Semantic Generalisation and the Inference of Meaning

Andrew D. M. Smith

Language Evolution and Computation Research Unit,
School of Philosophy, Psychology and Language Sciences, University of Edinburgh
andrew@ling.ed.ac.uk

Abstract. In this paper, a computational model of a successful negotiated communication system is presented, in which language agents develop their own meanings in response to their environment and attempt to infer the meanings of others' utterances. The inherent uncertainty in the process of meaning inference in the system leads to variation in the agents' internal semantic representations, which then itself drives language change in the form of semantic generalisation.

1 Introduction

Modern evolutionary linguistics is primarily occupied with understanding the apparently paradoxical universality and massive diversity found among human languages; much of the recent work within this paradigm uses A-life computational simulation techniques to explore these questions. It is important to note that the evolution of language should not just be considered in a genetic framework, but also in a cultural one; children acquire their language based on the linguistic information produced by those around them. Recently, researchers have shown that the cultural transmission of language itself leads to adaptive pressures which can explain some of the characteristic properties of language [4, 7, 15]. Cultural explanations of the structure of language mean, importantly, that there may well be less need to fall back on the existence of a somewhat ethereal language acquisition device [11].

Many A-life models of language evolution focus on the emergence of syntactic structure in their agents' language, but fail to explain the origin of the semantics on which their 'emergent' syntax is built. One of the most intriguing universal features of language is the ceaseless and inevitable nature of language change, driven by language variation within speech communities [18]; in this paper I present a model of conceptual development and negotiated communication between agents, in which agents can create meanings individually and still co-ordinate a joint language. I show, moreover, that the inherent uncertainty in the process of meaning inference in such a usage-based model of language [5], leads to variation in the agents' conceptual structures, which itself drives language change in the form of semantic generalisation.

2 Meaning Representation and Creation

A large proportion of recent A-life research into language evolution has been focused on the emergence of syntactic structure, or more specifically compositionality [1, 4, 7].

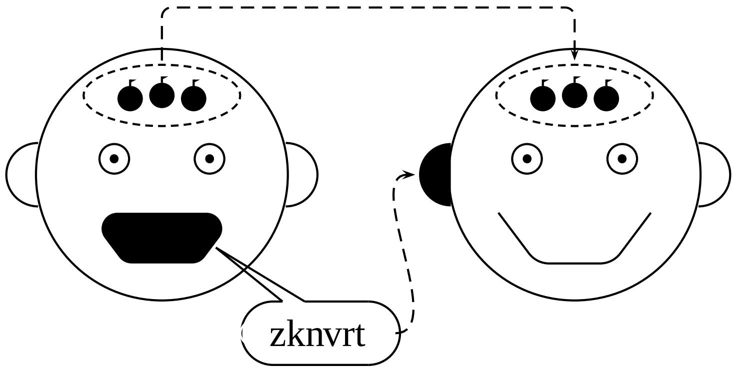

Fig. 1. A communicative episode which consists of the explicit transfer of both a signal ‘zknvrt’ and a meaning ‘three apples’ from speaker to hearer.

Although the precise implementations vary to some extent, in all compositionality models structure arises in the signal space as a direct consequence of the agents recognising and coding regularities between parts of signals and parts of meanings. There are a number of problems with models like these, particularly with respect to their semantic representations, which I will discuss here briefly.

Firstly, the agents are always provided with a structured meaning space, which is explicitly linked to an unstructured signal space. As Nehaniv [9] has pointed out, structure develops in the signal space precisely because the agents use signals coupled with already structured meanings; the signals are essentially parasitic on the meanings, and structure only emerges as the agents decipher the pre-defined semantic coding system.

Secondly, no treatment of meaning can avoid addressing the fundamental concepts of both *sense* and *reference* [6]. A basic model of sense relations would include at least, for instance, some notion of antonymy, the relationship between a word and its opposite[1], which is “one of the most important principles governing the structure of languages” [8, p.271], and might be expected also to include other notions such as *hyponymy*, which describes the relationship between a subset and a superset, as for instance between *cat* and *animal*. Reference, on the other hand, is defined in terms of the objects or actions in the external world to which a word refers in a particular context. It is very difficult to relate the ‘semantic’ structure in such compositionality models to sense relationships in the semantics of real languages in any way, and, even more problematically, there is no reference at all in the meaning systems, either because there is no external world in the experiment at all [1], or because the world is inaccessible to the agents [4, 7]. Without reference, and with only a very tenuous link to sense, it is clear that such representations of meaning are actually semantic in name only.

[1] There are many types of opposition in language, including gradable antonyms, such as WET/DRY, expressing meanings on a relative scale; ungradable binary antonyms, such as ALIVE/DEAD, expressing complementary propositions; and converses, such as ABOVE/BELOW, BUT , which refer to the same relationship from opposite points of view.

Thirdly, as I have argued elsewhere [13, 14], the lack of reference in the meaning representations of these models leads to an implementation of communication which is seriously flawed, because of its necessary reliance on *explicit meaning transfer* [13]. Figure 1 shows an example of such a communication episode, with the speaker on the left uttering a word *zknvrt*, which is received by the hearer on the right. Simultaneously, however, the meaning THREE APPLES, is transferred directly from speaker to hearer. During communication, the hearer is given explicitly not only the signal, but also both the meaning, and the information that it should make the appropriate association between the particular signal and meaning. Such a model not only sidesteps the very important Quinean [12] problem of how hearers interpret the meanings of unfamiliar words from a set which is in principle infinite, but also actually undermines the need for language at all: if meanings can be transferred accurately by telepathy, then the signals are not being used to convey meaning. There is little motivation, therefore, for the emergence of a system in which the agents spend their time and energy developing a communication system which encodes exactly the same information as another system which they are born with and which already works perfectly.

The semantic model in this paper, on the other hand, tries to avoid these pitfalls; the basic procedure of agent-based grounded meaning creation through discrimination games was originally presented by Steels [16], and has been used by others both in simulated worlds [2, 14], and on robots interacting with objects in a real environment [17, 19]. An agent is situated in a world made up of abstract objects, with which it interacts by playing discrimination games, which consist of four parts:

scene-setting: a set of objects, called the context, is chosen from the world and presented to the agent; one of these objects is chosen to be the target of discrimination.

categorisation: the agent cycles through the objects in the context, returning, for each, an association with one or more of its semantic representations.

discrimination: the agent tries to define a distinctive category for the target object, i.e. a category which both represents the target and does not represent any other object in the context.

adaptation: the agent modifies its internal conceptual structure in some way.

In the model used in these experiments, the objects in the world have a number of abstract, meaningless features, the values of which are normalised to lie in the range 0.0...1.0, and the agents have a specific sensory channel corresponding to each feature, on which they build a hierarchical semantic representation, or a discrimination tree [16]. Each node on the discrimination tree is a discrete category, corresponding to a particular contiguous segment of the continuous feature value space. Categories are created by splitting the sensitivity range of a node into two discrete, equally sized segments, each of which is therefore sensitive to half the range of the previous node. The trigger for the creation of new meanings is failure in the discrimination game; it is important to note, however, that there is no pre-definition of which categories should be created, nor is there any guarantee that the newly created categories will turn out to be useful in future discrimination games.

Because of this, agents develop different, though equally valid, representations of the same world. I quantify the similarity of two agents' meaning structures by averaging the similarity of the particular discrimination trees built on each of their sensory channels. If $k(t,u)$ is the number of nodes which trees t and u have in common, and $n(t)$ is the total number of nodes on tree t, then the similarity between any two trees t and u is:

$$\tau(t,u) = \frac{1}{2}\left(\frac{k(t,u)}{n(t)} + \frac{k(t,u)}{n(u)}\right) \tag{1}$$

I then obtain an overall measure of *meaning similarity* σ between two agents, by averaging τ over all their sensory channels. If a_{ij} identifies channel number j on agent i, and each agent has c sensory channels, then the meaning similarity σ between agents a_1 and a_2 is:

$$\sigma(a_1,a_2) = \frac{1}{c}\left(\sum_{j=0}^{c-1} \tau(a_{1j},a_{2j})\right) \tag{2}$$

If two agents a_1 and a_2 have identical conceptual structures, where $\sigma(a_1,a_2) = 1$, then I refer to their meanings as being *synchronised* [14]. Meanings in this model are expressed using the notation *j-r*, where j identifies the number of the sensory channel, and r is a sequence of 0s and 1s representing the path from the tree's root to the node in question; if the tree's root is on the left, and it branches to the right, then 0 signifies a traversing of the lower branch, and 1 the upper branch.

Importantly, the meanings are grounded in the world, created in response to the environment, and encode both sense and reference relations: the hierarchical nature of the tree means that meanings nearer the root of trees (and therefore with a shorter route r) are more general than those nearer the leaves, and that a node and its daughter nodes can be related by hyponymy; while the meanings also refer to the properties of objects in the world.

3 Communication and the Inference of Meaning

Avoiding the problems with previous models discussed above, the speaker's meaning is *not* transferred to the hearer, in contrast to associative learning models [4, 3], nor does the hearer know which object in the context is being referred to, in contrast to Steelsian guessing games [17]; the communication process is made up of three separate sections:

production: the speaker, having found a distinctive category in a discrimination game, chooses a signal to represent this meaning.
transfer: the signal is transferred to the hearer.
interpretation: the hearer interprets the signal from the context in which it is heard.

Each agent maintains a dynamic lexicon of associations between signals and meanings, for use both in production and in interpretation. Each entry in this lexicon contains a signal s, a meaning m, a count u of the pair $< s,m >$'s mutual association, and a representation p of the agent's confidence in the association between the pair $< s,m >$. A signal-meaning pair can be used both by being uttered by the speaker and by being understood by the hearer, so that u is the total number of communicative episodes in which the agent either uttered s to represent m, or interpreted s as representing m. An agent's confidence in a signal-meaning pair is based solely on the relative co-occurrence of signals and meanings, or the proportion of times in which s has been used that it has been associated with m. More formally, $p(s,m)$ can be expressed as:

$$p(s,m) = \frac{u(s,m)}{\sum_{i=1}^{l} u(s,i)}, \tag{3}$$

where l is the number of entries in the lexicon. Communicative success is based on referent identity; speaker and hearer communicate successfully by referring to the same object, but they are not obliged to use the same meaning to do so.

Because meanings are not explicitly transferred between agents, the hearer must infer the meaning of the signal from the context. It is important to note that the hearer does *not* know which object is the target object, and so it must try to come up with descriptions for all the objects in the context. The hearer plays a series of discrimination games and creates a list of possible meanings to consider, each of which describe only one of the objects in the context. The hearer has *no* other information, so all these meanings are equally plausible; it therefore associates each of them with the signal in the lexicon, modifying the values of u and p in the lexical entries accordingly. Having modified its lexicon, it chooses, from the list of possible meanings, the meaning in which it has the highest confidence p. This modification of the lexicon in context is the only way in which the agents learn; in contrast to similar language game models [17], they receive no feedback about the success of either their communication games or their lexicon development. I have shown previously [13] that communication is successful under these circumstances if the speaker chooses a signal that it thinks the hearer will understand. Of course, the speaker does not have access to the hearer's lexicon, as this would defeat the object of ruling out mind-reading, so it bases its decision on what *it itself* would understand, if it heard the signal in this context, without knowing the target object. This technique for choosing signals is a version of the obverter mechanism [10], modified so that the only lexicon an agent can investigate is its own, and which I therefore call *introspective obverter*.

4 Meaning Similarity and Meaning Variation

The introspective obverter algorithm allows the agents to develop successful communication systems; I have shown elsewhere that to achieve optimal communication, the agents should have synchronised meaning structures [14]. One effect of an optimal communication system, however, is that the agents quickly settle on a common language, which is consistently reinforced through continued use and is completely stable, in contrast to the fluidity of human language.

The development of perfectly synchronised meaning structures is, however, very unlikely, given the inherent randomness in the agents' meaning creation algorithms; so what happens to the agents' language when their meanings are not synchronised and each is trying to communicate their language to the other? I explore this by tabulating detailed extracts from their lexicons through the progress of a simulation. The agents firstly develop most of the meaning structure which enables them to succeed in the discrimination games, and only then do they begin to communicate with each other.

Agent 1				Agent 2			
Signal	Meaning	Usage	Confidence	Signal	Meaning	Usage	Confidence
bc	3-1	2	0.5	egla	1-11	9	0.45
yq	1-01	8	0.32	bc	3-1	2	0.4
tjop	3-01	7	0.32	tnip	3-10	6	0.35
...	...	...	...	...	...	...	...
egla	1-11	8	0.23	yq	1-01	7	0.25
tnip	3-10	6	0.17	tjop	3-00	6	0.2

Table 1. An extract from two lexicons, showing the high degree of coordination between the signal-meaning pairs.

Agent 1				Agent 2			
Signal	Meaning	Usage	Confidence	Signal	Meaning	Usage	Confidence
jch	2-000	9	0.69	jch	2-000	9	0.17

Table 2. The meaning exists in both semantic structures, and so context-driven disambiguation leads to an agreement over the meaning of the signal.

Extracts from two sample lexicons after four hundred communicative episodes, showing the three words in which each agent has the highest confidence, is given in table 1. Although the agents have different levels of confidence in the signal-meaning pairs, they have broadly settled on a common language, with most words referring to the same meaning for both agents.

Later on, the first agent uses a meaning 2-000 in a discrimination game, but has no word for this meaning in its lexicon. These are exactly the circumstances under which we allow lexical innovation to occur, and so it coins a new word jch and utters this to the other agent. The development of the new word jch in both agents' lexicons depends on a number of parameters, including naturally the number of times it is chosen to be uttered, but also crucially whether the meaning to which it is linked is present in both agents' semantic representations, and it is this semantic development of the word jch which I will now discuss.

If the meaning 2-000 does exist in both agents' semantic representations, then we find, after rolling the simulation on a few hundred episodes, that the agents' lexicons contain the extracts shown in table 2. The second agent has associated jch with many different meanings, having heard it in a number of different contexts. The meaning 2-000, however, has occurred in each of these contexts, and so the agent's confidence in this particular signal-meaning pair is higher than in any other; repeated exposure in different contexts has disambiguated the agent's set of possible semantic hypotheses.

If the meaning 2-000 does not exist in both agents' representations, however, then the process of coordination is not so smooth, as we see in table 3. The second agent has again associated jch with a large set of possible meanings, but this time meaning 2-000 cannot be among them, as it is not in this agent's semantic repertoire, and so it is instead most confident in the meaning $4-01$. The agents now each have a different

Agent 1				Agent 2			
Signal	Meaning	Usage	Confidence	Signal	Meaning	Usage	Confidence
jch	2-000	9	0.47	jch	4-01	6	0.11

Table 3. The meaning does not exist in both agents' semantic structure, and no agreement over the meaning of the signal is reached. As a consequence, both agents' confidence in their respective lexical association falls.

meaning associated with the word, and this has interesting consequences. The second agent now uses jch to represent meaning 4-01, and utters this in a context in which the first agent's meaning 2-000 is not a possible meaning. This leads the first agent to become less confident in its original association; when agents are using the word for different meanings over time, both agents' confidence in their respective original associations for the word fall. There is now a conflict between the agents' use of the word jch, which must in the end be resolved by one agent losing so much confidence in its preferred association that it chooses another meaning for the word.

Because of the hierarchical nature of the meaning creation process, it is more likely that the meanings which the agents have in common, and on which they can agree word associations, are the more general meanings, which are nearer the roots of the discrimination tree. When the conflict over the use of the word jch is resolved, therefore, it is likely that the eventual meaning to which it is attached is, other things being equal, a relatively more general meaning than it was originally coined to serve; its use in a variety of contexts by agents who have different semantic structures has led to its meaning becoming more generalised.

The agents' lexicons are of course dynamic, and do not stop developing because the conflict over one word has been temporarily resolved. If the first agent comes across a context in which the original meaning of jch, namely 2-000, is needed, it now has no word which it can use and so must coin another one; the generalisation process itself has led to the loss of words from some part of the lexicon, which, if they are needed again, leads inevitably to more innovation. Of course, if meaning 2-000 still does not exist in the other agent's semantic structure, then the same process of conflict over meaning, generalisation and innovation is likely to happen all over again. We find, therefore, that meaning uncertainty, which is inevitable when meanings must be inferred instead of given, leads to the development of a continuous cycle of language innovation and semantic generalisation and extension.

5 Conclusions

I have presented a model in which agents individually create meanings based on their interactions with their external environment, and develop a co-ordinated communication system without either feedback or the explicit transfer of meaning. If the agents have very similar conceptual structures, then they are able to develop optimal communication systems by inferring the meanings of unfamiliar words through their exposure in different contexts. On the other hand, variation in conceptual structure itself creates pressure leading to a cycle of innovation and semantic generalisation. Work is currently under way to explore this trade-off between semantic uniformity which helps communication and semantic variability which drives language change.

References

[1] J. Batali. The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In E. Briscoe, editor, *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, pages 111–172. Cambridge University Press, Cambridge, 2002.

[2] T. Belpaeme. *Factors influencing the origins of colour categories*. PhD thesis, Artificial Intelligence Lab, Vrije Universiteit Brussel, 2002.

[3] A. Billard and K. Dautenhahn. Experiments in learning by imitation - grounding and use of communication in robotic agents. *Adaptive Behavior*, 7(3/4):415–438, 1999.

[4] H. Brighton. Compositional syntax from cultural transmission. *Artificial Life*, 8(1):25–54, 2002.

[5] W. Croft. *Explaining Language Change: an evolutionary approach*. Longman Linguistics Library. Pearson, 2000.

[6] G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50, 1892.

[7] S. Kirby. Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, 5(2):102–110, 2001.

[8] J. Lyons. *Semantics (2 vols.)*. Cambridge University Press, Cambridge, 1977.

[9] C. L. Nehaniv. The making of meaning in societies: Semiotic and information-theoretic background to the evolution of communication. In *Proceedings of the AISB Symposium: Starting from Society — the application of social analogies to computational systems, 19–20 April 2000*, pages 73–84, 2000.

[10] M. Oliphant and J. Batali. Learning and the emergence of coordinated communication. *Center for Research on Language Newsletter*, 11(1), 1997.

[11] S. Pinker and P. Bloom. Natural language and natural selection. *Behavioral and Brain Sciences*, 13:707–784, 1990.

[12] W. v. O. Quine. *Word and Object*. MIT Press, 1960.

[13] A. D. M. Smith. Establishing communication systems without explicit meaning transmission. In Jozef Kelemen and Petr Sosík, editors, *Advances in Artificial Life: Proceedings of the* 6^{th} *European Conference on Artificial Life*, number 2159 in Lecture Notes in Artificial Intelligence, pages 381–390. Springer-Verlag, 2001.

[14] A. D. M. Smith. Intelligent meaning creation in a clumpy world helps communication. *Artificial Life*, 9(2):559–574, 2003. in press.

[15] K. Smith. Compositionality from culture: the role of environment structure and learning bias. Paper given at the 4^{th} International Conference on the Evolution of Language, Harvard University, 2002.

[16] L. Steels. Perceptually grounded meaning creation. In M. Tokoro, editor, *Proceedings of the international conference on multi-agent systems*. MIT Press, 1996.

[17] L. Steels and F. Kaplan. Bootstrapping grounded word semantics. In E. Briscoe, editor, *Linguistic Evolution through Language Acquisition: formal and computational models*, pages 53–73. Cambridge University Press, Cambridge, 2002.

[18] R. L. Trask. *Historical Linguistics*. Arnold, 1996.

[19] P. Vogt. *Lexicon Grounding on Mobile Robots*. PhD thesis, Vrije Universiteit Brussel, 2000.