Mehler, Alexander (2007): Evolving Lexical Networks. A Simulation Model of Terminological Alignment.
In: Benz, Anton, Christian Ebert & Robert Van Rooij (eds.): Proceedings of the Workshop on Language, Games, and Evolution at the 9th European Summer School in Logic, Language and Information (ESSLLI 2007), Trinity College, Dublin, 6-17 August, 57–67.

Evolving Lexical Networks.
A Simulation Model of Terminological Alignment

Alexander Mehler
Bielefeld University
Alexander.Mehler@uni-bielefeld.de

Abstract

In this paper we describe a simulation model of terminological alignment in a multiagent community. It is based on the notion of an *association* game which is used instead of the classical notion of a *naming* game (Steels, 1996). The simulation model integrates a small world-like agent community which restricts agent communication. We hypothesize that this restriction is decisive when it comes to simulate terminological alignment based on lexical priming. The paper presents preliminary experimental results in support of this hypothesis.

1 Introduction

The approach to interactive alignment in communication (Pickering and Garrod, 2004) postulates two mechanisms of alignment: (i) *priming* as a short-term mechanism of information percolation within the same or between different levels of representation and (ii) *routinization* as a long-term mechanism of expectation driven control of dialogue unfolding. So far, this mechanistic approach has not been tested by a simulation model of the build-up of routines and their underlying priming processes. This paper describes a multiagent model of the emergence of *intralevel* priming relations and, thus, serves as a preparatory study to such a simulation model. It focuses on semantic priming based on lexical associations and postulates that alignment is manifested by the harmonization of contiguity and similarity associations among interacting agents. Further, it hypothesizes that this alignment converges into a lexical network which constrains the lexical choices of interlocutors in order to maintain the success of their communication. As far as language communities are considered (beyond small groups of a couple of interlocutors), this leads to a dynamic understanding of lexical networks which are seen to converge into fluent equilibria in order to serve the communication needs of speakers and hearers.

The present research is in the line of approaches to circumvent models of linguistic conventions (Lewis, 1969) which rely on common ground in terms of knowledge that is shared and known to be shared (Clark, 2000). Recently, Barr (2004) has re-conceptualized simulation models of intra-generational learning in terms of such a circumvention. He argues against simulation models in which frequencies of successful communications are implicitly utilized as shared representations or in which all agents interact with each other. Such an unstructured community model is characteristic of many implementations of the naming game (Steels, 1996; Hutchins and Hazlehurst, 1995; Baronchelli et al., 2006a). Thus, in order to prevent that common ground simply results from an unstructured community model, it is required to be structured. Although we follow this conception, we depart from Barrs simulation model in two respects: Firstly, we consider sense relations of lexical units and, thus, do not implement a classical naming game. Secondly, we consider community models which are more realistic in terms of what is known about social communities. More specifically, we aim at a simulation model which obeys the principles of complex networks (Newman, 2003) and integrates a learning model of lexical associations subject to discourse processing in a multiagent setting. The simulation model integrates three areas: the theory of language evolution (Niyogi, 2006; Jäger, 2006), *Latent Semantic Analysis* (LSA) (Landauer and Dumais, 1997) as a learning theory of lexical associations and *Complex Network Theory* (CNT) which studies constraints of networking in agent *and* in linguistic networks.

The basic idea is to utilize the community model as the *independent* variable and the topology of the evolving linguistic network as the *dependent* one. We predict that if the community model has the *Small World* (SW) property (Newman, 2003) then the linguistic network — which is evolving subject to information flow within this community structure — is similar to *real* linguistic networks (Mehler, 2007). It turns out that this similarity is indicated by the SW-property. Thus, we predict that the SW-property of the community network and the simulated linguistic network correlate positively. The paper presents preliminary results in support of this hypothesis. It is organized as follows: Section 2 outlines a vector space model whose computation can be iterated in a multiagent setting and, thus, serves as a simplified reconstruction of LSA. Section 3 presents an algorithm for the generation of SW-like communities in accordance with well-known results of CNT. Section 4 describes an association game which is designed to simulate terminological alignment in agent communities. Section 5 presents results of a preliminary simulation experiment. Finally, Section 6 gives a conclusion and prospects future work.

2 Toward Distributed Semantic Spaces

Approaches to multiagent simulations of language learning and evolution include *inter*-generational simulations in which languages evolve subject to the bottleneck of generation change. The *iterated learning model* is a prime example of this approach (Kirby and Hurford, 2002). On the other hand, *intra*-generational models of language learning and change concentrate on the spread of linguistic forms, meanings, lexicons or even grammars in the same generation of learners (Niyogi, 2006). This paper belongs to the latter framework. However, we will uniformly speak of *multiagent simulation models of language learning* without differentiating inter- and intra-generational approaches. This is done in order to give a more general account of the type of language game we will introduce.

In the area of semantics, simulation models mostly implement a variant of the *signalling* or, more specifically, *naming game* (Steels, 1996). This game starts from a set E of expression units and a set M of meaning units where each agent $a \in P$ of a given population P learns a meaning relation $\|\|_a \subseteq E \times M$. Under this regime, a simulation is said to be successful after a number of iterations i if a "language" converges in P, that is, if each agent $a \in P$ has the same meaning relation and retains it after any further iteration. Generally speaking, the naming game relies on a *bipartite* graph model whose instantiations range from a state of complete semantic diversification (a single expression carries several meanings) to complete formal diversification (a single meaning is expressed by several expressions) where the "ideal" of a simultaneous unification (i.e., a one-to-one mapping) lies in-between. Note that under this perspective, E and M are seen to be unstructured (cf., e.g., Baronchelli et al. (2006a) who treat mappings of different objects as independent).

There are many reference points under which implementations of the naming game differ: Firstly, simulation models may (and most of them do) require that $\|\|_a$ is a bijective function so that neither synonymy nor homonymy can emerge in a's memory (Barr, 2004; Baronchelli et al., 2006a). Secondly, simulations may allow for growing sets E and M during iterations (Steels, 1996). Thirdly, the meaning relation $\|\|_a$ may be weighted by means of frequencies so that mappings between meanings and expressions can be conceived as probability functions — this approach leads to a Zipfian notion of signalling and its study in diversification analysis (Altmann, 1985). Fourthly, the topology of the community may vary so that agents are additionally endowed with a neighbor selection function (Barr, 2004).

Although some of these variants are considered in this paper, we depart from the classical setting of the naming game as follows: instead of a bipartite graph model, we use a unipartite one in which vertices denote signs whose association links are object of learning. Hashimoto (1997) proposes a pioneering approach towards multiagent learning of such associations. In such a model, the meaning relation $\|\|_a$ is defined as a homogeneous relation $\|\|_a \subseteq V \times V$ over a set V of signs where, initially, an agent a does not associate any signs, i.e. $\|\|_a^{t=0} = \emptyset$.[1] As $G(a,t) = (V, \|\|_a^t)$ defines a directed graph with vertex set V and edge set $\|\|_a^t$, the memory graph $G(a,t)$ of an agent a at time t can be conceived as a lexical network if V is restricted to lexical units. Now, convergence after a number of iterations means that all agents have learnt the same lexical associations and, thus, the same lexical network. As the number of lexical items and, correspondingly, their candidate associations can get very large, a threshold of community-wide conformity is needed instead. In the naming game of Barr (2004), e.g., four expression and meaning units are considered so that learning has to choose among 24 candidate mappings. If in contrast to this, a set of $|V| = n$ lexical items is considered, there are $2^{n(n-1)}$ directed candidate graphs which can be built out of these n items (for the sake of simplicity, the identity of link weights is disregarded). In order to master this complexity, we redefine the convergence criterium as follows: Let $0 \ll \sigma < 1$ be a threshold. A simulation of an association game (see Section 4) is said to converge after a number of $t \in \mathbb{N}$ iterations if the graph $G(P,t) = (V, E_t, \omega_t)$ with vertex set V and edge set

$$E_t = \left\{ (v,w) \;\middle|\; \frac{|\{G(a,t) = (V, \|\|_a^t, \omega_a^t) \mid a \in P \wedge (v,w) \in \|\|_a^t\}|}{|P|} \geq \sigma \right\} \quad (1)$$

has the SW-property according to the Watts-Strogatz-model and the preferential attachement model and does retain it for any further iteration — cf. Bollobás and Riordan (2003) for a formal account of the latter two models. This redefinition of convergence is in the line of approaches which refer to quantitative characteristics (e.g., indices as the cluster coefficient or distributions as some power law) of linguistic systems as reference points of judging simulation quality: as far as the simulating system approximates the corresponding characteristics of the simulated one it is said to be successful (Mehler, 2006). $\omega_t \colon E_t \to \mathbb{R}$ is an association measure (Bock, 1974) appropriately derived from the set of functions $\{\omega_a^t \colon \|\|_a^t \to \mathbb{R} \mid a \in P\}$. We call

$$\{(V, \|\|_a^t, \omega_a^t) \mid a \in P\} \quad (2)$$

[1] Note that we do not define $\|\|_a$ as a relation over E or M, but over a set of de Saussurean articulations of both sets. This opens the door to combine the naming game with the association game developed here.

a *distributed semantic space* which after t iterations of the association game is distributed over the population P.

The question arises how to model the association functions ω_a^t. This can be done according to the *Weak Contextual Hypothesis* (WCH) (Miller and Charles, 1991) which says that the similarity of the contextual representations of words contributes to their semantic similarity. The WCH is most prominently implemented by LSA as a model of contiguity and similarity associations (Landauer and Dumais, 1997). It implements semantic spaces (Rieger, 1978) as a geometric model of meaning in which signs are interrelated even if they do not co-occur, but are similar according to the WCH. That is, signs are mapped onto meaning points (e.g. feature vectors) whose geometric distance models their semantic similarity. Although LSA focuses on language learning, it does not deal with the evolution of linguistic systems, nor does it explore constraints which separate "natural" semantic spaces from implausible ones. Rather, LSA is a *single*-agent model leaving out the dynamics of *multiagent* communication.

Empirical evidence which allows deciding the cognitive plausibility of competing models of semantic spaces comes from *Complex Network Theory* (CNT) (Newman, 2003). Recently, Steyvers and Tenenbaum (2005) interpreted the SW-property of association networks as an indicator of efficient information storage and retrieval. They apply CNT which investigates network topologies in terms of their small world characteristics (Watts and Strogatz, 1998): Firstly, compared to random graphs, SW-graphs have a considerably higher amount of cluster formation. Secondly, compared to regular graphs, any randomly chosen pair of nodes of a SW-graph has, on average, a considerably shorter geodesic distance (the geodesic distance of two vertices in a graph is the length of the shortest path in-between). Steyvers and Tenenbaum (2005) show that lexical association networks as well as reference systems like WordNet share these properties. A central implication of these findings is that they question the cognitive plausibility and adequacy of LSA and related models which rely on semantic spaces in terms of *completely connected weighted graphs* as the underlying memory representation format — by refusing this memory model we view the signalling system to be learnt no longer to be unstructured.

Although Steyvers and Tenenbaum (2005) propose a growth model of semantic networks in accordance with CNT, they do not develop a multiagent simulation out of which such networks emerge. One might think that a candidate solution comes from CNT itself where lexical association networks are a prime object of study. But, actually, these co-occurrence networks are far to simple to grasp the kind of lexical associations underlying priming. The reason is that they consider any two items to be linked if they co-occur at least once (Ferrer i Cancho et al., 2007). In contrast to this, an approach is needed which is sensitive to the frequencies of co-occurrences. Although LSA clearly meets this requirement, it is nevertheless much to complex to serve as a learning model in a multiagent setting. This insufficiency is clarified by the following definition:

Definition 1 *A lexical association measure α is said to be* iteratively computable *if for any sequence $\langle x_1, \ldots, x_n, y\rangle$ of texts the following statement is true for any pair of lexical items $v, w \in V$ and some function β which is irrespective of co-occurrences in $x_1, \ldots, x_n$:*

$$\alpha(v, w, \langle x_1, \ldots, x_n, y\rangle) = \alpha(v, w, \langle x_1, \ldots, x_n\rangle) + \beta(v, w, y)$$

Iteratively computable association measures which are sensitive to certain sequences of texts to be processed are indispensable for modeling distributed semantic spaces. The reason is that in such simulation models single agents process different sequences of texts which unfold with the ongoing simulation experiment. Further, agents may process different texts simultaneously so that the set of texts being produced and processed in such experiments is partially ordered — other than predicted by the set-theoretical corpus model of LSA. This also holds for the Vector Space Model (VSM) whose iterative computation already fails because of its weighting scheme based on logarithmic damping.

In order to arrive at iteratively computable association measures we reconstruct the VSM in a way which prevents the usage of logarithms and standardization. Three functions are needed: (i) a function for the iterative memorization of lexical associations, (ii) a function for calculating priming relations of lexical items and (iii) a function for mapping priming relations of textual units:

- For a sequence $S = \langle x_1, \ldots, x_n \rangle$ of n texts, the association of two lexical items $v_i, v_j \in V$ is computed as

$$\alpha(v_i, v_j, S) = \sum_{k=1}^{n} \left(f_{ik} \frac{k}{F_{ik}} \right) \left(f_{jk} \frac{k}{F_{jk}} \right) = \sum_{k} k^2 \frac{f_{ik} f_{jk}}{F_{ik} F_{jk}} \quad (3)$$

  where F_{ik} is the number of texts out of $\langle x_1, \ldots, x_k \rangle$ in which v_i occurs; f_{ik} is the frequency of v_i in x_k. For a given k, $\frac{k}{F_{ik}}$ is the larger, the rarer v_i's text frequency F_{ik}; this effect is reinforced by f_{ik}. $f_{ik}\frac{k}{F_{ik}}$ resembles the TFIDF-weighting scheme of the VSM, but disregards any kind of, e.g., logarithmic damping as well as standardization. This is done in order to meet Definition 1.

- As we assume that agents do *not* memorize any single text, but only text frequencies, the cosine approach of the VSM in measuring lexical similarities is obsolete. Instead of this, a lexical item v_j is said to be primed by an item v_i for an agent a at time t to the degree of $\alpha(v_i, v_j, S_a^t)$ where S_a^t is the sequence of texts a has processed till iteration t. Accordingly, we utilize $\alpha(v_i, v_j, S_a^t)$ to build the lexical network $(V, \| \|_a^t, \omega_a^t)$ memorized by agent a at time t, that is, $\omega_a^t(v_i, v_j) = \alpha(v_i, v_j, S_a^t)$.

- Finally, a text x is represented as a fuzzy set of lexical items whose membership value is computed by a function of the degrees to which they are primed by the lexical tokens of x. These membership values model the degree by which items are primed by x *as a whole*, that is, in terms of *text priming* (Sharkey and Sharkey, 1992).

3 Artificial Small World Networks

The majority of models of language evolution or change starts from an unrealistic community model in which for an increasing number of iterations all agents tend to communicate to every other agent to the same degree. This model corresponds to a *Completely Connected Graph* (CCG) where the smaller the agent community, the less iterations are needed to get a CCG. Earlier simulation models of language evolution exemplify this situation as they are based on very small populations (Hutchins and Hazlehurst, 1995; Steels, 1996; Hashimoto, 1997). Conversely, if the number of iterations is small but the population large, agents have the chance to communicate only with a small number of other agents so that a random graph emerges according to the probability function used to choose speakers and hearers. Generally speaking, while in a CCG clustering is maximal and geodesic distances

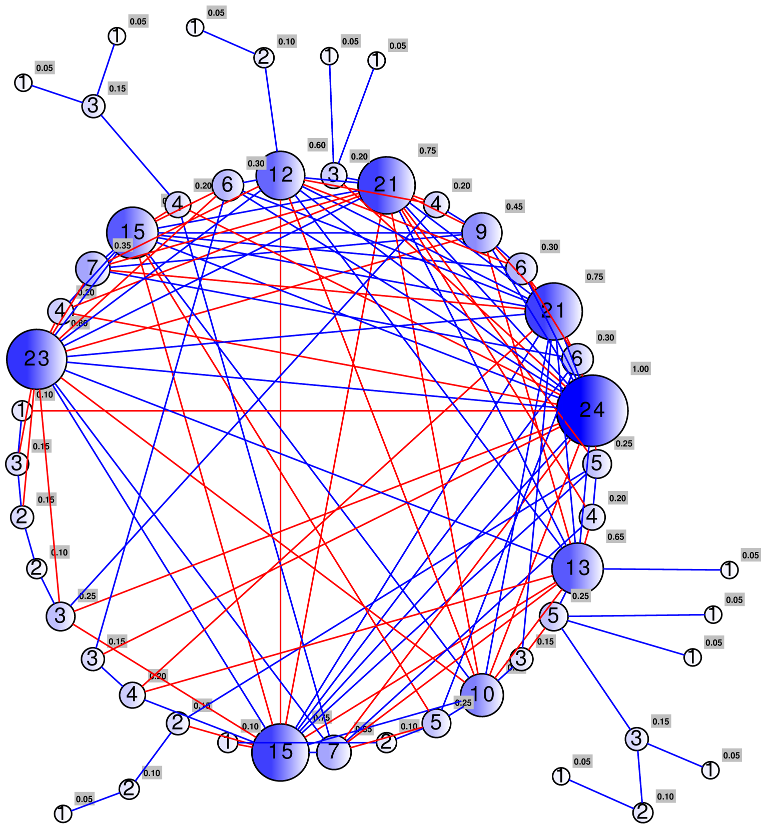

Figure 1: A sample Aswan $G = (V, E)$, $|V| = 50$, in which node centrality is signalled by vertex size. The SW-profile of G is as follows: cluster coefficient $C(G) = 0.25$, average geodesic distance $L(G) = 3.09$, assortativity $A(G) = 0.11$, power law-exponent $\gamma = 1.3$.

are minimal, random graphs assume short, but not minimal average geodesic distances in conjunction with small clustering values (Watts and Strogatz, 1998). Thus, whereas CCGs are unrealistic in terms of the topology of social networks, random graphs of the latter kind disregard, at least, the clustering of social groups.

Barr (2004) compares CCGs with a structured community model. Starting from randomly distributing agents over a plane Π each agent a is assigned a Gaussian neighbor selection function N_a which measures the probability by which another agent b is a neighbor of a as a function of their distance in Π. Under this regime, the more peaked N_a, the smaller the agents' neighborhoods, the longer the average geodesic distance of agents mapped by consecutive neighborhoods. As this model does not know short-cuts for linking "remote" agents, geodesic distances tend to be longer than in random graphs. At the same time, the less peaked N_a, the more likely two neighbors of a are neighbors on their own. Thus, Barrs model seems to approximate though not being equal to some regular graph of the same size. Regular graphs are known for high cluster values and long geodesic distances (Watts and Strogatz, 1998) — *obviously a likewise unrealistic model of social communities.*

What is missed so far is a community model which integrates efficient information flow by means of short-cuts with clustering as known from social groups together with highly skewed communication links and heterogeneous agent mixing (apart from homogeneously mixing CCGs). Although there are approaches to language evolution which consider related graph models, they either handle them as the *dependent* variable (as done by Gong and

Input: number N of vertices, exponent γ of power law, fraction p of local links. **Steps:** 1. **Preferential attachment:** Generate a power law $\mathsf{P}[\mathsf{N},\gamma] = \mathsf{a}\mathsf{k}^{-\gamma}$ and assign each vertex $\mathsf{v_i} \in \{\mathsf{v_1},\ldots,\mathsf{v_N}\} \in \mathsf{V}$ a degree $\mathsf{d_{v_i}}$. 2. **Random plane:** Randomly place N vertices $\mathsf{w_1},\ldots,\mathsf{w_N} \in \mathsf{W}$ on a plane and map each of the vertices $\mathsf{v_i} \in \mathsf{V}$ onto W such that the more central $\mathsf{w_j}$ (in terms of the sum of its Euclidean distances to every other vertex in W), the higher the degree of $\mathsf{v_i}$ mapped to $\mathsf{w_j}$. **Variant:** randomly place the N vertices $\mathsf{v_1},\ldots,\mathsf{v_N}$ on a plane. 3. **Minimal spanning tree:** Construct a *power law-conformant minimal spanning tree* (PLC-MST) which spans all vertices in V and is conformant to $\mathsf{P}[\mathsf{N},\gamma]$. 4. **Local transitivity-providing links:** link each vertex $\mathsf{v} \in \mathsf{V}$ to its nearest $\mathsf{pd_v} - \mathsf{d}'_\mathsf{v}$ neighbors where d'_v is the number of edges by which v has already been linked in the PLC-MST of Step 3. 5. **Remote short-cuts providing links:** randomly link each vertex to $\mathsf{d_v} - \mathsf{d}''_\mathsf{v}$ vertices where d''_v is the number of edges adjacent to v as generated in Step 3 and 4.

Table 1: An algorithm for the generation of small-world networks.

Wang (2005) who consider the emergence of agent communities as a by-product of language evolution, but disregard the SW-property of the linguistic network), or consider only a selected number of characteristics of social networks (as done, e.g., by Baronchelli et al. (2006b) who concentrate on the preferential attachement model of Barabási and Albert (1999)). What we need instead is a SW-like community model which combines several of these features and can be used as the *independent* variable in simulation experiments.

Table 1 presents an algorithm for the generation of such *Artificial Small-World Agent Networks* (ASWAN). It is based on Jin and Bestavros (2006) who generate SW-graphs according to the model of Watts and Strogatz (1998) *and* of Barabási and Albert (1999). Thus, their graphs combine high cluster values with short geodesic distances and skewed degree distributions. However, this model may generate disconnected graphs. In order to prevent this we generate SW-graphs according to Algorithm 1. This algorithm differs from Jin and Bestavros (2006) in that it constructs a minimal spanning tree which guarantees connectedness of the vertices randomly mapped onto a plane. Further, we consider the variant that the degrees d_v mapped to vertices v reflect their position in the plane so that the more central a vertex, the higher its degree. Figure 1 exemplifies a small-world graph generated by Algorithm 1. Figure 2 outlines the behavior of the algorithm. It shows that for increasing values of γ the cluster value of the network decreases, while the average geodesic distance increases. The next section utilizes Algorithm 1 as a generative model of small world agent networks.

4 Association Games

In this section we describe an association game as part of an intragenerational simulation model of terminological alignment in a multiagent community which has the small world-property. The basic notion is that of an *association game* which replaces the classical notion of a signalling game (cf. Section 2). The simulation model takes as input a small world agent community. It restricts which agents can communicate with each other. Further, the model iterates the association game defined as follows:

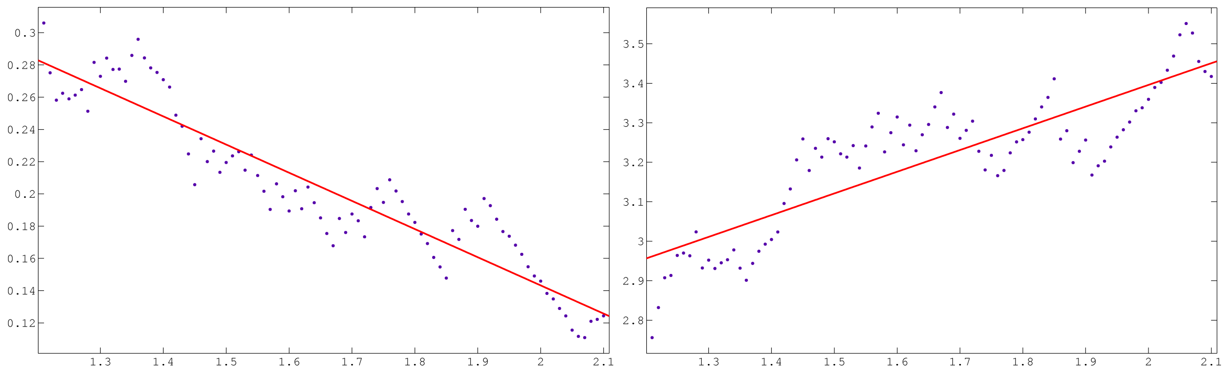

Figure 2: *Left:* C as a function of γ (determination coefficient of fitting $\approx .87$), $p = .6$. *Right:* L as a function of γ (determination coefficient of fitting $\approx .71$), $p = .6$. Numbers are averaged over 500 runs using 100 agents in each run.

Community Model Let $C(P) = (P, E)$ be a SW-community model of population P generated by Algorithm 1. Let further round $t \in \mathbb{N}$ be given. The sender a_S is picked at random who will then communicate to all his neighbors (in the sense of a "classroom situation"). That is, we do not pick the listener a_L at random among the neighbors of a_S. The reason is that in a SW-graph with node connectivity according to a power law the chance is high that a purely connected agent is selected as the sender while highly connected agents have a higher chance to be selected subsequently as the listener (cf. Baronchelli et al. 2006b). Under this regime, highly connected agents would be "instructed" more often by lowly connected ones than vice versa — in contrast to what is expected from their topological significance as a sort of hubs.

Memory Model Each agent $a \in P$ disposes of a semantic space $G(a, t) = (V, \|\|_a^t, \omega_a^t)$ where the lexicon V is shared by all agents $a \in P$. $\{(V, \|\|_a^t, \omega_a^t) \mid a \in P\}$ is a semantic space distributed over P at time t according to Section 2. Note that $G(a, t)$ and $G(b, t)$ will differ for two agents $a \neq b$ subject to the varying communication situations to which they participate.

Text Generation Model A lexical prime $v_+ \in V$ is randomly picked and used by the sender a_S to activate a subgraph in his semantic space $G(a_S, t)$. This subspace consists of a subset of the nearest neighbors of v_+ in $G(a_S, t)$. As an example, suppose that v_+ equals *color* so that its neighborhood consists of words like *red*, *yellow*, *sun* etc. Initially, primed lexical neighbors are picked at random. The resulting neighborhood is used to let a_S produce an output text x_t which consists of m tokens of the lexical items primed by v in $G(a_S, t)$. This procedure generates a multiset in which the same item may recur where tokens are selected randomly. Obviously, this is the entry point for a more elaborate text generation model (cf. Biemann 2007) which generates texts of varying length that obey some well-known quantitative text characteristics.

Update Model The output text x_t of the sender is then processed by the listener a_L who utilizes the lexical tokens of x_t to prime an item v_- in $G(a_L, t)$. Next, this item is compared with v_+ where the longer the geodesic distance $L(v_-, v_+)$ in $G(a_S, t)$, the less the communication success. That is, sender and listener play an association game in which the sender is masking which item he used to prime the tokens of his output text and where the listener has the task to find out which word the sender had initially in mind when producing

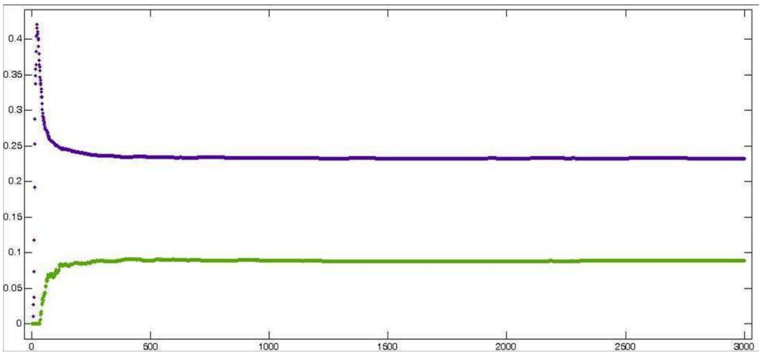

Figure 3: The cluster value of the summary lexical network as a function of the number of iterations. Upper curve: SW-community based run; lower curve: random graph-community based run.

x_t. We assume that the listener tells the sender the word, i.e. v_-, he has interpreted so that the sender can decide whether he was understood or not. The association game is successful if both sender and listener associate the same or related words with the same input text. In order to implement an update model which goes beyond decision theory, we call the communication of a_S and a_L *successful* if $L(v_-, v_+) \leq r$ for some $r \in \mathbb{N}$ as a further parameter of the model. Otherwise the association game is *unsuccessful*. In successful communication both sender and listener update their memories according to Equation 3, otherwise not. In a more realistic setting it is expectable that the sender is his "first" recipient, whereas the listener will process x_t even if he does not unmask the prime v_+ or one of its nearest neighbors correctly. Under this regime, a more appropriate update model is to use x_t to update the sender and listener memory *twice* if communication is successful (reinforcement learning), but only *once* if it is not.

The association game implements, so to speak, a variant of "*I spy with my little eye, something beginning with ...*".[2] Here, *terminological alignment* means that sender and listener align their lexical associations as they continually communicate so that they finally play the game more and more often successful. This is the local perspective. On the global level we expect that this game leads to a lexical network which has the SW-property subject to the SW-property of the community network $C(P)$ — without any reference to common ground shared by the whole community.

5 A Preliminary Experiment

This section exemplifies a simplified version of the association game: for a population of 100 agents we compare two competing community models $C_1(P)$, which is generated according to Algorithm 1 as an Aswan, and $C_2(P)$, which is a random graph. We consider a lexicon V of 500 items and set the threshold σ of the summary language network (see Equation 1) to 37.5%. Further, we compute $t = 3,000,000$ iterations of the association game and fall back to a decision-theoretical update model in which sender and receiver memory is updated after each iteration. Figure 3 compares both community models by example of the cluster coefficient C. It shows that $C_1(P)$, but not $C_2(P)$, converges into a lexical

[2] In German: "*Ich sehe was, was Du nicht siehst ...*".

network which shows clustering as known from SW-like lexical networks (Mehler, 2007). This picture is confirmed by the average geodesic distance and the higher speed by which a single connected component emerges in the lexical network of the SW-community $C_1(P)$. Evidently, these results have to be substantiated by a thorough parameter study.

6 Conclusion

In this paper we have introduced a novel type of language game. Based on the notion of alignment in communication, we defined an association game by which interacting agents may align their lexical associations in sequenced communication situations. Further, we argued for a more realistic and, thus, structured community model as input to simulation experiments which reflects the insights of computational sociology. In order to meet this requirement we proposed an algorithm for the automatic generation of such artificial communities. A first test of the model indicates promising results. Future work will focus on systematically testing the model and elaborating it in terms of a turn-taking model. This will be a further step towards a simulation model of dialogical alignment.

References

Gabriel Altmann. Semantische Diversifikation. *Folia Linguistica*, 19:177–200, 1985.

Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286: 509–512, 1999.

Andrea Baronchelli, Maddalena Felici, Vittorio Loreto, Emanuele Caglioti, and Luc Steels. Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, P06014, 2006a.

Andrea Baronchelli, Vittorio Loreto, Luca Dall'Asta, and Alain Barrat. Bootstrapping communication in language games: strategy, topology and all that. In *Proceedings of the 6th International Conference on the Evolution of Language*, pages 11–18, 2006b.

Dale J. Barr. Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28(6):937–962, 2004.

Chris Biemann. A random text model for the generation of statistical language invariants. In *Proceedings of HLT-NAACL-07, Rochester, NY, USA*, 2007.

Hans H. Bock. *Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten.* Vandenhoeck & Ruprecht, Göttingen, 1974.

Béla Bollobás and Oliver M. Riordan. Mathematical results on scale-free random graphs. In Stefan Bornholdt and Heinz G. Schuster, editors, *Handbook of Graphs and Networks. From the Genome to the Internet*, pages 1–34. Wiley-VCH, Weinheim, 2003.

Herbert H. Clark. *Using Language.* Cambridge University Press, Cambridge, 2000.

Ramon Ferrer i Cancho, Alexander Mehler, Olga Pustylnikov, and Albert Díaz-Guilera. Correlations in the organization of large-scale syntactic dependency networks. In *Proc. of TextGraphs-2 at NAACL-HLT'07, Rochester, New York*, 2007.

Tao Gong and William S-Y. Wang. Computational modeling on language emergence: A coevolution model of lexicon, syntax and social structure. *Language and Linguistics*, 6(1):1–41, 2005.

Takeshi Hashimoto. Usage-based structuralization of relationships between words. In P. Husbands and I. Harvey, editors, *ECAL97*, pages 483–492. MIT Press, 1997.

E. Hutchins and B. Hazlehurst. How to invent a lexicon: the development of shared symbols in interaction. In G. N. Gilbert and R. Conte, editors, *Artificial Societies: The computer simulation of social life*. UCL Press, London, 1995.

Gerhard Jäger. Convex meanings and evolutionary stability. In Angelo Cangelosi, Andrew D. M. Smith, and Kenny Smith, editors, *The Evolution of Language. Proceedings of the 6th International Conference (EVOLANG6), Rome*, pages 139–144, 2006.

Shudong Jin and Azer Bestavros. Small-world characteristics of internet topologies and implications on multicast scaling. *Computer Networks*, 50(5):648–666, 2006.

Simon Kirby and James R. Hurford. The emergence of linguistic structure: An overview of the iterated learning model. In Angelo Cangelosi and Domenico Parisi, editors, *Simulating the Evolution of Language*, chapter 6, pages 121–148. Springer, London, 2002.

Thomas K. Landauer and Susan T. Dumais. A solution to plato's problem. *Psychological Review*, 104(2):211–240, 1997.

David Lewis. *Conventions. A Philosophical Study.* Harvard U.P., Cambridge, Massachusetts, 1969.

Alexander Mehler. Stratified constraint satisfaction networks in synergetic multi-agent simulations of language evolution. In Angelo Loula, Ricardo Gudwin, and João Queiroz, editors, *Artificial Cognition Systems*, pages 140–174. Idea Group Inc., Hershey, 2006.

Alexander Mehler. Large text networks as an object of corpus linguistic studies. In Anke Lüdeling and Merja Kytö, editors, *Corpus Linguistics. An International Handbook of the Science of Language and Society*. De Gruyter, Berlin/New York, 2007.

George A. Miller and Walter G. Charles. Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1):1–28, 1991.

Mark E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

Partha Niyogi. *The Computational Nature of Language Learning and Evolution.* MIT Press, Cambridge, 2006.

Martin J. Pickering and Simon Garrod. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–226, 2004.

Burghard B. Rieger. Feasible fuzzy semantics. In K. Heggstad, editor, *COLING-78*, pages 41–43. ICCL, Bergen, 1978.

A. J. C. Sharkey and N. E. Sharkey. Weak contextual constraints in text and word priming. *Journal of Memory and Language*, 31(4):543–572, 1992.

Luc Steels. Self-organising vocabularies. In C. Langton and T. Shimohara, editors, *Proceedings of Artificial Life V, Nara, Japan*, 1996.

Mark Steyvers and Josh Tenenbaum. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1):41–78, 2005.

Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.