

Cumulative cultural evolution: Can we ever learn more?*

Paul Vogt¹²

¹ Induction of Linguistic Knowledge / Computational Linguistics section
Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

² Language Evolution and Computation Research Unit, School of Philosophy,
Psychology and Language Sciences, University of Edinburgh, U.K.
paulv@ling.ed.ac.uk

Abstract. This paper investigates the dynamics of cumulative cultural evolution in a simulation concerning the evolution of language. This simulation integrates the iterated learning model with the Talking Heads experiment in which a population of agents evolves a language to communicate geometrical coloured objects by playing guessing games and transmitting the language from one generation to the next. The results show that cumulative cultural evolution is possible if the language becomes highly regular, which only happens if the language is transmitted from generation to generation.

1 Introduction

Our knowledge seems to be ever more increasing, our social networks are getting more complex, technology is advancing all the time, etc. In short: many aspects of our culture seem to be getting ever more complex. This paper investigates some dynamics of this *cumulative cultural evolution* by exploring a recent model that simulates the cultural evolution of language.

Cultural evolution is often characterised in terms similar to Darwin's [1] evolution theory, see, e.g., [2, 3]. Simply put, this means that culture evolves based on the principles of variation, competition and selection. Darwinian explanations have been applied to various aspects of cultural evolution, such as linguistics [4].

Boyd and Richerson [2] have argued that social learning is favoured if it is less costly or more accurate than individual learning of such skills or knowledge. Furthermore, if the cost decreases or accuracy increases for learning skills socially over generations, such skills may improve accumulatively over time. Translating this to language evolution, this view supports a gradual evolution of complexity in languages, see, e.g., [5]. Recent computational studies on the cultural evolution of language have shown how languages may become increasingly complex in terms of, e.g., grammatical structures when it is transmitted from one generation to the next [6, 7]. One typical change that seems to have occurred in these models

* This study has been supported by an EC Marie Curie fellowship and a VENI grant funded by the Netherlands Organisation for Scientific Research (NWO).

is that languages change in order to become more learnable, as hypothesised in, e.g., [8]. If languages are learnt easier (or faster), this both decreases the learning costs and increases learning accuracy.

Using the model introduced in [7], this paper illustrates how the iterated learning of compositional structures from initially holistic languages can lead to a cumulative cultural evolution until an optimal (cognitive) platform is reached. Although this model has been used previously to study various aspects regarding the evolution of compositionality (e.g., [7, 9, 10]), this study is unique in the way the dynamics of the evolution is analysed. This is done using different parameter settings and measures highlighting how this model can contribute to our understanding of how cumulative cultural evolution can work. The next section will present the model. Section 3 will present and discuss experimental results, which show that indeed learning cost decreases and learning accuracy increases. Conclusions are provided in Section 4.

2 Iterated learning and language games

As the model of this paper simulates a transition from holistic languages to compositional languages, it is good to start with a definition of such languages:

Holistic languages are languages in which parts of expressions have no functional relation to any parts of their meanings. For instance, there is no part of the expression “bought the farm” that relates to any part of its meaning **has died**.

Compositional languages are languages in which parts of expressions do have a functional relation to parts of their meanings and the way they are combined. For instance, the part “John” in “John loves Mary” refers to a guy named **John**, likewise “loves” and “Mary” have their own distinctive meanings. In addition, this sentence has a different meaning in English when the word-order changes, as in “Mary loves John”.

Following Wray’s [5] hypothesis that modern languages were preceded by holistic protolanguages, Kirby and colleagues have shown how an initially holistic protolanguage can transform into a compositional language if the language is iteratively transmitted from one generation to the next through a *bottleneck* (which means that the next generation only observes a small subset of the language from the previous generation) [6]. This *bottleneck effect* is understood by realising that holistic languages are unstable over time when the next generation only observes a small subset, while compositional languages are stable (see Fig. 1). This, of course, is only possible if the learners in this model can extract compositional structures from their (possibly holistic) input.

The basic principle of this *iterated learning model* (ILM) [6] is that the population consists of two groups: adults and children. The children acquire the language by observing the adults’ speech directed to them. At the end of an iteration, during which the children have observed a part of the language, all adults are removed, all children become adults and new children are added to

Type	$G(n)$	Utter.	$G(n+1)$
Hol.	ab-00	ab-00	ab-00
	cd-01	cd-01	cd-01
	fg-10	fg-10	fg-10
	fb-11		??-11
Comp.	ab-00	ab-00	ab-00
	ac-01	ac-01	ac-01
	db-10	db-10	db-10
	dc-11		dc-11

Fig. 1. This figure illustrates why holistic languages (upper part) are unstable when a population of generation $G(n+1)$ only observes three of the four utterances (Utt.) from generation $G(n)$'s language (i.e. word-meaning mappings). In this case, if generation $G(n+1)$ wishes to communicate about meaning 11 (meanings in this example are to be read as bit-strings), then this generation will have to create a new word. If the language is compositionally structured, as in the bottom part of this figure, observing only three out of four instances would allow the next generation to reconstruct the entire previous language. Hence transmitting a compositional language through a bottleneck is more stable than transmitting holistic languages.

the population. This process is then repeated, such that language is transmitted over subsequent generations. In earlier ILMs (e.g., [6]), this transmission was entirely *vertical*, because adults only directed their speech to children, who only listened until they were adults themselves. With such a condition, the only way to achieve a bottleneck is by the experimenter setting a parameter regulating that children only see a fragment of the language of a given size.

It has been shown that when communication within one iteration is *isotropic*, which means that communication goes in all directions from adult to adult, adult to child, child to adult and child to child, then compositionality can evolve as a stable system without the need for the experimenter to impose a bottleneck [9].³ This is because when children need to speak, they may face the consequences of the bottleneck if they need to speak about previously unseen meanings (cf. meaning 11 in Fig. 1). This *implicit bottleneck* is then a more natural bottleneck and may in part explain why in horizontal models of language evolution compositionality evolves, see, e.g., [11]. In this model, which is a simulation of the Talking Heads experiment [12], the population plays a large number of *guessing games* to develop a language that allows the population to communicate about their world, which contains a number of coloured geometrical shapes.

It is impossible to present all details of the model in this paper; the interested reader is referred to [7, 9]. The guessing games are played by two agents: a speaker and a hearer. Both agents are presented a small number of objects randomly sampled from the world. These objects constitute the *context* of the game; the world contains a total of 120 objects (12 colours combined with 10 shapes).

³ In [9] this type of transmission was called *horizontal transmission*, but since in [3] this refers to transmission only within one generation, the term *isotropic* is preferred.

1	$S \rightarrow \text{greensquare}/(0,1,0,1)$	0.2
2	$S \rightarrow A/\text{rgb } B/s$	0.8
3	$A \rightarrow \text{red}/(1,0,0,?)$	0.6
4	$B \rightarrow \text{triangle}/(?,?,?,0)$	0.7

Fig. 2. This example grammar contains rules that rewrite a non-terminal into an expression-meaning pair (1, 3 and 4) or into a compositional rule that combines different non-terminals (2). The meanings are 4-dimensional vectors, where the first 3 dimensions relate to the RGB colour space (rgb) and the 4th relate to the shape feature (s). The question marks are wild-cards. Each rule has a rule score that indicates its effectiveness in past guessing games. Only sentences of 2 constituents are allowed in this grammar.

Each agent individually categorises the perceptual features of each object using a method based on the discrimination game, whose details are irrelevant to the scope of this paper. Suffices to say that each object is categorised such that it is distinctive from all other objects in the context. If distinctive categorisation fails, a new category is constructed for which the object’s perceptual features serve as exemplars. (Note that initially, each agent has no categories at all; these are all constructed by these discrimination games.) Categories are represented as prototypical points in a 4-dimensional space, each dimension relating to a perceptual feature, which are the red, green and blue components of the RGB colour space and a shape feature.

Once the agents have categorised the objects in the context, the speaker selects one object at random as the *topic* of the communication. This agent then searches its grammar for ways to encode an expression that conveys the topic’s meaning. The grammar (Fig. 2) consists of simple rewrite rules that associate forms with meanings either holistically (e.g., rule 1) or compositionally (e.g., rule 2 combined with rules 3 and 4). The grammar may be redundant in that there may be rules that compete to encode or decode an expression (cf. [11, 13]). The speaker searches for those (compositions of) rules that match the topic’s meaning and if more than one are found, he selects the rule that has the highest rule score. If the speaker fails to encode an expression this way, a new form is invented as an arbitrary string and is associated with the topic’s meaning or – if a part of the meaning matches some non-terminal rule – with the rest of this meaning.

In turn, the hearer tries to decode the expression by searching her own grammar for (compositions of) rules that match both the expression and a category relating to an object in the current context. If there are more such rules, the hearer selects the one with the highest score, thus guessing the object intended by the speaker. The hearer then points to this object, and if this is the object intended by the speaker, the speaker acknowledges success; otherwise, the speaker points to the topic allowing the hearer to acquire the correct meaning.

If the guessing game was successful, both the speaker and hearer increase the scores of the rules they used and lower the scores of those rules that compete with the used rules. If the game has failed, the scores of used rules are lowered

and the hearer acquires the proper association between the heard expression and the topic’s meaning. To this aim, the hearer tries the following three steps until one step has succeeded:

1. If a part of the expression can be decoded with a part of the topic’s meaning, the rest of the expression is associated with the rest of the meaning. For instance, if the hearer of the grammar shown in Fig. 2 hears the expression “redcircle” meaning (1,0,0,5), the part “red”-(1,0,0,?) can be decoded, so the hearer adds rule $B \rightarrow \text{circle}/(?,?,?,.5)$ to its grammar.
2. If the above failed, the hearer searches its memory, where she stores all heard or produced expression-meaning pairs, to see if there are instances that are partly similar to the expression-meaning pair just heard. If some similarity can be found, the hearer will break-up the expression-meaning pairs containing the similarities – following certain heuristics, thus forming new compositional rules. Suppose, for instance, the hearer had previously heard the expression-meaning pair “greensquare”-(0,1,0,1), and now hears the expression-meaning pair “yellow square”-(1,1,0,1). The hearer can then break up these pairs based on the similarity “square”-(?,1,0,1), thus forming rules $S \rightarrow C/r D/gbs$, $C \rightarrow \text{green}/(0,?,?,?)$, $C \rightarrow \text{yellow}/(1,?,?,?)$ and $D \rightarrow \text{square}/(?,1,0,1)$. Note that this is not the ideal break up, since it breaks apart the red component of the RGB colour space from the blue and green components and the shape feature. The next section shows that over time such mistakes diminish as a result of competition and selection.
3. If the above adaptations both fail, the heard expression-meaning pair is incorporated holistically, leading to a new rule such as $S \rightarrow \text{yellowcircle}/(1,1,0,.5)$.

At the end of these steps, the hearer performs a few post-processes to remove any multiple occurrences of rules and to update the grammar such that other parts of the internal language relates more consistently to the new knowledge. Full details of the model are found in [7, 9].

3 Cultural evolution of language

In order to see how improvements accumulate culturally, a series of simulations were carried out in which the above model was run for 20 iterations (or generations) of 100,000 guessing games each. In these simulations the population size was set to 50 (25 adults and 25 children). Ten simulations were run with different random seeds. Although most different runs were quite similar, a few revealed some noticeable differences (see [10] for a discussion on these differences). Therefore, only the results of a few single runs are presented here.

Before presenting the results, a few measures are defined:

Communicative success measures the number of successful guessing games within a time window of 50 games.

Compositionality measures the number expressions decoded or encoded using a compositional rule over a time window of 50 games.

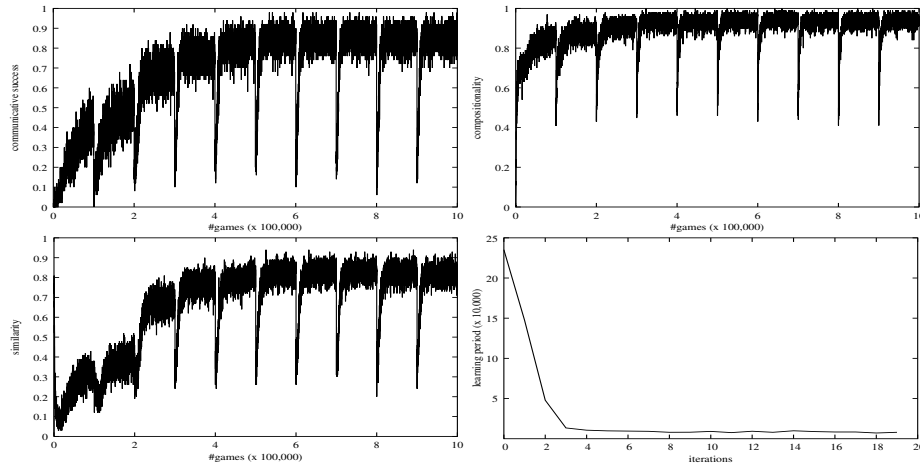


Fig. 3. The results of the first simulation. The graphs show communicative success (top left), compositionality (top right), similarity (bottom left) and learning period (bottom right). See the text for details.

Similarity measures the number of games in which both agents used the same syntactic structure within a time window of 100 games. A syntactic structure is considered similar if the words and the linguistic categories used are the same and in the same order. (A linguistic category is characterised by the dimensions that make up the conceptual space of a non-terminal node.)

Learning period measures the number of guessing games it takes within each iteration for communicative success to exceed 0.8 for the first time.

All these measures (except learning period) are normalised to a value between 0 and 1. In addition, the relative frequencies of rule types used during successive periods of 10,000 guessing games are analysed. As the agents can break up the 4-dimensional conceptual space in two conceptual spaces (or linguistic categories) of lower dimension without knowing how, 15 different rule types (including the holistic type) can develop. Only 5 rule types are inspected in this paper (all other had insignificant frequencies). These are:

- | | |
|--------------------------------------|----------------------------|
| I: $S \rightarrow r g b s$ | holistic rule |
| II: $S \rightarrow A / r B / g b s$ | red v. green, blue & shape |
| III: $S \rightarrow B / g b s A / r$ | green, blue & shape v. red |
| IV: $S \rightarrow C / r g b D / s$ | colour v. shape |
| V: $S \rightarrow D / s C / r g b$ | shape v. colour |

Figure 3 shows the results for the first 10 iterations of one simulation (the results did not change much during the final 10 iterations and are therefore not shown). The top left graph shows the evolution of communicative success, which rises to a more or less stable level of about 85-90% during the first 5-6 iterations. Each iteration is marked by a sharp decrease due to the drastic change of the population when all adults are replaced by the children and new language-

less children are introduced. Compositionality (top right graph) increased more rapidly in initial stages, but also kept on rising during the first 5-6 iterations until a stable level around 93% was reached. Similarity (bottom left) evolved more similar to communicative success and reached a stable level of around 80%.

All these three graphs clearly show that levels reached at the end of one generation are rapidly reached in the next generation, and – while there is still room for improvement – these levels further increase at a slower rate. It was expected that after a plateau was reached, learning speeds will continue to improve. This expectation was based on earlier visual inspection of similar graphs. However, closer analysis of the learning period reveals that this was not the case (Fig. 3 bottom right). Clearly, learning took quite a while in the first iteration, but rapidly decreased during the next 5 iterations and then remained more or less at the same level.

The graphs concerning communicative success and similarity clearly show some stagnation in development during the second iteration. Figure 4 (left) shows why this is the case: During the first two iterations there is a lot of competition between different rule types, as there are still a substantial number of holistic rules and rules of type II and III around, before rules that differentiate colours from shapes start to dominate in iteration 3. Recall that agents break apart holistic rules when they find a similarity in the expression and the meaning. Previous analysis of the model has shown that the probability of finding in two different games a similarity in any conceptual space, other than the colour and shape spaces, is 1.5 to 4 times larger than the probability of finding a similarity in the colour or shape space [7]. Most likely, compositions of type II and III are found. However, although such combinations are more frequently encountered, they are less efficient than rules of type IV and V (combining colour and shape), because rules of type II and III require almost 4 times as many different rules than those of type IV and V. Moreover, a meaning that only takes the red component of the RGB colour space cannot be applied with all possible meanings that can be constructed in the other 3 dimensions. (Note that this is mainly due to skew distribution of features in the RGB space.)

Once rules of type IV and V start to dominate, the other rules' relative frequencies start to diminish rapidly. In this simulation, there is first an almost

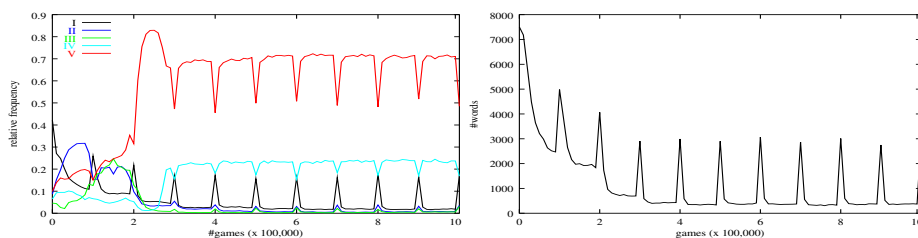


Fig. 4. The evolution of the relative frequency with which different rule types were used (left) and the number of different words used (right). Both were measured for successive periods of 10,000 guessing games.

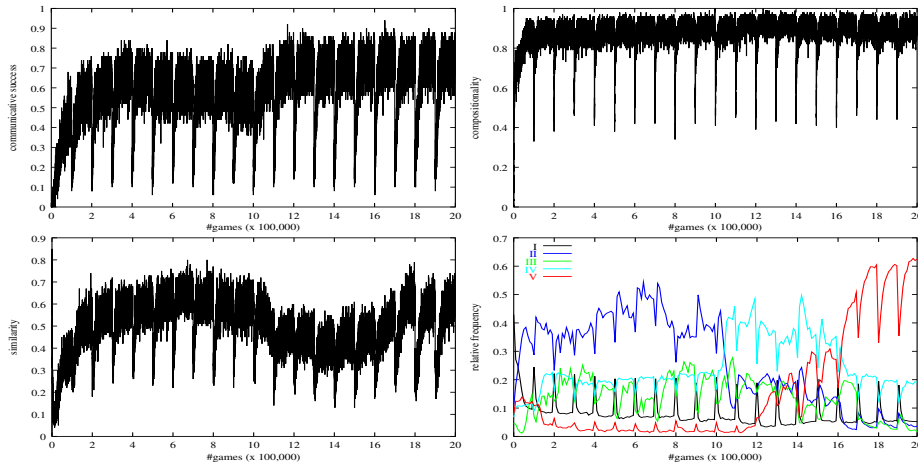


Fig. 5. The results of the second simulation. The figures show communicative success (top left), compositionality (top right), similarity (bottom left) and the relative frequency with which certain rule types were used during successive periods of 10,000 games (bottom right).

complete dominance of rule type V over rule type IV, but near the end of the third iteration, the relative frequency of type IV starts to become substantial too. Once these two rule types completely dominate (usage is then almost 100%), communicative success and similarity can further improve. Drops in the usage of these rules coincide with the replacement of adults with new children, after which the new children first invent quite some new words holistically when they lack the linguistic knowledge to encode an expression for some meaning.

This tendency is also reflected in figure 4 (right) which shows the evolution of the number of different words used in the language during successive periods of 10,000 games. The evolution first shows an immense increase of different words: up to around 7,500 words are used in the first 10,000 games. Ignoring the rise in words after the replacements of half the population, the number of different words then decreases sharply until it stabilises around 200 words. As has been observed in [14] for simulations evolving non-compositional vocabularies, the decrease of words coincides with an increase in communicative success. Interestingly, it has been suggested that the large number of words created early in development (which to a lesser degree is also true shortly after the start of a new iteration) increases the chances of finding similarities and thus give a boost toward the development of compositional structures [10]. This, in addition to the fact that the number of words created is proportional to the population size, explained the observation that the level of compositionality was substantially larger for populations of sizes 40-100 than for populations of less than 40 agents [10].

Comparing the 10 different runs revealed that 7 out of 10 yielded results similar to the run discussed; the other 3 evolved similar to the one shown in

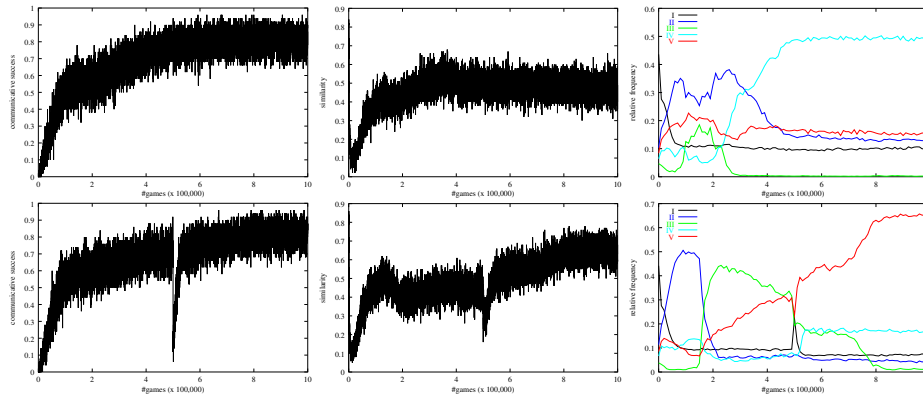


Fig. 6. The results of an experiment to compare two conditions of less iterations, each lasting for more guessing games. The top graphs show the communicative success (left), similarity (middle) and relative frequency of rules (right) for 1 iteration of 1 million games. The bottom graphs show the same measures for 2 iterations of 500,000 games.

Figure 5. This figure shows the results of another run with the same conditions, though plotted for the entire 20 iterations. Here the first 5 iterations are similar to the previous run, though communicative success reaches a lower level. Moreover, communicative success then slowly decreases, also slowing down the development of similarity. From the 11th iteration, communicative success rises again slightly to a seemingly stable level around 0.7 reached in iteration 13. Similarity, however, suddenly decreases from the 11th iteration and only starts to rise again from iteration 15.

The difficulties arising during this evolution are apparently caused by the strong competition between different rule types. Type II evolves as the most dominant rule during the first 10 iterations, followed by a period of dominance of rule type IV, which is finally taken over by rule type V. During this final period, the competition with other rules seems to weaken. Interestingly, the transition from rule type II to IV coincides with the decrease in similarity. This may be explained by realising that for some time type II has established itself in the language, though in competition with other types. When another type takes over, this type is in strong competition with the previous type, so that different agents are likely to favour one type over another, meaning that in a given situation one agent may use type II while another uses type IV. This, thus, lowers similarity, even though communicative success has improved during this period. This is related to the fact that well structured rules of type IV are taking over. Given the decrease in competition with other rules near the end and the higher frequency of type V rules, it is expected that the results would have improved further if the simulation was run longer.

So far, the simulations were carried out for 20 iterations of 100,000 games each. However, what will happen if we look at one iteration that is run for 1 million games? Figure 6 shows the results of a simulation with that condition.

In order to make an interesting comparison with the case where language is transmitted from one generation to the next, this figure also shows results of a simulation run for 2 iterations of 500,000 games each. Both simulations show a very similar evolution of communicative success (apart from the discontinuity in the second case). Similarity, however, tells a very different story. In the 1 iteration case, similarity does not further increase, but instead decreases slowly after approximately 300,000 games. When there are 2 iterations, similarity tends to increase further in the second iteration.

The rightmost graphs in Figure 6 show that in the 1 iteration case, there is first a lot of competition leading to many fluctuations, but after about 500,000 games, the system tends to become stable and the population never seems to get out of this system that has 4 rule types being used quite frequently (note the different scale on the y-axis compared to the one shown in Fig. 4). Although the communication system is successful to some degree, different agents have acquired different grammars, thus reducing similarity. In the second condition, there is also a lot of competition and fluctuation during the first iteration. Now before the system stabilises, the population is changed and the new agents quickly learn the well structured parts of the language from their ancestors and further improve this language.

This comparison thus shows that if there would be no population turnover and all agents are equally old (which is the case for the 1st iteration of all simulations), there comes a point in which there is hardly any improvement to gain. This is most likely because at some point, each agent has acquired a system that works sufficiently well, which *is* the case, because communicative success is fairly high. Only when a new generation starting from afresh enters the population, there is room for improvement. The ‘good’ structure – the one that combines colour with shape – is relatively to learn. Each colour can be combined with any shape. Hence, each time it is used in previously unseen situations, this structure allows successful interaction. Successful interaction increases the rule score, thus allowing this type of rule to win the competition with other rule types. A rule type that combines the red component with the other dimensions (i.e. rule type II or III) does not have that property, because if, for example, you have a category whose red component has value 1 (i.e. redness), only $\frac{5}{12}$ of all categories in the complementary dimensions can be combined with this one.⁴ So, the likelihood of applying this rule type successfully in unseen cases is lower than the odds for success when rule type IV or V is used.

4 Discussion and conclusions

This paper investigates the dynamics of cumulative cultural evolution in language using a computer model that simulates the evolution of compositional structures in language. This model, first introduced in [7], integrates the iterated learning model [6] with the Talking Heads experiment [12]. The results show

⁴ In the model 5 out of 12 colours have a red component of value 1.

that initially holistic languages can evolve into compositional languages by transmitting the language from one generation to the next. Unlike results obtained with the ILM in which language is transmitted vertically [6], no transmission bottleneck needs to be imposed when the language is transmitted isotropically, because then language learners face an implicit bottleneck [9].

The model is based on the idea that the evolution of language is a cultural analogy of Darwinian evolution [4]. In addition, the ontogenetic development of language is also similar to Darwinian evolution (cf. *Neural Darwinism* [15]) in that individuals acquire many variants of the language by communicating with different individuals and making errors in acquisition. These different variants then compete and are selected based on their effectiveness in communication. The same principles work at the cultural level, where the main elements subject to evolution are the word forms. So, there are two different levels of evolution interacting with each other (see [16] for an interesting discussion of such systems).

At the cultural level a set of conventions are formed that map expressions to objects in the agents' world. These expressions are segmented internally to form compositions with parts of the objects' meanings. These meanings are constrained by the visual input to the agents and the way these are processed in terms of features (in the model these features are similar for all agents). Initially, there are different (possibly ill-structured) rule types competing with each other. The simulations show that when a next generation learns from such a mixed language, they rapidly acquire the well-structured rule types and improve the structure of the rest of the language. The improvement primarily concerns the way meanings are segmented (i.e. the improvement from using rule type II or III to types IV or V). These improvements are triggered by a similarity in form, in addition to a similarity in meaning. Hence, there is a co-evolution of form (or syntax) and meaning (or semantics) in that on the one hand evolution of syntax is based on exploring similarities in form, though constrained by semantic structures and on the other hand evolution of semantics is based on exploring similarities in meaning, though constrained by syntactic structures [7].

As predicted by Boyd and Richerson [2], these simulations show a cumulative cultural evolution when the costs of social learning are low and accuracy is high. This can be concluded affirmatively because only when well structured compositional systems (i.e. those systems that have rules combining colours with shapes) develop, communicative success can further improve. Learning these well structured systems is less costly, because they are acquired more rapidly than unstructured ones, and they are also learnt more accurately, given the increase in similarity in language use of different individuals.

Prior to this study, it was anticipated – based on visual inspection of the data – that once a plateau of optimal communicative success was reached, learning speed will improve over subsequent generations. This, however, was not confirmed by this study. Once an optimal plateau was reached, learning speeds remained more or less constant over different generations. Explanations for this are that the agents have either reached a cognitive plateau on which no further improvements are possible, or because the evolved language reflects the struc-

ture of the environment so well that no pressures on further improvements are present.

Concluding, this study suggests that cumulative cultural evolution of language is possible if 1) the language is transmitted from one generation to the next, 2) the language evolves to well structured systems so it can be easily learnt by members of the population and 3) there is still room for improvement in the cognitive capacity of individuals or perhaps that there is a need for improvement because the structure does not reflect the environmental structure. These findings may well extend to other domains of cultural evolution.

References

1. Darwin, C.: *The Origin of Species*. John Murray, London (1959)
2. Boyd, R., Richerson, P.: *The origin and evolution of cultures*. Oxford University Press, Oxford (2005)
3. Cavalli-Sforza, L.L., Feldman, M.W.: *Cultural Transmission and Evolution: A quantitative approach*. Princeton University Press, Princeton, NJ (1981)
4. Croft, W.: *Explaining Language Change: An evolutionary approach*. Longman, Harlow (2000)
5. Wray, A.: Protolanguage as a holistic system for social interaction. *Language and Communication* **18** (1998) 47–67
6. Kirby, S., Smith, K., Brighton, H.: From UG to universals: linguistic adaptation through iterated learning. *Studies in Language* **28**(3) (2004) 587–607
7. Vogt, P.: The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence* **167**(1–2) (2005) 206–242
8. Deacon, T.: *The Symbolic Species*. W. Norton and Co., New York, NY. (1997)
9. Vogt, P.: On the acquisition and evolution of compositional languages: Sparse input and the productive creativity of children. *Adaptive Behavior* **13**(4) (2005) 325–346
10. Vogt, P.: Stability conditions in the evolution of compositional languages: issues in scaling population sizes. In Bourguine, P., Képès, F., Schoenauer, M., eds.: *Proceedings of the European Conference on Complex Systems, ECCS'05*. (2005)
11. Batali, J.: The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In Briscoe, E., ed.: *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge University Press, Cambridge (2002) 111–172
12. Steels, L., Kaplan, F., McIntyre, A., Van Looveren, J.: Crucial factors in the origins of word-meaning. In Wray, A., ed.: *The Transition to Language*, Oxford, UK, Oxford University Press (2002)
13. De Beule, J., Bergen, B.K.: On the emergence of compositionality. In Cangelosi, A., Smith, A., Smith, K., eds.: *The Evolution of Language: Proceedings of the 6th International Conference on the Evolution of Language*. (2006)
14. Baronchelli, A., Loreto, V., Dall'Asta, L., Barrat, A.: Bootstrapping communication in language games: Strategy, topology and all that. In Cangelosi, A., Smith, A., Smith, K., eds.: *Proceedings of Evolang 6*, World Scientific Publishing (2006)
15. Edelman, G.M.: *Neural Darwinism*. Basic Books Inc., New York (1987)
16. Tamariz, M.: Evolutionary dynamics in language form and language meaning. In Cangelosi, A., Smith, A.D.M., Smith, K., eds.: *The Evolution of Language: Proceedings of the 6th International Conference on the Evolution of Language*, World Scientific Press (2006)