Published in *Advances in Artificial Life*, Floreano, D., Nicoud, J-D. and Mondada, F. (eds), Proceedings of the 5th European Conference in Artificial Life, ECAL'99, Laussanne, Switzerland, September 1999. Springer. LNAI vol. 1674

Modelling the Evolution of Linguistic Diversity

Daniel Livingstone and Colin Fyfe

Applied Computational Intelligence Research Unit,
Dept. CIS, University of Paisley, Paisley, PA1 2BE, UK
{livi-ci0,fyfe-ci0}@paisley.ac.uk

Abstract. Some recent Artificial Life models have attempted to explain the origin of linguistic diversity with varying conclusions and explanations. We posit, contrary to some existing Artificial Life work, that linguistic diversity should naturally emerge in spatially organised populations of language learners, and this is supported by our experimental work and by recent literature.

1 Introduction

There exists a rich literature on linguistic diversity, much focussed on tracing the roots and relationships between languages and regional variations in dialect. This division between language and dialect is not clear-cut. Two different languages may be mutually intelligible while two dialects of the one language may possess a low degree of mutual intelligibility [1]. *Dialect continua* make the distinction between languages and dialects harder still. These chains of dialects can cover large areas and cross many national boundaries, each dialect intelligible to speakers of neighbouring dialects. Yet dialects at the ends of the chains may be mutually incomprehensible.

Artificial Life inspired research has shown how populations of agents can negotiate a language and how communication schemes can evolve. Generally, these attempt to negotiate a single language common to all agents. There exists a small body of Artificial Life based work which studies linguistic diversity. These have tended to ignore the simple hypothesis that linguistic diversity may be the non-adaptive result of spatially distributed learning. The model presented here investigates this.

2 Dialect in Models of the Evolution of Language

A small number of Artificial Life papers address the issues of language diversity and innovation directly. Maeda and Sasaki, [2], consider the effects of contact between different linguistic groups, demonstrating the subsequent language re-organisation.

Kirby, [3], studies the evolution of Universal Grammar and shows, implicitly, how spatial organisation can affect the evolution of languages. Hashimoto and Ikegami, [4], show how pressure for greater expressive and parsing ability leads to increasingly complex grammars which are able to generate and interpret wider ranges of signals.

Arita and Taylor [5] argue that spatial distribution can explain the origin of linguistic diversity. Inheritance and mutation are important factors for generating diversity within this model, where learning reduces diversity. Yet in human language it is the cultural, not genetic, factors which determine which language is learned. In Arita and Koyama [6] an innate language is used to study the evolutionary dynamics of vocabulary sharing. One conclusion is that mutation rate is important in defining the class of vocabulary sharing. This is not surprising with a heritable vocabulary, but what this represents linguistically is not explained. Steels and Kaplan [7] demonstrate stochasticity, from errors in language use, as a source of change in language. The formation of multiple, distinct, dialects is not covered in this study in which the agents are not spatially organised. While Nettle and Dunbar [8] cite examples and use simulation to demonstrate that dialect has adaptive value in a social context, we will show that adaptive explanations are not required for evolution of linguistic diversity.

3 A Model for Studying the Evolution of Dialect

We adapt an earlier model [9,10] in which the emergence of dialects was observed but not studied, focussing now on the evolution of dialects. Genetic evolution of agents is removed, and learning modified to take place between successive generations of agents. Our model of language is a greatly simplified one, in which agents learn to map signals sent by others to internal states. These internal states could be said to relate to 'meanings' or to external observations.

A language agent is modelled by a fully connected Artificial Neural Network, (ANN) which has two layers of nodes, with N inputs (the internal state, ***x***) and M outputs (the signal, ***y***). The output layer has an additional bias node.

The internal state is a sparse bipolar vector (only one node set to +1). This is fed forward to determine the agent's signal for that state, each output being thresholded to a bipolar value, (1). Signals are arbitrary bipolar vectors. For interpretation of signals, competition is applied at the internal state layer, so any signal fed back from the language layer corresponds to only one internal state. For M output neurons, there are 2^M possible signals in the language, and for N input neurons there are N possible states. In experiments detailed here, M = N = 3. Agents have three possible internal states and eight possible signals, providing redundancy in the signalling ability.

$$y_j = \sum_{i=0}^{N-1} x_i w_{ij} \tag{1}$$

$$\text{then } Y_j = 1 \text{ if } y_j \geq 0, \; Y_j = -1 \text{ if } y_j < 0$$

$$\Delta w_{ij} = \eta(x_i - x_i')y_j \tag{2}$$

[11] suggests using transmission behaviour to train language reception and reception behaviour to train language production. Accordingly, the signal from a teacher is presented at the output layer of the learner and fed back to produce an internal state. The error between the teacher's and the learner's internal states is used for learning (2), performed when the learner misclassifies the signal.

The first generation of agents are initialised with random weight values and then provide training for one another for *t* training rounds, as in the following algorithm:

```
For each agent (in random order)
   Pick another agent to be teacher
   Generate a training signal from teacher
   Train pupil on signal and state
```

When each generation is created, the agents are born with zero valued weights. The new agents learn first from agents in the parent generation for *t* training rounds, and then further negotiate language amongst themselves for *t*/2 rounds.

4 Experimental Results

Two sets of experiments were run. In the first set, agents inhabit a one-dimensional world. The second set was performed in a dimensionless world, in which every agent is equally likely to interact with every other agent.

4.1 Evolution of Dialects in a 1-D World

Agents are organised linearly and teacher agents are selected according to a normal distribution curve centred on the current learner agent. An agent may not be it's own teacher. To visualise the languages, the three bits of a signal are used to set red, green and blue colour values of pixels. Plotting the pixels for each agent for each state, three columns are drawn each generation. Adding the columns for each generation to those of the previous, three bars are formed. A number of simulations were run, the results of one are shown in Fig. 1. The use of different signals can be seen to spread and recede over time.

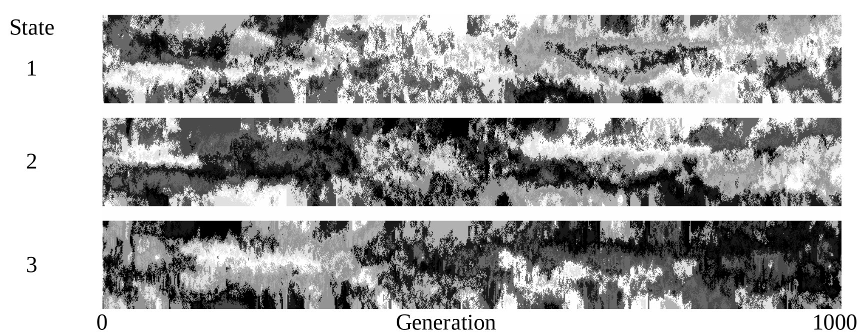

Fig. 1. The 'evolution' of signals representing three states in a spatially distributed population of 120 agents. Each bar shows the evolution of signal use for one of the three internal states

A measure of the difference between signals can be found using the Hamming distance. With a small difference it is possible that both signals are interpreted the same, due to redundancy in the signals. The Hamming distance between signals of neighbours, plotted over the whole population, can indicate language boundaries, and is shown for the final generation of the experiment above (Fig. 2 left). There is a high degree of change over the population. Despite this, signals used by adjacent agents for a particular meaning rarely vary by more than one bit. This could indicate a dialect continuum across the population. In a number of places, the signals used for two or all three of the meanings change together. These might indicate linguistic discontinuities.

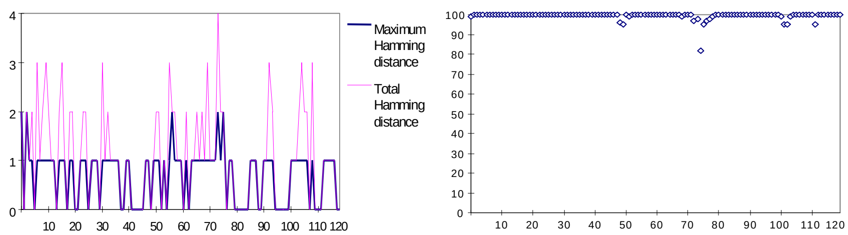

Fig. 2. The maximum (for any one signal) and total (over three signals) Hamming distances between signals used by adjacent agents in a spatially distributed population (left) and the percentage of communicative successes over the spatially distributed population (right)

Fig. 2 (right) plots the communicative success of the agents. Despite the high amount of diversity in the signals used, most agents achieved perfect scores. The lowest scoring agent correctly interpreted 82 percent of signals, showing that a dialect continuum exists from one extreme end of the population to the other.

4.2 Evolution of Dialects without Spatial Organisation

In these experiments agents interact with each other with uniform probability. No distinct dialects are observed, and populations quickly negotiate a global language. Signal redundancy allows the languages to use two or more signals for each state, where each signal is associated with only one state (Fig. 3). Eventually even this diversity disappears, and a one-to-one signal-state mapping established. Using signal uncertainty as a measure, we have determined that there is significantly less diversity under these conditions (mean uncertainty of 0.75 bits). In contrast, where agents are spatially organised no significant diversity loss is observed or measured after 10,000 or even 100,000 generations (mean uncertainties of 2.14 and 2.05 bits respectively).

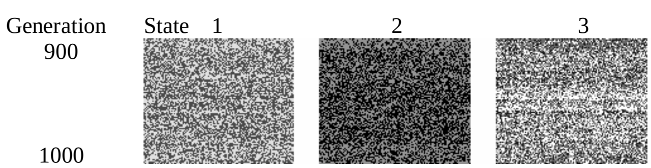

Fig. 3. The evolution of signals without spatial distribution

5 Conclusions

While adaptive explanations of language variation show how dialect may be exploited, our model shows a linguistic system that maintains a high level of diversity without adaptive benefits. Inheritance and mutation are also absent in our model and no errors occur in information transmission, so these are also not required to maintain diversity. Stochasticity occurs in the random selection of agents and in choosing a state to signal, and is not completely removed. The role of stochasticity for linguistic innovation is not disputed, and in our model a population without diversity cannot develop it. There are qualitative similarities between spatial organisation of dialects in our model and the organisation of human dialects. The dialects in our model form a continuum, connecting dialects that are, at the extremes, not mutually intelligible.

We conclude that sustained linguistic diversity can emerge from the interactions of agents, as a consequence of learning and the distribution of agents. Support for this comes from recent work [12] which uses evidence from global linguistic data, suggesting that linguistic diversity is related to *ecological risk*. Ecological factors determine the social networks formed by societies, which in turn determine linguistic diversity. This is the same as our conclusion – that linguistic diversity emerges as a result of constraints on interactions between language learners. We are currently working to determine the effects different parameters have on diversity.

References

1. Crystal, D.: The Cambridge Encyclopedia of Language, Cambridge University Press (1987)
2. Maeda, Y., T. Sasaki, et al.: Self-Reorganizing of Language triggered by "Language Contact". ECAL 97, MIT Press (1997)
3. Kirby, S.: Fitness and the selective adaptation of language. In Approaches to the Evolution of Language, J. R. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), Cambridge University Press (1998) 359-383
4. Hashimoto, T. and Ikegami, T.: Evolution of Symbolic Grammar Systems. ECAL 95 (1995)
5. Arita, T. and Taylor C. E.: A simple model for the evolution of communication. Evolutionary Programming V, MIT Press (1996)
6. Arita, T. and Koyama Y.: Evolution of Linguistic Diversity in a Simple Communication System. Artificial Life IV, UCLA, MIT Press (1998)
7. Steels, L. and Kaplan F.: Stochasticity as a Source of Innovation in Language Games. Artificial Life VI, UCLA, MIT Press (1998)
8. Nettle, D. and Dunbar R. I. M.: Social Markers and the Evolution of Reciprocal Exchange. Current Anthropology **38**(1) (1997) 93-99
9. Livingstone, D. and Fyfe C.: A Computational Model of Language-Physiology Coevolution. Computing and Information Systems, University of Paisley **5**(2) (1998). 55-62.
10. Livingstone, D. and Fyfe C.: Modelling Language-Physiology Coevolution. In The Emergence of Language, J. R. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), Cambridge University Press (in press)
11. Oliphant, M.: Formal Approaches to Innate and Learned Communication: Laying the Foundation for Language, University of California, San Diego (1997)
12. Nettle, D.: Explaining global patterns of language diversity. Journal of Anthropological Archaeology **17** (1998) 354-374.