Spaces of Imitation

Elpida S. Tzafestas
Institute of Communication and Computer Systems
National Technical University of Athens
Zographou Campus 15773, Athens, GREECE.

Abstract

This paper presents an imitation mechanism and a study of its behavior in spatial grid-based configurations. The imitation mechanism replicates external signals without associating with objects, as in higher-level imitation; it is therefore a model of proto-imitation where agents imitate unconditionally the agents they happen to interact with. We study the mechanism in 2D space to understand how it influences the emerging spatial structures. Our results indicate that in the absence of "adults", i.e agents that act as tutors and do not imitate others, the whole grid eventually stabilizes to one common language for all agents. This is more rapid when various agent mobility schemes are introduced. Complex local micro-languages may emerge permanently, for example if a fair number of adults is present or temporarily if new words have to be invented from time to time, for example to account for new environmental stimuli or cultural factors. All our results suggest that a purely reactive proto-imitation mechanism can suffice to produce various language configurations in space, without the need for agents to reason explicitly on conditions and possible outcomes of language adoption.

1. Introduction

Research in imitation spreads in various disciplines (for an overview see [1]) and generally centers around two major themes: the role of imitation in social interaction and communication and the mechanism by which imitated responses are produced. Functional studies related to social behavior and communication are common both in psychology and theoretical biology and rely on the implicit assumption that imitation is mainly a means to (learn to) communicate or interact socially [2][3]. Studies related to the neural mechanisms behind imitative response generation investigate mainly phenomena of neonatal or early infant imitation and are tackling questions such as the degree to which neonatal imitation is goal-directed, motivated and selective [4]. Other important issues we retained from the literature as specifications for modeling are:

- Imitation should start as a reactive or impulsive process and subsequently catalyze itself through the social interaction itself. This self-catalysis may be through direct training by adults; adult turn-taking in imitation is one such way of training [3].
- Imitation is a multi-level process found in many different forms of varying degrees of complexity in a vast number of animal species [5]. However, even animals that can imitate cannot learn to communicate at a human-like level of complexity and this is an instance of the major question in cognitive science "what makes humans different from animals" [6].
- When used in a communicative context, human-level imitation manages to co-exist with diverse, arbitrary sign systems, which is apparently contradictory with the effect of imitation [7].
- Disturbed versions of imitative mechanisms may be found in a spectrum of developmental disorders [8] and this is an indication that imitation cannot be studied outside a development context, at least functionally [9].

We are therefore developing a model that allows proto-imitation outside explicit communication but may allow emergence of communication in the medium or long term ontogenetically. This model separates response imitation from response association to external meaning so as to make room both for species that can imitate but not associate meaning and for disabled humans that cannot associate well or even proto-imitate well. We work at the response imitation level to show that a wealth of phenomena may later emerge and especially those related to direct associations.

The usual agent model found in the literature (see for example [2]), supposes the existence of M objects with one signal associated with each object. The term "object" can denote anything from an individual or an inanimate object to an action or an event, in short any external thing that can be referred to. The signals are supposed to transfer information about the objects and can take values in any physical medium that an agent is able to use; however, the typical case is to think of signals as vocalizations. An imitative process is one that allows one agent to learn to use another agent's signal to refer to the same object. As such, imitation allows to two or more agents to communicate by using the same signals for the same objects. One common formalization of this imitation process [7] in a population of agents is through the use of a MxL language matrix per agent where each entry denotes the probability for the agent to refer to object i (i=1,2,…,M) by using the signal j (j=1,2,…,L). This setup presupposes that associations between objects and signals should exist for imitation to take place and that the role of imitation is to make the associations of different agents converge to one common language. This assumption is partly due to the adoption of the language domain as the experimental field of imitation *par excellence*.

Our own model of imitation of fellow agent responses does not assume any prior association to any external object. Instead we model the way an infant agent recognizes and reproduces an observed signal coming from an adult agent and standing for the response to a perceived object without having access to the object itself and thus without associating with it. Our model is a *functional* model of a neural structure that generates responses to match and replicate an external stimulus, i.e. the signal received. We opted for a study with the aid of a functional model of the neural structure rather than the structure itself, because we did not want to constrain our results within the possibilities of a given structure. We are rather seeking the organizational properties that such a structure should have to allow imitation.

We begin therefore by adopting the view that the ontegenetic development at the neural level follows the same principles as Darwinian evolution at the population level ([10]) and that any novel responses should be generated internally and selected within the environment rather than be directly "taught" by it. By applying this view to the proto-imitation level of low-level reproduction of external signals, we get a model where continuous internal generation of responses combined with an environmental selection (i.e. reinforcement through actual response matching to external signal) allows various imitative phenomena within a population of agents. Association of signals to objects can appear later, if at all. We use real-valued signals as in [2] instead of the more common discrete or symbolic signals, because we feel we should not ignore the continuous nature of real-world signals (especially vocalizations) and because we want to explicitly model the distance between agent responses.

An agent possesses a number of internal "frequencies" or eigenfrequencies (these could correspond to real vocal parameters or to neural patterns) that have varying degrees of affinity to a given signal: because frequencies as well as signals take real values in [0,1], affinities of frequencies will also take real values in the same interval. The response to a signal is the eigenfrequency with the highest affinity. At each step, new eigenfrequencies are generated proportionally to the affinity of the previous ones. The highest matching frequencies reproduce massively, while the lowest ones vanish and are replaced by newly generated random eigenfrequencies. An exploration factor is also defined, which is the maximum percentage of random eigenfrequency replacement independently of affinity. The overall affinity of an agent to an external signal is the average affinity of all its frequencies, thus it is internally generated and not externally imposed/designed in any way. This measure expresses ***how "well" an agent recognizes and can reproduce a signal*** and may therefore serve as a basis for subsequent emergence of communication. This model is summarized in Fig. 1. Initial results given in [11] show that this primitive mechanism can allow agents to build a common language with the agents they interact with and hence that the communicative value of the mechanism can be a subsequent discovery rather than a prerequisite of the mechanism.

```
// Frequencies: f[] - size K, f[i] in [0,1]
// External Signal: x in [0,1], Meaning/Object: n

// 1. Matching step: Compute affinity of each
frequency
foreach frequency f[i] (i=1 to K)
{     diff = |f[i]-x|
      if (diff <= T[i]) // a threshold
            affinity[i] = (1-diff) (in [0,1])
            else affinity[i] = 0;    }
// The (indirect through imitation) response
// of the agent to the external meaning
//object n is:
language[n] = f[i] with max affinity[i]
Agent's total_affinity =
      avg(affinity[i],i=1 to K); in [0,1]

// 2. Selection step:
//    Reproduce frequencies proportionally to
//    affinity
foreach frequency f[i] (i=1 to K)
{     // Pop = Number of clones of f[i]
      pop = (affinity[i] * K/total_affinity);
      for j=1 to pop
            //add new frequency in the range
            //    [f[i]-T[i],f[i]+T[i]]:
                  newf[j] = f[i]-
T[i]+(random()*2*T[i])); }
// 3. Normalization step:
//    Inject new random frequencies
//    if less than K frequencies created
//   (newf array)

// 4. Mutation step:
//    Replace randomly K1 of the frequencies:
      K1 = random()*ExplorationFactor*K
```

Fig. 1. The eigenfrequency-based functional model of generative imitation.

We have run spatial simulations of the imitation mechanism in two-dimensional grids consisting of 100 positions. Each position may have one or more agents according to the initial conditions and the evolution of the experiment. In each simulation cycle, every agent encounters another agent according to the specific interaction scheme of the experiment. Each encounter involves the encountered agent acting as a sender of one or more signals and the target agent acting as a receiver that imitates the received signals. This process uses two parameters, the number M of actual external objects referred to by the adult agents and the imitation factor which is the maximum number of signals received and imitated on every encounter (imitation factor $\leq$ M). All adult vocalizations as well as all infant frequencies are initialized randomly in the range [0,1] and M is set to 2. Final affinity with one or more agents is defined as the average affinity for all external objects. Finally, because an agent cannot associate objects and vocalizations, its current signal (word) for a particular object is the last emitted one.

2. Sedentary agents

In the first class of experiments we have experimented with sedentary agents placed in regular 2D grids of the usual type found in the literature. We have used a 10x10 grid and we have run experiments for the periods indicated in the particular result graphs shown.

2.1 Regular grids

Firstly we have experimented with regular 2D grids where each position in the grid is occupied by a single agent interacting only with its eight neighbours.

To avoid possible biases in the results, we also experimented with grids where each position is occupied by the same number of agents (here 3) or by varying numbers of agents (here, uniformly distributed between 2 and 5).

Figure 2 gives the results of a typical run of the latter case. The agents develop very quickly high affinities with members of their own position ("family") and with immediate neighbours, but overall affinity with random agents of the grid does not initially change significantly. However, because the imitation mechanism never actually stops, the frequencies of agents change continuously by slowing adapting to their interacting partners and affinity spreads steadily and therefore in the long run all agents develop higher overall affinities and converge to perfect affinity (one common language) with the whole population, including all immediate neighbours. Thus given sufficient time and if nothing else comes in the way, the imitation mechanism can produce a common agent language in the 2D space.

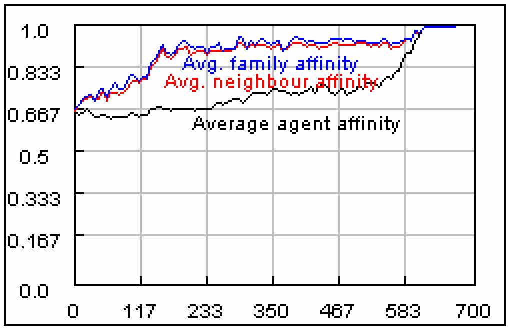

Fig. 2. Sample run of 278 agents distributed on the grid (at least one agent in each position) for 700 steps. The graph shows the average affinity with agents in the same position (family), with immediate neighbours around and with all population of the grid.

2.2 Families

We proceeded then to experiment with more intricate relationships between agents, namely with the introduction of families. In this setup, all agents co-existing on the same position of the grid are assumed to form a real family, so that a high number of agent interactions will be with members of their own family rather than their neighbours from the outside. This interaction model needs an additional **openness** parameter, which is defined as the probability to interact with neighbours rather than with the family. Extreme experiments where openness is very low are possible, in which case the average affinity with neighbours is markedly lower than the average affinity with family. As a result almost every position of the grid may have a different micro-language with lower affinity with neighbouring micro-languages.

Figure 3 gives the results of a typical run. As before, the agents develop very quickly high affinities with members of their own family and with immediate neighbours, but overall affinity with random agents of the grid actually drops because of the abrupt initial development of very high affinities with family and neighbours. However, for the same reasons as before, in the long run all agents develop higher overall affinities and converge to perfect affinity (one common language) with the whole population, including all immediate neighbours. Note that the overall convergence is slower in this case, because of the high inertia of the agents interactions with members of their own family.

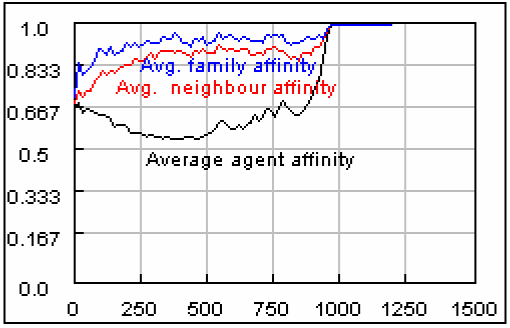

Fig. 3. Sample run of 284 agents forming families and distributed on the grid (at least one agent in each position) for 1500 steps. Openness = 50%. Graph as before.

3. Mobile agents

The next class of experiments involved mobility of agents in the grid. Building on the previous interaction models, we designed new interaction models in which an agent may migrate elsewhere or may instead of interacting with one of his immediate neighbours interact with a random agent from the grid, as if he had moved (momentarily or permanently) elsewhere on the grid.

3.1 Nomadic families

Nomadic families are otherwise normal families that migrate elsewhere on the grid from time to time. Migration is controlled by an additional **exploration** parameter, which is defined as the probability to migrate to a random position on the grid. When a family migrates it may join another family and the two will merge. As a result of such joint motion, some positions on the grid may at times be desert (not inhabited).

Figure 4 gives the results of a typical run. Now the agents develop very quickly high affinities with members of their own family as well as with immediate neighbours and the overall population, before converging in the long run to perfect affinity (one common language) with the whole population. Note how quick the overall convergence is in this case, due to the random interactions with agents from all over the space. Note also that if the migration probability is relatively high, then there is no discrimination between immediate neighbours and remote agents, so the affinities of agents with these two categories will not differ statistically.

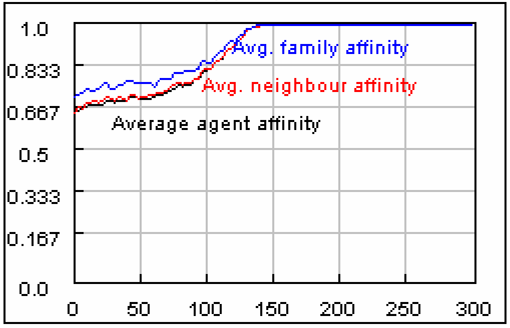

Fig. 4. Sample run of 293 agents forming nomadic families and distributed on the grid (at least one agent in each position) for 300 steps. Openness = 50%, migration probability = 30%. Graph as before.

3.2 Families with independence

In families with a degree of independence, individuals interact with members of their family or with random agents from the grid, rather than with their family or with their neighbours. Such families may be sedentary or nomadic as a group.

Figure 5 gives the results of a typical run. As before, the agents develop very quickly high affinities with members of their own family as well as with immediate neighbours and the overall population, before converging in the long run to perfect affinity (one common language) with the whole population. The overall convergence time is as quick as in the previous case, due to the random interactions with agents from all over the space and as before the affinities of agents with immediate neighbours and all agents on the grid do not differ statistically. However, convergence in this case happens abruptly as a “punctual” event rather than steadily as in the case of nomadism without independence.

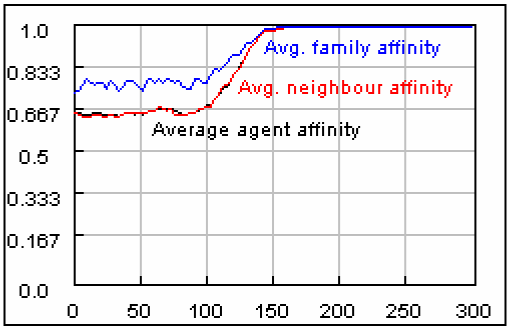

Fig. 5. Sample run of 387 agents forming sedentary families with independence and distributed on the grid (at least one agent in each position) for 300 steps. Openness = 50%. Graph as before.

3.3 Panmixia with or without family

In the final mobility model we have in principle no notion of family, but all individuals may interact freely with any other agent anywhere on the grid. A variant of this model is when members of families may individually migrate and become members of other families elsewhere on the grid.

Figure 6 gives the results of a typical run of sedentary families with independence and individual migration possibility. As before, the agents develop very quickly high affinities with members of their own family as well as with immediate neighbours and the overall population, before converging in the long run to perfect affinity (one common language) with the whole population. Overall, the only induced change by the addition of individual migration possibility to a setup of families with independence is to make the progression to perfect affinity a little less abrupt.

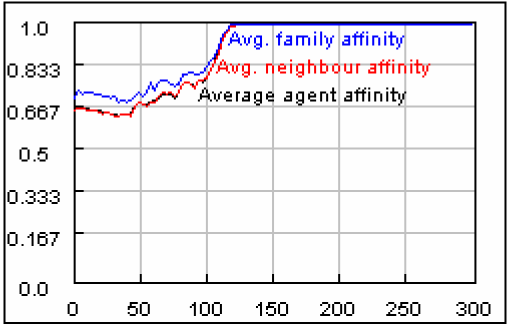

Fig. 6. Sample run of 284 agents forming sedentary families with independence and distributed on the grid (at least one agent in each position) for 300 steps. Openness = 50%, individual migration probability = 30%. Graph as before.

Figure 7 gives the results of two typical runs for the cases of independent agents without family and both without and with individual migration probability. As before, the agents develop very quickly high affinities with all other agents, before converging in the long run to perfect affinity (one common language) with the whole population. As expected there is no difference between the affinities with agents in the same, a neighbouring and a remote position and the convergence is very rapid. Not surprisingly, the case with individual migration probability is far quicker than all the previous cases.

4. Presence of adults

In all the previous cases, and given sufficient time, the population has the potential to stabilize in perfect or near perfect affinity among agents. However, this case is only marginally realistic; in general, we expect on the one hand to have some agents that somehow resist to change, either because they are too old to change or because they are stubborn or because they are “leaders” in the sense that they want to define themselves the interaction context or for other analogous reasons. This may be implemented as agents that simply persist to their own language without imitating anyone. This is again an extreme case, because in reality even such resistent agents are open to members of their own family or other specific agents regarded as kin. However, for the purpose of our modeling scenario, it is safe to explore the various extremities of the setup in order to delimit the bounds of the spatial/social phenomena that can emerge and can be obtained by the imitation model.

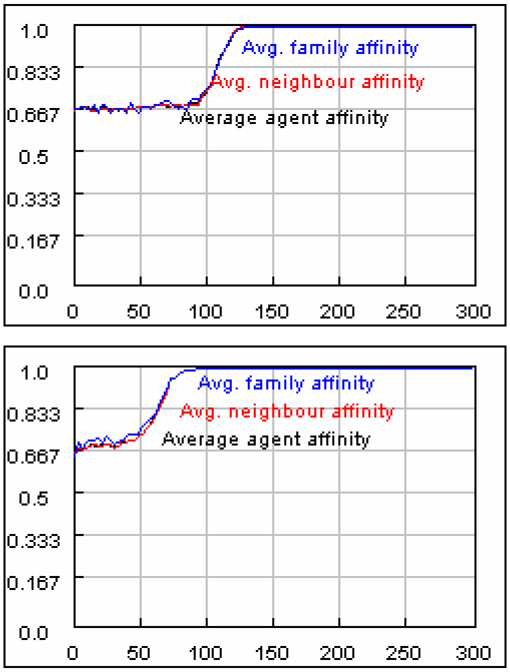

Fig. 7. (top) Sample run of 294 agents with independence but without family and distributed on the grid (at least one agent in each position) for 300 steps. Openness = 50%. (bottom) Sample run of 282 agents with independence and individual migration probability but without family and distributed on the grid (at least one agent in each position) for 300 steps. Openness = 50%, individual migration probability = 30%. Graphs as before.

Figure 8 gives the results of a typical run of sedentary families with adults. As before, the agents develop very quickly high affinities with members of their own family as well as with immediate neighbours and the overall population, and the pattern of convergence is more or less the same as that without adults but much quicker and in the long run converging to imperfect affinity (not shared language) with all classes of agents (family, neighbours, remote agents). Notice the presence of perfect affinity in the largest part of the grid, except around and in the vicinity of adult agents.

Figure 9 gives the results of a typical run of nomadic families with independence and adults. As before, the agents develop very quickly high affinities with members of their own family as well as with other agents (without discriminating between immediate neighbours and remote agents), in the long run converging to imperfect affinity (not shared language) with all classes of agents (family, neighbours, remote agents).

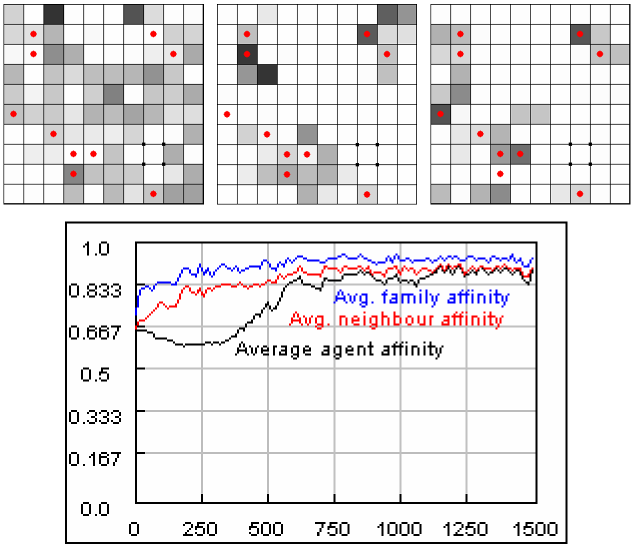

Fig. 8. Sample run of 287 agents forming sedentary families with adults and distributed on the grid (at least one agent in each position) for 1500 steps. (top) Average neighbour affinity per position at t=100, t=700, t=1500. The color coding is such that the 0-1 affinity scale is mapped linearly to the black-white gray shade scale so that 0=black and 1=white. Positions with at least one adult are depicted with a (red) dot (bottom) Graph as before.

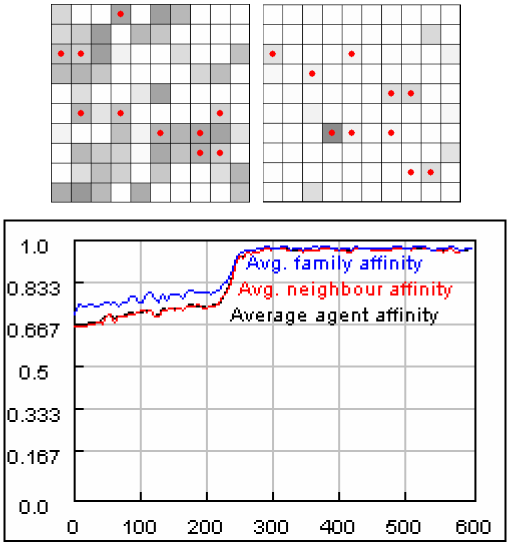

Fig. 9. Sample run of 273 agents forming nomadic families with independence and adults and distributed on the grid (at least one agent in each position) for 600 steps. (top) Average neighbour affinity per position at t=100, t=600. Color coding as before. Positions with at least one adult are depicted with a (red) dot (bottom) Graph as before.

5. Changing wor(l)ds

Another realistic scenario involves introduction from time to time of new objects for which new words have to be invented. Again we examine the extreme case where occasionally a new word has to be invented from scratch, rather than the more realistic case where the new word may be more or less derived from previous words and independently of the fact that any new word has to fit the actual phonetic environment of the overall language. This can be implemented in our population very simply by occasionally reinitializing all individual agents' words (signals) for one of the objects (as if a new word had to be devised for a new object that replaces one of the old ones).

Figure 10 gives the results of a typical run of sedentary families with language renewal as described. As expected, the agents tend to develop very quickly high affinities with members of their own family as well as with immediate neighbours and the overall population, with convergence curves being "reinitialized" every time word renewal takes place. Note that, as time goes by, re-convergence is quicker because affinity is slowly built and maintained in high levels despite such local disturbances. As a result, all events necessitating re-adaptation are handled more quickly and more efficiently. This is a crude demonstration of the Baldwin effect [12] within the imitation mechanism in the sense that adaptation makes subsequent adaptation more easy, because subsequent "generations" of frequencies within the agent have a better potential to adapt easily and quickly to new signals. Apparently, because of language renewal, affinities are high enough but not perfect thus language between agents is not fully shared (see fig. 10 bottom).

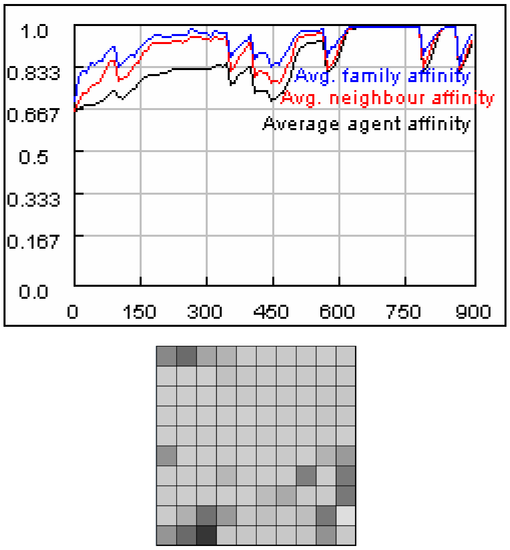

Fig. 10. Sample run of 285 agents forming sedentary families and distributed on the grid (at least one agent in each position) for 900 steps with occasional language renewal. (top) Graph as before. (bottom) Average agent word (signal) per position at t=900. The color coding is such that the 0-1 signal scale is mapped linearly to the black-white gray shade scale so that 0=black and 1=white.

6. Conclusion

We have presented a proto-imitation mechanism with the initial potential to produce stable mono-language populations if given enough time resources. We have shown how various mobility and interaction schemes may result in the long run in convergence to a unique language shared among agents and how the convergence speed and pattern differ across the various schemes. We have also shown that imitation-proof agents may completely change the emerging spatial patterns and trigger the formation of micro-languages or persistent localized linguistic communities. Moreover, convergence may be perturbed by eventual needs for language renewal, because of cultural innovation or for other similar reasons. Consequently, a sole mechanism of reactive (hardwired) proto-imitation may produce many different language configurations within a delimited physical space for variations of basic parameters such as mobility factors, agent change inertia and environmental perturbations. Higher-level mechanisms are not necessary to produce this wealth of configurations. Further investigation into proto-imitation will be in the direction of a deeper examination of the internals of the presented and other similar mechanisms as well as in the direction of imitation coupled with development in a causal way.

References

[1] Hurley, S., Chater, N.: Perspectives on imitation in animals and humans. MIT Press, Cambridge (2005)

[2] Suzuki, J., Kaneko, K.: Imitation games. Physica D 75, 328-342 (1994)

[3] Eckerman , C.O., Stein, M.R.: How imitation begets imitation and toddlers' generation of games. Devel. Psychol. 26, 370-378 (1990)

[4] Gergely, G. Bekkering, H., Király, I.: Rational imitation in preverbal infants. Nature 415, 755 (2002)

[5] Caldwell, C.A., Whiten , A.: Evolutionary perspectives on imitation: Is a comparative psychology of social learning possible? Anim. Cogn. 5, 192-208 (2002)

[6] Oliphant, M.: Rethinking the language bottleneck: Why don't animals learn to communicate?. Presented at the 2nd Intern. Conf. on the Evolution of Language (1988)

[7] Krakauer, D.C.: Selective imitation for a private sign system. J. Theor. Biol. 213, 145--157 (2001)

[8] Stieglitz Ham, H., Corley, M., Rajendran, G., Carletta, J., Swanson, S.: Imitation of meaningless gestures in individuals with Asperger syndrome and high-functioning autism. J. Autism Dev. Disord. 38, 569-573 (2008)

[9] Karmiloff-Smith, A.: Development itself is the key to understanding developmental disorders. Trends in Cognitive Sciences 2, 389-398 (1998)

[10] Edelman, G.M.: Neural darwinism. Basic Books, New York (1987)

[11] Tzafestas, E.S.: On Modeling Proto-Imitation in a Pre-Associative Babel. From Animals to Animats 10, Springer-Verlag NLAI 5040, 477-487 (2008)

[12] Baldwin, J.M.: A new factor in evolution. Amer. Natur. 30, 441-451 & 536-553 (1896)