COMPUTER MODELS OF LANGUAGE DIVERSITY
A Determination of Social and Individual Benefits

René J. Jorna, Niels R. Faber
Dept. of Social Sciences,Frisian Academy (KNAW), Doelestraat 8, Leeuwarden, The Netherlands
Dept. of Business and ICT, University of Groningen, Nettelbosje 2, Groningen, The Netherlands
r.j.j.m.jorna@rug.nl, nfaber@fryske-akademy.nl

Keywords: Multilingualism, multi-actor simulation, language competition, language vitality.

Abstract: In studying benefits and costs of language diversity the use of computer models is rare. There may be various reasons for this situation, e.g. a) the complexity of the interpretation of language diversity, b) the difficulty in operationalizing factors and dimensions of diversity, c) the absence of realistic actor models, that is to say of actors that can communicate in a realistic way and d) the underperformance or restrictedness of existing software programs. They may be only a few of the reasons to abstain from using computers models to analyze, explain and predict language diversity. How valid these arguments may be, it is strange that for weather analyses and forecasts and for tide and streaming conditions many computer models have been developed and are successfully applied. Tides and weather are also complex, many factors and dimensions are also applicable, and software programs and computational power can always be enhanced. The only big difference that we can think about in modeling and comparing natural (tide and weather) and artificial (language and communication) systems is the presence or absence of a realistic model of plausible cognitive actors and the interaction of these actors. It is also possible to defend the line of reasoning that by definition humans systems and their social structures cannot be modeled in software. The existence of cognitive science and social simulation since the 60s of the last century to our opinion shows that the rejection of this kind of modeling is not justified. In our presentation we will propose a pseudo specification of plausible cognitive actors that have a memory, have goals, have communication instruments, i.e., a language, and are able to exhibit social behavior (Helmhout, 2006; Wijermans, 2008)..

1 INTRODUCTION

The debate about the value of mastering and speaking more than one language has its ups and downs. One the one hand it is argued that if as many people as possible speak just one language, communication between people will be more smoothly and less misunderstanding takes place. On the other hand it can be argued that more languages are needed, because language expresses to a large extent the way people look at their (regional) part of the world. Language incorporates historical and cultural heritage. From the 1980s on, in the heydays of focusing on economic growth and uniform economic regulations in the EC, the idea was that if enough people would learn and speak English including non-native speakers, European profits and markets would flourish. This could mean a neglect of other languages than English. However, ten years ago (since the end of the 90s) a renewed interest in regional and minority languages started in the EC. Not that the interest in languages in the EC was gone in the past, but the entrance of new member states in the EC with their diverse languages and the growing importance of China with a completely different language structure diminished the dream or nightmare of one European (English) Language.

The foregoing arguments, pro or contra, can be seen as mainly ideological in nature, but there are also some factual observations pro or contra multilingualism. We mention three. In the first place, our human psychological (cognitive) structure prevents us from using and changing (multiple) languages as if we are changing trousers or buying new cars. Language is close to ourselves. Psychological research from the last twenty years showed that the best time to learn a new second language is if we are between 3-5 years old. Complete multilingualism not only in language comprehension, but also in language production (and pronunciation) is nearly impossible if you start after age 12.

In the second place, sociological research has shown that using a language is subject to the same social comparison mechanisms as are income, housing and birth. With using a certain language one may show others that one has a higher social status or is of a higher birth. For example, in the past in middle Java three languages were used with respect to communication. Depending on the social status of the other person, one should use language 1 and was forced not to use language 2.

In the third place, it is a linguistic fact that because of increasing communication and globalization less and less languages survive. Of the 6000 languages that existed hundred years ago, more than 2000 are being extinguished or have only a few speakers left. This may be regrettable from a linguistic point of view and also from the perspective of language diversity, it is a fact that is worth studying.

Taken into account the (cognitive and social) psychological and sociological facts about language learning, its use and status, the question is whether we can understand the mechanisms and characteristics of language dominance or extinction, of language use and of multilingualism? One way of studying these mechanisms is by doing intensive and extended empirical research. Although interesting and necessary, empirical research with respect to multilingualism has two shortcomings: a) the difficult separation of the levels of aggregation of individual, group and societal, and b) the long time-horizon over which language learning or multi-language use take place and change. We, therefore, suggest another line of research: computer simulation models. In our case this means the use of multi-actor simulation models (MAS-models) consisting of cognitive plausible actors. In the remainder of this article we start with a limited inventory of the simulation models that are available. We then complete the shortcomings of the models with components that we think are necessary. Then we will give design and technical specifications of a multi-actor model of multi-language use with cognitive plausible actors. We will look at these models in terms of ability to answer the various questions. Finally, we continue with formulating some questions that a computer simulation model might answer. We emphasize that at this moment we cannot present results of computer simulation runs.

2 BACKGROUND: EXISTING SIMULATION MODELS OF LANGUAGE COMPETITION

Various initiatives have been formulated to model language extinction and revitalization (Fernando, Lissa & Goldstein 2007), bilingualism and language competition (Castello, Toivonen, Eguiliz & San Miguel, 2006) and corrections or completions of the Abrams-Strogatz model (Staufferl, Castello, Eguiliz & San Miguel, 2007).

A basic and often used model was formulated by Abrams and Strogatz (2003). They focused on languages as fixed, and as competing with each other for speakers. They also assume a highly connected population with no spatial and social structure in which all speakers are monolingual. In their assumptions they also mention that two languages cannot coexist stably. They plotted their model to data from 42 regions in various parts of the world and they found that status is most relevant, linguistically. Many other models were developed based on their dynamics of language death.

Fernando et al. (2007) devised a mathematical model to understand language extinction. By what characteristics and mechanisms do languages fade away? This dynamic modeling of languages not only looks at the mechanisms, but also studies interventions to raise the number of speakers again. Interventions may concern explicit teaching of minority languages at school or using multimedia, television and radio in the minority language to stimulate the endangered language. Fernando's et al. model reveals that it is possible to preserve a minority language, but that long-term interventions will be necessary. As parameters in their model they use linguistic and societal circumstances (e.g., the responsiveness of children to learning a language as a function of the proportion of conversations heard in that language, the relative importance of conversations heard in the family and elsewhere, and the amplification of spoken to heard conversations of the high status language because of the media). However they stress that the lack of existing phenomenological data concerning language use, transfer and distribution of minority languages limit the validity of the model they developed and used for explanation

Castello's et al. (2007) developed an agent based model to study the competition between two socially equivalent languages. What are the effects of bilingualism and social structure? They use small world networks, like community structures, to look at linguistic domains and their boundaries. Their work showed a final scenario of dominance of one language and extinction of the other and they were able to replicate the intuition that the average time for extinction of a language is a function of the size of the community. The larger the community, the longer it takes for a language to become extinct.

Staufferl et al. (2007) used the well-known Abrams & Strogatz model to compare it with agent based Monte Carlo simulations. Again they used networks of agents to populate their models. Key words in their research are dynamics of language, complex systems and social interactions. Starting with computer models of language competition that mainly appeared in in physics journals, they compare their model with the one in the publication in 2003 by Abrams–Strogatz. The main goal of Staufferl et al. work is to check to what extent the results of the mean-field approximation of the Abrams–Strogatz model are confirmed by agent-based simulations with many individuals.

Many other models can be discussed with sometimes very elegant mathematics (Wang & Minett, 2005), but we argue that in these models at least three components are missing. In the first place it seems that a language has status. To us, this seems to imply that a language is an entity that can exist without some (human) actor knowing and using it. We oppose this stance, and instead see language as an artifact, i.e., a human made construct, that is the result of individual knowledge and social interaction. Status in this way is not ascribed to the language as such. Instead status is ascribed to an actor when the actor uses a particular language. This implies other ways of theorizing, and thus working with what Simon (1969) called: using the sciences of the artificial. Simon explained that many human made artifacts should be studies in terms of complexity, architecture and coherence.

In the second place, because humans are the sign and symbol developers, learners and users, every simulation model should take into account the fact that humans are cognitive or information processing systems. This means that every simulation, also an agent based simulation, should be based on plausible cognitive actors (Helmhout, 2006; Jorna, Faber, & Hadders, 2009; Wijermans, 2011).

In the third place, because humans are besides cognitive systems also social actors, simulation models should include social parameters, such as coordination mechanisms, e.g., standardization, authority and mutual adjustment (Jorna, et al., 2009; Gazendam & Jorna, 1998).

In the next sections we will first go into the details of the constituents of cognitive actors, after which we describe the functional and technical specifications of our simulation model and the questions it may answer.

3 SIMULATION MODELS OF MULTILINGUALISM USING PLAUSIBLE COGNITIVE ACTORS

As indicated in the previous section, current simulation approaches towards language diversity, including agent-based initiatives, have advantages and specific limitations. The limitations concern both the explanation and prediction of the development of situations in which multiple languages are used. We argued that current approaches ascribe autonomous behavior to a language. Phenomena such as competition between two or more languages are consequently described as naturally occurring processes. They are not based on the fact that language is an artifact and, therefore, completely human based. What is lacking in the used models is the central role of the actors who use language to communicate and to implicitly define their own identity. We argue that language, language learning and language diversity is not just an autonomous system with its only its own dynamics. Language is not just out there as is the rain or are mountains. Language and language diversity is the result of a social system, its dynamics, and the cognitive actors that comprise the social system.

In contrast to the agent-based approach presented by Costello et al. (2007) and Staufferl et al. (2007), we suggest the use of a multi-actor simulation in which also cognitive plausible actors are modeled. In this way, we are able to separate language and language diversity from the social system in which they emerge. Being able to make this separation, enables us to study the social dynamics (i.e., the causes) that underlie changes in language and dynamics in situations of language diversity (i.e., the effects). This section continues more detailed conceptual specifications of a MAS approach.

Cognitive actor approach: The elementary building block of our approach is a cognitively plausible actor. This plausible cognitive actor extends the agent concept. In the field of Artificial Intelligence, an agent is defined as “a computer system, situated in some environment, that is capable of flexible autonomous action in order to meet its design objectives” (Wooldridge, 2002, p.15). The agent is subsequently ascribed a fixed set of properties consisting of autonomy, social ability, reactivity, and pro-activeness (Bellifemine, Caire, & Greenwood, 2007, p.3-4). This means that an agent is able to operate without human intervention, has the ability to communicate with other agents or humans, is able to respond to its environment, and acts in the environment in pursuit of its own agenda (i.e., its own set of goals). Additional properties an agent might have are mobility, truthfulness, benevolence, rationality, and the ability to learn (Bellifemine et al., 2007, p. 4). The properties that are ascribed to an agent however, do not specify, nor indicate how these characteristics are realized in the construction of an agent. We, therefore, need an extended agent concept, in which human information processing in terms of cognitive architecture, mental representations and mental processes are implemented and realized.

In distinguishing the terms agent and actor, we follow Helmhout's (2006, p. 19) stance. Although he indicates that in general the distinction is not always clear, he provides a criterion that is relevant for our purpose. Namely, an actor is cognitively plausible, based on insights of cognitive science, and therefore more autonomous than an agent. Helmhout connects the agent concept to the notion of agency, or the idea that an agent serves a person. This implies that although the agent portrays the properties as described above in its behavior, they are initialized by his principal (i.e., the person to whom the agent belongs). An actor in contrast is cognitively plausible and more autonomous in the sense that it acts on its own behalf and does not perform a service for a real person.

A cognitive plausible actor is constructed around a subsystem that allows it to operate on a functional and intentional level of description (Dennett, 1978, 1987; Jorna, 1990; Helmhout, 2006). Although Dennett also distinguishes a physiological/physical level on which functions and intentions run, this level is not important, here. The functional level provides learning and cognitive mechanisms (Newell & Simon, 1972; Anderson & Lebiere, 1998). The intentional level enables an actor to possess beliefs, desires, and intentions (Helmhout, 2006). Gazendam & Jorna (2002) argue that rationality and intelligence reside at the intentional level.

Helmhout (2006) presents a detailed model of a plausible cognitive agent (see Figure 1).

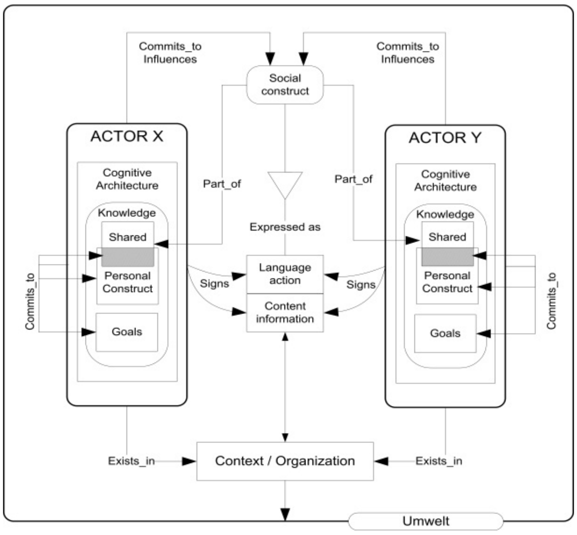


Figure 1*: Multi-actor simulation model with cognitive plausible actors (Helmhout et al., 2009*

In this model two actors X and Y are presented. The actors live in larger social environment (Umwelt) in which organizations, societies and institutions exist. The actors use social constructs (social structures) and signs (e.g., organized in languages), by which they communicate and exchange content information. Actor X and Y themselves are modeled according to standard theory within cognitive science. They have an architecture consisting of memory, perception and action/motor components. The content of memory consists of goals and personal constructs, formulated in terms of mental representations. By default there is a connection between the representations and the language or multiple languages the actors use. This means that an actor has one or more languages at his disposal by which an actor makes sense of reality to himself and to others. From a more general perspective a cognitive actor can be specified consisting of three main components: cognition and perception and action. One way to model the cognitive process components is using classical cognitive science ((Card, et al, 1983; see Figure 2).

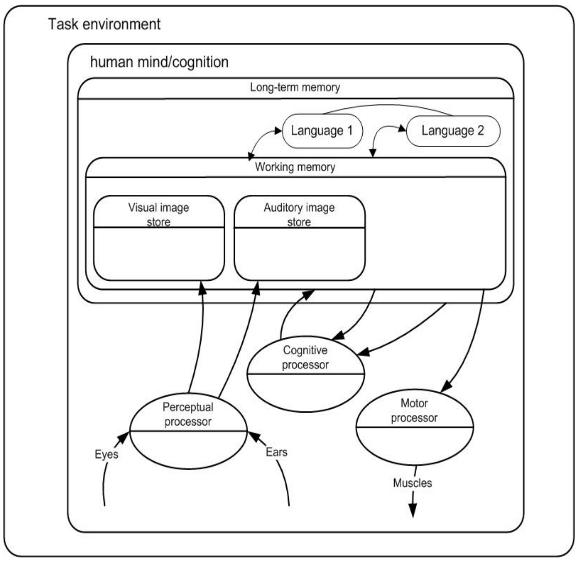


***Figure 2**: Model van Actor: Human Information Processing / Cognitive System (based on Card et al. 1983)*

A cognitive system, the actor, is a rational actor consisting of architecture, mental representations and processes on representations. Long term and working memory and storages are architectural components of cognition. Within memory we have representations consisting of beliefs, desires and intentions, which can be the expressions of goals, rules or facts (existing of declarative and procedural memory). The perception component enables the actor to perceive its environment, and forms the input system of the actor. The action component enables the actor to act, to speak, to interact or to write within its environment. It provides the actor with the ability to change its environment according to its own wishes. The cognitive processor is the subsystem that provides the actor with rationality and intelligence, although they are also present in the memory system. Helmhout (2006) identifies three functional elements that form the minimum set of functions of which the cognition component needs to consist, namely memory, learning, and goal-directed behavior (Helmhout, 2006, p. 28). Memory enables the actor to retain and search experiences. Learning gives the ability to adopt elements that are stored in memory in order to improve its behavior. Goal-directed behavior allows the actor to pursue a specific goal. In combination with memory and learning, goal-directed behavior provides the actor with means to show problem-solving behavior. With respect to the simulation of multilingualism in actors we have to formulate supplementary characteristics of a cognitive actor.

Cognitive actors have the ability to understand and speak a language (see Figure 2), consisting of a lexicon, grammar, syntax, semantics and pragmatics. In Figure 2 we incorporated two languages, but more are possible. The language component is a variable in the simulation model. An actor can have one or more languages, the mastery of all languages can be equal or varied and an actor the ability to strengthen his or her language. With respect to the languages, L1 can be dominant (has more speakers) over L2, can have more status than L2, or L1 and L2 are linguistically similar.

Language prerequisites: A central aspect of multi-actor systems is the ability of actors to communicate with each other. In a multi-agent system, the language that is used for communication in terms of vocabulary and grammar is predetermined as we saw in the models discussed in section 2. In these models the design of agents and the multi-agent system environment fix the ways agents are able to communicate (e.g., Bellifemine et al., 2007; Glaser, 2002; Padgham & Winikoff, 2004). Additionally, communication patterns between agents are commonly specified in protocols (Bellifemine et al., 2007). Both the predetermination of language and the fixation of communication protocol limit the ways in which agents communicate with predictable patterns. Furthermore, a multi-agent system principally is monolingual. In case two agents with different languages are placed within a multi-agent system they will not be able to communicate. Language is not positioned at Dennett's (1987) functional level as a part of the information processing make-up of an agent. It is only an external, predetermined protocol.

A **multi-actor** system for language diversity, however, needs to surpass language abilities that are embedded in actors during design. Instead of externally given, language becomes part of the cognitive structure, i.e., part of mental representations. Whenever an actor wants to pass a message to another actor, it needs to make a decision. In addition to constructing the message in terms of content, the actor needs to choose the language that is used to formulate the message in. In other words, language moves from being a function of perception and action to content within the cognition component. We therefore need to go into the details of aspect of the cognitive system that deals with language. The question is what are the minimal requirements of an actor with regard to communication with one or multiple languages?

Memory, goals, and communication: Having positioned language and its use at Dennett's functional and intentional level, what remains is to give a specification of what actors should look like in the envisioned MAS for language diversity. The basic components of perception, action, and cognition have already been discussed. All three components play a role in communication in the context of language diversity. The roles of action and perception are straightforward. These components enable the actor i) to transmit a message to another actor, and ii) to receive a message sent by another actor. The role of the cognitive component might be less obvious.

As said, the cognitive component has three functions, namely memory, learning, and goal-directed behavior. In relation to language, memory is the component that holds the actor's knowledge about one or more languages. This component enables the actor to produce messages from the knowledge the actor wishes to transmit to another actor. Reversely, memory enables the actor to understand a received message, using the knowledge s/he has about the language to make an interpretation. In formulating and understanding a message, retrieval time within memory is identified as an important determinant. From cognitive science, it is known that retrieval time of certain knowledge improves when this knowledge is accessed more frequently. Access time of more frequently accessed knowledge improves by a continuous re-organization of storage in memory. The memory component of the actor should reflect this behavior in that the arrangement of the actor's memory needs to be self-organizing.

The memory component is strictly bounded to the actor's abilities to communicate. The actor is not able to compose messages of which s/he has no knowledge stored in memory. This is where the learning function has a central role. The learning function provides the actor with the ability to extend his/her knowledge of a language; it enables the actor to extend and improve vocabulary and grammar on the language, and develop his/her ability to use the language in utterances and understanding. Also, learning provides the ability to expand the actor's knowledge about languages. This expansion extends into several directions. First, the actor is able to learn a new language in addition to the language already known to him/her. Second, learning enables the actor to develop insights about multiple languages s/he knows, such as for instance similarities and/or distinctions between languages. The latter mechanism provides the actor with tools to develop his/her understanding about the languages s/he knows. A more profound understanding of a language might lead to an improvement in the actor's mastery of the usage of a language.

For the function of goal-directed behavior, which is offered by the actor's cognitive system, the relation to language is perhaps less clear. We argue that the use of a particular language in communicating with another actor is not a direct goal of an actor. Instead, we state that the use of a particular language is instrumental to the goal the actor intends to achieve; it reduces to the vehicle used to transport the message to his/her communicative partner. Various determinants are envisioned that dictate the choice for a particular language to arrange the communication with the other actor. For instance, the status an actor is attributed from his discussion partners when using a certain language has been used by Abrams & Strogatz (2003) in their model of language competition. Status however, in this way is uni-dimensional; an actor only gains status when using a particular language. Alternatively, an actor might receive status for being able to use two or more languages. To our knowledge, this is a direction that has not been explored from the perspective of language diversity and simulation, yet. Furthermore, status in the context of language has often been assumed stable. We oppose this assumption, and hypothesize that status changes. The rational is that status is the product of social interaction. To us, this implies that status is affected by similar processes that for instance affect language (for languages develop / change over time) and language diversity (e.g., language dominance and extinction). In this way status becomes a variable instead of parameter in the simulation. The effects of incorporating such an approach towards status are unknown and require further study. Another factor that determines the choice for a specific language we envision to be efficiency. The efficiency of a language is determined by i) the ability of the sending actor to use the language, and ii) the estimation of the sending actor of the ability to understand the language of the receiving actor. An actor will from the viewpoint of efficiency choose that language that facilitates communication as a whole, ensuring efficient formulation and interpretation of the language at both ends (i.e., at sending and receiving ends).

4 SPECIFICATION OF THE SIMULATION DESIGN

With the provided dimensions that shape the actors for a simulation of language diversity, as provided in the previous section, this section discusses three topics. First, we extend the discussion from the previous section, presenting a preliminary specification of the design of the simulation we see. The specification is given at two levels. We provide a brief specification of the simulation environment.

Subsequently, a more detailed specification is provided concerning the make-up of the multilingual actors we aim to position in the simulation environment. The second topic concerns three experiments we aim to develop. These experiments all focus on the goal to understand underlying mechanisms and characteristics of language dominance, extinction, and multilingualism. The main question is what happens to two equivalent languages in a context of multilingualism? Will such a situation be able to exist for a long term, or does in the end one language dominate and all other languages disappear? This brings us to our third topic. The content-driven questions we want to ask the simulation. We will formulate three questions

4.1 Preliminary simulation specification

The multilingual, actor-based simulation we discuss in this paper centers on the issue of understanding what happens in situations of language diversity within a given population of actors. In order to realize such a simulation, the primary prerequisite a simulation environment - given the existence of multiple actors - should offer, is a platform for communication. Regardless of any components individual actors are equipped with, the simulation environment needs to have some basic mechanism that ensures the transmission of messages between actors. Furthermore, the environment should be designed in such way that actors can exist (i.e., software actors should be able to behave according to their internal arrangements). These two requirements are part of the Abstract Architecture Specification of the Foundation for Intelligent, Physical Agents/Actors (FIPA; e.g. Bellifemine, Caire, & Greenwood 2007). The Abstract Architecture Specification provides a specification for actor messages, a message transport service, an actor directory service, and a service directory service (Bellifemine et al., 2007, p. 17). The structure of actor messages is prescribed in the FIPA-ACL (Agent Communication Language). Message content specification is provided in the FIPA-SL (Semantic Language). Additionally, message content can be described in more detail using ontologies. The message transport service (see FIPA Agent Message Transport Service Specification) ensures sending and receiving of messages between actors in the actor platform. The actor directory service maintains information about the actors residing in the actor platform at any time. Actors that are added to the actor platform need to register at the actor directory service; actors leaving the platform withdraw from the directory service. The service directory service concerns a repository of all services (i.e., actor's behaviors) offered by the various actors in the actor platform including a description of the service. All mentioned components are mandatory in an actor platform (Bellifemine, et al., 2007).

Within or on top of the actor platform, the actors reside. The previous section identifies various components that comprise the actors for the intended simulation. For communication purposes, actors require perception and action components. These components need to connect to the message transport layer of the actor platform in which the actors will be placed. For now, meeting the FIPA specifications for communication suffices for the perception and action components. Figure 3 presents the generic components of an Agent Platform as specified by the FIPA. In this figure, two agents are displayed. These agents communicate with each other through the Message Transport Service. The Agent Management System holds a record of the agents that are active within the platform. The design of the cognitive component is less straightforward. Bellifemine et al. (2007) identify four distinct actor architectures to shape the cognitive component: logic based, reactive, BDI, and layered (or hybrid) architectures (ibid. p. 4). For our purpose, the BDI architecture seems fitting. A logic-based architecture does not for instance provide a learning function, nor does a reactive architecture (ibid, p. 4-5). The BDI architecture seems promising, for instance the Procedural Reasoning System (PRS; Georgeff and Lansky, 1987). This architecture offers memory and provides a learning mechanism (beliefs about the world are updated using observations the actor makes). Also it displays goal-directed behavior.

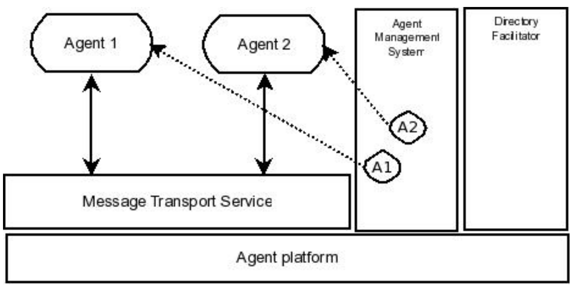


Figure 3: *A generic model of an agent platform (in this case an agent can be and an actor)*

In the previous section, we already identified three elements that are required to build the cognitive component, namely memory, learning, and goal-directed behavior. In terms of actor subsystems, these functions translate into various building blocks. The memory and learning functions indicate that the actors should be equipped with some sort of storage, which allows for long-term storage, retrieval and retention of words or concepts. Particularly difficult in relation to multilingualism is the question in which language the concept is stored in memory? Here we assume that whenever an actor stores a concept in memory, the actor is able to articulate the concept in at least the language in which the actor learns it. Normally this is the actor's dominant language. However, the actor can have more languages at his disposal. This assumption follows from the idea that an actor's understanding of concepts primarily is embedded in his dominant language. From this assumption follows that an appropriate memory structure requires to 1) be able to store the concept in the language in which it is learned and 2) provide entries for retrieval in at least the dominant language and the language in which the concept was learned. The function of goal-directed behavior demands the actor to be equipped with a subsystem that is able to store various goals the actor will pursue. Helmhout (2006) uses a goal stack in combination with a goal handler for this purpose in the RBot architecture. The BDI architecture places goals in the goals storage facility. The concrete design of the subsystem that fulfills the actor's goal-directed behavior function should for the simulation discussed here at least steer the actor to communicate with other actors. Furthermore, a mechanism is required that handles the status that is ascribed to the actor for using a specific language, as discussed in the previous section.

The languages an actor has in its memory need to be considered separately from the actor communication language that is available from the agent platform. The actor communication language enables actors residing on the platform to communicate with each other in a standardized way. We think of this standard way of communication as the physical means that enable actors to send and receive messages. In this way, the message transport service is similar to the physical level of communication that humans use to communicate, i.e., humans communicate using their vocal and hearing organs and the air as a medium. Similar to an actor platform, the usage of these physical organs and medium does not provide an indication regarding the content of the messages that are sent back and forth between the humans, let alone the language that is used to formulate the messages in. Perceiving the message transport service in this way, allows the formulation of a language on top, which the actors use to retain facts about their world, reason with and formulate messages in. Hence, the messages that are send between actors follow the actor communication language specification to ensure proper delivery within the actor system, and contain messages in some (perhaps other) language that makes sense to the actor.

Figure 4 displays how messages in the language of an actor (L1) are formulated and translated to a message that can pass the message transport service. L1 is the language that is understood by the actor. In order to communicate through the message transport service, a message needs to be translated to the actor communication language (ACL). The interpreter and translator processes ensure the transformation of the incoming message to a message in L1 (interpreter) and vice versa (translator). In this approach, the FIPA message structure specification is maintained; the used language only shapes the message content (i.e., the payload within the message structure). In essence, the interpreter and translator processes empty and fill the message content with the message that is formulated in L1. In case an actor is able to communication in multiple languages, the interpreter and translator processes are able to make the transformation from the actor communication language to any of the languages the agent understands and vice versa.

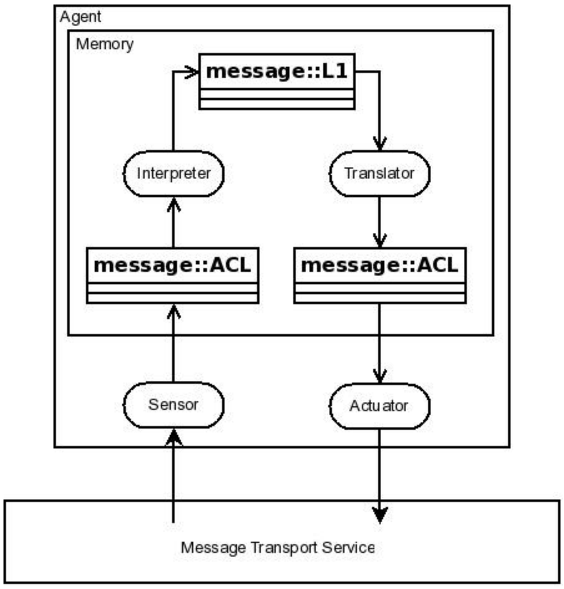


Figure 4: *The position of language in the actor*

An actor's memory stores concepts, and relation between these concepts; together, these concepts and relations form an ontology (Bellifemine et al, 2007). We argue that each concept is expressed in one or more languages, and at least in the actor's dominant language. Figure 5 provides a possible storage structure for storing concepts in memory. Each concept has one or more names, and each name is expressed as a word. The word is part of a lexicon, which is part of a specific language.

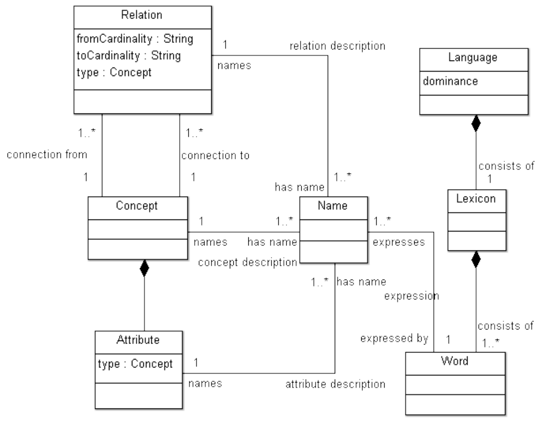
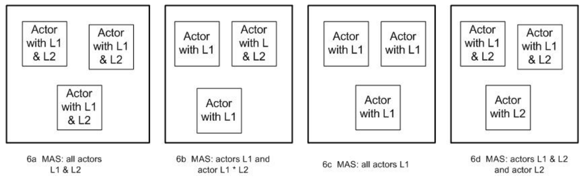


***Figure 5:** Concepts and possible structure for storing concepts in memory of an actor*

Furthermore, concepts consist of attributes, and concepts can be related. An attribute is of a certain concept type. For instance the attribute height of the concept house has a numerical type. Both concepts house and numeric need to exist in the actor's memory in order to reason. The relation also is typified. Additionally, relations have some cardinality with the concepts they connect. Cardinality indicates how many concepts are part of the "from" and "to" sides of the relation. Both attributes and relations are named, following the same naming logic as the naming of a concept. In the described structure, concepts, attributes, and relation can carry multiple names, expressed by multiple words, originating from different languages.

4.2 Experiment design and questions to be asked to the simulation

A simulation model with cognitive actors in relation to the study of multilingualism requires four components of which we discussed three in the foregoing sections: a conceptual specification of the issue, an extended description of a cognitive actor based on cognitive science and a functional and quasi technical specification of the setup of the simulation model. Besides the validity discussion which we will not address here, a simulation model needs to give answers to questions. What are the questions? We will formulate three questions. To look at these questions the following experimental setup is suggested. In figure 6 four multilingual situations are presented. In 6a all actors use L1 and L2, in 6b one actor uses L1 and L2 and the others use L1, in 6c all actors only use L1 and in 6d one actor uses L1 and the other actors use L1 and L2. At the moment we are not making all kinds of differentiations with respect to the status of the language, the quality of the languages and all kinds of societal and institutional arrangements



Starting from this very simple world model with a few actors and the basic design of languages, we can think of at least three questions that will be investigated in the simulation runs. The first question is under what conditions or characteristics of the actors using L1 and L2 will L1 become dominant over L2? This question is about language survival or death and concerns the situation in figure 6a. The second question is how many actors are necessary to keep L2 existing? In this case the situations in figure 6a (all multilingual) and figure 6d (some are monolingual) are compared. The third question is whether the communication speed is different when we compare bilingual actors in a society with a society of monolingual actors? Many other questions can be formulated, but our first intention is to have the simulation run according to our specifications.

5 CONCLUSION AND DISCUSSION

In this paper, we presented an alternative direction for studying multilingualism, using multi-actor simulations that consist of cognitive plausible actors. The choice for usage of such simulation models originates from our position that language is a result of individual cognitive and social human behavior. This perspective opposes the position that language is an entity that exists by itself. We also showed that the suggested approach deviates from simulation models that are currently used to study multilingualism.

In addition to existing knowledge about multi-agent or multi-actor systems, we suggested two structures to equip actors with means to deal with language. First, the internal functioning of language processing of an actor has been explained. On top of an existing multi-actor platform, the suggested structure enables the actor to use a certain language to communicate with other actors. The Interpreter and Translator components offer the appropriate transformation of messages to messages that can be handled by the actor platform. Secondly, we introduced a memory arrangement for actors to deal with language in relation to ontologies. Current approaches in multi-actor modeling state that actors use ontologies to retain information about specific domains of interest. We extended this approach. The concepts and relations comprising an ontology are expressed in a certain language. We assume that a concept or relation always is retained in an actor's memory in the actor's dominant language, for this is the language that provides the actor with the most elaborate meaning. In addition, concepts and relations can reside in memory in a different, non-dominant language, which allows the actor to send and receive messages in a non-dominant language.

Multi-actor simulations for language diversity and multilingualism open possibilities in multiple directions. A first effect of our choice is that a common attribute of language status, which holds a central position in for instance the Abram-Strogatz (2003) model, is repositioned. No longer is status attributed to language itself. Instead, our perspective enables the assignment of status to the user of a particular language. Furthermore, this positioning of status at the level of the language user allows for connecting status to him/her being multilingual. A second effect is that a mechanism for language production is made available that can be used to investigate and predict the dynamics of a specific language and situations of multilingualism. Within the multi-actor simulation, actors are the carriers of languages. Dynamics at the level of the actor (e.g., message generation and interpretation) and at the level of a group of actors (e.g., language interaction and status ascription to language use) provide insights into the generative processes of language utterance. In this way, the mechanisms that underlie language use and change can be studied. We argue that only when multi-actor simulation models based on cognitive plausible actors, a more profound understanding of language extinction or revitalization and multilingualism is obtained.

REFERENCES

Abrams, D.M., & Strogatz, S.H. (2003). Modelling the dynamics of language death. *Nature, 424*, pp. 900.

Anderson, J.R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah (NJ): Lawrence Erlbaum Associates.

Bellifemine, F.L., Caire, G., & Greenwood, D. (2007). *Developing multi-agent systems with JADE*. Chichester: John Wiley & Sons.

Card, S.K., Moran, T.P., & Newell, A. (1983). *The psychology of Human-Computer Interaction*. Hillsdale (NJ): Lawrence Erlbaum Associates.

Castelló, X., Toivonen, R., Eguíluz, V.M., & San Miguel, M. (2007). *Modelling bilingualism in language competition: the effects of complex social structure.* The Fourth Conference of the European Social Simulation Association ESSA 2007, Toulouse, France.

Dennett, D. (1978). *Brainstorms: philosophical essays on mind and psychology*. Cambridge (MA): Bradford Books/MIT Press.

Dennett, D. (1987). *The intentional stance*. Cambridge (MA): The MIT Press.

Fernando, C., Valijärvi, R.L., & Goldstein, A. (2010). A model of the mechanisms of language extinction and revitalization strategies to save endangered languages. *Human Biology*, *82*(1), pp. 47-75.

Gazendam, H.W.M. & Jorna, R.J. (1998). Theories about Architecture and Performance of Multi-agent Systems. University of Groningen: *SOM Research Report 98 A02*.

Gazendam, H., & Jorna, R.J. (2002). Transaction cost economics and plausible actors: a cognitive reappraisal. *Semioticon: the semiotic frontline*. Retrieved from www.semioticon.com

Georgeff, M.P., & Lansky, A.L. (1987). Reactive reasoning and planning. In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87)*, pp. 677-682, Seattle, US.

Glaser, N. (2002). *Conceptual modelling of multi-agent systems: the CoMoMAS Engineering Environment*. Dordrecht: Kluwer Academic Publishers.

Helmhout, J.M. (2006). *The social cognitive actor*. (Doctoral dissertation, University of Groningen, Groningen). Ridderkerk: Labyrinth Publishers.

Jorna, R.J. (1990). *Knowledge representation and symbols in the mind*. Tübingen: Stauffenburg Verlag.

Jorna, R.J., Faber, N.R., & Hadders, H. (2009). Multi-actor systems, cognitively plausible actors and organizational knowledge. In W.R. King (ed.) *Knowledge Management and Organizational Learning* (pp. 279-300). Dordrecht: Springer.

Newell, A., & Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs (NJ): Prentice-Hall.

Padgham, L, & Winikoff, M. 2004. *Developing intelligent agent systems: a practical guide.* Chichester: John Wiley & Sons.

Simon, H.A. (1969). *The sciences of the artificial*. Cambridge (MA): The MIT Press.

Stauffer, D., Castelló, X., Eguíluz, V.M., & San Miguel, M. (2007). Microscopic Abrams-Strogatz model of language competition. *Physica A*, *374*, pp. 835-842.

Wang, W.S.Y., & Minett, J.W. (2005). The invasion of language: emergence, change and death. *Trends in Ecology and Evolution, 20*, pp. 263-269.

Wijermans, N. (2011). *Understanding crowd behaviour*. (Doctoral dissertation, University of Groningen, Groningen). Groningen: University of Groningen.

Wooldridge, M. 2002. *An introduction to multi-agent systems*. Chichester: John Wiley & Sons.