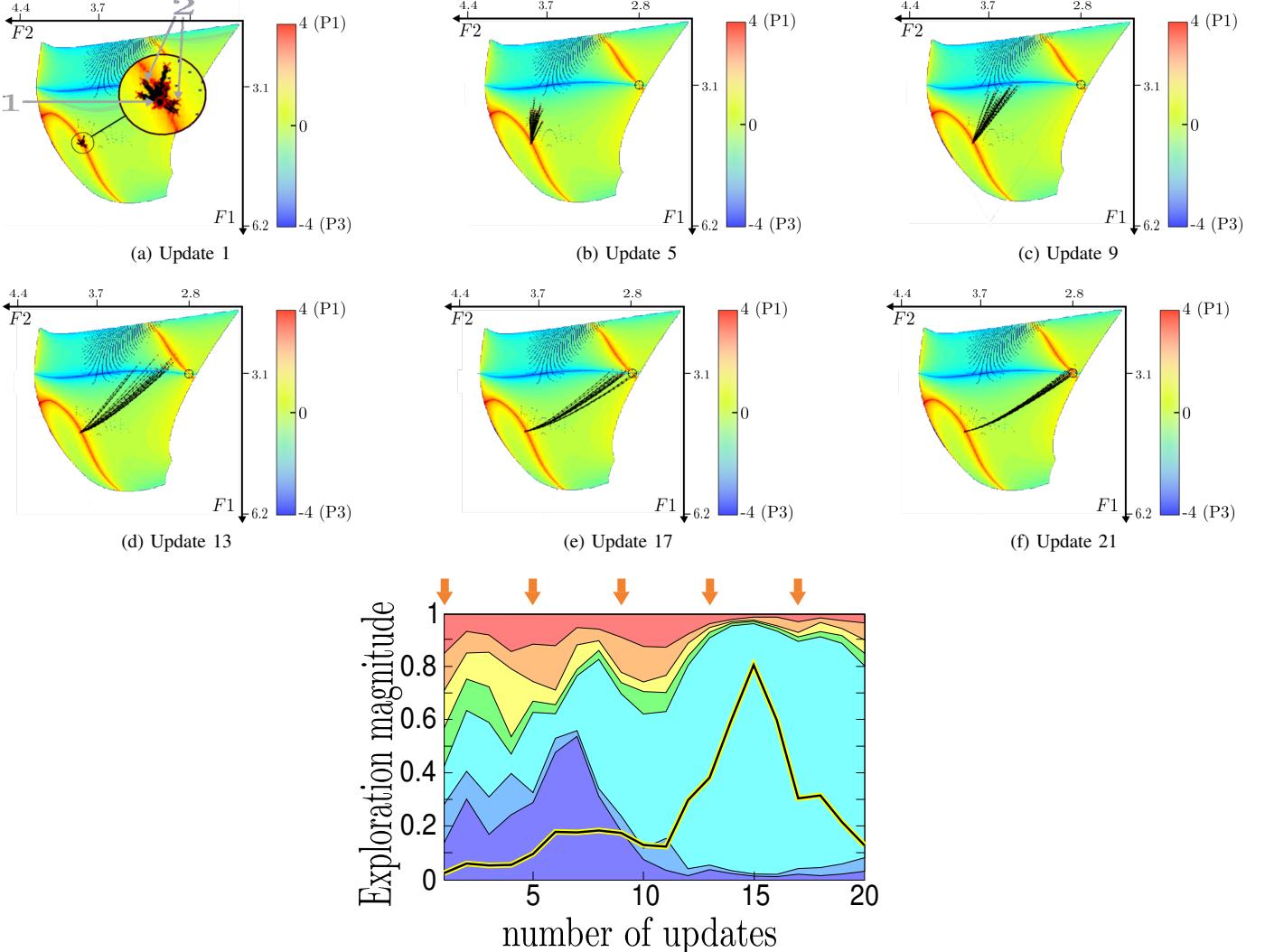
For this aim, we have developed a computational model using an articulatory synthesizer, movement generation and auditory perception processes, coupled with the stochastic optimization algorithm  $PI^{BB}$ . We have run simulations where the system iteratively optimizes the reaching of various auditory vowel goals and we have performed various analyses on the underlying results: extraction of exploration magnitude and sensitivity analysis during the exploration process, as well as the order of articulator recruitment. These results show that the order of recruitment was dependent of the auditory goal to be reached and that, on average on various auditory goals, the jaw is predominantly and firstly recruited by the optimization process. Moreover, a sensitivity analysis suggests that the order of recruitment is determined by the relative influence of each articulator on the cost function at a given time step.

Our hypothesis seems simpler than the aforementioned previous propositions in the sense that it relies on a general optimization process allowing to find adequate motor commands in order to reach various sensory goals. We do not claim that other factors, such as prelinguistic behaviors, the social environment, or curiosity-driven learning, do not play a role in this process. Actually, social guidance and curiosity-driven learning could both be mechanisms accounting for the selection of auditory goals that are set by the experimenter in the model presented in this paper. As a whole, we agree with Oller's proposition ([26]) that a number of forces probably act in favor of canonical babbling, thus explaining its robustness. Here we have proposed an original one which has the advantage of relying only on very general learning mechanism and also generalizes on other developmental aspect such as the proximodistal law of arm control.

## ACKNOWLEDGMENT

This work was partially financed by ERC Starting Grant EXPLORERS 240 007.

The authors would like to thank Louis-Jean Boë for the design of Fig. 1 (vocal tract by Sophie Jacopin).



(g) Relative exploration values where the update number of the snapshots are signalized by the orange arrows.

Fig. 11. Snapshots from a  $PI^{BB}$  exploration to learn to produce the vowel \u (the small circle at coordinates ( $F1 = 3.1, F2 = 2.8$ )). Sub-figures 11a to 11f show the deployment of the  $K$  concurrent samples at particular updates of the optimization process. The black lines represent the formant's trajectories, they all start from the same position (1 in sub-figure 11a) which is produced with all the articulators at their rest position. The end of the trajectories are represented by a red cross (2 in sub-figure 11a) and used to compute the costs. Superimposed to these trajectories is a heatmap, identical in all the sub-figures, representing the relative influence on the cost function of  $P1$  over  $P3$  as computed by Equation (7). For each tuple of  $(P1, P3)$  values regularly sampled in the range  $[-2.5, 2.5]^2$  with steps of 0.01, we first compute the resulting formant values  $(F1, F2)$  returned by the synthesizer, fixing the five other articulatory parameters to a constant value allowing the production of the target vowel \u. Then we compute the ratio given by Equation (7) for each of the  $(P1, P3)$  tuples. The heatmap shows the value of this ratio for each  $(P1, P3)$  tuple and plotted at their resulting positions in the  $F1 - F2$  plane. In red areas  $P1$  has a greater influence on  $J$  than  $P3$ , while blue areas correspond to the reverse situation and green to neutral areas where both parameters have similar contribution. Sub-figure 11g shows the exploration magnitudes of each articulator during the optimization process, using the same conventions as in previous similar figures. The orange arrows correspond to update index of the sub-figures 11a to 11e.

## REFERENCES

- [1] M. A. Arbib. Interweaving protosign and protospeech: Further developments beyond the mirror. *Interaction Studies*, 6:145–171, 2005.
- [2] G. Baldassarre. What are intrinsic motivations? a biological perspective. In *Development and Learning (ICDL), 2011 IEEE International Conference on*, volume 2, pages 1–8, Aug 2011.
- [3] A. Baranes and P.-Y. Oudeyer. Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems*, 2012.
- [4] A. Barto, S. Singh, and N. Cheneatz. Intrinsically motivated learning of hierarchical collections of skills. In *Proc. 3rd Int. Conf. Dvp. Learn.*, pages 112–119, San Diego, CA, 2004.
- [5] N. Bernstein. *The Coordination and Regulation of Movements*. Pergamon, 1967.
- [6] B. I. Bertenthal and C. von Hofsten. Eye, head and trunk control: The foundation for manual development. *Neuroscience and Biobehavioral Review*, 22:515–520, 1998.
- [7] N. E. Berthier, R. Clifton, D. McCall, and D. Robin. Proximodistal structure of early reaching in human infants. *Exp Brain Res*, 1999.
- [8] R. Carlson, B. Granström, and D. Klatt. Vowel perception: The relative salience of selected acoustic manipulations. *STL-QPSR*, 34:19–35, 1979.
- [9] A. A. Ghazanfar and D. Y. Takahashi. The evolution of speech: vision, rhythm, cooperation. *Trends in cognitive sciences*, 18(10):543–553, 2014.

- [10] A. A. Ghazanfar, D. Y. Takahashi, N. Mathur, and W. Fitch. Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Current Biology*, 22(13):1176–1182, 2012.
- [11] J. Gottlieb, P.-Y. Oudeyer, M. Lopes, and A. Baranes. Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in cognitive sciences*, 17(11):585–593, 2013.
- [12] J. R. Green, C. A. Moore, M. Higashikawa, and R. W. Steele. The physiologic development of speech motor control: lip and jaw coordination. *Journal of speech, language, and hearing research : JSLHR*, 43(1):239–55, feb 2000.
- [13] J. R. Green, C. A. Moore, and K. J. Reilly. The sequential development of jaw and lip control for speech. *Journal of speech, language, and hearing research : JSLHR*, 45(1):66–79, feb 2002.
- [14] F. H. Guenther, S. S. Ghosh, and J. A. Tourville. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and language*, 96(3):280–301, 2006.
- [15] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, 9(2):159–195, 2001.
- [16] N. J. Hodges, S. Hayes, R. R. Horn, and A. M. Williams. Changes in coordination, control and outcome as a result of extended practice on a novel motor skill. *Ergonomics*, 48(11-14):1672–1685, 2005.
- [17] A. J. Ijspeert, J. Nakanishi, and S. Schaal. Movement imitation with nonlinear dynamical systems in humanoid robots. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 2, pages 1398–1403. IEEE, 2002.
- [18] J. M. Iverson, A. J. Hall, L. Nickel, and R. H. Wozniak. The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. *Brain and language*, 101(3):198–207, 2007.
- [19] D. Klatt. Prediction of perceived phonetic distance from critical-band spectra: A first step. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'82*, volume 7, pages 1278–1281. IEEE, 1982.
- [20] P. MacNeilage and B. Davis. Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology*, 11:696–700, 2001.
- [21] P. F. MacNeilage. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21:499–511, 1998.
- [22] P. F. MacNeilage and B. L. Davis. On the origin of internal structure of word forms. *Science*, 288:527–531, 2000.
- [23] S. Maeda. Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. *Speech production and speech modelling*, 55:131–149, 1989.
- [24] C. Moulin-Frier, J. Diard, J.-L. Schwartz, and P. Bessière. COSMO (“Communicating about Objects using Sensory-Motor Operations”): a Bayesian modeling framework for studying speech communication and the emergence of phonological systems. *Journal of Phonetics*, 53:5–41, 2015.
- [25] C. Moulin-Frier, S. M. Nguyen, and P.-Y. Oudeyer. Self-organization of early vocal development in infants and machines: The role of intrinsic motivation. *Frontiers in Psychology*, 4(1006), 2013.
- [26] D. K. Oller. *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- [27] P.-Y. Oudeyer, A. Baranes, and F. Kaplan. Intrinsically motivated learning of real-world sensorimotor skills with developmental constraints. In *Intrinsically motivated learning in natural and artificial systems*, pages 303–365. Springer, 2013.
- [28] P.-Y. Oudeyer, A. Baranes, F. Kaplan, and O. Ly. *Intrinsically Motivated Cumulative Learning in Natural and Artificial Systems*, chapter Developmental constraints on intrinsically motivated skill learning: towards addressing high-dimensions and unboundedness in the real world. Springer, to appear.
- [29] P.-Y. Oudeyer and F. Kaplan. What is intrinsic motivation? a typology of computational approaches. *Frontiers in Neurorobotics*, 1, 2007.
- [30] P.-Y. Oudeyer, F. Kaplan, and V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286, 2007.
- [31] M. Schlesinger, D. Parisi, and J. Langer. Learning to reach by constraining the movement search space. *Developmental Science*, 3:67–80, 2000.
- [32] J. Schmidhuber. A possibility for implementing curiosity and boredom in model-building neural controllers. In J. A. Meyer and S. W. Wilson, editors, *Proc. SAB'91*, pages 222–227, 1991.
- [33] M. Schroeder, B. Atal, and J. Hall. *Frontiers of Speech Communication Research*, chapter Objective measure of certain speech signal degradations based on masking properties of human auditory perception, pages 217–229. London Academic Press, 1979.
- [34] J. Schwartz, L. Boë, N. Vallée, and C. Abry. The Dispersion-Focalization theory of vowel systems. *Journal of Phonetics*, 25(3):255–286, 1997.
- [35] D. Southard and T. Higgins. Changing movement patterns: Effects of demonstration and practice. *Research quarterly for exercise and sport*, 58(1):77–80, 1987.
- [36] F. Stulp, L. Herlant, A. Hoarau, and G. Raiola. Simultaneous on-line discovery and improvement of robotic skill options. In *International Conference on Intelligent Robots and Systems (IROS)*, 2014.
- [37] F. Stulp and P.-Y. Oudeyer. Adaptive exploration through covariance matrix adaptation enables developmental motor learning. *Paladyn. Journal of Behavioral Robotics*, 3(3):128–135, September 2012.
- [38] F. Stulp and P.-Y. Oudeyer. Emergent proximo-distal maturation through adaptive exploration. In *International Conference on Development and Learning (ICDL)*, 2012.
- [39] F. Stulp and O. Sigaud. Robot skill learning: From reinforcement learning to evolution strategies. *Paladyn. Journal of Behavioral Robotics*, 4(1):49–61, September 2013.
- [40] B. Vereijken, R. E. v. Emmerik, H. Whiting, and K. M. Newell. Free (z) ing degrees of freedom in skill acquisition. *Journal of motor behavior*, 24(1):133–142, 1992.
- [41] A. Warlaumont. A spiking neural network model of canonical babbling development. In *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on*, pages 1–6, 2012.
- [42] A. Warlaumont. Salience-based reinforcement of a spiking neural network leads to increased syllable production. In *Development and Learning and Epigenetic Robotics (ICDL), 2013 IEEE International Conference on*, pages 1–6, 2013.
- [43] A. S. Warlaumont. Prespeech motor learning in a neural network using reinforcement. *Neural Networks*, 38:64–95, 2013.
- [44] A. S. Warlaumont. Modeling the emergence of syllabic structure. *Journal of Phonetics*, In press.