

Robots will dominate the use of our language

Jürgen Brandstetter and Christoph Bartneck

Adaptive Behavior

1–14

© The Author(s) 2017

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1059712317731606

journals.sagepub.com/home/adb



Abstract

Robots are able to influence the usage of human language even after the interaction between the human and robot has ended. Humans influence each other in the usage of words, and hence, the robots they program indirectly affect the development of our society's vocabulary. Most human–robot interaction studies focus on one robot interacting with one human. Studying the dynamic development of language in a group of humans and robots is difficult and requires considerable resource. We therefore conducted a social simulation of a human–robot communication network based on a real-world human–human network, allowing us to study how the centrality of the robots' owners influences the propagation of words in the network and what influence the number of robots in the network has on achieving a fixation state. Our results show that robots owned by highly connected people have less effect on the dynamics of language than robots owned by less connected people. Highly connected people interact with many others and therefore are more strongly influenced by a greater number of people and their robots. We have found that 11% of the humans owning a robot is sufficient for the robots to dominate the development of the language resulting in 95% of the humans using or adopting their words.

Keywords

Human–robot interaction, language, persuasion, simulation

Associate Editor: Tom Froese

1. Introduction

As spoken language is such an important part of human communication, researchers want to understand how its vocabulary spreads through the population, how it changes over time and how it affects people. One way of studying such dynamic behaviours is by computer simulations. The general idea of using simulations to understand the dynamics of language change is not new. Steels (1995), for example, used such simulations in 2006, and Baxter, Blythe, Croft, and McKane (2009) showed the usefulness of simulations for the study of language change. However, over the years, the simulations have become more complex and include more and more parameters. For example, Pierrehumbert, Stonedahl, and Daland (2014) developed a simulation that included many different human biases and learning rates. Most of these simulations focused on human–human interaction.

Robots and other speech-enabled technology are expected to dramatically expand their presence in our society. It is therefore necessary to include robots in simulations of the development of language. Robots have characteristics that are fundamentally different from human speakers, such as their ability to

communicate through the Internet instantaneously without the human speakers noticing. They would be able to update their vocabulary within seconds. Moreover, robots can be completely consistent in the use of their language. To be able to study the effect of robotic communication partners, a new simulation model needs to be developed.

Such new simulations need to consider the technological abilities of robots. Open questions include what happens when robots are connected to the cloud? Will they just use the cloud as storage or will they synchronise with each other? Should the synchronisation be limited to the robots owned by a single person or family? Or will all the robots in the world synchronise with each other? The latter scenario opens further questions about how the companies providing cloud synchronisation services deal with the information they receive from the people using the robots. These questions do

HIT Lab NZ, University of Canterbury, Christchurch, New Zealand

Corresponding author:

Jürgen Brandstetter, HIT Lab NZ, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand.

Email: j@brandstetter.io

of course go beyond what this study can achieve, but we believe that it is necessary to start asking these questions in this age of sharing our most private moments and even thoughts with companies such as Facebook and Google.

As mentioned above, humans adapt their language to each other. For a new word, such as ‘to google’, to gain general acceptance it needs to gain a critical mass of users which may take years. Robots, on the contrary, can update their vocabulary through the Internet within a fraction of this time. Even though a future where robots learn words from humans and ‘tell’ them to other robots might exist at some point, at this point in time, as far as we know, no widespread voice-activated device or system learns new words by itself. The so-called ‘learning’ is only used when a user clearly states that the device should use another word for an object. For example, one can tell Siri to change the name of the owner. This adaptation remains local, and any other update or adaption is done by the company providing occasional updates to the dictionary and pushing it to the robots. This update pattern of getting updated by a centralised system will be the base for the robots in this study.

From the work of Brandstetter, Beckner, Sandoval, and Bartneck (2017), we learnt that robots are able to influence the words humans use. This influence was shown to last even after the completion of the interaction between the humans and the robots. Furthermore, they showed how much influence a robot and a human have on their communication partners. This knowledge of a one-to-one interaction, however, does not tell us much about the overall influence robots might have when many of us own robots and constantly communicate with them.

We therefore developed a social simulation to explore the dynamics of how robots can influence the use of language in a mixed human–robot communication network. It is important to note that each owner of a robot is linked to a single robot node and that the robot node is only linked to its human owner. The robots can, however, talk to an Internet server instantaneously (see Figure 1). This network design resembles the most likely scenario in which we have robots in our home that mainly talk to their owners but rarely to other humans in the network. Such social robots are already available in the market. Future scenarios may include robots that are able to also operate outside the home and even go shopping for us. But such a usage scenario is still more fiction than science. We will therefore focus on a network design that is likely to represent human–robot interaction in the upcoming decade.

Some users in the human network will own a robot, and the characteristics of these owners might have an influence on how effectively the robots influence the dynamics of the language. These robot owners can be considered ‘influencers’. Marketing experts have long discovered that ‘influencer marketing’ is a powerful

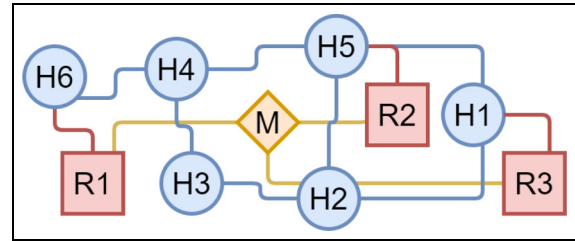


Figure 1. Basic human–robot network graph. Circles represent humans and squares robots. The diamond is the master server. The connecting lines indicate communication channels. A human can communicate only with one robot or with other humans. Robots communicate with the human and the main server.

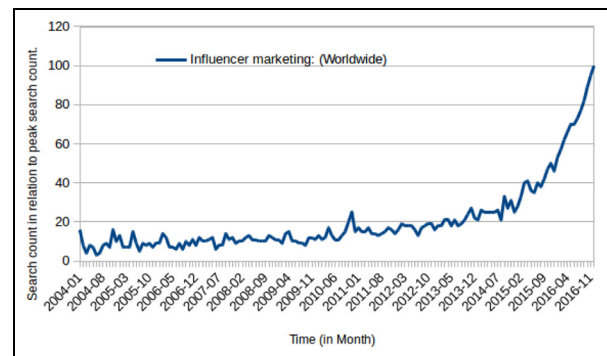


Figure 2. Google Trends on the term ‘influencer marketing’.

tool to influence the market (Gillin, 2009). Influencer marketing has caught the public attention as it is visible from Google Trends (see Figure 2). The idea behind influencer marketing is to find out who are the influencers in the society or group and manipulate them specifically. New tech tools, for example, are given to famous tech bloggers who in turn report on them. If these bloggers can be persuaded to like a product, their followers are more likely to like it too and maybe even buy it (Pavlika, 2016). This phenomenon is common not only for bloggers but also with social media influencers on Instagram or Snapchat, which in return has sparked specific influencer marketing companies like ‘The Social Club’¹ that focuses on identifying the influencers and forming marketing relationships with them.

In recent times, the company Lord & Taylor used influencer marketing to sway fashion bloggers. They managed to make 50 fashion bloggers wear the same dress and show them to their audience, and as hoped, the dress sold out quickly (Griner, 2015). In comparison with a traditional advertisement that addresses a wider audience, influence marketing focuses on a very selective target group (Griner, 2015; Hall, 2016; Talavera, 2015).

It is a nontrivial task to find out who are the major influences in a social network/community. However, the most common strategy to find influential nodes is

called ‘centrality measurements’ (Liu, Xiong, Shi, Shi, and Wang, 2016; examples of centrality measures are Closeness-Centrality, Betweenness-Centrality, Page-Rank). In most of the centrality measurement types, node properties like the number of connections a node has and how often one crosses a node when travelling from one side of the network to the other are used to measure the importance of nodes. With that knowledge, very specific nodes can be targeted and their selection might influence the dynamics of the language. Depending on the centrality of the robots’ owners, the persuasive power of the robot might change. A robot owned by a highly connected human speaker might exert a stronger influence on the development of the language compared with a robot that is owned by a less well-connected human speaker, meaning that decision of a well-connected human speaker to adopt a certain word in the Naming Game makes others change to the same word.

The persuasive power of the robots might depend not only on the characteristics of their owners but also on their quantity. If the majority of the humans own a robot, then the robots might be more influential compared with when only a handful of users own them.

During many simulations, observations and experiments (Baronchelli, Gong, Puglisi, & Loreto, 2010; Blythe & Croft, 2012; Castellano, Fortunato, & Loreto, 2009, p. 595; Cangelosi & Parisi, 1998; Centola, 2010, p. 1194, 2013, p. 6; Centola, 2010, p. 1194; Centola & Baronchelli, 2015; Fagyal, Swarup, Escobar, Gasser, & Lakkaraju, 2010; Lu, Korniss, & Szymanski, 2008, 2009), it became clear that the development of language within a network follows three stages (see Figure 3). These stages split an S-curve function into three regions. The first stage is the innovation or emergence stage, where a new word finds its critical mass. In this stage, individual speakers need to find like-minded others to form a critical mass (Centola, 2013, p. 6). In the second stage, called the diffusion, selection or propagation stage, the minority starts to become a majority and the social pressure to switch to the new form increases. The final and third stage is the Fixation Stage. During this stage, the previous minority is now the majority and the new word becomes the standard form, until, of course, a new minority emerges and overthrows the old one (Fagyal et al., 2010). The question is what role the robots may play in achieving such a fixation state.

2. Research questions

In summary, we are interested in the following research questions:

Q1. How do the centrality characteristics of the robots’ owners influence the dynamics of the language and the resulting persuasive power of the robots?

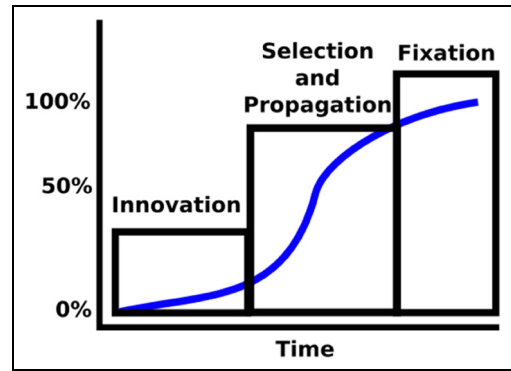


Figure 3. A consensus-building process follows an S-curve that can be separated into innovation, selection and fixation sections (Fagyal et al., 2010, p. 2).

Q2. How does the start rank calculated by the centrality measures influence the dynamics of the simulation?

Q3. How many robots are needed to get the language used in the network into a Fixation Stage?

3. Literature

This section will describe the two major research areas upon which the experiment is build: The Naming Game and Graph Knowledge.

3.1. Naming Game

The so-called Naming Game is a well-studied linguistics experiment developed by Baronchelli, Felici, Loreto, Caglioti, and Steels (2006). Its goal is to see how people influence each other’s language. More specifically, it looks into the propagation of a new word through a network of people. The game itself has been explored in areas of simulation as well as experiments. A great overview of the Naming Game and all its different implementations, experiments and attributes can be found in the excellent ‘Statistical physics of social dynamics’ paper by Castellano et al. (2009). Furthermore, different network topologies were explored. One of the reasons the Naming Game is a favourite for such simulations is its simplicity (Baronchelli et al., 2010; Beuls & Steels, 2013; Centola & Baronchelli, 2015; Lu et al., 2008, 2009; Xie et al., 2011).

The basic principles of the Naming Game are that all nodes in the network have to find an agreed name for the same object, noting that there is only this one object present in the game. Each conversation has exactly one speaker and one listener, and each node can be both a speaker and a listener. Each node has a dictionary of words that he or she or it will use during the conversations. When a speaker has no applicable word in his or her or its dictionary for the object, he or she or it will

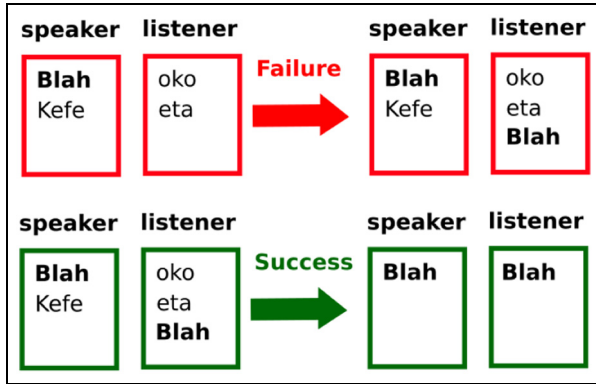


Figure 4. Naming Game when only humans talk to each other. Image taken from Lu et al. (2009).

generate a random word and add it to his or her own dictionary and tell it to the listener. If the listener does not recognise this word (failure), he or she will add it to its own dictionary, but if he or she or it recognises the word (success), the listener and the speaker will both forget all other words except the one word both know (to see how a successful and failed game look like, see Figure 4).

To choose a speaker/listener pair from the network, the following approach is used. First, any random node from the network is chosen to be a speaker. Second, one of the speaker's neighbours is randomly chosen to be the listener. With this selection process, each node has an equal chance to be a speaker; however, depending on the connections of a node, some nodes will listen to more speakers than others.

3.1.1. Naming Game adjustments. As the Naming Game is in our experiment played not only between humans but also between humans and robots, we had to adjust the behaviour when humans and robots communicate with each other. Thus, we have made two changes.

First, we adjusted the update rule of the robots so they behave differently to humans. The robots are assumed to be connected to a central server through the Internet, and their vocabulary remains static until an update is pushed to them from the server. Once a new word has entered the collective vocabulary, it remains in there. This results in an extremely consistent behaviour for all robots as they will all continue to use the exact same word. To model this behaviour in the Naming Game, we changed the update rule so that in case a robot is a listener, it will not add the new word it hears from the human to its dictionary. We refer to the robot's behaviour as being 'consistent' and for the remainder of this chapter will refer to it as such. The consistent behaviour of the robots leads to them only ever using the exact same word, which we refer to as the robots' word. Our simulation did not examine how the

frequency of vocabulary updates influences the development of the language in the simulation. The simulation is based on a static vocabulary for all robots.

Second, we adjusted the update rule of the humans when adding new words to the dictionary. In the original Naming Game, every success will remove all previously known words except the success word. But we have learned from Brandstetter et al. (2017) that the influence by humans and robots is not equal. More specifically, the influence of a human towards another human is 61.12% with a standard deviation of 20.53, whereas the influence of a robot towards a human is 38.92% with a standard deviation of 18.65.

To incorporate that into a success situation, we changed the update rule as follows. When a listener hears a word from a robot or human, he or she accepts that word according to the probability just lined out. If he or she does not accept it, he or she will simply add the word to the dictionary instead of replacing all words in the dictionary.

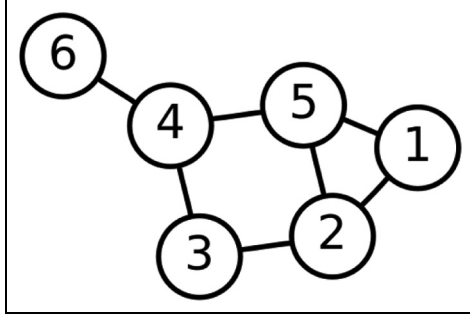
3.2. Graph Knowledge

In many cases where human-made systems are simulated, an underlying network structure exists (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006; Newman, Barabasi, & Watts, 2011). For example, transportation is built on a network structure called a street network. Companies or friendships can be mapped as social networks. Facebook and Twitter are based on such social networks. Depending on the network structure and how nodes communicate with each other, changes in the network can have different outcomes depending on its structure. For example, if we knew that a plague had broken out in New Zealand, we could cut the network connections. In this case, we would cancel all trips from and to New Zealand to prevent the spread of the plague to the rest of the world. As New Zealand is an isolated island, cutting its connections is easy and a spread could be relatively easily stopped. If such a plague occurred in Central Europe, it might be impossible to contain it.

As mentioned above, one of the goals of this study is to understand how the properties of the robots' owners in a graph/network affect the communication in the network. To understand network properties of robot owners, we need to have a look at graph properties and position properties, such as centralities (Barrat, Barthlemy, Pastor-Satorras, & Vespignani, 2004). A simple graph generally consists of two types of elements: a node and a link, also called a vertex and an edge (see Figure 5 for an image of a simple graph). A typical graph, which all social beings experience, is the social network they live in. In this network, the nodes represent humans, and the links represent the relationships between them.

Table 1. A table of the shortest distance from one node in network 5 to each other node.

Node	6	5	4	3	2	1	Sum
6	0	2	1	2	3	3	11
5	2	0	1	2	1	1	7
4	1	1	0	1	2	3	8
3	2	2	1	0	1	2	8
2	3	1	2	1	0	1	8
1	3	1	2	2	1	0	9

**Figure 5.** Simple undirected graph.

Author: <https://en.wikipedia.org/wiki/User:AzaToth>

When one wants to influence other nodes in a graph, it is important to know how much influence each node in the network can have. This influence can be measured by calculating how many steps each node is away from other nodes and by seeing how well-connected each node is. In Figure 5, we see that it takes in the optimal case three steps to go from node 1 to node 6.

To define how much influence a given node has, we use centrality measurements. A centrality measure assigns every node in the network a number of influence. Let us look at an example for a simple network to understand how these measurements are calculated. Looking at Figure 5, we can see that some nodes are closer to other nodes and some are farther away. We can now measure how many steps it takes from one node to another node. Table 1 shows how many steps it takes from any node in the network to any other node, by using the shortest distance between any two nodes. From that, we can see that node 5 is only seven steps away from all other nodes, whereas node 6 is 11 steps away from all other nodes. That gives node 5 the highest centrality rank and node 6 the lowest rank. We can now sort all nodes by their centrality rank, which is 5, 4, 3, 2, 1, 6. Further down in our experiment, we will use the centrality rank as start position to whom owns a robot. To do so, we use percentiles. For example, the 0% percentile means that people own robots by starting at node 5. When using the 25% percentile, people start owning a robot at node 3. So, if we want to assign three robots and our start position is the 25% percentile, then nodes 3, 2 and 1 will get a robot assigned.

In the previous example, we have used a very primitive way of calculating the centrality rank; however, depending on the type of centrality measure, the same nodes can be assigned a different centrality rank. The centrality measurements used in this study are Closeness-Centrality, Betweenness-Centrality and Page-Rank, and the following sections will show and describe how those measurements work. The reason why we chose those three types of centrality measures comes from the fact that first they are well established and tested measures (Liu et al, 2016). Second, their results of assigning ranks to nodes are different from each other. And finally, NetLogo, as the tool of choice, can only deal with the following four types – betweenness, eigenvector, page-rank and closeness – whereby, eigenvector and betweenness do not show big enough difference in assigning ranks to nodes.

The following will describe how the centrality is measured and what is needed to do so. First, we need to define the graph and functions we use. A graph is defined as $G = (V, E)$, where V are all the vertices (nodes) and E are all the edges (links) (Hatfield, Cacioppo, & Rapson, 1993; Page, Brin, Motwani, & Winograd, 1998; Sabidussi, 1966a). Next, we define $d(y, x)$ as a function that describes the shortest distance between one node (x) and another node (y), that is, how many steps it takes from node x to node y .

3.2.1. Betweenness-Centrality. Betweenness-Centrality describes how often a node x is used for the shortest path ($d(s, t)$) between all nodes (s) to all nodes (t). We can describe the total number of shortest paths ($d(s, t)$) between s and t as $\sigma_{st} = \sum d(s, t)$, and the number of shortest paths between s and t that go through x as $\sigma_{st}(x) = \sum_{x \in d(s, t)} d(s, t)$. The Betweenness-Centrality is described as $g(v) = \sum_s \sum_t (\sigma_{st}(v) / \sigma_{st})$ (Freeman, Borgatti, & White, 1991).

When using Betweenness-Centrality in an application, this information can be used to see how often a node x can influence information travelling from node s to node t . The colours in image 1 in Figure 6 show a visual representation of nodes with a high to low Betweenness-Centrality rating.

3.2.2. Closeness-Centrality. Closeness-Centrality describes how many steps it takes a node x to reach all other nodes. Each new measurement is started from point x .

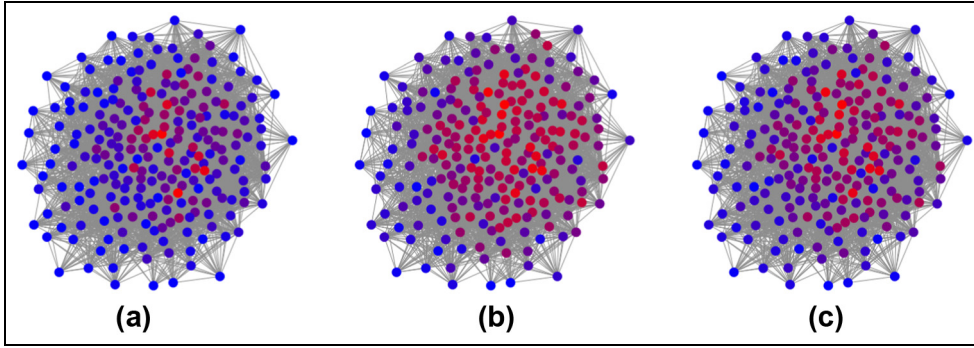


Figure 6. Colour-coded centrality ranks of the ‘Classroom Network’ network. Red means high centrality value and blue means low centrality value. (a) Betweenness-Centrality, (b) Closeness-Centrality and (c) Page-Rank.

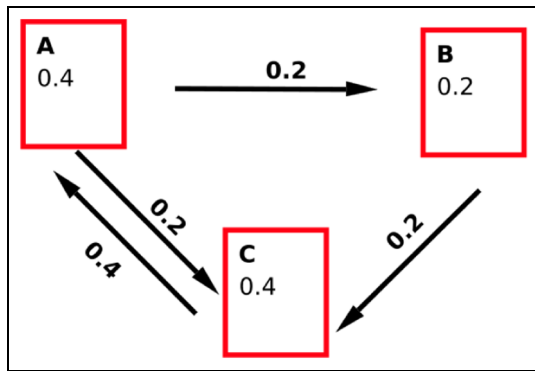


Figure 7. An example of node ranks according to the Page-Rank centrality measure (Page et al., 1998).

We can describe the Closeness-Centrality for node x as $C(x) = 1/(\sum_y d(y, x))$, where y is every node in the graph. When we do this for each node in the graph, we can see which one has the shortest distance to all other nodes (Sabidussi, 1966b). The colours in image 2 in Figure 6 show a visual representation of nodes with a high to low Closeness-Centrality rating.

3.2.3. Page-Rank. The Page-Rank measure is based on the idea that links that point to an observed node describe the importance of the observed node. But instead of simply counting the connections a node has, Page-Rank takes into account how many connections the nodes it is connected to have. Figure 7 shows how the rank of a node influences the rank of a neighbouring node.

Let us define B_u as the sum of nodes pointing to u and N_u the sum of nodes u point to. The rank of a node $R(u)$ is then described as $R(u) = \sum_{v \in B_u} (R(v)/N_v)$ (Page et al., 1998). It is important to note that the Page-Rank approach is recursive and will eventually stabilise itself. The colours in image 3 in Figure 6 show a visual representation of nodes with a high to low Page-Rank rating.

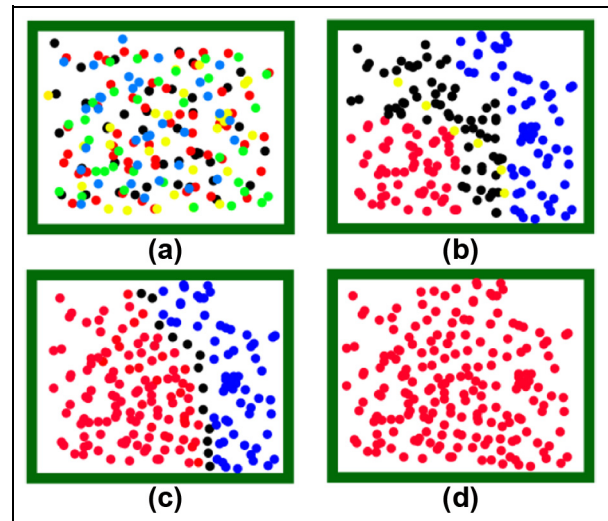


Figure 8. A network that establishes a consensus over time: (a) many clusters with no overall consensus, (b) Regional consensus starts to develop (c) two major clusters with some nodes accepting both words/idea (Lu et al., 2008) and (d) Full consensus in the whole network is established.

3.3. Network structure and the Classroom Network

When running experiments where consensus building has a strong focus, it is important to also look at the network that is used during the process. Centola and Baronchelli (2015) studied how a consensus-building process can look like and how the network structure influences it. For example, when playing the Naming Game, the network can end up in one of three stages. First, the whole network agrees on a common word. Second, no consensus is found at all and speakers learn and forget words all the time. Third, regions of synonyms stay in place, and only speakers at the borders of the regions learn and forget words over time. Figure 8 visualises what this can look like.

To find out why some words are accepted by the whole network and some are only used in some regions, as mentioned above, Centola and Baronchelli (2015)

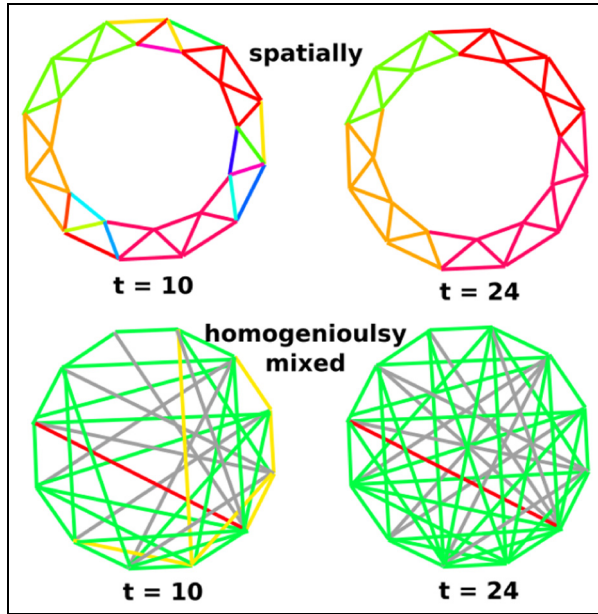


Figure 9. Spatially embedded and homogeneously mixed networks (Centola & Baronchelli, 2015).

looked at the network topology for a cause. In their experiments, they used three different topologies of networks: a spatially embedded network, a random network, and a homogeneously mixed network (see Figure 9). Through this experiment, they were able to show that network topology had an effect: The

homogeneously mixed network found a full consensus in which over time everyone used the same word; the random network found no consensus at all, and the spatially embedded network was found to form regions of words.

As we have learnt that the network topology has an effect on the outcome of the experiment, it is important to use an accurate social network. To get to this network, we have two choices: first, using a model that can create social networks, like Watts and Strogatz's small world networks model (Watts & Strogatz, 1998), or using an observed real-world network. Our choice went to a real-world network called the 'Classroom Network' (Stehlé et al., 2011). This network is available online² and can easily be imported into simulation software. The Classroom Network itself describes a full day of interaction amongst the students of 10 school classes. The network has $|V| = 236$ nodes (members), and the sum of all connections is $|E| = 5899$. The total count of communications between any two nodes is also known and sums up to 37,351 contacts per day (Stehlé et al., 2011). The structure of the network can, among others, be seen in Figure 10.

The advantage of this network is that it is based on real-world observations which gives our simulation a high ecological validity. The results of our simulation are therefore more applicable to the real world. Artificial networks are able to generate more extreme and biased networks that might be useful to test some questions, but the network structure itself was not in

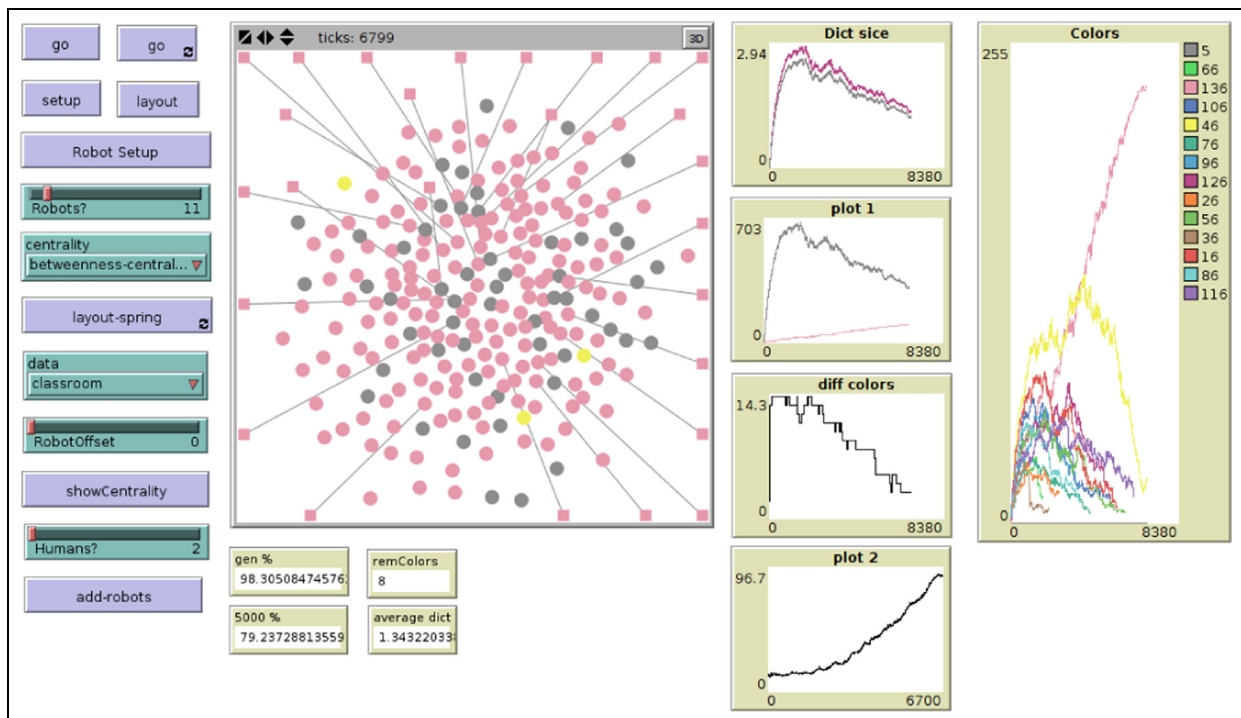


Figure 10. The user interface with the Classroom Network loaded.

the main focus of our study. Moreover, it is difficult to transfer the results of simulations done on artificial networks back onto reality.

4. Experimental setup

To find out in what situations the robots are most effective, we have divided the experiment into two stages. In the first stage, we compare different centrality measurements (Betweenness-Centrality, Closeness-Centrality and Page-Rank) against each other and against randomly owned robots. In stage 2, we explore how the connectivity of humans and robots within the network influences the robots' power to sway humans to use the robot's word.

Stage 1: In this stage, we distribute the robots according to one of the three centrality measures, Betweenness-Centrality, Closeness-Centrality and Page-Rank, or Random. The centrality types are used, first to find the best to least connected human in the network (see section 'Graph Knowledge' on how to rank nodes according to centrality measures), and then in assigning the robots deterministically from the best connected humans on upwards. We also change how many humans own a robot, which can be 0% (only humans), 3%, 6%, 9%, 11%, 26%, 44% and 85%. It is important to mention that the number of humans never changes. Only the number of robots owned by humans can change.

Stage 2: In this stage, we change the start position of the centrality rank calculated for stage 1. While in stage 1, we allocated the robots deterministically from the human with highest centrality rank (0% percentile) to lowest centrality rank (see section 'Graph Knowledge' on how to rank nodes according to centrality measures). In this stage, we used different starting points: the 0%, 25%, 50% and 75% percentile of the ranked nodes (e.g. the 50% percentile means the person who is in the middle between the highest and lowest ranked nodes).

4.1. NetLogo

To develop our simulation, we chose NetLogo, which is a special agent-based simulation framework developed by the Northwestern University (Wilensky, 1999). The idea behind NetLog is to give the researcher all the tools necessary to simulate agent-based models such as birds flocking, traffic movement and communication between people. It provides the researcher with functions like generating networks, loading network files, calculating centrality measures, exporting results, automating different simulation settings and visualising the network effect. To run simulations with huge datasets, NetLogo also functions in a headless mode. This mode

provides the same functionality but can be run on the command line and without a graphical user interface to allow running multiple settings simultaneously on multiple machines. In the case of this experiment, we used six computers to divide the workload, therefore having a speed-up of about six times.

4.1.1. NetLogo setup. To run our simulation with NetLogo, we had to implement our own setup, which is based upon scripting NetLogo and using user interface elements to support different test parameters (see Figure 10 which shows a screenshot of how our simulation setup looks like). In total, three major parameters are used to run and change our simulation: the Centrality parameter, the Robot Starting Point parameter, and the Number of Robots parameter. All parameters get adjusted during the different simulation runs. Other parameters such as the robot/human update rule are fixed and implemented as described in the 'Naming Game Adjustments'.

Centrality parameter: With the 'Centrality' parameter, the type of centrality (Betweenness-Centrality, Closeness-Centrality, Page-Rank or Random) used to distribute the robots can be selected.

Robot Starting Point parameter: The 'Robot Starting Point' parameter describes the centrality rank percentile at which the human will own a robot. For example, if the rank is 0%, then the starting point is at the highest rank (0% percentile). If the rank is the 50% percentile, the starting point is at the middle. For a more detailed description on how the ranking via centrality works, look at section 'Graph Knowledge'.

Number of Robots parameter: The 'Number of Robots' parameter describes what percentage of humans will own a robot: 0% means no robots at all and 50% means every second human owns a robot.

4.1.2. Running the simulation. When running the simulation, our NetLog setup sets up the network properties and then runs the simulations until every human in the network is using the same word or multiple words keep existing, and saves the measured points for each run.

It is important to understand that each run of the simulation can have a slightly different outcome. This unpredictability comes from the fact that each agent is randomly picked. To actually draw a conclusion from the simulation, each simulation must run multiple times. And as a final step, the average of all simulations is taken to draw a conclusion.

Initialization: Before any decision for the distribution of robots is made, the Classroom Network is loaded as an undirected network with 236 nodes ($N = 236$). Next, the robots are distributed according to the rules of stages 1 and 2.

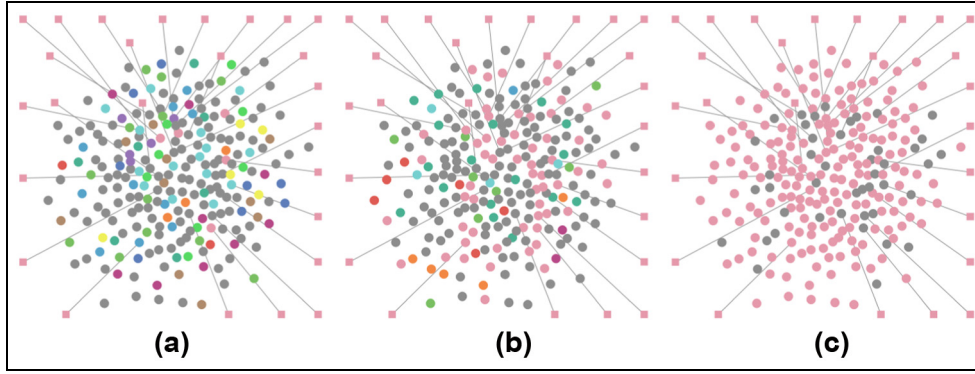


Figure 11. The distribution of the different words after 100, 4000 and 7500 iterations when having 11% of robots in the network: (a) after 100 iterations, (b) after 4000 iterations and (c) after 7500 iterations. The square dots represent robots.

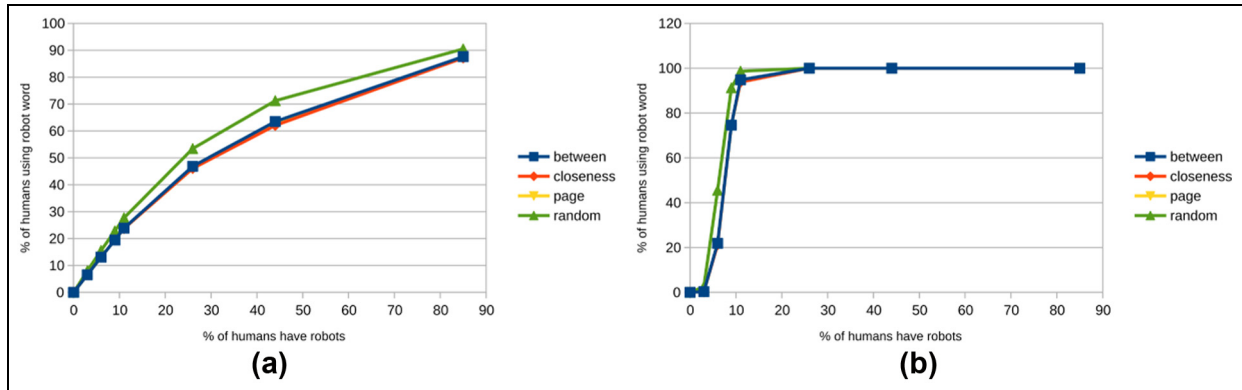


Figure 12. Difference between the three centrality types and random: (a) snapshot after 1000 iterations and (b) snapshot after the simulation ended.

Run: During the run step of the simulation, NetLogo will follow these phases: First, one random node from the whole network is picked. This node will be the talker node and can be a human or a robot. Second, one of the talkers' neighbours is randomly picked. Again, the neighbour can be a human or a robot and is the listener node. Third, the talker selects a word according to the Naming Game rules and tells it to the listener. Finally, the talker and listener will update their internal dictionary according to our adjusted Naming Game rules (see section 'Naming Game Adjustments'). After all four phases are over, the program will start from the beginning. This will go on until all human nodes in the network use the same word, or if multiple words keep existing. The simulation will end at iteration number 100,000. To get a visual perception of how the network changes during different steps, see Figure 11. To see what the different centrality measures look like, see Figure 6.

Saving Data: To measure the effect of the robot during the interaction and not just the end result, we have added some measurement points. First, the program measures how many steps it took to finish each run,

meaning when all humans used the same word. Next, the distribution of the robot word in percentage for the 1000th, 2500th, 5000th and 37,351th steps is measured, as well as measuring the particular step at which the human decided to use the same word (this does not have to be the word introduced by the robots). In total, each setup ran 6000 times. For example, if 6% of all humans have robots and Betweenness-Centrality is chosen to distribute the robots, combining all possibilities we have 32 cases, which ends up having to run 192,000 simulations.

5. Results

Our first question (Q1) was to find out if different centrality measures have a different effect on the robots' power to sway humans to the robot's word. For that, we compared the three centrality types (Closeness-Centrality, Betweenness-Centrality and Page-Rank) with each other and with the random setting (stage 1). Figure 12 shows a plot of a snapshot of all types after 1000 steps and when the simulation ended.

Table 2. ANOVA between the three centrality types and random that measured how high the robot word rate was at 5000 iterations.

	Estimate	SE	t value	p value	Mean	SD
Betweenness-Centrality	59.1216	0.1929	306.455	<2e – 16	67.57	38.642
Closeness-Centrality	–0.4284	0.2728	–1.570	0.116	67.08	38.641
Page-Rank	–0.3378	0.2728	–1.238	0.216	67.18	38.71
Random	6.4982	0.2728	23.818	<2e – 16	75.00	35.80

We performed an analysis of variance (ANOVA) in which the centrality type was the independent variable and the distribution rate of robot word after 5000 iterations was the dependent variable. The plots already reveal what the ANOVA confirmed. We could not measure any significant ($F(3, 191, 996) = 307, p < .001$) difference between the centrality types (see Table 2). The residual was $\min = -65$, $\text{median} = 24$ and $\max = 41$.

We suspected that the reason for not having a statistical significant difference between the three centrality types is that each centrality type assigns similar ranks to the nodes. For example, the person with most connections will have highest rank in all three types. To see whether our hypothesis is correct, we used a Granger causality test. The test showed that there was no significant difference between the ranks in any two centrality types (Betweenness-Centrality vs Closeness-Centrality: $F(-1, 233) = 0.3095, p < .5785$; Closeness-Centrality vs Page-Rank: $F(-1, 233) = 0.3025, p < .5829$; Page-Rank vs Betweenness-Centrality: $F(-1, 233) = 0.0078, p < .9297$).

However, through a visual inspection, we noticed a difference between the random setting and the three other settings. Looking at our Figure 12, we could see that the random setting was consistently stronger than the other settings. This effect was not statistical significant, but it hints at a potential follow-up study.

Our second question (Q2) was to find out whether changing the start rank calculated by the centrality measures would change the dynamics of the simulation. Looking at the results of this simulation, we can see in Figure 13(a) that the different starting points indeed have an effect. This effect is also supported by an ANOVA ($F(3, 20, 962) = 1151, p < .001$). For the ANOVA, we look at a setting with 11% of robots and the robot word rate of 5000. Of special interest is that low centrality ranks create a stronger effect, meaning less well-connected nodes have greater effect.

Next, we have a look at the random effect and the centrality starting position to see whether the random effect is still more effective when we move the centrality starting point. When looking at the following two graphs of Figures 14 and 13a where 9% of all humans have a robot, at the centrality position of 50%, the Betweenness-Centrality had the same effect as the

random condition (45% vs 45%). However, at the centrality position of 75%, the Betweenness-Centrality surpassed the random condition (55% vs 45%).

In our third question (Q3), we wanted to know how many robots are needed to get into the Fixation Stage. To analyse this, we will have a look at the comparison of humans only versus humans plus robots. In Figure 15, we can see that the adaption follows in both cases an S-curve. However, it is clear that the introduction of robots accelerates the adaption rate.

Figure 12(b) shows that eventually all humans will use the robots' word as long as at least 9% of the humans own a robot. If fewer robots are available, then the network does not end in a fixation state determined by the robots. If more than 11% of robots are available, the robots' word will always be adopted by 90% of all humans. A more context-oriented analysis can be drawn by looking at Figure 13(b). This figure is a snapshot at iteration 37,351, which is the amount of communications the people in the classroom-network had during one day. It shows again that 11% of humans need a robot so that all humans use the robot word in the end. Changing the start position to the 45% percentile or higher will result in only 9% of robots needed to get into the Fixation Stage.

6. Discussion

In this study, our main goal was to find out whether robots could manipulate humans to use the robots' chosen word. This manipulation was effectively achieved by making the robots consistent, so that all robots would always choose the same word for a particular object. The study uses the idea of Baronchelli et al. (2006), who created a linguistics game called the Naming Game to study how language develops, to see how effective robots are and how positioning them in a network affects their power to sway humans to use their word. We chose to simulate a game based on a real-world network called the Classroom Network (Stehlé et al., 2011), and to generate a plausible result, we adjusted the update rule of the Naming Game with the data of language manipulation gathered in Brandstetter et al. (2017).

In our first question (Q1), we wanted to know how positioning robots according to the different centrality

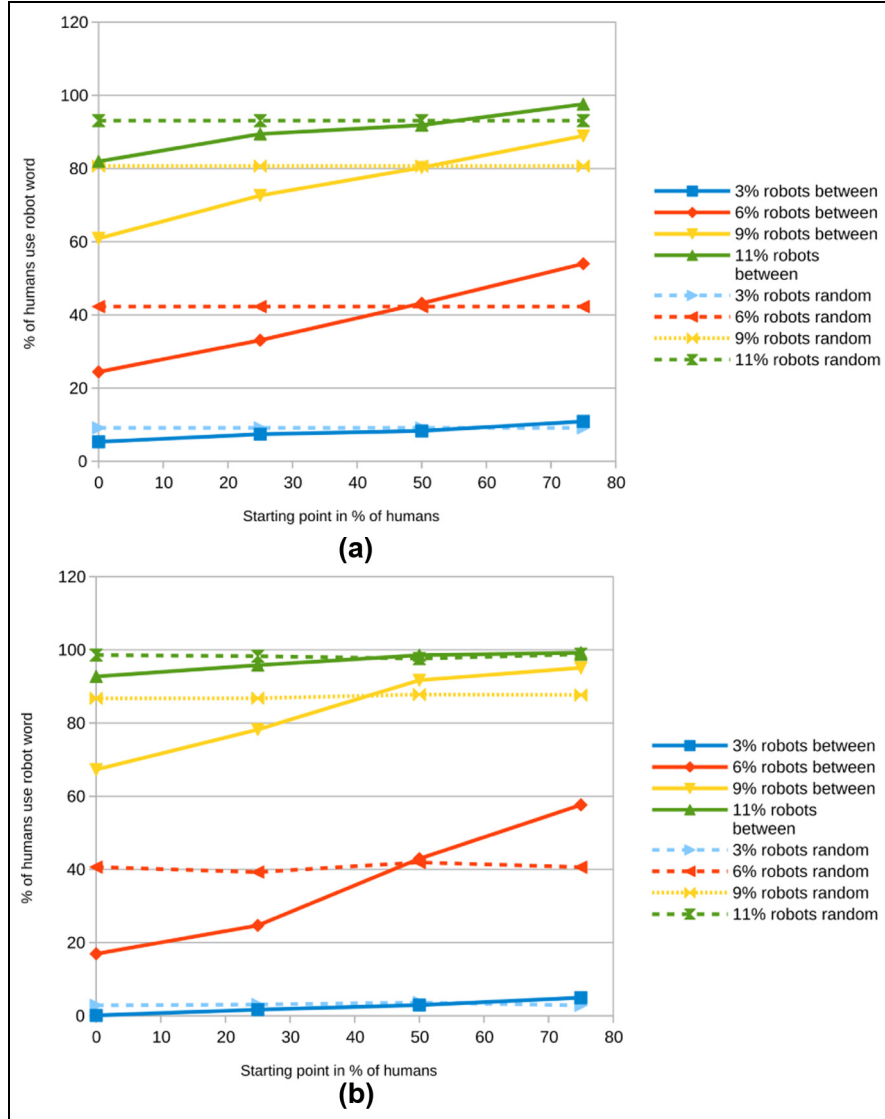


Figure 13. Betweenness-Centrality and random distribution starting at ranks 0, 25, 50, 75: (a) stage 2 setting, at iteration 5000, and (b) stage 3 setting, at iteration 37,351.

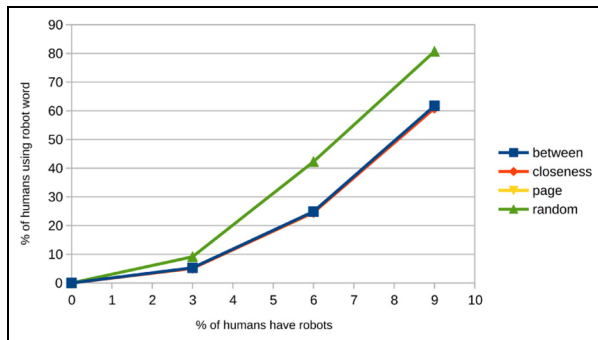


Figure 14. Centrality snapshot of maximum 9% of robots, at iteration 5000.

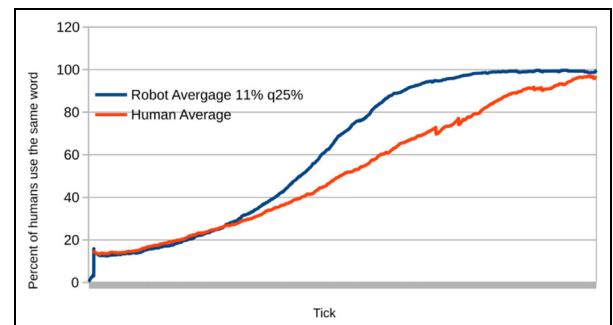


Figure 15. Humans only compared with 9% of robots with start position at 75% percentile.

measurements influences the dynamics of the simulation. In the first simulation setting, we always started with the human with the highest centrality value, that

is, the person who is best connected. When we compared the three centrality types, we could not find a significant difference. This comes from the fact that all

three centrality measures chose similar nodes as the most influential one in the network. From that point on, we discarded two types and only used Betweenness-Centrality to rank the nodes in the network. More interesting than the comparison of the three centrality types is that the random setting does not only seem different; it seems to have a stronger effect than the other three centrality types.

In our second question (Q2), we wanted to know whether the start position of the centrality measure had an influence on the robots' power to sway humans to use the robot's word. As Q1 showed that there was no difference between the three centrality types, but there was a difference between the random setting, we decided to run the simulation with different start positions.

After the analysis, we could indeed see a significant difference between the centrality start positions: The lower the rank of a person in the network, the stronger the effect of the robots. This somewhat counterintuitive outcome might come from the fact that highly connected nodes also have more people influencing the node itself. That means that a person with a smaller personal network gives each connection a higher priority, and therefore, a robot connection has more weight for a less connected person than for a highly connected person. This conclusion also shines light on the fact that the random distribution in Q2 was consistently stronger than when giving the robots to the strongest nodes.

This explanation is also in line with the linguistic knowledge of convention building. It turns out that new words tend to come from the less connected people, and highly connected people tend to adjust (Pierrehumbert et al., 2014). The explanation for this is again that highly connected people hear more options for the same word or idea.

In our last question (Q3), we wanted to know how many robots (in percentage) were needed to get into the Fixation Stage (where more than 90% of humans use the robot word). Our analysis showed that only 9% of robots in the network are needed to have a good chance to be able to achieve the Fixation Stage when the starting position is higher or equal to the 50% percentile centrality rank position. But when the robot count is higher than 11%, the robots most likely reach the Fixation Stage no matter at what centrality rank position they started. In Figure 13(b), the effect can be seen after 37,351 steps, which is the communication count measured in the Classroom Network. Given the considerable penetration of smartphones in our society³ that is well above 9% and that do already talk to users through their voice agents, it is easy to see what important role machine speakers will play. Smartphones will lead the way and robots are likely to extend their influence. Machines will have an influence on the development and usage of our language. Let us hope that they will not force us into a 'Newspeak' (Orwell, 1949), as political 'Bots' are already known to influence the

public opinion (Forelle, Howard, Monroy-Hernández, & Savage, 2015; Howard & Kollanyi, 2016; Suárez-Serrato, Roberts, Davis, & Menczer, 2016).

In conclusion, notwithstanding factors like bias, distribution and awareness that must also play a role, it can be said that when there are enough robots, putting those robots in the hands of the 'right' people will result in a very good chance that the humans will adopt the robots' words.

7. Future work

As we have completed our study on the Classroom Network, it would now be interesting to see whether we would get similar results using other networks. Testing this with other real-world networks would still be preferred over artificial networks.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was made possible through the support of a sub-award under a grant to Northwestern University from the John Templeton Foundation (Award ID 36617).

Notes

1. <https://thesocialclub.co/>
2. <http://www.sociopatterns.org/datasets>
3. https://en.wikipedia.org/wiki/List_of_countries_by_smartphone_penetration

References

- Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2006, P06014–P06014. doi:10.1088/1742-5468/2006/06/P06014
- Baronchelli, A., Gong, T., Puglisi, A., & Loreto, V. (2010). Modeling the emergence of universality in color naming patterns. *Proceedings of the National Academy of Sciences*, 107, 2403–2407. doi:10.1073/pnas.0908533107
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 3747–3752. doi: 10.1073/pnas.0400087101
- Baxter, G. J., Blythe, R. A., Croft, W., & McKane, A. J. (2009). Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change*, 21, 257–296. doi: 10.1017/S095439450999010X

- Beuls, K., & Steels, L. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PLoS ONE*, 8, e58960. doi:10.1371/journal.pone.0058960
- Blythe, R. A., & Croft, W. (2012). S-curves and the mechanisms of propagation in language change. *Language*, 88, 269–304. doi:10.1353/lan.2012.0027
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424, 175–308. doi:10.1016/j.physrep.2005.10.009
- Brandstetter, J., Beckner, C., Sandoval, E. B., & Bartneck, C. (2017). Persistent lexical entrainment in HRI. In *Proceedings of the 2017 ACM Association for Computing Machinery/IEEE International Conference on Human-Robot Interaction, HRI '17* (pp. 63–72). New York, NY: ACM. doi:10.1145/2909824.3020257
- Cangelosi, A., & Parisi, D. (1998). The emergence of a 'Language' in an evolving population of neural networks. *Connection Science*, 10, 83–97. doi:10.1080/095400998116512
- Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, 81, 591–646. doi:10.1103/RevModPhys.81.591
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329, 1194–1197. doi:10.1126/science.1185231
- Centola, D., & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, 112, 1989–1994. doi:10.1073/pnas.1418838112
- Centola, D. M. (2013). Homophily, networks, and critical mass: Solving the start-up problem in large group collective action. *Rationality and Society*, 25, 3–40. doi:10.1177/1043463112473734
- Fagyal, Z., Swarup, S., Escobar, A. M., Gasser, L., & Lakkaraju, K. (2010). Centers and peripheries: Network roles in language change. *Lingua*, 120, 2061–2079. doi:10.1016/j.lingua.2010.02.001
- Forelle, M. C., Howard, P. N., Monroy-Hernández, A., & Savage, S. (2015). Political bots and the manipulation of public opinion in Venezuela. doi:10.2139/ssrn.2635800.
- Freeman, L. C., Borgatti, S. P., & White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, 13, 141–154. doi:10.1016/0378-8733(91)90017-N
- Gillin, P. (2009). *The new influencers: A marketer's guide to the new social media*. Sanger, CA: Quill Driver Books.
- Griner, D. (2015). *Lord & Taylor Got 50 Instagrammers to wear the same dress, which promptly sold out – Adweek*. Retrieved from <http://www.adweek.com/news/advertising-branding/lord-taylor-got-50-instagrammers-wear-same-dress-which-promptly-sold-out-163791>
- Hall, J. (2016). *The influencer marketing gold rush is coming: Are you prepared?* Retrieved from <http://www.forbes.com/sites/johnhall/2016/04/17/the-influencer-marketing-gold-rush-is-coming-are-you-prepared/#49a2c08a2964>
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, 2, 96–100. doi:10.1111/1467-8721.ep10770953
- Howard, P. N., & Kollanyi, B. (2016). *Bots, #Strongerin, and #Brexit: Computational propaganda during the UK-EU referendum*. Retrieved from <https://ssrn.com/abstract=2798311>.
- Liu, J., Xiong, Q., Shi, W., Shi, X., & Wang, K. (2016). Evaluating the importance of nodes in complex networks. *Physica A: Statistical Mechanics and its Applications*, 452, 209–219. doi:10.1016/j.physa.2016.02.049
- Lu, Q., Korniss, G., & Szymanski, B. K. (2008). Naming games in two-dimensional and small-world-connected random geometric networks. *Physical Review E*, 77, 016111. doi:10.1103/PhysRevE.77.016111
- Lu, Q., Korniss, G., & Szymanski, B. K. (2009). The Naming Game in social networks: Community formation and consensus engineering. *Journal of Economic Interaction and Coordination*, 4, 221–235. doi:10.1007/s11403-009-0057-7
- Newman, M., Barabasi, A. L., & Watts, D. J. (2011). *The structure and dynamics of networks*. Princeton, NJ: Princeton University Press.
- Orwell, G. (1984). 1984. New York, NY: Harcourt Brace & Company.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The PageRank citation ranking: Bringing order to the web. *World Wide Web Internet and Web Information Systems*, 54, 1–17. doi:10.1.1.31.1768
- Pavlika, H. (2016). *What can we expect for influencer marketing in 2017?* Retrieved from <http://www.adweek.com/socialtimes/holly-pavlika-collective-bias-guest-post-influencer-marketing-2017/648472>
- Pierrehumbert, J. B., Stonedahl, F., & Daland, R. (2014). *A model of grassroots changes in linguistic systems*. arXiv: 1408.1985. Retrieved from <http://arxiv.org/abs/1408.1985>.
- Sabidussi, G. (1966a). The centrality index of a graph. *Psychometrika*, 31, 581–603. doi:10.1007/BF02289527
- Sabidussi, G. (1966b). The centrality of a graph. *Psychometrika*, 31, 581–603. doi:10.1016/j.socnet.2005.11.005
- Steels, L. (1995). A self-organizing spatial vocabulary. *Artificial life*, 2, 319–332. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8925502>.
- Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J. F., & . . . Vanhems, P. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE*, 6, e23176. doi:10.1371/journal.pone.0023176
- Suárez-Serrato, P., Roberts, M. E., Davis, C., & Menczer, F. (2016). *On the influence of social bots in online protests* (pp. 269–278). Cham, Switzerland: Springer. doi:10.1007/978-3-319-47874-6
- Talavera, M. (2015). *10 Reasons why influencer marketing is the next big thing*. Retrieved from <http://www.adweek.com/socialtimes/10-reasons-why-influencer-marketing-is-the-next-big-thing/623407>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442. doi:10.1038/30918
- Wilensky, U. (1999). *NetLogo*. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL 2009. Evanston, IL. Retrieved from <http://ccl.northwestern.edu/netlogo/>
- Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., & Szymanski, B. K. (2011). Social consensus through the influence of committed minorities. *Physical Review E*, 84, 011130. doi:10.1103/PhysRevE.84.011130

Author biographies



Jürgen Brandstetter is currently CTO at Osnova, where he works on artificial intelligence (AI)-driven mathematics tutoring. He holds a PhD in human interface technology from the HIT Lab NZ (University of Canterbury). His field of research is in human robot interaction with a focus on persuasive robots. He has a master's degree in computer science from the University of Vienna. He was also a visiting scholar at the University of Oxford and worked at Microsoft Research. He also has an entrepreneurial spirit that reflects in being a cofounder in multiple tech-startups.



Christoph Bartneck is an associate professor and director of postgraduate studies at the HIT Lab NZ of the University of Canterbury. He has a background in industrial design and human-computer interaction, and his projects and studies have been published in leading journals, newspapers and conferences. His interests lie in the fields of human-computer interaction, science and technology studies and visual design. More specifically, he focuses on the effect of anthropomorphism on human-robot interaction. As a secondary research interest, he works on bibliometric analyses, agent-based social simulations, and the critical review on scientific processes and policies. In the field of Design, he investigates the history of product design, tessellations and photography.