

Economics and Language | *Abstracts*

FROM SIGNALS TO SYMBOLS: GROUNDING LANGUAGE ORIGINS IN COMMUNICATION GAMES.

Angel Alonso | Universidad Complutense. Madrid

1. Introduction.

Language is a characteristic feature of human nature. Presently, the study of language has become a multidisciplinary enterprise. Due to the insights into human nature that economics provides, research in this field has become relevant for the study of language and language origins (Wärneryd, 1994) . This paper shows that some of Hockett=s design features of human language (see page 2) have its origin in strategies applied in game theory.

After being a taboo question for more than a hundred years , the topic of language origins has been readressed in the past thirty years. In fact, the Societé Linguistique de Paris banned its discussion in 1866 as a topic of scientific research after the Berlin Academy of Sciences initiated its study in 1769. However, the New York Academy of Sciences decided finally to reopen the question of language origins and organised a conference whose proceedings were published in 1976, more than a century after the Paris Academy=s decision. 1976 is considered the starting point for modern research on language origins. Since 1996 several conferences on language origins have been organized, where the topic has been lively discussed from a variety of new approaches. One of these approaches is based on game theory and evolutionary game theory. Therefore concepts from these theories may become crucial for the question of language origins and language evolution as they may contribute substantially to our understanding of how a language-ready brain initiated language. Therefore we will discuss some results of game theory and evolutionary game theory which are interesting for the study of language origins.

2. Arguments for grounding language origins in communication games

At present, studies on game theory, economic behavior and probability theory carried out by economists may shed light on the question of language origins and language evolution. We claim that economic theory is relevant to language origins and language evolution for several reasons. Firstly, the study of language origins needs to establish which scenarios trigger the emergence of language, once the human brain has attained a certain level of complexity. These scenarios are not supposed to be like those imagined by Kipling in his *Just so stories*, which are often found in language origins research. The required scenario must be a plausible scenario in which language properties emerge. In this paper, we claim that this plausible scenario was envisaged by Adam Smith in his *Inquiry into the Nature and Causes of the Wealth of Nations* (1776). .According to Smith=s account, division of labor, goods exchange and language are causally related . The division of labor produces a diversity of goods that could be exchanged. Goods exchange creates the necessity of a contract, and contracts require concerted or coordinated actions among the contracting individuals. In coordinating actions , actors are involved in communication games where they convey information . Then a symbolic language subserves the communication of information . Other communication systems used by nonhuman animals are set up by signals. Consequently, they differ significantly from symbolic systems. Contrary to signals, symbols arbitrarily represent

selected features of objects and events, those that speakers more or less consciously think are useful for some purpose in the community where they live.

Secondly, some essential features of the human symbolic communication system - Hockett's design features (Hockett, 1958,1960, 1961) - arise in a natural way from this scenario involving certain strategies and cost-benefits analysis used in games. These design features are :

- (1) Interchangeability (or parity) : when two participants in a communicative act cooperate yielding an exchange of messages ;
- (2) specialization : when the communicative act does not involve the expenditure of energy. In this sense, language is specialized , i.e. costless , or cheap ;
- (3) semanticity : speakers coin meaningful symbols because they share certain semantic conventions tied with objects and situations ;
- (4) prevarication : when speakers use symbols (sentences) deceitfully, not conveying information corresponding to real events or situations ; and
- (5) cultural transmission, i.e., language (words and noncompositional syntax) is learned in a way where imitation and memory play a significant role. Cultural transmission from a generation to another makes language evolve in a form (the evolutionary bottleneck) that may account for both crosslinguistic diversity and linguistic universals.

Modern economic theory has taken up communication in a way that we believe supports both Smith's scenario on language origins that yield some of its design properties. Thus, communication games have been dealt with by Wärneryd (1988) and Blume (2001, 2004), among others. These studies support the claim that coordination between speakers / players is attained by language (A cheap talk @), so reinforcing Smith's setting. The property of semanticity may be understood as a Nash equilibrium (Rubinstein, 2000.). Finally, cultural transmission of language finds an explanation because an imitated action is the best strategy (Axelrod, 1981; Schlag, 1998).

3.Conclusions.

In this paper we claim that the language origins scenario is provided by labor division, in which some essential design features arise in a simple way from communication games. Therefore, although language requires a brain with a certain level of complexity (a language-ready brain) that allows for mental operations, language originated as a social technique for communication , so it is not an innate biological function of the brain. Thus language evolves from a synergy of brain and society.

IN WHAT SENSE?

Asunción Álvarez | Dpt. Of Philosophy, King's College London | Asuncion.Alvarez@Kcl.Ac.Uk

One of the milestones in the history of the analytic philosophy of language was the introduction of the notion of sense into the theory of meaning by Gottlob Frege. In Frege's account, meaning is explained in terms of a twofold semantics whereby any given linguistic term has a reference - the object (if any) denoted by the term - and a sense - the cognitive "mode of presentation" under which the reference appears to the thinker. According to Frege, senses would not be subjective mental representations, but objective entities available to all thinkers, which would enter into the semantic description of a *public* language.

As Frege himself was aware of (cf. the famous "Aristotle" footnote in "On Sense and Reference), however, it could be objected that the view of sense as a mode of presentation is hard to reconcile with its role within a public language. Frege settled the issue by arguing that, whereas such variations in sense could be tolerated in everyday communication, they were to be avoided in

scientific discourse and should not occur in a logically perfect language. The question nonetheless remains of a successful communication is established between speakers of a language who have different senses of the same term.

In this paper, I will try to answer this question by applying the game-theoretic approach that Ariel Rubinstein employed in his *Economics and Language* to explain certain semantic and pragmatic issues in language.

Frege argued that the sense of a term in a given language is grasped by all competent speakers of that language. However, he also claimed that there are degrees to the grasp of senses by thinkers; and that there are also degrees of competence which thinkers (or speakers) can have. These theses jointly allow us to sketch out a tentative picture of how sense works. If we establish that any given term within a given language may well have not a unique sense, but a (possibly infinite) set of possible senses, then the definition of competence for usage of that term will require that the speaker grasp only the senses within an unspecified (possibly fuzzy) subset within the total set of possible senses. The size of this subset will vary according to the type of competence under definition: for the term "electron", for example, the subset required for competence in everyday language will be fairly small; whereas competence in modern physics will require a much larger subset. Alternatively, a distinction can be made between the multiple senses of an expression and its "total sense", which would be the sum or set of all the possible senses of the expression. The degree to which a thinker grasps the (total) sense of a given expression could thus be described in terms of the subset of senses which he does grasp.

Having defined this structure for senses, I will discuss how a game-theoretic approach might serve to explain how in successful communication the various subsets involved overlap so as to allow for an evolutionarily stable equilibrium to be attained.

DELIBERATION, PREFERENCE UNCERTAINTY AND VOTING RULES

David Austen-Smith & Timothy J. Feddersen | Northwestern University

dasm@kellogg.northwestern.edu & tfed@kellogg.northwestern.edu

A deliberative committee is a group of at least two individuals who first debate about what alternative to choose prior to these same individuals voting to determine the choice. We argue, first, that uncertainty about individuals' private preferences is necessary for full information sharing and, second, demonstrate in a very general setting that the condition under which unanimity can support full information revelation in debate amounts to it being common knowledge that all committee members invariably share identical preferences over the alternatives. It follows that if ever there exists an equilibrium with fully revealing debate under unanimity rule, there exists an equilibrium with fully revealing debate under any voting rule. Moreover, the converse is not true of majority rule if there is uncertainty about individuals' preferences.

THE ECONOMICS OF BABEL

Manuel F. Bagüés | Universidad Carlos III. Madrid. manuel.bagues@gmail.com

Languages, and more generally coding systems, tend to be subject to a trade-off between achieving accuracy in their description and maintaining its informative power. This trade-off will exist as long as the use of additional symbols remains costly. An optimal coding system must therefore achieve an equilibrium between capturing the essential of reality while being as limited as possible. Any economist/philologist who has ever addressed a non-economist/non-philologist on an economic/linguistic topic is probably well aware of this problem. In order to make the message understandable very often a sacrifice in the rigour of the discourse must be made.

Coding systems also exhibit well known network externalities. The profit of using a certain coding system is increasing with the number of users. Given the existence of these positive externalities, if individuals choose non cooperatively their coding systems an excessive heterogeneity will arise. For instance, when economists choose a certain jargon to describe some economic concepts they will not take fully into account the negative external effect of this jargon not being understandable to non-economists. The negative consequences of language heterogeneity are hard to overestimate. Language affinity constitutes a key significant explanatory variable of trade flows. Several studies have also found a negative correlation between the rate of growth and the index of linguistic variety in certain areas (see for instance Alesina et al, 2002).

From a dynamic point of view coding systems must face an additional problem. In a changing world, an efficient coding system must be able to adapt to variations in the realities that it aims to transmit. However, if introducing changes into a coding system requires some "investment" by the reformers while the pay-offs extend to all future users, changes in coding systems will be adopted at a suboptimal rate. The intuition is that since any generation introducing a reform will incur all the costs but appropriate only part of the benefits, some changes that would be efficient from a (intertemporal) social perspective will not be adopted. In fact, there are many well known cases of societies whose failure to introduce changes in their coding systems in response to changes in reality brought them severe consequences. The Chinese reluctance to move away from the ideographic system of characters once the print was discovered has been pointed out as one of the key elements that would explain the divergence in the economic evolution of China and Europe after the XVth century. While in Europe Gutenberg's "discovery" of the print fostered the diffusion of knowledge, the large number of different ideograms that are required to reproduce a book prevented that China experienced a similar phenomenon. Similarly, although hindu-arabic numerals were known in Europe since the beginning of the eleventh century, European countries failed to adopt the new system for several centuries. A big part of this retard could be related to the limited appropriability of the full profits of adoption by reformers. Other anecdotal examples may include the survival of the graphic distinction between letters "j" and "g" in Spanish language, as Nobel prize laureate Juan Ramón Jiménez argued, or the maintenance of genders in most languages.

In this paper we use an overlapping generations framework to show that while the above static inefficiency -derived from the existence of network effects- may induce the creation of intragenerational agreements to reduce the excessive heterogeneity, these cooperative agreements will however aggravate the dynamic inefficiency.

This is, the main contribution of this article is to show while imposing a standardization of the set of coding systems might be efficient from a static perspective, it will aggravate the dynamical inefficiency, as it will raise the opportunity cost of introducing reforms. This is, a homogeneization of the set of coding systems may induce as an undesirable fossilization of linguistic structures. The intuition for this result is that when the set of coding systems becomes more homogeneous, the cost for each individual of deviating from the standard becomes higher.

Note that the theoretical construction that is given in this paper is far from inclusive. Language is just partially a mechanism of communication, and many other aspects are not covered in this framework. Still, our model delivers several testable empirical implications which seem to be consistent with several well known linguistic stylized facts.

NATURAL MEANING AND CONVERSATIONAL IMPLICATURES

Anton Benz | University of Southern Denmark. benz@sitkom.sdu.dk

We argue that conversational implicatures are based on the natural information that is carried by the utterance of a signal in a conventional signalling system. In a first step we clarify the distinction

between natural, semantic, and conventional meaning in a Lewisian (1969) framework of embedded signalling games. By embedded signalling games we mean signalling games which occur as sub-games of larger games. Throughout we restrict attention to games of pure coordination. In (1957) Grice introduced his famous distinction between natural and non-natural meaning. He saw this distinction as an improvement over the distinction between conventional and non-conventional meaning of language. Lewis defined conventions as certain solutions to games of complete coordination. He showed that, indeed, the conventional meaning in his sense is always non-natural meaning (Lewis, 1969, Sec. 4.5). Both philosophers aimed at a reductionist definition of communicated meaning: Grice in terms of recognised intentions and Lewis in terms of coordinated behaviour. According to the standard theory of the neo-Gricean approach (Levinson, 1983, Ch.3), non-natural meaning comprises both what is said (semantics) and what is implicated. This entails that, in general, the meaning of a signal in a signalling game is more than the pure semantic meaning of the signal.

In order to work out the interaction between different types of meaning in the production of conversational implicatures we look at the sub-class of signalling games which are defined by pure semantics. We assume that semantics is given by a function that maps linguistic forms F onto their (underspecified) semantic meanings. This function defines a hearer's strategy in a signalling game where the speaker chooses a form F in order to express a meaning M , and the hearer a meaning M_0 as interpretation of F . They have success if $M = M_0$. We call this game the semantic interpretation game.¹ The game is solved by a pair h_S, H_i of a speaker's strategy S and the hearer's strategy H (defined by pure semantics) if it forms a signalling system.

We show that there is a uniform definition of natural meaning that can explain Grice classical examples for natural meaning and the enrichment of semantic meaning which builds the basis for conversational implicatures. This enrichment is the effect of pragmatic constraints that we have to add to the semantic interpretation game. We assume that games are represented in extensive form. If G is the set of all (complete) branches of a game, then we define the natural information of a signal a relative to G as the set of all branches which contain a . Now let's assume that this game is a semantic

interpretation game which has a solution h_S, H_i . We reconstruct the distinction between natural and semantic information as the difference between the natural information of the signal F relative to G and $H(F)$. Grice (1957) illustrated his distinction between natural and non-natural meaning by an example which involves communicating by showing a photo and showing a painting. It resembles games of partial information (Parikh, 2001) and play a role similar to super-conventional games in (v. Rooij & Sevenster, 2006).

SIMULATING PROCESSES OF LANGUAGE EMERGENCE AND COMMUNICATION

Timo Honkela, Ville Könönen, Tiina Lindh-Knuutila and Mari-Sanna Paukkeri

Adaptive Informatics Research Center. Laboratory Of Computer And Information Science

Helsinki University Of Technology. marisa@james.hut.fi

1 Introduction

We have conducted extensively research on simulating processes of language learning and emergence (consider, e.g., [3, 4, 15] and learning and communication in multiagent systems (consider, e.g., [9, 11, 10])). In this paper, we discuss how the emergence of subjective models of the world can be simulated using different approaches in learning and what is the role of communication and language. We consider, in particular, the role of unsupervised learning in the formation of agents' conceptual models, the original subjectivity of these models, the

communication and learning processes that lead into intersubjective sharing of concepts, and a game theoretical approach to multiagent systems including reinforcement learning processes.

An intelligent agent usually has a purpose or a goal, probably set by the designer of the agent, and the agent tries to act rationally for satisfying the goal. For making rational decisions, the agent has to model its environment. In the following, we consider some specific issues related to the modeling.

1.1 Complex and non-stationary environments

In the easiest case, the environment in which the agent is located, is static, i.e. all properties of the environment remain constant for the whole life-time of the agent. However, often the situation is not so simple. The properties of the environment may vary with time, i.e. the environment is called non-stationary. The non-stationarity may be due to the environment itself or the limited resources of the agent. For example, in many real problem instances, the learning agent is not capable of sense the real state of the environment and thus some relevant properties of the environment remain hidden. Another example of non-stationarity are multiagent systems. In these systems, properties of the environment for each individual agent depend on the actions of all agents located in the real environment. Thus for acting rationally, agents should model also other agents in the system.

Pfeifer and Scheier [14] stress that the behavior of an agent is always the result of system-environment interaction. It cannot be explained on the basis of internal mechanisms only. They illustrate that the complexity that we as observers attribute to a particular behavior does not always indicate accurately the complexity of the underlying mechanisms. Experiments with very simple robots that merely react to stimuli in their environment have shown that rather complex behavior can emerge.

1.2 Subjective versus intersubjective

Moore and Carling [13] state that “[l]anguages are in some respect like maps.

If each of us sees the world from our particular perspective, then an individual’s language is, in a sense, like a map of their world. Trying to understand another person is like trying to read a map, their map, a map of the world from their perspective.” In many computational approaches to semantics and conceptual modeling, an objective point of view has been used: it is assumed that all the agents have a shared understanding and representation of the domain of discourse. However, Moore and Carling’s statement emphasizes the need for explicit modeling of the other’s point of view. This requires modeling of the subjective use of language based on examples, and, furthermore, to model intersubjectivity, i.e., to have a model of the contents of other subjective models.

As an example related to the vocabulary problem, two persons may have different conceptual or terminological “density” of the topic under consideration. A layman, for instance, is likely to describe a phenomenon in general terms whereas an expert uses more specific terms.

1.3 Learning agents

Different machine learning techniques can be utilized for adding adaptivity to agent-based systems. There are basically three major learning paradigms in machine learning: supervised learning, unsupervised learning and reinforcement

learning (RL). In supervised learning, there exists a teacher having knowledge of the environment, in the form of input-output pairs, and the learning system for which the environment is unknown. The teacher provides samples from the environment by giving correct outputs to inputs and the goal of the learning system is to learn to emulate the teacher and to generalize the samples to unseen data. In unsupervised learning, contrary to supervised learning, there exists no external teacher and therefore no correct outputs are provided. RL is located between supervised and unsupervised learning: correct answers are not directly provided to the learning system but it learns features of the environment by continuously interacting with it. The learning system takes

actions in the environment and receives reward signals from the environment corresponding to these action selections.

2 Unsupervised learning of conceptual systems

In the following, we consider the unsupervised learning of conceptual systems. We first study the modeling an individual agent that learns to create a conceptual space of its own and learns to associate words and expressions with conceptual space. The notion of conceptual space is taken from Gärdenfors who also presents the basic motivation and framework for dealing with conceptual spaces [1]. After considering one individual agent, we consider a multiagent system in which a shared conceptual system is formed in a self-organized manner.

2.1 One agent

The self-organizing map (SOM) [7, 8] is an unsupervised learning model which is often considered as an artificial neural network model, especially of the experimentally found ordered “maps” in the cortex. The SOM can be used as a central component of a simulation model in which an agent learns a conceptual space, e.g., based on data in which words are “experienced” in their due contexts [16]. This approach has been used, for instance, to analyze the collection

of Grimm fairy tales. In the resulting map, there are areas of categories such as verbs and nouns. Within the area of nouns a distinction between animate and inanimate nouns emerges [5]. The basic idea is that an agent is able to form autonomously a conceptual mapping of the input based on the input itself.

2.2 Community of agents

The basic approach how autonomous agents could learn to communicate and form an internal model of the environment applying self-organizing map algorithm was introduced, in a simple form, in [2]. Later we developed further the framework that would enable modeling the degree of conceptual autonomy of natural and artificial agents [4]. The basic claim was that the aspects related to learning and communication necessitate adaptive agents that are partially autonomous. We demonstrated how the partial conceptual autonomy can be obtained through a self-organization process. The input for the agents consists of perceptions of the environment, expressions communicated by other agents as well as the recognized identities of other agents [4]. A preliminary implementation of a simulated community of communicating agents based on these ideas did not succeed to fully demonstrate the emergence of a shared language [6]. When language games [17] was included in the simulation model, it resulted in a simple language emerging in a population of communicating autonomous agents [12]. In this population, each agent was able to create their own associations between the conceptual level and the emerged words, although each agent had a slightly different conceptual representation of the world. The learning paradigm for the conceptual learning was fundamentally unsupervised, but the language learning tested has been so far supervised, i.e. the communicating agents are provided feedback of the outcome of the game as well as the “right answer”. The reinforcement learning and the unsupervised learning models for language games remain to be implemented.

3 Game theoretical approach for multiagent systems

As discussed in Section 1, an intelligent agent usually has a goal. For studying agent based systems formally, e.g. for developing learning algorithms for intelligent agents, it is useful that the goal can be expressed mathematically. Traditionally approach is to define an utility function for the agent, i.e. there is a scalar value connected to each possible action measuring the fitness of the action choice for satisfying the goal of the agent.

3.1 One agent

In single-agent systems, achieving the rational behavior is simple: the agent always selects an action with the highest utility value. There exists a vast number of learning methods that utilize, in one way or another, the rational decision making in agent-based systems.

3.2 Community of agents

The branch of science studying decision making in single-agent systems is called decision theory. However, when there exists multiple active decision makers (agents) in the same environment, decision theory is not suitable for achieving rational behavior any more. Game Theory is an extension of decision theory to multiagent systems. In game theory, agents explicitly model the dependency of their utility functions on the actions of all agents in the system. The goal of the agents is to find equilibrium actions, i.e. the actions that maximize their utility values assuming the all agents will use the same equilibrium. Theoretically, game theory provides a perfect tool for modeling multiagent systems. In practice, there are many problems with game theory. For example there can exist multiple equilibria and the agents should coordinate which one they will select. For calculating an equilibrium, the agents should not only model their own utility function but also the functions of all other agents in the system. This is often intractable and therefore some other methods, e.g. Bayesian techniques, should be used for creating more coarser models of other agents. By using these "external" models, the game theoretical problem reduces to the simpler problem solvable by using decision theory.

4 Discussion

Language does not need to be viewed plainly as a means for labeling the world but as an instrument by which the society and the individuals within it construct a model of the world. The world is continuous and changing. Thus, the language is a medium of abstraction rather than a tool for creation and mediation of an accurate "picture" of the world. The point of view chosen is always subject to some criteria of relevance or usefulness. This is not only true for the individual expressions that are used in communication but also concerns the creation or emergence of conceptual systems. It makes sense to make such distinctions in a language that are useful in one way another. In this paper, we have discussed the role of unsupervised learning in the formation of agents' conceptual models in a non-stationary environment and a game theoretical approach for multiagent systems. In the final version of our paper, we will consider in more detail the relationship between unsupervised and reinforcement learning in the context of concept formation and communication within a community of agents in a non-stationary environments.

EVOLUTON OF HORN STRATEGIES

Kris De Jaegher | Utrecht School Of Economics, Utrecht University. K.Dejaegher@Econ.Uu.Nl

Horn's rule (1984) is the general rule in pragmatics saying that (un)marked expressions get an (un)marked interpretation. For instance, if I argue that every ten minutes, a man gets mugged in New York, then normally you will interpret this as meaning that this concerns a different man every ten minutes. Yet, strictly speaking, I could make the same statement to imply that every ten minutes, the same very unlucky New Yorker gets mugged. Thus, strictly speaking, my statement is ambiguous. The reason that you give my statement the first interpretation rather than the second is because you estimate the first interpretation to be far more likely. But why do I not make my statement less ambiguous? Presumably, disambiguation is costly. If I have to clarify what I mean in a unambiguous way, this will take me more time. And as you are unlikely to make the second interpretation, I can leave my statement ambiguous without danger.

Some efforts have been made to account for Horn's rule by means of noncooperative game theory (Parikh, 1991, 2000, 2001), using signalling games. Yet, from the perspective of noncooperative game theory, there is no reason why Horn's rule would apply. It may very well be that you and I both know that it would be more efficient if I would not have to indicate that I do not mean that

one and the same man gets mugged in New York every ten minutes. But if I believe that you will interpret the sentence 'Every ten minutes, a man gets mugged in New York' as meaning one and the same man, then I will have to point out that I do not mean one and the same man each time I make the more plausible statement. And since I use my signals in this way, if I do not specify that I do not mean one and the same man, you will interpret it as meaning one and the same man. Simply, there are multiple Nash equilibria to the signalling game. Parikh argues that the fact that the signalling equilibrium that follows Horn's rule is Pareto efficient will make it a focal point (Schelling, 1960), and will cause it to be played. Yet, as pointed out by Van Rooij (2004), equilibrium selection based on Pareto efficiency is not a standard game-theoretic argument.

Additionally, there have been some recent efforts to check whether evolutionary game theory cannot yield selection of signalling equilibria that obey Horn's rule. As pointed out by Van Rooij (2004), all signalling equilibria, including equilibria that do not obey Horn's rule, are evolutionary stable equilibria (ESS), and therefore the ESS concept does not bring one any closer to a solution. Under replicator dynamics, Benz, Jäger and Van Rooij (2005) show for a specific game that the basin of attraction of an equilibrium that obeys Horn's rule is larger than the one of an equilibrium that does not, and that starting from a strategy profile where all strategies are played with equal probability, under replicator dynamics an equilibrium obeying Horn's rule evolves. Yet, these authors do not give much intuition for why this result is obtained.

The purpose of this paper is to show that the evolution of an equilibrium that selects Horn's rule follows straightforwardly from the fact that a separating equilibrium must at some point have evolved from a pooling equilibrium (otherwise, it does not make sense to study the evolution of signalling). In a pooling equilibrium, the optimal action taken by the receiver will correspond to the most likely event. In a separating, the absence of a signal is therefore already interpreted as referring to a frequent event.

From there on, the interpretation of a signal as referring to an infrequent event can evolve. From a pooling equilibrium, for a separating equilibrium to evolve where a signal refers to a frequent event, the receiver would somehow have to change his initial interpretation of absence of a signal as referring to the frequent event. It is this reversal of meaning that makes it difficult for a separating equilibrium that does not obey Horn's rule to evolve from a pooling equilibrium. The paper formalizes this point, applying replicator dynamics to a simple asymmetric signalling game. The paper starts by setting out the model in Section 2. Section 3 suggests exceptions to Horn's rule when taking the action best for the frequent even in the infrequent state of the world is risky. The paper ends with a conclusion in Section 4.

SEMANTIC VALUE AND MONETARY VALUE

Andrew Jorgesen | Dept. of Philosophy. University of St. Andrews. akj1@st-andrews.ac.uk

The claim is this: reflecting on the analogy with monetary value can tell us a lot about semantic value (i.e. meaning). The analogy helps us address the metaphysical problem of locating meaning, saying what sort of thing it is and what sort of thing can have it. Money tokens are bearers of monetary value and are passed around from person to person and exchanged for goods and services. The tokens have a monetary value, and so do the goods and services, but neither category *is* the monetary value. We know this because during times of inflation or deflation the tokens and goods and services can remain the same but their buying power (respectively, cost) changes. If the same arrangement held true of semantic value, we would say that linguistic expressions and objects had semantic values, but that neither was semantic value: in particular, objects were not the semantic values of expressions. This would be a useful way of looking at things for inferentialist

approaches to meaning that don't identify, as the direct reference theorist does, reference with meaning, or make reference a key ingredient of meaning.

Monetary value seems to be holistic and "inferentially articulated" in the way some think semantic value might be. The value of a pound floats freely against the other currencies in the sense that a pound has the same monetary value as sometimes more sometime fewer Euros, for example. And all other currencies are similar. There does not seem to be a direct external check on monetary values. There is no gold standard, and even if there were that wouldn't provide an *external* ground for monetary value in as much as it too is subject to laws of supply and demand (more gold could always be found or people could come no longer to desire it). The sense in which this is "inferential" is that the value of money tokens is a function of their relations with other tokens rather than a matter of approximation to something external to the system. Even goods and services, though of interest in themselves, don't have a 'true' *monetary* value outside their role in the system.

The analogy also sheds light on how subjective assessments of value can be bound up in a common scheme and made commensurate. This affords holistic approaches to language like Quine's an account of linguistic communication that might otherwise seem impossible. Dummett argues that holistic accounts of meaning like Quine's where meaning and belief are interlinked, cannot explain how two speakers could communicate. If they had different beliefs, some of their expressions would differ in meaning, and this means that, given holism, all of their expressions would differ in meaning. Hence because communication requires a mutual understanding of what an utterance means, communication would not be possible.

A similar argument should show that trade and exchange aren't possible. So, seeing how they are possible has valuable lessons for the philosopher of language. The argument that trade is impossible would go as follows. Trade, like communication, requires a mutual agreement as to what something is worth. But what something is worth to the buyer depends on their personal circumstances (what they already have, their beliefs and attitudes etc.). Similarly for the seller. So, only in the highly unlikely situation where buyer and seller are in the same personal circumstances will trade be possible, but of course then it will not take place at all since they are in the same circumstances with the same attitudes and beliefs. The appropriate response is that though buyer and seller differently value the money tokens and goods or services prior to the transaction, the fact of the transaction itself effects a reconciliation and commensuration of their respective assessments of value. Prior to the transaction, the money is worth more to the seller than what they want to sell and worth less to the buyer than what they want to buy with it. But the fact of exchange means actually the item and the money are worth the same (whether the buyer or the seller was correct about the true worth of the item depends on the effects of other exchanges involving other people and governed by principles of supply and demand).

If meaning is akin to monetary value as I have described it, the appropriate response to Dummett's communication objection is that the fact of communication, of passing a linguistic token between speaker and audience, effects an equation of semantic value between the meaning the utterance had in the speaker's scheme and the meaning it had in the audience's. Precisely what this value is also depends on the effects of other speech transactions involving other speakers. The equation of value at the point of communication and the need to be consistent within one's own scheme will have consequences for one's communicative engagements with others, sometimes calling for a revision of the semantic value previously accorded to other utterances. The need to interact with others through exchange and communication exerts a rational pressure on agents to make their assessments of value accord with other people's. In practice people's subjective assessments of value, both monetary and semantic, will be heavily constrained and tend to approximate societally accepted values because being drastically out of step with others can cause problems. What is needed to meet Dummett's objection then is a dynamic conception of meaning in which subjective meanings evolve over time to approximate societally accepted ones.

Semantic values can be the objects of negotiation just as monetary value can. This gives the inferentialist a means to capture the important externalist insights of Hilary Putnam. Where there is negotiation the question also arises whether all parties to the negotiation are equal. If not, the subjective assessment of value of the more powerful party counts for more in determining the accepted value of a token than the weaker party's. This recalls Putnam's "division of linguistic labour" – his thesis that the referents of expressions are determined by experts, though we will say it is the semantic values of tokens and objects that is fixed by the experts. Of further interest to philosophers of language will be whether analogues of the forces of inflation and deflation, and supply and demand exist in the linguistic domain.

If meaning is akin to monetary value, a pairing of expressions with their extensions (even one allowing for indexicality and other phenomena) like that offered by model theoretic semantics will not be appropriate because it will fail to capture the dynamic nature of meaning. But economists can help philosophers of language by showing how this dynamic, inferential, holistic value is mathematically tractable.

Projected conclusion: The point of the exercise is to think analogically and see what that can tell us about meaning. Even if the analogy is not perfect (because, say, there is nothing in the monetary case corresponding to the level of semantically significant sub-sentential expressions), it can be revealing. If there is no problem of exchange analogous to the problem of communication, then it is of interest to see whether it genuinely does arise in the semantic case and if so why. It is also useful to see what comes of thinking of referents and expressions as both alike *having* semantic value rather than referents *being* the semantic value of expressions.

HONESTY NORMS, EMOTIONS AND GAMES

Raúl López-Pérez Institute for Empirical Research in Economics, Zürich. rlopez@idea.uab.es

This paper assumes that (some) people get disutility when they unilaterally deviate from norms commending (1) to send informative and true messages and (2) to achieve surplus-maximizing allocations. The model is consistent with abundant qualitative experimental evidence showing that, under certain conditions, communication increases both cooperation in social dilemmas and coordination in games like the Battle of the Sexes and the Stag Hunt.

BELIEFS ATTRIBUTION SENTENCES AND INTERACTIVE EPISTEMOLOGY

Ittay Nissan | Hebrew University. alizaet@013.net

The causality in what is referred to as "causal theories of reference" is mental causality. These theories are based upon sentences that attribute to people beliefs, desires and behavior conforming to a certain mechanism, a certain decision rule that connects their beliefs and decisions to their behavior. Specifically, the dynamics that the causal theories of reference point to as what constitutes the foundations of the process in which a linguistic expression gains its reference, are dynamics in which people's beliefs about other people's beliefs have an influence on their behavior. When people talk they attribute to each other not only the ability to use language in a competent way, but also the belief that their partners for conversation have this ability and the belief that their partners for conversation have beliefs about their own ability to use language in a competent way and so on.

In 1984, Robert Stalnaker argued, taking the picture that was drawn in the previous paragraph under consideration, that in order to escape circularity, we shouldn't take the objects of beliefs attribution sentences, the propositions which are the content of the beliefs attributed, to be

sentences in the speaker's language (or in some kind of "language of thought"), since in doing so we will create a situation in which the analysis of beliefs attribution sentences is based on a concept of reference which is analyzed using those same sentences. Stalnaker suggested taking the propositions to be sets of possible worlds, i.e. of possible ways reality might be, and argued that this point of view gives the propositions "just the structure that is motivated by the pragmatic account of the functional role of representation".

Stalnaker sketched the outlines of "the pragmatic account" without going into details, but in situations of the kind that the causal theories of reference are based upon, situations where beliefs of people about other people's beliefs effect their behavior, filling up the details of "the pragmatic account" is not a trivial task. In economics, the field of "Interactive Epistemology" deals in modeling situations of this kind. It is possible to distinguish between two kinds of models in this field: semantic models and syntactic models. In semantic models the knowledge and the beliefs of the players are described as sets of "states of the world" and in syntactic models the knowledge and the beliefs are described using tools taken from the field of modal logic.

In the 20th century 90^s, as a result of what seemed to be a conceptual ambiguity in the semantic models, researchers started to offer explications to the concept of "the states of the world". As a part of this program, Robert Aumann suggested, in two papers from 1999, to reduce the semantic modal that he had developed in 1976 to a syntactic one by defining a state of the world in the semantic modal as a closed, coherent and complete list of formulas in the syntactic model. Some of those formulas are formulas that have the interpretation of sentences of the form "x believes that p", i.e. of beliefs attribution sentences.

I want to argue that Aumann's position should be seen as an argument against Stalnaker's suggestion. The objects of beliefs attribution sentences, according to Stalnaker, are sets of "possible worlds" (or "states of the world"), but what exactly are those possible worlds? It seems that in situations of interactive epistemology, we must give them a linguistic structure, in order to avoid conceptual difficulties. Since the situations which stand at the base of the causal theories of reference, which were one of Stalnaker's original motivations to suggest his analysis, are situations of this kind, it seems like Stalnaker's position is facing a problem.

The last argument can be seen, I would like to argue, as evidence on behalf of the position that takes language to be conceptually prior to rationality. In 1982, Donald Davidson argued in favor of this position. In many ways, my argument is an extension of Davidson's argument, but unlike Davidson, my emphasis is on evaluating the different analysis to beliefs attribution sentences according to their level of "suitability" to the use that the economical models make of these sentences. My assumption here is that the analysis that will provide the most stable basis to these models, other things being equal, will be the preferred one. Since my argument's starting point is the economical models, I must not ignore the issue of sentences that attribute beliefs with a certain probability to people, and the relation between this kind of "quantified beliefs" attribution sentences to "qualified beliefs" attribution sentences.

Holding the philosophical position according to which language is conceptually prior to rationality is not neutral to the economic research and might point to promising research directions and to other fields of knowledge which can enrich, by integration, the economic point of view.

WHY EVOLUTION DOES NOT ALWAYS LEAD TO AN OPTIMAL PROTO-LANGUAGE: AN APPROACH BASED ON THE REPLICATOR DYNAMICS?**Christina Pawlowitsch | Department of Economics, University of Vienna;
christina.pawlowitsch@univie.ac.at**

Sender–receiver models in the style of Lewis (1969), Hurford (1989), or Nowak and Krakauer (1999) can be used to explain how meaning is acquired by signals in situations of cooperative interaction. Importantly, meaning here is not an *ex-ante* concept, but arises as an equilibrium property of a game. A strategy of this game is a pair of a sender and a receiver matrix, where the sender matrix associates signals to events that possibly become the object of communication, and the receiver matrix associates events to possibly received signals. An equilibrium strategy, then, can be interpreted as a proto-language, that is, a set of event–signals relations that facilitate communication over a finite number of events.

This game generally admits a multiplicity of Nash equilibrium components, where two (or more) events share the use of one signal or where two (or more) signals are associated with the same event, which leads to a situation where some of the potential of communication is left unexploited. Wärneryd (1993) as well as Trapa and Nowak (2000) show that only the strict Nash strategies, where each event is bijectively linked to one signal and where the inverse of this mapping is used to associate signals with events – which therefore guarantee the full potential of communication – are evolutionarily stable. Evolutionary stability implies asymptotic stability in the replicator dynamics. Simulations with this model in the style of a replicator dynamics as reported in Nowak and Krakauer (1999), however, typically give rise to suboptimal proto-languages where more than one event is linked to the same signal whereas another signal remains idle. In view of Wärneryd (1993) and Trapa and Nowak (2000) – which implies that such a strategy cannot be evolutionarily stable – this raises the question whether this reflects generic behavior of the replicator dynamics for this model; and, if so, what are the properties of a strategy that can protect itself from being driven out by this dynamics despite the fact that it cannot be evolutionarily stable.

This paper gives answers to these questions in terms of neutral stability and its dynamic consequences. It, first, provides a complete characterization of neutrally stable strategies for this game, showing that in such a situation, indeed, there can be two (or more) events that are linked to the same signal or two (or more) signals that are linked to the same event – as long as the degree of ambiguity is not too high. What is not compatible with neutral stability are situations where two (or more) events are simultaneously linked to two (or more) signals, or where some signals remain idle in the presence of events that are never possibly successfully received. The reason for this is that these cases leave too much room for shifting around elements in the sender and receiver matrices such that one mutation would be enough for increasing the communicative potential of the strategy at hand with respect to other strategies without reducing its communicative potential with respect to itself.

The second part of this paper is concerned with a qualitative analysis of the long-run behavior of the replicator dynamics of this model. This is done by exploiting the dynamic implications that are dormant in its static game theoretic analysis, which, together with neutral stability, centrally revolves around the double symmetry of this game. Building on a result by Bomze (2002), it can be shown that the replicator dynamics of this model generically converges to a neutrally stable strategy, but not necessarily to an evolutionarily stable strategy. This is exactly what is reflected in the simulations reported in Nowak and Krakauer (1999). Evolution does not always lead to a communicative optimum, but can be trapped in a situation where some of the potential of communication is left unexploited. The implications of this for equilibrium selection in games of

perfectly coinciding interests with a multiplicity of optimal equilibria are that – without any further ex-ante semantic assumptions – the mere possibility to send and to receive simple signals does not guarantee that a trial and error process always will lead agents to a perfectly coordinating equilibrium. The results presented here crucially rely on the assumption that the events that possibly become the object of communication occur with equal frequencies and are equally important as well as that all available signals induce the same costs. Furthermore, it is assumed that, in every period, there is a large number of individual agents who are identical with respect to their strategy sets and payoff functions. In particular, there is no spatial structure according to which individual agents could be distinguished.

Language diversification is one of the central linguistic facts that need explanation (Hurford, 2003). First examples with spatially structured populations suggest that this model might help us to understand the evolution of locally coexisting proto-languages, providing the basis for explaining how the diversification of more complex linguistic structures could evolve.

DYNAMICAL PERSUASION WITH A CONSTRAINT NUMBER OF ARGUMENTS IN A ONE-DIMENSIONAL SPACE

Morgane Tanvé | Centre D'Économie De La Sorbonne. Morgane.Tanve@Univ-Paris1.Fr

This paper examines dynamical persuasion with a constraint number of arguments, in a one-dimensional space. This model intends to demonstrate in which extent the number of authorized arguments limits the aimed audience. Many papers on persuasion have shown that individuals who are uncertain about the quality of an information or an opinion source will infer that the source is of higher quality when its reports conform to the individuals' prior expectations. Though, when people try to build up an opinion on a given subject, they would try to ask a trusted expert¹. This expression « trusted expert » refers in fact to two notions. On the one hand, an expert is someone who has a precise information on a given subject. On the other hand, a trusted expert is someone who certainly has an accurate information on a given subject, but whose information cannot be too far from the prior expectations of the audience. This second perspective is a subjective notion which falls within the competence of people who try to build up an opinion. That second aspect gives rise to a literature on information bias in media.

A precision is however noteworthy. We do not consider, in this paper, that the speaker who tries to convince the greatest number of individuals knows the realization of the state of the world. The speaker does not try to send an information but rather his opinion. This paper is more related to the models in political economics on ideology. In order to convince the greatest number of individuals of his opinion, the speaker must arbitrate between the both aspects previously described. On the one hand, appearing as a trusted person in showing arguments which, at the beginning, do not disturb the audience he tries to convince. On the other hand, appearing as an expert and so producing a number of arguments testifying that he has an enlightened opinion. This trade-off is done in a chronological way. The speaker will produce arguments which are large enough at the beginning, in order to appear as a trusted individual, i.e. as an individual who shares the idea of his « electorate ». Then, as people interested by his opinion update their belief, the speaker will clarify his opinion in order to establish his expertise. In our view, the expression « large argument » means arguments which can be used as support to several opinions. We envisage an argument as a function of two variables : On the one hand, the opinion which is stood up ; on the other hand, the precision, i.e. the strength with which the opinion is defended.

We try to capture the idea according to which the vaguely a message is, the more possible interpretations it can have. In this way, a vague argument or an unprecise one can be heard by several individuals insofar as such an argument can corroborate several points of view, even divergent ones. An unprecise argument attracts a large audience but a relatively weak support. On the contrary, an accurate argument limits the possible interpretations by explaining clearly the

opinion of the speaker. Actually, the proofs which allow to back up the accurate defended argument create contestation among individuals who do not agree with each of the stated elements. Consequently, an argument can be approved only by an audience all the more limited as the argument is precise. Nevertheless, the aroused support by the listening of an accurate argument is high. Modelling an argument as a function of two variables seems to reflect, in our view, the trade-off previously described, namely appearing as a trusted person (preference for large arguments) and as an expert (preference for accurate arguments). The originality of this paper hangs on the formalization of what is an argument. An argument is a function of two variables, namely the defended opinion and the precision about this opinion. In a first approach, we consider that the opinion of the speaker is exogenous. Furthermore, we assume that the speaker has to choose q arguments among a number a of possible arguments. These a available arguments are of different accuracy. We have motivated the negative correlation between accuracy and audience. In order to take this link into account, we assume that an argument is a probability distribution over the opinion space. In other words an argument is seen as a mixed strategy over an interval with no gaps. Actually, to end of realism, we consider that a message must be coherent in its purpose. Either the speaker defends precisely his opinion (pure strategy on his opinion, i.e. unit probability), or he does not defend clearly this one (mixed strategy over an interval of the opinion space). In the latter case however, the speaker cannot defend any other opinion, but points of view more or less close of his one, point of view fixed exogenously in a first draw. The speaker will more ponderate opinions which are close to his one. Consequently, this leads us to restrict the probability distributions to symmetrical ones (symmetry in relation to the opinion defended by the speaker). Schematically, an argument is an isosceles triangle with middle of the basis being the opinion stood up by the speaker. Opinions of the n players ($n \geq 1$ listeners and 1 speaker) are uniformly distributed over the interval $[0; 1]$. In this way, the listener i has a prior expectation $x_i = i/n$. The game takes place in as much stages as the imposed number q of arguments. In the first stage, the speaker sends a message. We assume that only individuals whose prior expectation received a positive probability via this message update it. This is done in a Bayesian way with respect to the probability distribution of the message. This process takes place q times. We show that the speaker uses a funnel-shaped process. The first sent messages are large, i.e. unprecise regard to the stood up opinion. However, as the argumentation goes, the speaker sends more and more precise messages. We also find that the weaker q (authorized number of arguments), the more restricted the beginning audience choosed by the speaker. In this way, he uses immediately accurate arguments.

This paper offers a new perspective for persuasion by exploring a formalization of argumentation via mixed strategy. Nevertheless, although this paper tries to formalize a dynamical persuasion, it does not wonder about the reasons for which people could update their belief by hearing a lambda speaker. The precision of arguments used by the speaker rest on the more or less rich presentation of proofs. These proofs have only a validity in the defence of a purpose and not in an absolute way. These issues - Why do individuals update their belief by hearing some speaker ? How to define the accuracy of an argument ? As a function of a number of proofs or as a function of pertinence of these ones ? - remain open and offer extensions for future studies.

REGISTERED PARTICIPANTS

| | | |
|-------------------------|---|---------------------------|
| Brice Corgnet | Universidad Carlos III (Madrid) | bcorgnet@est-econ.uc3m.es |
| Miranda del Corral | UNED | mirandiki@gmail.com |
| Edward Erwin | | eerwin@miami.edu |
| Oswaldo Feinstein | Facultad Latinoamericana de Ciencias Sociales | ofeinstein@yahoo.com |
| José Luis Ferreira | Universidad Carlos III (Madrid) | jlfer@eco.uc3m.es |
| Maria Fretescu | | mariafretescu@yahoo.com |
| Roderick A. Jacobs | University of Hawai'i | rjacobs@townsquare.net |
| Alicia Jiménez | International Monetary Fund | ajimenez@imf.org |
| Luis Jiménez | | ljimenez888@yahoo.com.ar |
| Hubert J. Kiss | Universidad de Alicante | berti@merlin.fae.ua.es |
| Svetlana Moisei | | svetlanamoisei@yahoo.com |
| Carmelo Núñez | Universidad Carlos III (Madrid) | cnunez@eco.uc3m.es |
| Iara Vigo de Lima Onate | University of Stirling | ivdo1@stir.ac.uk |
| Jorge Sainz | Universidad Rey Juan Carlos | Jorge.sainz@urjc.es |
| Ana Sanz | CSIC (Madrid) | anasanzf@gmail.com |
| Ana Vasiliu | JspurC | vasiliu@jspurc.org |
| Stefano Versace | Freie Universität Berlin / Università degli Studi di Milano | ste_vers@yahoo.it |