Warlaumont, A. S. (2014). Reinforcement-modulated self-organization in infant motor speech learning. In J. Mayor & P. Gomez (Eds.), Computational Models of Cognitive Processes, pp. 113–124. World Scientific. doi: 10.1142/9789814458849_0009

Reinforcement-modulated self-organization in infant motor speech learning

Anne S. Warlaumont

Cognitive and Information Sciences, School of Social Sciences, Humanities, and Arts, University of California, Merced, 5200 North Lake Rd.
Merced, CA 95343, USA
E-mail: awarlaumont2@ucmerced.edu
annewarlaumont.org

Neural network models of early motor speech vocal learning are reviewed, with a focus on those models that utilize reinforcement to modulate what would otherwise be self-organized learning. It is argued that such a mechanism likely plays a role in bringing about the changes observed in prespeech vocalizations produced by human infants. Such models complement well the already popular purely self-organized learning models that focus on effects of exposure to sounds from the ambient language.

Keywords: Reinforcement, Self-organization, Speech-language development, Infancy, Babbling, Self-organizing map, Spiking neural networks

1. Introduction

Connectionist models can for the most part be grouped into two broad categories, those that learn under supervision, meaning that they receive explicit feedback about what their correct outputs should be, and those that learn in an unsupervised or self-organized fashion, meaning that they are given no information about what their outputs ought to be. The focus of the present paper is on a third way of providing feedback to connectionist models, reinforcement, and the valuable role that it may play in motor speech learning. I will argue that reinforcement is particularly useful in this domain because it relaxes some unrealistic assumptions made by supervised approaches to motor learning while at the same time providing structure that is lacking in motor exploration compared to perceptual learning. Reinforcement can be incorporated through a very simple modification to standard unsupervised learning approaches: increasing the learning rate when reinforcement is received. After reviewing some supervised and unsupervised approaches in the domain of early motor speech development, I will present two recent examples of the reinforcement-modulated learning approach from that same domain.

2. Background

There are many ways in which infant vocalizations change, apparently as a consequence of learning, across the first year or two of life, as the repertoire expands in pitch, amplitude, and vocal quality, comes to include syllabically patterned vocalization, and eventually comes to include specific word forms[1]. A number of modelers interested in providing neural network accounts for some of these changes have focused their efforts on showing how unsupervised learning might contribute to the development of vocal imitation abilities. For example, one typical approach is to model the infant as having two layers of neurons, a sensory layer and a motor layer, connected via self-organizing Hebbian connections. The sensory and motor layers themselves can be treated as self-organizing maps or can be given static connections to sensory input feature vectors and motor actuators. Training typically proceeds by producing random motor outputs, running those outputs through a vocalization synthesizer, hearing the resulting sounds, then allowing self-organized learning to take place. Such models have demonstrated the ability to learn to imitate sounds[2–5] so that, for example, when a model hears the vowel /a/, activation propagates from the sensory representation of that sound to the motor system and the model then produces a similar sound.

However, development of the ability to imitate sounds that are already in the infant's repertoire is only one feature of prespeech vocal development. Other work has addressed how infants might come to learn to produce particular sounds when a representation of a specific speech sound is activated. One approach that has been taken is to provide a neural network with input representing a target speech sound and then to have the network produce a motor output in response, training the network using supervision, where the correct motor commands to produce the input sound are given directly to the model[6]. A problem with this supervised approach is that it assumes that the infant somehow has access to explicit information regarding the correct motor commands that go with a particular speech sound. This seems very unlikely to be the situation faced by human infants. In fact, that information is precisely the knowledge that the model is trying to learn. So although the model performs well, and proves that it is possible to encode the sensory-motor relationship for speech sounds in neural network weights, its learning is not as realistic as it should be.

A different approach to learning associations between speech sound representations and motor commands is to have a model produce random vocal babble and when it happens by chance to produce a sound from a particular speech sound category allow the network to learn an association between the randomly generated motor commands and the speech sound category[7]. This approach makes more realistic assumptions about the knowledge infants possess prior to learning. The approach could be classified as an unsupervised model (ignoring the issue of learning which sounds are speech sound targets in the first place).

Another important feature of early vocal development is that infants often produce sounds seemingly rather spontaneously, not in direct imitation of an adult nor apparently targeting a particular adult speech sound. And these spontaneously produced sounds change as the child gets older, coming to display more of the characteristics of adult vocalization. There are at least two ways in which such changes in spontaneously produced vocalizations can be effected. One way is for infant vocalizations to be influenced by targets, as in the DIVA model. By having a model receive auditory input from another speaker and by having this affect vocalization production (so that spontaneous vocalizations are now not purely random, but even during learning are affected by external input), a model can end up producing sounds that resemble that external input's[8,3]. This seems plausible, and would also be classified as a type of unsupervised learning. Figure 1 illustrates how some of the mechanisms reviewed here relate to each other and to those described below.

3. Reinforcement-modulated vocal learning with a self-organizing map

Another way to generate meaningful changes in spontaneously produced vocalizations is for the model's own vocalizations to be differentially reinforced and for that reinforcement to guide learning. One simple example of this can be found in a study in which my colleagues and I used a self-organized map (SOM) to control muscle activations in a vocal-tract-based vocalization synthesizer[9]. Typically, self-organizing maps are used in perceptual tasks. A single layer of neurons is connected to a vector of inputs. The neural network is exposed to inputs from a training corpus (in models of speech perception, these could be acoustic representations of various vowels in a language, for example) and with each input, adjusts its neurons' receptive fields (i.e. weighted connections between inputs and neurons) in

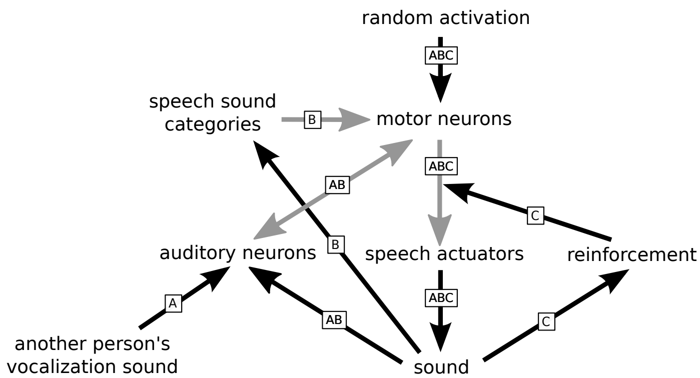

Fig. 1. Summary of key operations of some previous neural network models of early motor speech development. The gray arrows represent neural connections that can undergo self-organized learning. Black connections typically have been held constant over the course of learning in these models. Connections marked "A" are incorporated in models that self-organize auditory-motor connections on the basis of babbling and adult input. They learn to imitate adult sounds and their spontaneous vocalizations can come to resemble heard adult sounds. Connections marked "B" are those that incorporated in models that learn to associate certain motor patterns with speech sound categories (so far models have tended to assume that those categories' acoustic features are known beforehand). Connections marked "C" are incorporated in the models discussed below that learn through purely motor babble with learning modulated by reinforcement that is provided on the basis of a vocalization's acoustics.

a self-organizing fashion. SOMs in these cases capitalize on the structure inherent in the input, i.e. the kinds of vectors out of all possible vectors that are represented and their similarity to one another. When using a self-organizing map for motor learning, when neurons are not connected to perceptual inputs but rather to motor outputs, such as muscles, this inherent structure may not be present. In particular, models of early motor speech development such as many of those mentioned above have tended to have training trials with random motor "babble" at their core. Muscles are randomly activated and the resulting sounds are observed. In models where the goal is to form sensory-motor associations, this is not a problem, but if the goal is, without any auditory examples, to change the motor primitives themselves and in turn change the types of sounds that are produced during babble, there is a problem in that random motor output provides no interesting structure to self-organize around. Our study addressed this issue by introducing reinforcement when the model produced desirable sounds and by having learning occur only when reinforcement had been given. That is, learning was modulated by reinforcement.

Like the other models reviewed above, our model also learns by babbling (see Fig. 2A). A learning trial begins by randomly activating the SOM neurons. These activations then propagate via weighted connections to various muscles involved in speech. Using these muscle activation settings, a vocal tract simulator[10] generates a vocalization sound. The acoustic properties of this sound, such as its fundamental frequency (f0) and formant frequencies, can be estimated automatically, or a person can listen to the sound and rate it. We took the automatically estimated fundamental frequency (which can be undefined, often indicating that the sound does not contain voicing) and the first (F1) and second (F2) formant frequencies and set criteria for reinforcement based on these values. We tried several different sets of reinforcement criteria, including the following: (1) no criteria (all sounds were reinforced), (2) sounds must have a defined f0, (3) sounds must have a defined f0 and have F1 and F2 values that are similar to published values for American English speakers, (4) sounds must have a defined f0 and have F1 and F2 vowels that are similar to published values for Korean speakers. Based on the particular reinforcement criteria used in a given simulation, the model's vocalization was either reinforced or not reinforced. If reinforced, then the standard SOM learning algorithm was executed[11], with node activation, which again was randomly determined at the onset of the trial, determining the winning node and the weights from that and neighboring nodes being adjusted to become more similar to the muscle activations just produced. If the vocalization was not reinforced, then no learning took place on that trial.

We evaluated the sounds produced toward the end of the learning period, and compared those sounds across the different reinforcement criteria conditions. These results are shown in Figure 2B and 2C. The reinforcement of all sounds (Condition 1) corresponds to what happens when trying to self-organize around random motor output and as expected results in no increase in defined f0s. In contrast, under all the other conditions, wherein sounds had to have defined f0 in order to be reinforced, we found that toward the end of the learning period, the models produced many more vocalizations with defined f0 compared to early in the learning period. Similarly, when the model was reinforced for producing American English vowels, toward the end of learning the vocalizations it produced were more closely matched to American English vowels compared to when the model was reinforced for producing Korean vowels and vice versa. Taken together, the two results suggest that the reinforcement-modulated learning adaptation

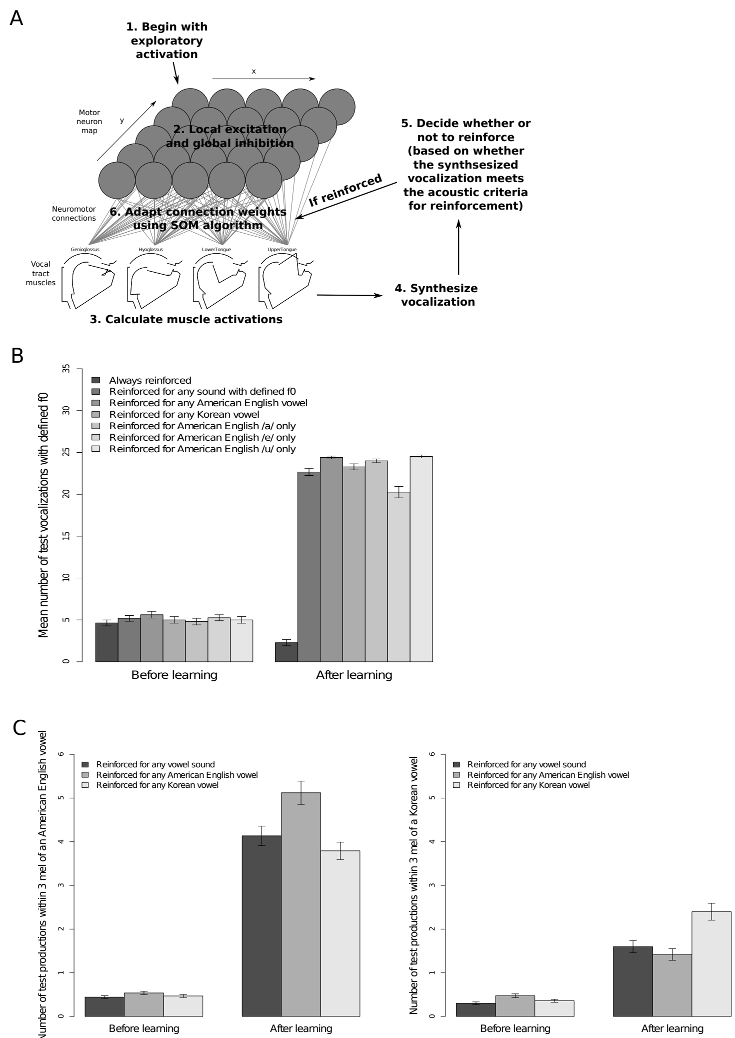

Fig. 2. A: Schematic description of the reinforcement-modulated SOM model[9]. B: Frequency of spontaneous vocalizations having defined f0 early and late in learning when the model is reinforced using various criteria. C: Spontaneous vocalizations' similarity to American English and Korean vowels early and late in learning when the model is reinforced for no particular language, American English-like vowels, or Korean vowels.

to the SOM is approach can work for modeling the development of spontaneous vocalizations coming to increasingly resemble adult speech.

4. Reinforcement-modulated vocal learning with a spiking neural network

In another study, I attempted to use the same principle of reinforcement-modulated learning to model how canonical babbling might come to emerge within infants' spontaneous vocalizations [12]. Around 6–7 months, infants begin to systematically produce canonical babbling, which is the act of producing prespeech vocalizations that have speech-like syllabic timing. Prior to that vocalizations tend to have either sloppier syllabic structure (as in gooing and marginal babbling) or no syllabic structure[1]. It is a bit trickier to model the emergence of canonical babbling than the emergence of phonation or of specific vowel types since babbling is an inherently temporal phenomenon. This may be part of the reason it has not been addressed by any of the other models mentioned above. My approach was to use a spiking neural network, specifically Izhikevich's[13] network, to generate the temporal dynamics of articulatory movement (Fig. 3). The network consists of a thousand neurons, both excitatory and inhibitory, that are connected to each other at random. Each neuron has a voltage parameter that increases under the influence of excitatory neuron input and decreases under the influence of inhibitory neuron input. At 1 ms intervals the voltage of each neuron is calculated and when a neuron's voltage exceeds a particular threshold, the neuron's voltage spikes and then immediately drops back down. When the voltages of all the neurons in the model are summed together, the resulting time series appears rather complex, with oscillations at a range of timescales[14]. By taking a subset of 100 of these neurons and calling them motor neurons, we are thus equipped with a naturally time-varying signal that can be given (after a bit of smoothing) to a vocalization synthesizer as muscle activations. The network learns using a spike-timing dependent plasticity (STDP) mechanism, which can be thought of as a type of Hebbian learning. Izhikevich's model extends previous spiking neural network models by making the rate of STDP, i.e. the rate of learning, dependent on the concentration of the reinforcement-related neurotransmitter dopamine (see also Florian's[15] and Farries and Fairhall's[16] related models). In the babbling model, a human "caregiver" listened to each sound produced by the model and decided whether or not to reinforce it, with the goal of the caregiver being to get the model to produce high-quality canonical babble more often (see Fig. 3B). When the listener decided to reinforce the model, a surge of dopamine was given to the neural network. As can be seen in Figure 3C, the network appears to have learned, through this reinforcement-modulated STDP mechanism, to produce high quality babble, doing so in a rather gradual fashion.

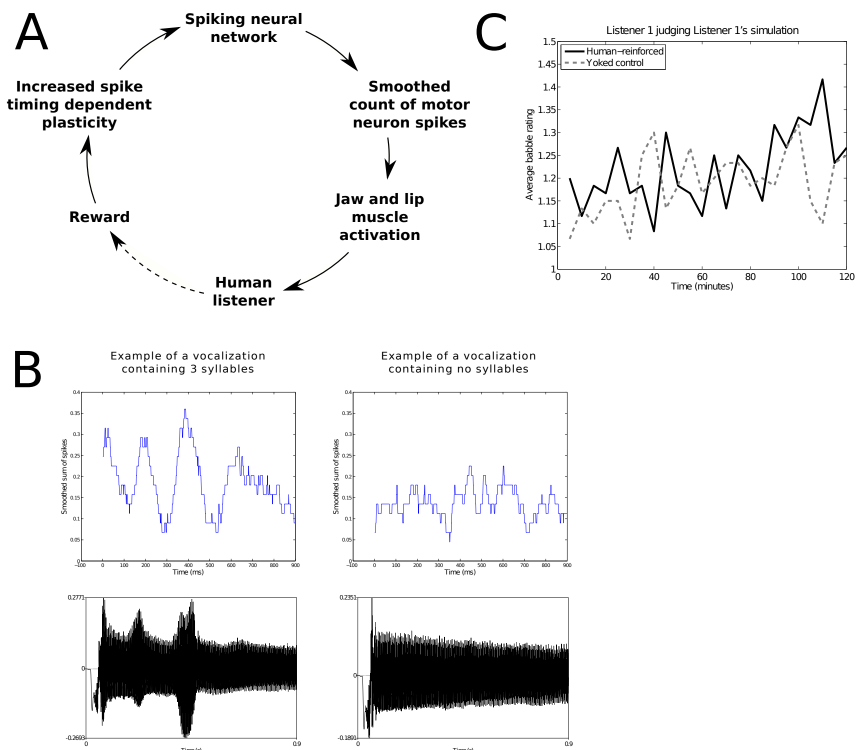

Fig. 3. A: Schematic illustration of the cycle of learning from producing vocalization with a spiking neural network. B: Examples of muscle activations produced by the model (top) and corresponding acoustic waveforms (bottom). The sound on the left would be classified as canonical babble whereas the sound on the right would not. C: Change in the tendency of the model to produce high-quality canonical babble over the course of learning, compared to a control simulation.

5. Conclusions

Taken together, these two models (the SOM model and the spiking neural network model) suggest that reinforcement-modulated learning can be a powerful mechanism for generating learning effects on spontaneous behavior, at least in the vocal domain. It is a general mechanism that can be applied to quite different types of self-organizing neural networks. That said, one might ask whether this reinforcement-gated learning is realistic. From the neurobiological standpoint, there is certainly evidence that dopamine plays a role in increasing STDP during motor learning[17]. And studies of human infant vocal learning indicate, at a behavioral level, that social reinforcement of infant vocalizations increases an infant's rate of vocalization[18] and that in naturalistic contexts mothers do selectively reinforce their infants for vocalizing[19]. Others have proposed that infants might derive intrinsic reinforcement from some of their own actions, such as sounds that expand the infant's current repertoire [20] or stimuli with appropriately moderate complexity[21]; intrinsic reinforcement could perhaps also be generated when infants produce sounds that resemble those they've heard produced by others in their environment. These ideas intrinsic sources of reinforcement still need to be tested in studies with human infants.

Another argument that supports the idea that reinforcement-modulated learning plays a role in this domain is that even deaf infants exhibit many of the changes in spontaneous vocalizations exhibited by hearing infants, such as increase in duration of vocalization, expansion of the ranges of pitches and vocal qualities produced, and even emergence of canonical babbling (although canonical babbling does emerge later for deaf infants)[1]. For profoundly deaf infants, the contagion mechanism of change in spontaneous vocalization modeled by Oudeyer[8] and suggested by Westermann and Miranda[3] is presumably not operational, at least not via the auditory domain, because these infants do not have access to auditory speech input. They do have the ability to feel the vibrations and other tactile sensations generated when they vocalize, and have access to visual and tactile input from other individuals. These tactile and visual sources of information could presumably provide quite useful reinforcement, both social and intrinsic, contingent on the infants' vocal behaviors. Social reinforcement could, as it does for hearing infants, take the form of a touch, smile, look of interest, etc. from a caregiver. Intrinsic reinforcement could result from the deaf infants' deriving more pleasure or interest first of all from vocal motor acts that result in vibration at the larynx and second of all from those that result in sounds with salient rhythmic patterns either at the larynx or through contact between the tongue, lips, and other structures. These social and/or intrinsic reinforcements, coupled with vocal motor exploration and reinforcement-modulated self-organized learning, could explain why deaf infants exhibit many of the spontaneous vocalization advances that hearing infants exhibit.

It is worth saying a few words about why supervised learning approaches are not ideal for this application. Supervised learning methods have been applied in the past in models of speech-language development, and this work has certainly been informative[6,22]. However, in trying to model development of motor learning of the sort that is the focus here, the desired outputs are muscle activations (or other effector values) that we should not assume infants innately know. Using a supervised approach in this context would require knowing ahead of time what the muscle activations or effector values ought to be. In fact, even if the goal is not to model human infants but to control a simulated or robotic vocal tract, it will often be the case that good muscle activations or articulator positions for producing a particular type of sound are unknown. This was the case in both the studies highlighted in this paper—in neither case did the author(s) know ahead of time what muscle activations should be set to in order to create the desired vowel sounds or syllable patterns. Autoencoder networks[23,24] provide a solution to the problem of of unknown targets in perceptual learning problems, so might be expected to also be appropriate for motor learning of the sorts focused on here. Unfortunately this is not the case because of the fundamentally different nature of the learning problem, where, rather than trying to recognize, categorize, and predict sensory inputs, we need to generate outputs that before learning are unknown. Although effects of perception on production are not the focus here, it is possible that autoencoder like processes are involved in infants' perception of their own and others' vocalizations, which could in turn have an impact on vocal motor learning by affecting what infants find reinforcing or by affecting vocal contagion.

It is probably true that multiple mechanisms (particularly contagion and reinforcement-modulated learning) are involved in the changes in spontaneous vocalization observed in human infancy. Thus, an important future direction will be to build models that incorporate multiple mechanisms, and to assess the specific contributions of each and the ways in which different mechanisms might interact. Additionally, although the focus here has been on modeling changes in spontaneous vocalizations, it seems likely that mechanisms such as reinforcement-modulated learning could be integrated nicely into existing models of phonetic learning, such as the DIVA model[7]. Of course, direct comparison to real human infant data will also be important, and it would be helpful to study human infant vocalizations with the goal of seeking more information about these various mechanisms' roles.

Finally, it seems reasonable to expect that reinforcement-modulated self-organized learning might be useful in other action and motor control domains. In fact, the neurobiological work with rats cited above as evidence for reinforcement-modulated motor learning was done in the context of learning a forelimb reaching skill[17]. In modeling work, reinforcement modulation of learning has already been utilized in other action-centered applications such as spatial navigation[25] and decision making[26]. Perhaps it will prove fruitful to pursue a similar approach in modeling language production at higher levels of analysis, such as production of words or sentences, and in modeling additional motor behaviors such as development of limb control for reaching and gesturing.

6. Acknowledgments

I would like to thank Gert Westermann, Kim Oller, Eugene Buder, Robert Kozma, Mike Goldstein, Thomas Hannagan, Bob French, Pierre-Yves Oudeyer, Clément Moulin-Frier, Mary Beckman, Chris Kello, Jeff Rodny, and the anonymous reviewer for helpful conversations and advice.

References

1. D. K. Oller, *The emergence of the speech capacity* (Lawrence Erlbaum Associates, Mahwah, NJ, 2000).
2. Y. Yoshikawa, M. Asada, K. Hosoda and J. Koga, *Connection Science* **15**, 245 (2003).
3. G. Westermann and E. R. Miranda, *Brain and Language* **89**, 393 (2004).
4. I. Heintz, M. Beckman, E. Fosler-Lussier and L. Ménard, Evaluating parameters for mapping adult vowels to imitative babbling, in *Proceedings of the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, (Brighton, UK, 2009).
5. A. S. Warlaumont, G. Westermann and D. K. Oller, Self-production facilitates and adult input interferes in a neural network model of infant vowel imitation, in *AISB 2011 Computational Models of Cognitive Development*, eds. D. Kazakov and G. Tsoulas (Society for the Study of Artificial Intelligence and the Simulation of Behaviour, York, UK, 2011).
6. H. Kanda, T. Ogata, T. Takahashi, K. Komatani and H. Okuno, Continuous vocal imitation with self-organized vowel spaces in recurrent neural network, in *2009 IEEE International Conference on Robotics and Automation*, 2009.
7. F. H. Guenther, S. S. Ghosh and J. A. Tourville, *Brain and Language* **96**, 280 (2006).
8. P.-Y. Oudeyer, *Journal of Theoretical Biology* **233**, 435 (2005).
9. A. S. Warlaumont, G. Westermann, E. H. Buder and D. K. Oller, *Neural Networks* **38**, 64 (2013).
10. P. Boersma, *Functional phonology: Formalizing the interactions between articulatory and perceptual drives* (Holland Academic Graphics, The Hague, 1998).
11. T. Kohonen, *Proceedings of the IEEE* **78**, 1464 (1990).
12. A. S. Warlaumont, A spiking neural network model of canonical babbling

development, in *Proceedings of the 2012 IEEE International Conference on Development and Learning (ICDL)*, 2012.
13. E. M. Izhikevich, *Cerebral Cortex* **17**, 2443 (2007).
14. C. T. Kello, J. Rodny, A. S. Warlaumont and D. C. Noelle, *Critical Reviews in Biomedical Engineering* **40**, 501 (2012).
15. R. V. Florian, *Neural Computation* **19**, 1468 (2007).
16. M. A. Farries and A. L. Fairhall, *Journal of neurophysiology* **98**, 3648 (2007).
17. K. Molina-Luna, A. Pekanovic, S. Röhrich, B. Hertler, M. Schubring-Giese, M.-S. Rioult-Pedotti and A. R. Luft, *PLoS ONE* **4**, p. e7082 (2009).
18. M. H. Goldstein, A. P. King and M. J. West, *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8030 (2003).
19. J. Gros-Louis, M. J. West, M. H. Goldstein and A. P. King, *International Journal of Behavioral Development* **30**, 509 (2006).
20. C. Moulin-Frier and P.-Y. Oudeyer, Curiosity-driven phonetic learning, in *Proceedings of the 2012 IEEE International Conference on Development and Learning (ICDL)*, 2012.
21. C. Kidd, S. T. Piantadosi and R. N. Aslin, *PLoS ONE* **7**, p. e36399 (2012).
22. M. S. Seidenberg and J. L. McClelland, *Psychological Review* **96**, 523 (1989).
23. D. Mareschal, R. M. French and P. C. Quinn, *Developmental Psychology* **36**, 635 (2000).
24. G. E. Hinton, *Trends in Cognitive Sciences* **11**, 428 (2007).
25. H. Voicu, R. Kozma, D. Wong and W. J. Freeman, *Connection Science* **16**, 1 (2004).
26. R. Ratcliff and M. J. Frank, *Neural Computation* **24**, 1186 (2012).