TOWARDS A FIXED WORD ORDER IN A SOCIETY OF AGENTS: A DATA-DRIVEN BASELINE PERSPECTIVE

GUY DE PAUW

CNTS - Language Technology Group
University of Antwerp - CDE
Antwerp, B-2610, Belgium
guy.depauw@ua.ac.be

In this paper we present a computational model for the emergence of a fixed word order in the language of a society of babbling agents. In contrast to the majority of research modeling the emergence of syntactic principles, in which a very strong governing role is set aside for the semantic aspect of communication, this paper touches on the possibility that one of the most basic grammatical principles, i.e. fixed word order, can develop without direct reference to concepts of shared meaning. With its rigid data-driven approach, the system described in this paper introduces a hitherto unexplored baseline perspective to the research efforts modeling the emergence of grammatical properties in language.

1. Introduction

The last decade has seen many researchers propose computational models as a way to provide some insight into how early hominids may have evolved into language users. Very interesting results have been achieved in the computational modeling of the emergence of structural properties and compositionality in language (Hashimoto & Ikegami, 1996; Kirby & Hurford, 2001; Batali, 2002; Steels, 2004). While sharing the same general research objective, these models differ from one another in the incentives they define for raising the complexity of communication into the realm of compositional language and what kind of cognitive capacities are assumed to be able to do so.

Even though computational modeling provides an interesting account of how different properties of compositional language may have emerged, some fundamental issues are apparent, as agents often find themselves equipped with a very intricate set of linguistic and cognitive principles that are inherently biased towards compositionality. While there is some evidence that these kinds of cognitive capacities might have been available to early hominids, a more problematic aspect underlying many of these computational models, is the concept of explicit meaning sharing in inter-agent communication (Kirby & Hurford, 2001; Batali, 2002). Following the 'gavagai' argument in Quine (1960), the referent of a signal should indeed not be considered as something that is unambiguously available to participants in a conversation.

In this paper, we describe a computational model that approaches the problem from an entirely different angle. It minimizes the semantic value of the agents' communicative attempts and avoids the problematic concept of shared meaning altogether. Experiments show that even in this overly restrictive setup, convergence towards fixed word order can emerge out of initial chaotic attempts. With its rigid data-driven perspective to language evolution, the baseline model presented in this paper raises some interesting questions as to the role of distributional properties in (i) the emergence of governing principles in compositional language and (ii) the previously unexplored role of statistical conventionalization effects in the evolutionary computational models of language.

2. Stripping Down Data-Driven Grammar Optimization

The system described in this paper is essentially a stripped down version of an agent-based system called GRAEL that is geared towards data-driven optimization and induction of grammars for natural language (De Pauw, 2002). The instantiation of GRAEL described in this paper discards the notion of a pre-defined language. The agents are equipped with the ability to create mental constructs and process agent-patient type relationships. They also have the capacity to record the frequency of linguistic events, a psycholinguistically realistic cognitive concept (Juliano & Tanenhaus, 1993).

We also assume the agents have acquired the naming insight. This translates in the following limited lexical knowledge attributed to the agents: a lexicon of 40 single-entity names, including the 10 agents in the society and 30 objects with an associated [±animate] subcategorization classification. They also have a repertoire of 4 attributes expressing singular properties for [+animate] entities (e.g. {**hungry** `tiger`}) and 20 terms that express relationships between a [+animate] agent of an action and its patient (e.g. {**eat** {`hungry tiger`},`agent8`}).

In accordance with Wray (1998) who proposes that the naming insight occurs at the cross section between compositional and holistic language, we simultaneously provide the agents with a set of 200 randomly generated holistic utterances expressing entire SVO relationships in one single term, but allowing for wildcards (e.g. {**eat** `[+animate],any`}) to express general concepts.

3. Something to talk about

The experiments described in this paper take place in a closed world of 10 agents and the 30 objects named in the lexicon. We assume the conversations to be a form of gossip, as suggested by (Dunbar, 1996), and suggest that the referents for these utterances are not visually available to the listener. This entails that the agents only have language to convey meaning without the help of gesturing or gaze following. When the speaker tries to convey a situation to the listener, he can only do so by uttering words in a specific word order. For the hearer, the only extra-linguistic tool for disambiguation of the utterances is the knowledge that the internal structure of the state-of-affairs expressed by the speaker does not change over time and the internal agent-patient relationships remain static.

Each agent in the society is provided with a single, static state-of-affairs to talk about, represented as a 9cell situation matrix (Figure 1, left). There are 8 participants in each state-of-affairs and the [+animate] participants have some attribute assigned to them (represented as circles in Figure 1). Relationships are defined between neighboring cells of the table, which can hold between the atomic entities in the cells. The concept *Agent6 and Agent7* for example could be expressed by `{and 4,1}`. Relationships can also be used as entities in a relationship themselves. This allows the situation *Agent6 and Agent7 chase a hungry tiger eating a pig* to be expressed by `{chase {and 4,1},{eat {hungry 5},8}}`.

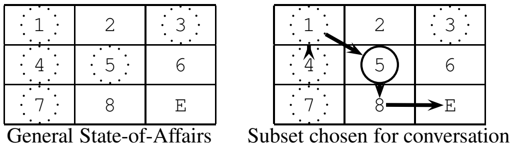


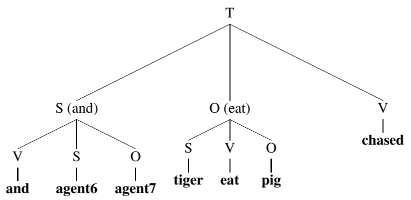
Figure 1. State-of-Affairs: Situation Matrix

In each communicative attempt a subset of the state-of-affairs is discussed. An agent will choose a cell from the table and follow a random path to the end-state `E`. The path through the situation matrix on Figure 1 (right) for example expresses the above example *Agent6 and Agent7 chase a hungry tiger eating a pig*, starting at cell `4` and ending in cell `E`. For a more detailed overview of how these state-of-affairs are internally represented, we refer you to De Pauw (2002).

4. Processing Utterances

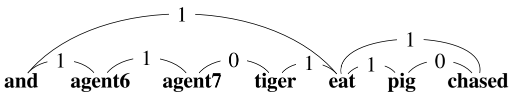
The concept of the situation matrix implies that the agents have the cognitive capacity to process structured mental representations. We assume that the same mental capacity also allows the agents to make a linguistic distinction between the agent and patient of an event. We define six different possible ways to order the agent (**S**ubject), action (**V**erb) and patient (**O**bject) of the relationship: **SVO**, **SOV**, **VSO**, **VOS**, **OSV** or **OVS**. Attributes are represented as either **VS** or **SV**.

Different relationships can be expressed using different orderings. Let us suppose the agent decides to generate the fully compositional (i.e. not using holistic tokens) utterance expression for the event `{chase {and agent6,7},{eat tiger,pig}}`. He therefore needs to express three different constructs, which each can be produced using a different ordering, as exemplified in the following structure:



The utterance the agent uses to express the complex meaning is consequently *And agent6 agent7 tiger eat pig chased*. Although this method does not establish totally free word-order, there are still many different possible word orders available to express the mental structure.

After generating the sentence, the speaker will store the frequencies of the component bigrams of the utterance in memory. The basic frequency of a bigram '*a b*' is defined as its total number of occurrences, divided by the total observed number of occurrences of '*a*'. The speaker however is allowed to refine the notion of a bigram by referring to the structural properties of his utterance. Adjacent bigrams are not counted if they belong to different nodes in the structure, while bigrams of non-adjacent head-nodes in a structure are. This is illustrated in the following example:



In subsequent productions, the agent will still consider the six aforementioned possibilities, but will calculate for each ordering the probability by multiplying the frequency values of its component bigrams. The ordering with the highest probability will consequently be the one the speaker prefers to use in his utterance.

Let us now turn to the **hearer** agent. He observes this sentence, but has no access to its meaning, nor its underlying structure. He simply records the frequency of the bigrams, so that he can use these frequencies later to apply a preferred ordering to new utterances. Structural information is not available to the hearer, so the bigrams are simply recorded as they are sequentially observed, without structural properties influencing the bigram count.

After processing the sentence, the hearer will have some vague notion of what exactly the speaker is talking about. The hearer agent will then pick one of the objects from the utterance and repeat it to the speaker, who will then turn back to his situation matrix, find the object the hearer wants to know more about and track a new path to the end-state, after which a new utterance is generated.

During each language game, the hearer will ask 500 such questions, so that in the end, plenty of information about that particular state-of-affairs has been conveyed. This obviously does not constitute a realistic model of conversation. A more realistic number of questions per conversation could be used if we proportionally raise the average number of conversations during an agent's lifespan, but this would negatively affect computational complexity.

Previously, we also mentioned a limited set of 200 randomly generated holistic utterances expressing general SVO situations (e.g. {`eat [+animate],any`}). A speaker agent can chose to use these holistic utterances to express (part of) the meaning. We do place an important restriction on the use of holistic tokens based on its probability: a holistic token can only be used if its probability is higher than the sum of the bigram probabilities between the names of its atomic concepts. The idea behind this is that it is only useful to use a holistic token, if it does not overly generalize the situation and miss out on salient information.

The probability of a holistic token will benefit from its general nature: a holistic token might therefore be more widely applicable. The atomic names on the other hand can be used less generally, but in a wider range of combinations. The experiments will show how these two conflicting forces work alongside each other: even though the experiments show compositionality is usually preferred, holistic artifacts do sometimes survive in the language of the society.

5. Experiments with GRAEL

We have performed ten different experiments, with roughly the same results. We single out a typical experiment for discussion. Each experiment was performed on a 10-agent society over the course of 5000 language game runs. During each run, each agent in the society is assigned as a hearer and as a speaker once. The GRAEL society is generation-based, but does not employ fitness functions, nor any type of genetic transmission of information. At some points in the society new agents are introduced, while other agents die at random intervals (lasting at least 200 runs, but with an upper-bound limit of 1000 runs on their life-span). This allows for a dynamic population size, which is only restricted by imposing a lower-bound society size of 5 agents.

5.1. *Quantitative Evaluation*

To check whether the agents are indeed developing some shared notion of word order, we halt the society after every 1000 runs and extract the agents. We then place (the same set of) 100 pre-defined meanings in their mind, for which they are asked to render sentences. The degree to which the sentences that are produced are the same among the agents, expresses the society's convergence on a word order.

To illustrate this convergence, we use the convergence diagrams in Figure 2. These diagrams display the 100 meanings, starting off with 4 attributes, followed by 36 3tuple relationships (4-21 have a [+animate] object, 22-39 a [-animate] object, all have [+animate] subjects), followed by 60 complex meanings. These meanings are ordered in increasing complexity of up to 15 relationships per utterance.

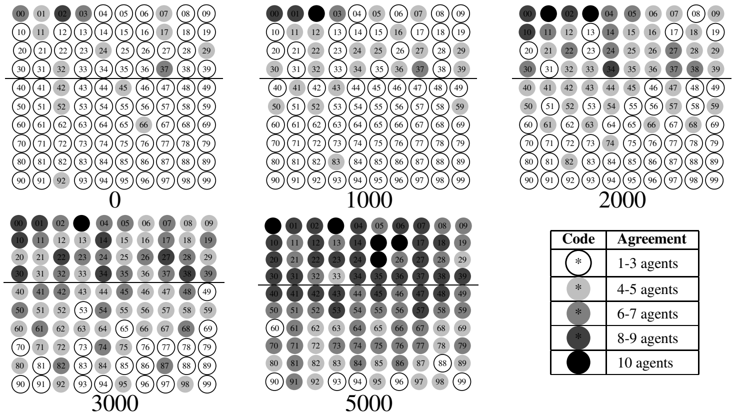


Figure 2. Convergence Diagram: shared word order in the society

The first diagram labeled '0' in Figure 2 displays the situation at the start of the GRAEL society. The 4 attributes (meanings expressed by at most two tokens) are trivially found by at least half of the population. The word order for some of the simple relations is also shared by several agents, either by a very lucky ordering of the words, or by several agents choosing the same holistic utterance to express a complex meaning, thereby reducing the randomness effect of word ordering.

After about 1000 language game runs, the word order for the attributes is shared by almost all agents. Looking at the diagram after 2000 runs, we notice that there seems to be a further tendency towards convergence for the word ordering of the more complex relationships as well. There is a noticeable difference between the meanings that express a relation between two **[+animate]** objects (meanings 4-21) and meanings that express a relationship with a **[-animate]** patient (meanings 22-39). On the whole, the word order seems less random for the former. This may be explained by agents having a much easier time distinguishing word order between two classes of objects, whereas it is not transparent from the word order alone which object is the agent and which is the patient.

The convergence diagram after 3000 runs shows that this positive trend of convergence continues. All simple meanings now display at least some limited degree of convergence, with many word orders shared by more than half of the population. It is mainly the young agents in the society that account for the randomness that is still present in the word ordering mechanism of the society. As convergence continues however, the newborn agents will be met with more consistent word order patterns, which helps them to pick up general tendencies more quickly. This is in line with the notion of the transmission bottleneck coined in Kirby and Hurford (2001).

At the end of the experiment, many complex meanings are being expressed in the same word order by more than half of the population. There is less disagreement on how to order sentences in which relations themselves are used as agent or patient, so that these patterns can be picked up faster by the newborn agents. Convergence has improved on some meanings, while other (mostly complex) meanings seem to have become a bit more randomized (e.g. (60)) compared to the earlier situation at 3000 runs.

The experiments show that introducing an amplifying function of statistical processing in language, allows the society to evolve from chaotic word order to a generally shared, but not finitely converging concept of word order. It appeared indeed during many experiments that the society is drifting from one local maximum of convergence to the next. We do not consider this a problem, as a non-converging model of language evolution constitutes a more realistic one.

5.2. *Holistic Artifacts*

Holistic elements disappeared over time in about half of the GRAEL experiments, with only a few of them surviving in the other half. This can be explained by the large amount of different meanings that need to be expressed by the agents, allowing for high probability values of atomic names, which can easily overtake that of the holistic token.

In four experiments two holistic meanings had been maintained throughout the society over time. Analysis of the data did not reveal any reason as to why these particular phrases had been maintained and why this occurred in some experimental runs and not in others. We do not consider the survival of holistic items a problem, as Wray (1998) suggests holistic utterances are always present in natural language, even today, as living fossils of an earlier protolanguage.

6. Discussion

In this paper we introduced GRAEL as a computational model for the emergence of a fixed word order without using the problematic tool of shared meaning. Simply on the basis of general cognitive mechanisms like the ability to build mental constructs and the ability to register frequencies of utterances, we have shown how fixed word order can evolve in a society of babbling agents. This provides a previously unexplored data-driven point of view to the problem of modeling language evolution. Even though we have made some unrealistic assumptions in the implementation of GRAEL, most particularly the marginalization of the aspect of meaning in communication, we propose that simple distributional effects of language use can account for the emergence of inherently formal and syntactic constraints on language. The almost trivial effect of statistical conventionalization may indeed explain the convergence towards a fixed word order in our experiments and we tentatively propose that this “artifact” of language use may actually be a strong amplifying function in language evolution itself. Further experimentation in the same vein however is warranted to support this claim.

There are however some obvious problematic issues in our interpretation of the data we have purposely ignored up to now. When we observe agents using the same word order to express the same meaning, we might consider them to be able to understand each other. But we are faced with the empirical problem that this type of convergence may to a large extent be due to statistical conventionalization, rather than true understanding between agents. Whether or not GRAEL then constitutes a realistic model becomes an uncomfortably moot point. This raises an interesting question: the simple statistical point of view adopted in this paper is an aspect of computational models of language evolution that has up to now been largely ignored. If it is indeed possible to model the emergence of fixed word order solely on the basis of distributional properties of language use, we need to reflect on the exact nature of the governing role of semantics in this type of modeling and start situating advances in the field against the background of a quantifiable baseline. This paper has attempted to offer an example of such a baseline and hopes to instigate a more general discussion on the signal-meaning relationship in the research field of computational modeling of language evolution.

Acknowledgments

The research described in this paper was financed by the FWO (Fund for Scientific Research).

References

Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (p. 111-172). Cambridge University Press.

De Pauw, G. (2002). *An agent-based evolutionary computing approach to memory-based syntactic parsing of natural language.* Doctoral dissertation, University of Antwerp, Antwerp, Belgium.

Dunbar, R. (1996). *Grooming, gossip and the evolution of language.* London: Faber.

Hashimoto, T., & Ikegami, T. (1996). The emergence of a net-grammar in communicating agents. *Biosystems*, *38*, 1-14.

Juliano, C., & Tanenhaus, M. (1993). Contingent frequency effects in syntactic ambiguity resolution. In *Fifteenth annual conference of the cognitive science society* (p. 593-598). Hilsdale, NJ.

Kirby, S., & Hurford, J. (2001). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the evolution of language* (p. 121-148). London: Springer Verlag.

Quine, W. V. O. (1960). *Word and object.* Cambridge: MIT Press.

Steels, L. (2004). Constructivist development of grounded construction grammars. In *Proceedings annual meeting association for computational linguistics conference* (p. 9-16). Barcelona.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, *18*, 47-67.