Bart de Boer & Paul Vogt (1999) Emergence of Speech Sounds in Changing Populations In: Dario Floreano, Jean-Daniel Nicoud & Francesco Mondada (eds.) Advances in Artifical Life, Lecture Notes in Artificial Intelligence 1674, Berlin Springer Verlag, pp. 664–673

Emergence of Speech Sounds in Changing Populations

Bart de Boer, Paul Vogt

AI-Lab
Vrije Universiteit Brussel
Pleinlaan 2
1050 Brussel
{bartb,paul}@arti.vub.ac.be

Abstract: This paper shows that realistic and coherent vowel systems can emerge from scratch in a population of agents that imitate each other under human-like constraints of production and perception. The simulation is extended so that populations can change; old agents can be removed, and new agents can be added. In these circumstances vowel systems can also emerge and be preserved. It is shown that sometimes an age structure in the population can improve preservation of the vowel systems.

1 Introduction

The study of human languages is not only concerned with describing the huge variety of phenomena that are encountered in human languages, but also with finding explanations for similar phenomena that recur in languages that are neither related historically nor geographically. Such phenomena are called language *universals*, or, because there are always exceptions, universals tendencies. This paper is concerned with explaining the universal tendencies of human vowel systems. Although the human vocal tract is capable of producing a wide variety of different vowels (at least 45 different basic vowel qualities [10]) one finds that all human languages use only a very small subset of these. Furthermore, these subsets are not chosen randomly, but they exhibit remarkable regularities. When one looks at the vowel systems of widely different languages, such as, for example the 451 languages in the UPSID (UCLA Phonological Segment Inventory Database [12, 13]) one finds that the maximum number of vowel qualities in any language is 15 in Norwegian [10] while the minimum number is 3 (although there are languages that are reported to have only two vowels [10]). One also finds that certain vowels, for example [i], [a] and [u] occur very frequently (in 87%, 87% and 82% of the languages in UPSID, respectively) while others, such as [y], [œ] and [ø] occur only rarely (in 5%, 2% and 3% of the languages, respectively). But one also finds that the structure of vowel systems tends to be regular. For example, if a language contains a front, unrounded vowel of a certain height, for example [ɛ] (which occurs in 41% of the languages in UPSID) one also tends to find the rounded back vowel of the corresponding height, in this case [ɔ], which occurs in 36% of the languages, but in 73% of the languages that also contain an [ɛ]. Many such regularities can be found and typologies of vowel systems have been based on them [2, 16].

A possible explanation of these phenomena is that they occur because humans have an innate, cognitive disposition towards using certain speech sounds and combinations of speech sounds. Innate distinctive features, rules and markedness constraints are proposed. Distinctive features are properties (often binary) of speech sounds that are used to distinguish them from each other. Rules determine in what way abstract distinctive features can be combined into phonemes (the basic abstract speech sounds that can distinguish meanings of words), and how the phonemes are transformed into an actual speech signal. Some features are more marked than others, meaning that they are less likely to occur. There are a number of problems with this theoretical framework, but the most important of these is that it does not actually explain anything at all. The features, rules and markedness constraints are inferred from observations of human sound systems. It would thus be circular to use them to explain the very phenomena from which they have been inferred.

An independent explanation of the structure of vowel systems can be derived from functional criteria, such as acoustic distinctiveness and articulatory ease. Lindblom and Lilencrants [11] have shown that optimization of acoustic distance between the members of a vowel repertoire results in predictions of the most frequently occurring vowel systems in human languages. Subsequent improvements on their model (see e.g. Schwartz *et al.* [15]) have confirmed their results. There are two problems, however. The first problem is that the number of vowels in the system has to be fixed beforehand. The second problem is that the model depends on explicit optimization. It shows that human vowel systems are in general optimized for acoustic distinctiveness, but it does not give an explanation of how the optimization takes place in human languages. It is clear that human language users do not optimize the sound system of their language explicitly. In fact, when learning a language, children imitate their parents as accurately as possible, more accurately in fact than strictly necessary for successful communication. This can be observed from the fact that children do not only learn the language of their parents, but also their dialect. Although people speaking different dialects can often understand each other perfectly, they are nevertheless aware of subtle distinctions in pronunciation. The optimization model is incomplete. It explains *why* vowel systems are the way they are—they are optimized—but not *how* they have become optimized.

The work described in this paper is based on the theory that self-organization in a population of language users drives sound systems towards optimality. It is based on the theories of Luc Steels of language as a complex dynamic system [17, 18, 19] and is also related to other work on the origins and the evolution of language [7, 9]. In the theories of Luc Steels language is considered as much to be a phenomenon of a population as it is knowledge of individuals. This approach does not consider language in terms of abstract ideal knowledge of an individual, nor does it deal with idealized speaker-hearer interactions. Rather, it stresses the fact that language is an open, distributed system, where no member of the population has complete knowledge of the language, let alone control, where interactions can be messy and incomplete, where new words, meanings and constructions can enter the language, and where individual speakers of the language can enter and leave the population. Coherence in the language is maintained through self-organization. Change, which is viewed as an inherent property of language, is driven by speech errors, by contradicting drives to minimize articulatory effort and to maximize communicative success and by conscious creation of new words and expressions. As language is seen as a complex dynamic system, the only way to explore it is with computer simulations.

In the work described in this paper a population of agents that can produce, perceive and learn speech sounds in a human-like way is modeled. The agent's goal in life is to imitate the other agents in the population as well as possible. The work presented here is not the first work that tries to explain the universal tendencies of human vowel systems through the interactions of agents in a population. First Glotin [2] and later Berrah [1] have both built computer simulations of populations of agents that imitate each other's speech sounds. However, in their simulations, the number of vowels in each agent is fixed beforehand and the main process that shapes the vowel systems is still explicit optimization. The interactions between the agents only serve to make the vowel systems in the population coherent, not optimized.

In previous papers, [3, 4] it has been shown that the proposed model really results in coherent and realistic vowel systems. Meanwhile the research has progressed to larger populations and more realistic results (compare, for example the figures in [3] with the figures in this paper.) Many (extremely boring) experiments have been done to test the sensitivity of the system to parameter changes and small changes of algorithm, but the self-organisation has been shown to be remarkably robust. The interested reader is referred to [5]. The main focus of this paper will be on investigating the results of population dynamics. Human languages are spoken in open populations. Old speakers of the language can die and children have to learn the language. With the computer simulations it can be investigated what happens in the case of a changing population. This is a test for the model: will the agents preserve vowel systems, just as happens in human language, and will it still be possible to have vowel systems emerge from scratch? Also socio-linguistic hypotheses that are hard to test in reality can be investigated with this model: what happens with different population replacement rates and is there an advantage in having speakers of different ages learning with different rates? These questions are treated in the section 3.

2 The Simulation

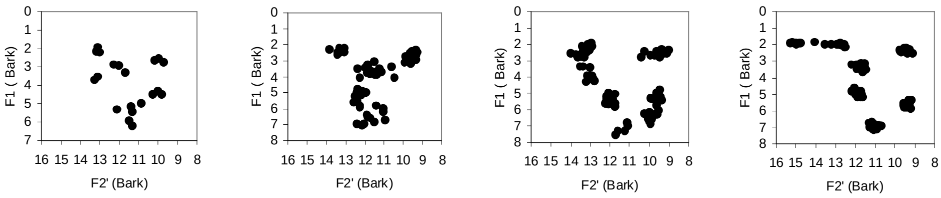
The simulation models the interactions of agents in a population. Each agent is equipped with a realistic articulatory synthesizer for producing vowels, a model of human perception of speech signals for calculating the distance between different signals and an associative memory in which they can store the associations between articulatory prototypes, acoustic prototypes, the number of times the prototypes have been used and the number of times they have been successfully used. Whenever an agent hears an acoustic signal, it calculates the distance between this signal and all its acoustic prototypes and considers the prototype that is closest to the signal as the one that it has recognized. Vowel signals are represented as the frequencies of the first four peaks (formants) of the acoustic spectrum of the vowel. Whenever an agent wants to produce a vowel, it takes the vowel's articulatory prototype and synthesizes its formants with the articulatory synthesizer. Some noise is added (by shifting the formant frequencies randomly) so that an articulatory prototype is never exactly realized the same way acoustically. The distance between a signal and the acoustic prototypes of vowels is calculated in a two-dimensional space with the logarithm of the frequency of the first formant as one dimension and the logarithm of the frequency of the *effective second formant* as the second dimension. The distance is calculated as a weighted Euclidean distance where the effective second formant frequency is weighted as 30% of the first formant frequency.

The effective second formant frequency is a non-linearly weighted sum of the second, third and fourth formant frequencies. It is based on the observation that due to the limited resolution of human perception of high frequencies, signals that have multiple peaks at high frequencies are perceived as identical to signals that have only one peak at high frequencies. The position of this single peak is a function of the position of the peaks in the original signal. In this paper it is calculated by a function that is based on Mantakas *et al.*'s [14] weighting function. It would have been better to calculate the distance in the original 4-dimensional acoustic space, but all available models of human vowel perception were based on the effective second formant.

The interactions between the agents have been called *imitation games*. For each imitation game, two agents are selected from the population at random. One of the agents will play the role of *initiator*, the other will play the role of *imitator*. Basically, the initiator selects an acoustic vowel prototype form its repertoire at random and synthesizes the acoustic signal for this vowel. The imitator listens to this signal and finds its acoustic prototype that is closest to it. It then synthesizes the articulatory prototype that is associated with this acoustic prototype. The initiator listens to this signal and finds *its* closest acoustic prototype. If this is associated with the articulatory prototype it originally produced, the imitation game is successful. If not, it is a failure. This information is communicated non-verbally to the imitator.

In reaction to the imitation game, both the initiator and the imitator update their repertoire of vowels. The use and success (if the game was successful) counts of the vowel prototypes that were produced are increased. Whenever the success/use ratio of a vowel drops below a certain value, (0.7 throughout the paper) and its use count is sufficiently high (5 throughout the paper) it is removed from an agent's repertoire. The imitator can make subsequent changes to its repertoire. If the imitation game was successful, the imitator shifts the articulatory prototype it used so that its acoustic realization matches the signal that was heard more closely. This is done in order to increase the coherence in the population. If the imitation game was a failure, and if the vowel that the imitator used was successful, the imitator adds a new vowel to its repertoire that closely matches the signal that was heard. The articulatory prototype of this vowel is determined by a hill-climbing heuristic: the agent makes a first guess, produces it, listens to itself, makes small articulatory adjustments and iteratively improves its first guess. The vowel is added because the other agent probably had two vowel prototypes at a location where this agent had only one, thus causing confusion. If the success/use ratio of the vowel that was used is low, however, it is moved closer to the perceived signal in an attempt to improve it. Finally, both the imitator and the initiator merge vowels that come too close together in acoustic or articulatory space.

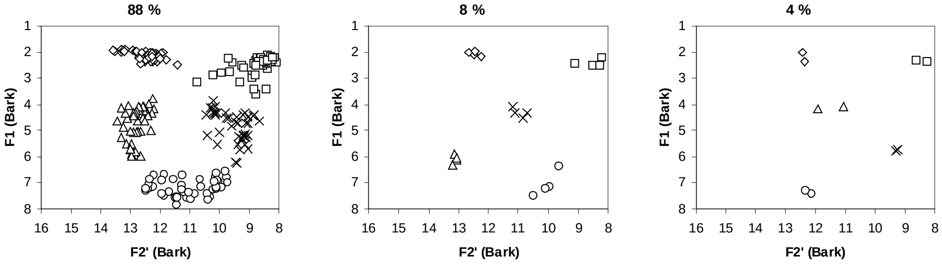
When these interactions are iterated a large number of times in the population, realistic vowel systems emerge. The process of emergence is illustrated in figure 1. In this figure, the acoustic vowel prototypes of all the agents in a population of twenty agents are plotted in the acoustic space formed by the first formant (on the vertical axis) and the effective second formant (on the horizontal axis) in the logarithmic Bark frequency scale. Note that the scales on the axes increase from top to bottom and from right to left, respectively. This has been done in order to get the vowel prototypes in a



configuration that corresponds to the way phoneticians usually plot vowels, with front vowels towards the left and high vowels towards the top of the graph. Note that due to articulatory limitations, the acoustic space that can be reached by the agents is limited to a triangular region with the apex at the bottom of the graph. Whenever there is a cluster in the graph, this means that most agents in the population have an acoustic prototype near this position. The number of clusters in the graph indicates the number of vowel prototypes of each agent.

The frames in figure 1 show the vowel systems in the population after 50, 500, 2500 and 15 000 imitation games from left to right. It can be seen that initially, the vowel prototypes are spread randomly through the available acoustic space. After 500 imitation games, some structure becomes visible. Vowel prototypes cluster together and multiple clusters start to form. After 2500 imitation games, the clusters are dispersed more evenly throughout the available acoustic space and after 15 000 imitation games, the clusters have become compact and more or less evenly dispersed. When this stage has been reached, the agents' vowel systems remain relatively stable. Although they never stop changing completely, the possible changes are relatively small, consisting mostly of small shifts of the clusters and the occasional merging of existing clusters or emergence of a new cluster if space is available. The acoustic noise setting for this graph was 10%.

In order to get an idea of the realism of the emerging vowel systems, several measures can be defined, for example the energy of the vowel systems as defined by Liljencrants and Lindblom [11], the average success of imitation or the average number of vowels per agent. These will be used later in this paper. However, direct inspection and classification of the agents' emerging vowel systems is also a good way to get an idea of the realism of the emerging systems. An example of the classification of emerging vowel systems with five vowels is given in figure 2. This figure was formed by doing 25 000 imitation games in 100 populations of 20 agents. The acoustic noise was set to 15%. Of the emerging vowel systems, 49 consisted of five vowels (the



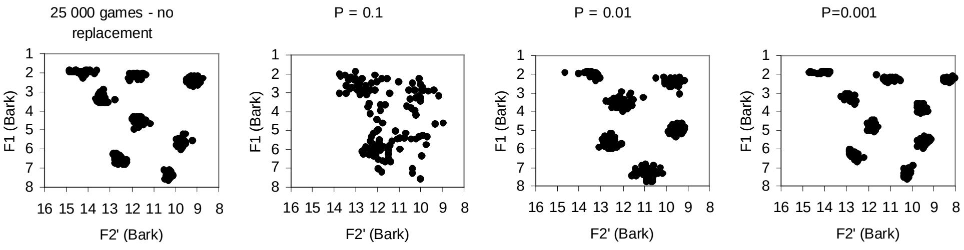
number of vowels in the final vowel systems is not determined beforehand). From each of these populations, one random agent was selected. The vowel systems of these agents were classified according to the relative position of the vowels in the system in the same way as Crothers [2] has classified vowel systems of human languages. It can be seen that the different types occurred (from left to right) in 88%, 8% and 4% of the cases. When the results are compared with those of human languages one finds a very good match. Schwartz *et al.* [15] found that of the languages in UPSID, 89% were of the leftmost type, and the middle and rightmost type both occurred in 5% of the cases. Equally good matches were found for other numbers of vowels, although there were some problems for very low and very high numbers of vowels (see [5, chapter 6] for more details).

Apparently the model makes realistic and coherent vowel systems emerge in a population by no other process than local interactions between the agents under articulatory and acoustic constraints.

3 An Open Population

In the real world, populations of language users are not static. People can die, and children get born. Old speakers continually leave the population and new speakers enter it. This usually does not disrupt the language. Research on the social dynamics of language using traditional methodologies is extremely difficult, as the factors influencing language change are very hard to identify, let alone that they can be controlled. In the artificial setting of the imitation, game, carefully controlled experiments with changing populations can be performed. From this more insight can perhaps be obtained in the social dynamics of real language. A number of experiments are presented that give an idea of the possibilities of the application of computer simulations to the study of sound change. Other parts of language, most notably the emergence of lexicon and semantics, have already been studied by other researchers (see e.g. [20]).

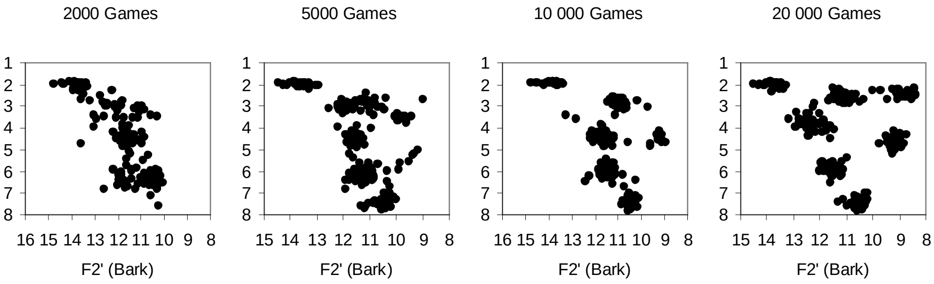
The population is changed by removing and adding agents at random. Making this process random was deemed to be the most realistic (as in human populations birth and death are basically also random processes) and the least likely to introduce artifacts due to regularities in the replacement. In every imitation game in which it participates, an agent has a certain probability of being removed. Also, in every imitation game there is a small probability of an agent being added to the population. Of course, the newly added agent's vowel repertoire is still empty. It will have to learn the sound system of the population to which it is added. The probabilities of adding new agents and removing old ones will have to be equal so that the number of agents in the


25 000 games - no replacement

P = 0.1

P = 0.01

P=0.001

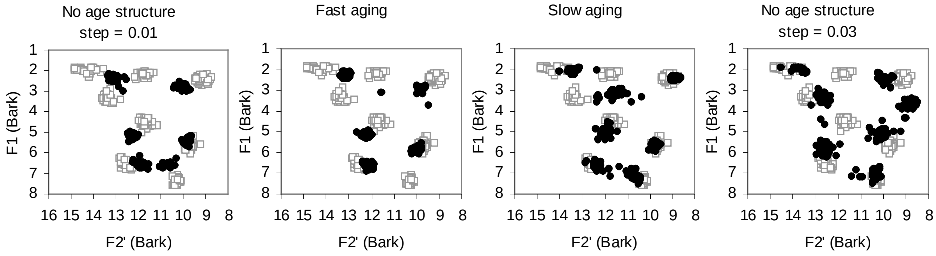


population remains relatively stable. However, depending on the value of the probabilities, the rate of replacement in the populations can be high or low.

A comparison between different rates of replacement is shown in figure 3. The leftmost frame of this figure shows the original vowel system that was used as the starting configuration for all other experiments presented in this section. This vowel system was obtained from running 25 000 imitation games in a population of 50 agents with an acoustic noise of 10%. The second graph from the left shows the vowel system of a population of agents in which the probability of adding or removing agents was 0.1. It shows the situation after 2500 imitation games. With these parameter settings, it was unlikely that any agent from the original population still remained in the final population. The third graph from the left shows a vowel system from a population with a probability of adding or removing agents of 0.01. As it is unfair to compare populations that have undergone unequal numbers of replacements, this population's vowel system is shown after 25 000 imitation games. The rightmost frame shows the vowel systems of a population with a replacement probability of 0.001 after 250 000 imitation games. From these graphs it is clear that the higher the replacement rate, the worse the agents are able to preserve the original vowel system. This has nothing to do with the absolute number of replacements. These are approximately equal for all graphs shown. Rather, the time in which agents stay in the population (and therefore the time they have to learn the vowel system) determines how accurately a vowel system is preserved. This does not mean that the systems with a high replacement probability do not become stable. After a while they settle in a state in which they contain fewer prototypes and the prototype clusters are more dispersed.

In fact, coherent and realistic vowel systems can even emerge from scratch in an open population. This is illustrated in figure 4. It shows a similar emergence of a vowel system as figure 1. The replacement probability was 0.01 and the acoustic noise was 10%. The emergence is slightly slower than in figure 1, but this is due to the fact that the population is larger—50 agents instead of 20. There are fewer vowel prototypes per agents and the prototype clusters are slightly bigger. Nevertheless, a coherent, successful and realistic vowel system emerges even though at the end of the run, no agent that was present at the beginning is present anymore.

In the experiments shown so far, there is no cost for articulatory effort. If an agent wants to imitate a signal it has heard, it can talk and listen to itself an unlimited number of times in order to perfect its vowel prototypes. In reality, such articulatory effort probably has a certain cost. Therefore an experiment was done in which the number of times an agent can talk to itself for improving a new sound is limited to 10. The



interesting finding of this experiment was that it then becomes advantageous to introduce an age structure in the population, so that younger agents can change their articulatory vowel prototypes more quickly than older agents. The older agents thus provide a stable target to which the younger agents can move their vowels. The results are shown graphically in figure 5. The original vowel systems are shown in gray, the final vowel systems, obtained after 15 000 imitation games are shown in black. In the leftmost and rightmost frames, the populations do not have an age structure. The difference between the two figures is that the step size with which the articulatory vowel prototypes were improved was 0.01 in the leftmost and 0.03 in the rightmost frame (0.03 was the value that was used in all other experiments presented so far). In the two middle frames, young agents used a step size of 0.03 and old agents tended towards step size of 0.03. The agents' step sizes changed in the following way:

$$\varepsilon_t \leftarrow \varepsilon_{t-1} + \alpha(0.01 - \varepsilon_{t-1}) \qquad \textbf{(1)}$$

Where ε_t is the step size at time *t* and α is a parameter that determines the speed with which the agents change their step sizes. The value of α was 0.1 for the left center frame and 0.01 for the right center frame. The difference in performance between the different systems in the figure are a bit hard to judge, therefore table 1 gives the measures of the system. The success is the average over the communicative success of the agents in the population. The energy of the vowel systems is calculated according to [11] as the sum over the reciprocal of the squared distance between all the vowels in an agent's vowel system. The size is the average number of vowels in the agents' vowel systems and the similarity is the communicative success that can be achieved between an agent that has the original vowel system and an agent that has the vowel system that emerges at the end of the simulation run.

Population:	$\varepsilon_0 = 0.01$ $\varepsilon_\infty = 0.01$ $\alpha = 0$	$\varepsilon_0 = 0.03$ $\varepsilon_\infty = 0.01$ $\alpha = 0.1$	$\varepsilon_0 = 0.03$ $\varepsilon_\infty = 0.01$ $\alpha = 0.01$	$\varepsilon_0 = 0.03$ $\varepsilon_\infty = 0.03$ $\alpha = 0$
Success:	0.8531 ± 0.040	0.7701 ± 0.035	0.7930 ± 0.041	0.8041 ± 0.041
Energy:	4.04 ± 0.59	5.55 ± 1.10	5.10 ± 0.95	3.83 ± 0.72
Size:	4.77 ± 0.40	5.66 ± 0.63	5.61 ± 0.58	5.14 ± 0.50
Similarity:	0.7347 ± 0.023	0.8231 ± 0.028	0.8235 ± 0.032	0.7884 ± 0.032

Table 1: Measures of the populations with age structure.

The statistics were obtained from 100 runs of each system. Shown are the averages for every measure and their standard deviations. Using the Kolmogorov-Smirnov test it can be shown that the vowel system sizes and the similarity measures in the populations with age structure are significantly higher (at the 1% level) than in the populations without age structure, indicating that the original vowel system is preserved better. The fact that the success is lower for the two systems with age structure is not important, as success always tends to be a little lower for larger systems.

These experiments show that the fact that old agents learn less well than young agents can sometimes be beneficial.

4 Conclusion and Future Work

The experiments presented here have shown that a) coherent and realistic vowel systems can emerge from scratch in a population of agents that imitate each other under constraints of production and perception and b) that the same simulation can be used to investigate the consequences of population dynamics on the emergence and the preservation of sound systems. The first result is of direct relevance for phonetics. It shows that the agents do not need innate predisposition towards certain structures in order for these structures to emerge. Traditionally it has been assumed that the fact that certain phonetic and phonological structures are found in unrelated human languages is proof for the theory that these structures are innate. This research shows that they can as well be the result of self-organization in a population. The second result is not directly relevant for linguistics, but it does show that the methodology can be used successfully for investigating phenomena as complex as dynamics of populations of language users. It incidentally also shows that the self-organizing emergence is extremely robust and happens even if the population is changing. Finally, the results of the simulations with age structure show that the fact that older language users learn less quickly than younger ones might not just be an unfortunate consequence of aging, but might help to preserve sound systems (and possibly other aspects of language) across the generations.

Of course, the simulation is still very simple. From the point of view of phonetics, one of the main things that needs to be investigated is whether it also works with more complex sounds. If one wants to investigate sound change seriously, one needs to take interactions between different speech sounds into account. Also, most of the more interesting universal tendencies of human sound systems have to do with consonants and combinations of sounds.

From the point of view of population dynamics, more complex experiments with spatial distribution of agents, or with populations with different sound systems that come into contact with each other can be conceived. Some of these experiments have already been done for other aspects of language (see e.g. [20]).

In any case, the use of computer simulations of populations of agents that learn and use realistic speech sounds has proven to yield interesting results. The experiments done thus far are only a beginning, but they show that the technique is promising and can help to shed new light on the complexities of human language.

5 Acknowledgements

This research was done at the artificial intelligence laboratory of the Vrije Universiteit Brussel. It is part of an ongoing research project into the origins of language and intelligence. Funding was provided by the GOA 2 project of the Vrije Universiteit Brussel. Part of the work was done at the Sony computer science laboratory in Paris, France. My thanks also go to an anonymous reviewer for providing useful suggestions for improving the paper.

References

1. Berrah, Ahmed Réda (1998) Évolution Artificielle d'une Société d'Agents de Parole: Un Modèle pour l'Émergence du Code Phonétique, Thèse de l'Institut National Polytechnique de Grenoble, Spécialité Sciences Cognitives.
2. Crothers, John (1978) Typology and Universals of Vowel systems, in: Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.) Universals of Human Language, Volume 2 Phonology, Stanford: Stanford University Press pp. 93–152
3. de Boer, Bart (1997) Generating vowel systems in a population of agents. In Phil Husbands and Inman Harvey (eds*.) Fourth European conference on Artificial Life*, Cambidge (MS) MIT Press pp. 503–510
4. de Boer, Bart (1998) Emergence of Sound Systems through Self-Organisation. In Han La Poutre' and Jaap van den Herik (eds.) *NAIC '98 Proceedings*, Amsterdam pp. 37–46
5. de Boer, Bart (*in preparation*) *Self-organisation in vowel systems*, Vrije Universiteit Brussel Ph. D. Thesis (to be defended 4-6-1999)
6. Glotin, Hervé (1995) La Vie Artificielle d'une société de robots parlants: émergence et changement du code phonétique. DEA sciences cognitives-Institut National Polytechnique de Grenoble
7. Hurford, Jim (to appear) Social transmission favours linguistic generalization. In Chris Knight & Jim Hurford (eds.) The evolution of language, Cambridge: Cambridge University Press.
8. Jakobson, Roman & Morris Halle (1956) Fundamentals of Language, the Hague: Mouton & Co.
9. Kirby, Simon (to appear) Syntax without natural selection: how compositionality emerges from vocabulary in a population of learners. In Chris knight & Jim Hurford (eds.) The evolution of language. Cambridge: Cambridge University Press.
10. Ladefoged, Peter & Ian Maddieson (1996) The Sounds of the World's Languages, Oxford: Blackwell.
11. Liljencrants, L. & Björn Lindblom (1972) Numerical simulations of vowel quality systems: The role of perceptual contrast, Language 48 pp. 839–862.
12. Maddieson, Ian (1984) Patterns of sounds, Cambridge University Press.
13. Maddieson, Ian & Kristin Precoda (1990) Updating UPSID, in: UCLA Working Papers in Phonetics 74, pp. 104–111
14. Mantakas, M, J.L. Schwartz & P. Escudier (1986) Modèle de prédiction du 'deuxiéme formant effectif' F2'—application à l'étude de la labialité des voyelles avant du français. In: Proceesings of the 15th journées d'étude sur la parole. Société Française d'Acoustique, pp. 157–161.
15. Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée & Christian Abry (1997), Major trends in vowel system inventories, Journal of Phonetics 25, pp. 233–253

16. Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée & Christian Abry (1997), The Dispersion-Focalization Theory of vowel systems, Journal of Phonetics 25, pp. 255-286
17. Steels, Luc (1995) A Self-Organizing Spatial Vocabulary, Artificial Life 2(3), pp. 319–332.
18. Steels, Luc (1997b) The Synthetic Modelling of Language Origins, Evolution of Communication 1(1): pp. 1–34
19. Steels, Luc (1998b) Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation, in: James R. Hurford, Michael Studdert-Kennedy & Chris Knight (eds.) Approaches to the Evolution of Language, Cambridge: Cambridge University Press pp. 384–404
20. Steels, Luc & Frédéric Kaplan (1998) Spontaneous Lexicon Change. In: *Proceedings of COLING-ACL 1998*, Montreal, pp.1243–1249.
21. Vanvik, A (1972) A phonetic-phonemic analysis of Standard Eastern Norwegian, in: Norwegian Journal of Linguistics 26, pp. 119–164