Journal of
Thai Interdisciplinary Research
Volume 13, Number 1, Pages 33–39

Multi-language communication protocol model based on conceptual spaces and language games

Somjin Juntarajessadakorn[1,2,*], Vatinee Nuipian[3], and Phayung Meesad[4]

[1]Department of Information Technology, Faculty of Information Technology, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand
[2]Department of Information Technology, Faculty of Science, Nakhon Pathom Rajabhat University, Nakhon Pathom, Thailand
[3]Department of Computer Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand
[4]Department of Information Technology Management, Faculty of Information Technology, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

Abstract

Interacting artificial intelligent systems need to understand human language. This paper demonstrates how artificial systems can learn and communicate with each other in three different languages: Thai, Chinese, and English. In this work, we extend a conceptual model based on Language Games and Self- Organizing Map neural networks. In addition, the model proposed includes an important feature that is a dynamic radius for multi-languages communicative interactions between autonomous agents. We present that each innate agent has their own mother tongue or first language according to their parents' nationality, primarily stores the primitive words in their own language list, and then maps it with other languages. Our proposed model shows the visualized simulation of the potential ways of a communication protocol during Language Games in which autonomous agents can learn and transfer knowledge of different languages between pairwise agents. Finally, the effectiveness of the communication protocol model is evaluated with the 100% success rate to indicate the co-occurrence of word and object convergence matching all three languages.

Keywords: language games, conceptual space, Self-Organizing Map, multi-language

1. Introduction

Language Game was originally introduced by Ludwig Wittgenstein [1] as a concept mapping an object reality to language where a speaker utters a word associated with an object, and a listener can infer that object. The notion of Language Game is long used in cognitive science, particularly in artificial intelligence (AI) since Luc Steels [2], the pioneer, introduced the work of robotics on the grounds of Language Game and many other types of research continuing until now [3-8]. It is unclear whether nowadays AI autonomous systems i.e. intelligence agents can totally learn and communicate in as complicated a way as natural human languages. However, this research area also still inspires many scientists to expose their new principles and techniques to form human concept models for autonomous agents, since most of the applications of AI interact with the human need to understand via natural human language. Nevertheless, many research studies based on Language Game have claimed how autonomous agents communicate in the way of human concepts. As such [9], Language Games and Conceptual Spaces theory [10] are used to simulate meaning negotiation in multi-agents, while other research [11, 12] combining prototype [13] and random set theory [14] study a formalization of concepts based on human concept use. However, it is said that there are many approaches that utilize the combination techniques based on Language Game and Conceptual Spaces in which those research studies have not been clarified here.

According to previous works [9], in the simplified version of the Language Games, agents perform pairwise games by communicating only in English in order to reach agreement on the name to assign to a single object. In this work, we propose an extension of a conceptual model framework, which is based on Language Games and Self-Organizing Map neural networks (SOM) [15] with a dynamic radius for multi-language communicative interactions between autonomous agents. The communication protocol is provided in three different languages: Thai, Chinese, and English. We assume that each agent has their own mother tongue or first language according to their parents' nationality. As a human, a child expects to inherit a set of language definitions initially similar to those of their parents. Further, they acquire knowledge of their daily physical world by interacting with their family.

[*]Corresponding author; e-mail: somjin@npru.ac.th

DOI 10.14456/jtir.2018.6

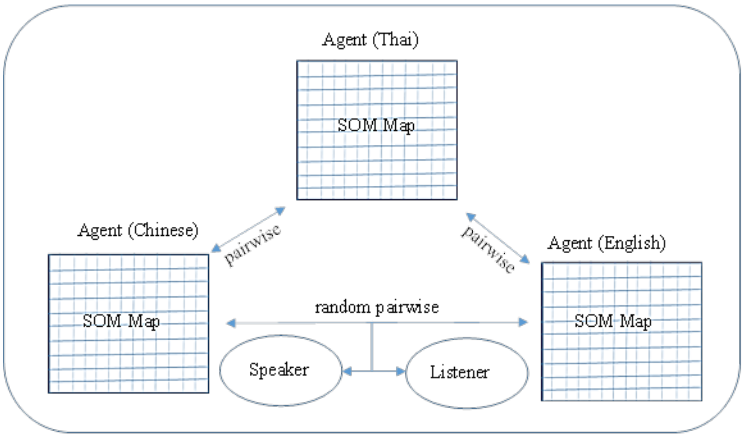

Figure 1 Conceptual framework of the communication protocol model

Therefore, innate agents primarily store their primitive words in their language list. We employ this concept in our proposed model in the following sections.

2. Research objectives

Our objectives of this paper are to propose a model for autonomous agents to learn and communicate with each other, in particular, with natural human language in a multi-language communication protocol model, and evaluate the convergence of the success rate of all languages learning and transferring in the multi-agent simulation framework.

3. Methods

3.1 Language Games

In a typical Language Game, the two agents find agreements on object and word pairing until they meet a consensus on what the object should be named. A Language Game is a communication process in a multi-agent system. Vogt and Coumans exposed in their work [16] that Language Games can be discriminated into three different styles: observational game, guessing game and selfish game. In the observational play, the speaker and the listener both know the topic of the game, such as colour. The speaker utters a word denoting the subject, for instance, the name of the colour. The game is successful if the listener knows the uttered word in the right meaning, but the game fails otherwise. During the guessing game, only the speaker knows the topic of the game. The speaker utters a word denoting the subject, and the listener must guess which subject the speaker means. In contrast, in the selfish game, the agent has to infer the meanings of words from their co-occurrences in different contexts or meanings. However, it is concluded that observational games and guessing games are faster than selfish games. Thus, an observation game is selected, and employed in our proposed model.

3.2 Conceptual Spaces

Conceptual Space is a metric space in which the main feature dominates on the ground of a geometrical framework for representation of the concepts based on the number of quality dimensions. Each dimension is associated with a metrical structure that allows a way of calculating the semantic similarity among concepts by using metrical distances. The ideas are not independent of each other but can be structured into domains, for instance, shapes and colours. In this work, Self-Organizing Map (SOM) was applied in which a concept is represented as a node of colour in a conceptual map (i.e., SOM map as in Figure 1). The conceptual map has three quality dimensions of red, green and blue. The distance between nodes is calculated by the Euclidian technique together with the searching technique; best-mapping unit (BMU), as part of the learning tools for the agreement process. Each node includes an array of concepts where each concept holds a single word of three different languages in a series of hash maps associated with a weight of belief.

In the process of concept mappings, each agent needs to formulate an innate conceptual map with the prototypes only in the initial state. The prototypes of 10 colours of gray, blue, green, light blue, red, pink, yellow, white, black and orange, including phonetic alphabets (Table 1) of their first language (i.e., each agent was born with only one nationality) are defined to create the innate conceptual map at the beginning of the game play. The SOM forms a conceptual map that is a repository of the physical world (the colours) for the internal cognition of the agent with a weight of belief in the association between the lexicon and the conceptual map. The concept's belief is an integer value from 0 to 100. Figure 2, an instance, shows the concept maps of a pairwise Thai agent and English agent in possible game plays. If a game succeeds, the belief will be increased by one, with a maximum of 100. In contrast, if a game fails, one, with a minimum of zero, decreases the belief.

From Figure 2 concept maps of pairwise Thai and English agents in possible game plays. First, the random topic is shown to the Thai speaker, the process of BMUs the search in their concepts. In this case the word "som" is found (with belief = 15) and shown to the English listener. Later they search in the

Table 1 Phonetic alphabets of ten colours in three languages

Named colors in rgb: 0.0f - 1.0f formats	Thai	Chinese	English
0.25f, 0.25f, 0.25f	thea	huse	gray
0.0f, 0.0f, 1.0f	na-ngein	lan se	blue
0.0f, 1.0f, 0.0f	kheiyw	luse	green
0.0f, 1.0f, 1.0f	fa	qian-lan	light-blue
1.0f, 0.0f, 0.0f	daeng	hong	red
1.0f, 0.0f, 1.0f	chomphu	fen	pink
1.0f, 1.0f, 0.0f	lụxng	huangse	yellow
1.0f, 0.9f, 1.0f	khaw	baise	white
0.0f, 0.0f, 0.0f	dam	heise	black
1.0f, 0.25f, 0.0f	som	chengzi	orange

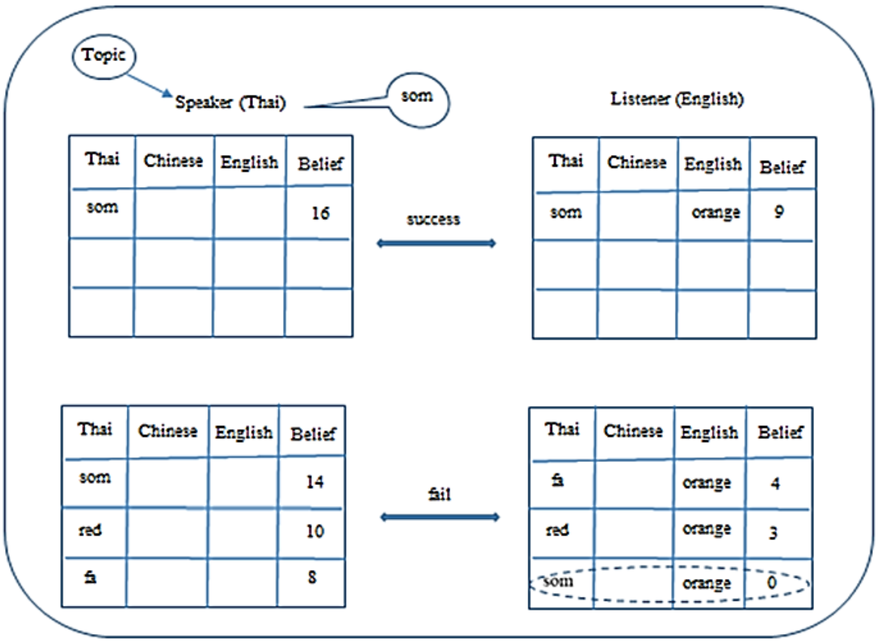

Figure 2 Concept maps of a pairwise Thai agent and English agent

same way as the Thai speaker. In their memories, the native words (e.g., orange) is primarily stored in hash maps which are reserved to keep three languages with a belief. In a win game, the English listener found the word "som" in their concepts; both the speaker and the listener remove other words and increase their belief (as shown in the upper part of the figure here); the counter belief of the Thai speaker and the English listener are 16 and 9 respectively. In a loss game, as shown in the lower part of this figure, the listener updates their knowledge by adding the word "som" into their memories, and the speaker decreases their belief by one (e.g., belief = 15-1 = 14)

In this research, an observational game is applied when both speaker and listener know the topic of game with some weight of belief associated with the pairing between lexicon and conceptual memory in advance. The proposed model based on Language Game algorithms can be described as follows:

1) The two agents are chosen randomly from groups of agents (one agent one nationality), one agent is assigned the role of speaker and the other of listener.

2) The topics are initially generated in three languages (Thai, Chinese and English), and then they are chosen randomly in the game according to the speaker's nationality.

3) Th speaker searches for a word in their belief spaces in SOM map with technique of the best-matching unit (BMU), and then utters this word in their native language to the listener.

4) The listener searches this topic in their memory, using the same technique as the speaker. In the searching process, each agent's memory primarily store their native words (i.e., first language according to their nationality) in hash maps, which are spaces for obtaining three languages. If the uttered word is found, the game is considered successful and both the speaker and the listener remove other words. So the

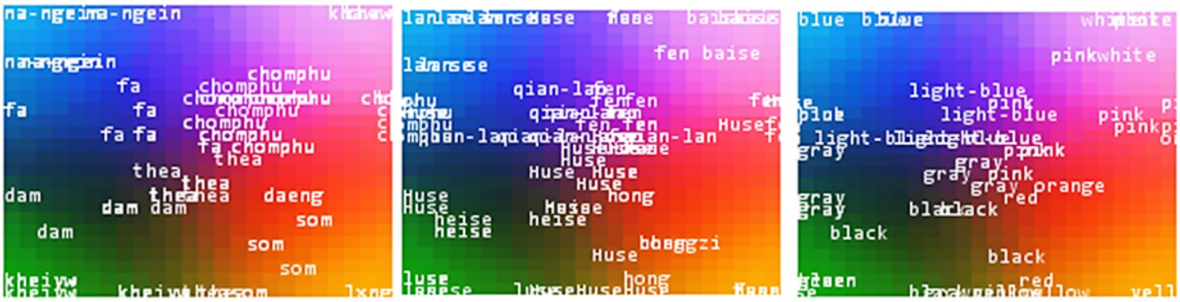

Figure 3 SOM maps of the speaker in three languages: Thai, Chinese and English in 500 games

Figure 4 SOM maps of the listener in three languages: Thai, Chinese and English in 500 games

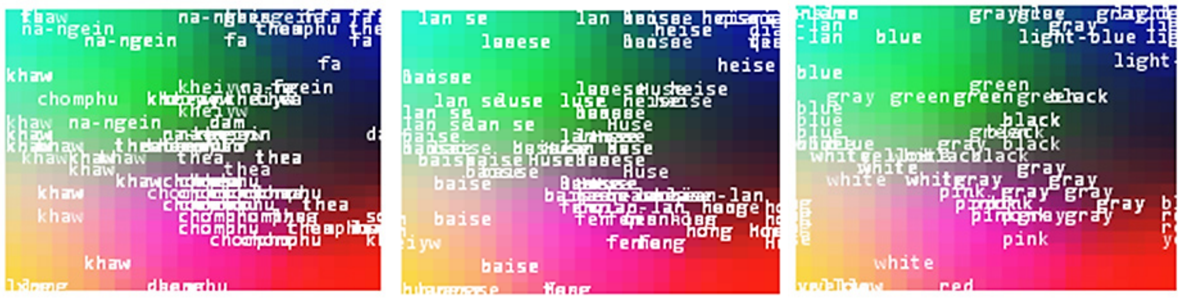

Figure 5 SOM maps of the speaker in three languages: Thai, Chinese and English in 1000 games

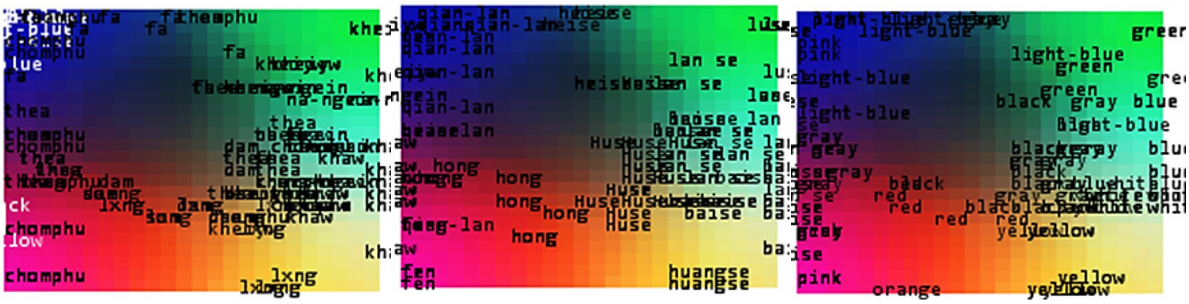

Figure 6 SOM maps of the listener in three languages: Thai, Chinese and English in 1000 games

uttered word stores two mapping languages from both the speaker and the listener.

5) If the previous step fails, the listener updates their knowledge by adding this uttered word into their memories, and the game is considered unsuccessful.

6) In a successful game, both speaker and listener increase their belief counters by one. In a failed game, only the speaker decreases their belief counter by one. The minimum belief is set at zero and the maximum belief is set at 100.

3.3 Experiments

After the SOM is trained 1,000 times with prototypes, a random topic (in a colour vector) is shown to the speaker and the listener. The process of searching for the best words is undertaken; the speaker first examines in their memory, which is stored as a hash collection. If the speaker finds the mapping node in this group, they keep this word and associate it with the node. For effective communication, the speaker will search for a word within radius R in their SOM. This radius is set to expand (e.g., r+= 0.2f; 0< r < 30.0f) until the speaker can find a word in the SOM map, and is mapped as the BMU. This node may contain many concepts. Each concept stores a word associated with the word's beliefs. Noticeably, a counter defines the weight of expectation; each time the game succeeds, the bar is increased by one. Therefore, the word is sorted by the weight of belief. Later the speaker utters this word to the listener as the topic that was shown to the speaker is also shown to the listener. The listener performs searching in the same way as the speaker does in their BMUs. If the listener finds this word, the game is considered successful and other words are removed from their memory. If the listener cannot find a word, the game is unsuccessful. The listener must add the speaker's word to their BMUs and the speaker decreases their belief counter by one for this word. In case of a successful game, the speaker and listener add their belief counters for that word by one.

3.4 Evaluation

The success rate indicates the effectiveness of the communication protocol of the proposed model. The formula, based on moving average approach, is used for evaluating the performance, and is calculated every 10 games after the listener finishes searching for a word as follows:

$$\text{Success Rate} = \frac{1}{m}\sum_{i=0}^{m}\left(\frac{1}{n}\sum_{j=t-n+1}^{t} x_j\right)_i \quad (1)$$

where x_j is the number of successful communications evaluated after every 10 games, n is the number of sliding windows, m is the total number of agents, and t is the number of games.

4. Results and discussion

The Language Game together with Conceptual Space approach, implemented on self-learning neural networks, i.e., SOM with a dynamic radius, can be used as a tool to communicate and share words of the spontaneous agreement process in artificial agents. We analysed the comparative performance,- the effects of the total number of agents and the number of game plays. The success rate of 50 agents in all languages is calculated and concurrently plotted on the discrete graph every 10 games where the total number of the games are: 500 games (Figure 7a), 1000 games (Figure 7b), 1500 games (Figure 7c) and 2000 games (Figure 7d) respectively. The charts show that the mapping between lexicons and memories slowly reaches agreement through 500 games (e.g., success rate = 0.7 approximate), and increases during 1000 games (e.g., success rate = 0.9 approximate), and more until the rate reaches the highest rate (success rate = 1.0), and then stabilizes. Moreover, Figures 8a–8d depict the success rate of 100 agents (double top up) in the same manner. Noticeably, all charts show a rather lower rate to reach agreements compared to those success rates of 50 agents. In conclusion, the evaluation results showed that the number of games yield fast convergence to reach agreement (Figures 7a–7d), while the number of pairwise agents has significant effects on the agreement in a society of agents (Figures 8a–8d). As in the analogy, like human beings, the more people there are the more problems they have.

During the purposed model simulation, SOM is considered an artificial neural network model of the brain. In the artificial brain, there is a conceptual space represented in two dimensions of SOM maps. Each SOM map consists of 24 x 24 nodes for simulated protocol communication in which each pair of speaker-agent and listener-agent learn and transfer their words in their languages. Figure 3 and Figure 4 show the SOM maps and lexicons of the paired speaker and listener in three languages: Thai, Chinese, and English (ordering from the left to the right hand-side), when they play 500 games. Similarly, Figure 5 and Figure 6 show the SOM maps of the other random paired speaker and the listener-agents when they play 1000 games. It is noted that each agent has a different form of nodes which is a kin to the form of the human brain: each one is different. In this work, the simulation of SOM maps demonstrate a clear picture of node mapping in three languages on cloned SOM maps, in which there is one language per map separation. However, all agents have the same number of clusters under the influence of prototypes, i.e., clustering in ten colours.

5. Conclusions

The contribution of this paper is the proposition of a simulation framework: a communication protocol model, in order to ground the symbols used during multi-language interactions between autonomous agents. To achieve that, our proposed model uses a Self-organizing Map with a dynamic radius technique based on the Conceptual Spaces approach and the notion of Language Game. Finally, we have shown that the number of games played and the number of pairwise agents affect the way in which final convergence is reached.

This work is a part of our dissertation. In future work, we will simulate agents simultaneously playing two roles, both speaker and listener, within a group of a speaker and multiple listeners. In addition, we will investigate word ambiguity and word variability to create effective communication as close to natural human languages as possible, and also to improve the emergence of the communication protocol model.

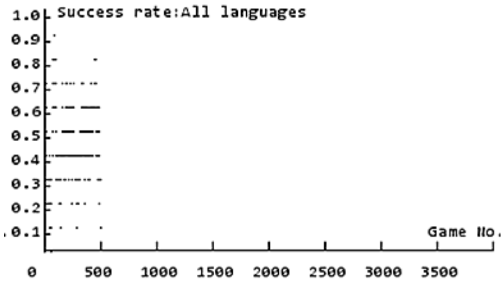

Figure 7a Success rate in 500 games

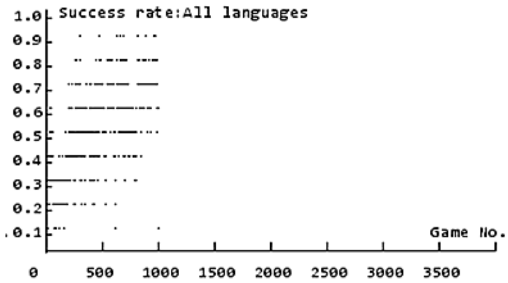

Figure 7b Success rate in 1000 games

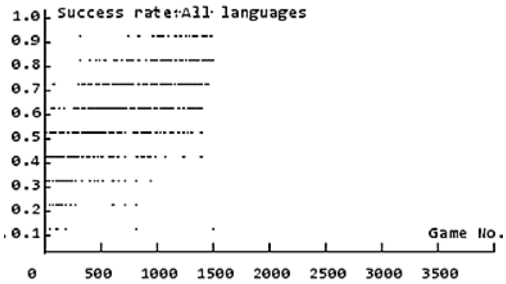

Figure 7c Success rate in 1500 games

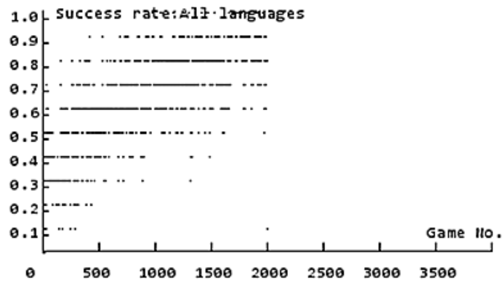

Figure 7d Success rate in 2000 games

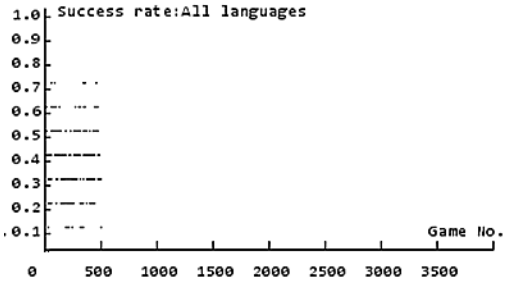

Figure 8a Success rate in 500 games

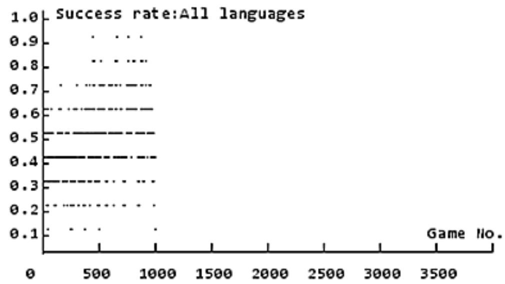

Figure 8b Success rate in 1000 games

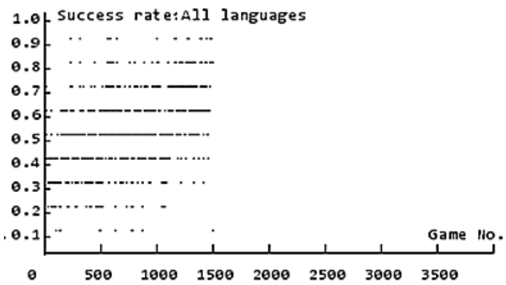

Figure 8c Success rate in 1500 games

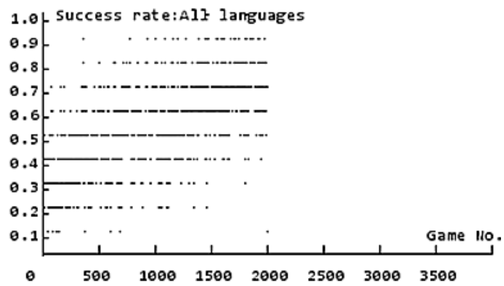

Figure 8d Success rate in 2000 games

References

[1] Wittgenstein L. **Philosophical investigations.** John Wiley & Sons; 2010.

[2] Steels L, Vogt P. Grounding adaptive language games in robotic agents. **Proceedings of the fourth european conference on artificial life**; 1997 Jul (Vol. 97).

[3] Steels L, Kaplan F. Stochasticity as a source of innovation in language games. **Proceedings of the Conference on Artificial Life VI (Alife VI)(Los Angeles)**; 1998 Jun 368-76.

[4] Steels L. Language as a complex adaptive system. **In Parallel Problem Solving from Nature PPSN VI;** 2000 (pp. 17-26). Springer Berlin/Heidelberg.

[5] Steels L. Language games for autonomous robots. **IEEE Intelligent systems**; 2001 Sep; 16(5):16-22.

[6] Steels L. Grounding language through evolutionary language games. **Language Grounding in Robots**; 2012:1-22.

[7] Steels L. Meaning and Creativity in Language. **In Creativity and Universality in Language**; 2016 (pp. 197-208). Springer International Publishing.

[8] Steels L. Human language is a culturally evolving system. **Psychonomic Bulletin & Review** 2017 ;24(1):190-3.

[9] Lindh-Knuutila T, Honkela T, Lagus K. Simulating meaning negotiation using observational language games. **Symbol grounding and beyond** 2006: 168-179.

[10] Gärdenfors P. **Conceptual spaces: The geometry of thought**. MIT press; 2004.

[11] Lawry J, Tang Y. **Uncertainty modelling for vague concepts: A prototype theory approach**. Artificial Intelligence 2009;173(18):1539-58.

[12] Eyre H, Lawry J. Language games with vague categories and negations. **Adaptive Behavior** 2014 ;22(5):289-303.

[13] Geeraerts D, editor. **Cognitive linguistics: Basic readings**. Walter de Gruyter; 2006.

[14] Goutsias J, Mahler RP, Nguyen HT, editors. **Random sets: theory and applications**. Springer Science & Business Media; 2012.

[15] Kohonen T. Essentials of the self-organizing map. **Neural networks** 2013;**37**:52-65.

[16] Vogt P, Coumans H. Investigating social interaction strategies for bootstrapping lexicon development. **Journal of Artificial Societies and Social Simulation** 2003;**6**(1).