

# THE EMERGENCE OF SEMANTIC ROLES IN FLUID CONSTRUCTION GRAMMAR

REMI VAN TRIJP

*Sony Computer Science Laboratory Paris, Rue Amyot 6,  
Paris, 75005, France  
remi@csl.sony.fr*

This paper shows how experiments on artificial language evolution can provide highly relevant results for important debates in linguistic theories. It reports on a series of experiments that investigate how semantic roles can emerge in a population of artificial embodied agents and how these agents can build a network of constructions. The experiment also includes a fully operational implementation of how event-specific participant-roles can be fused with the semantic roles of argument-structure constructions and thus contributes to the linguistic debate on how the syntax-semantics interface is organized.

## 1. Introduction

Most linguists agree that there is a strong connection between the semantic representation of a verb and the sentence types in which the verb can occur. Unfortunately, the exact nature of the syntax-semantics interface is still a largely unresolved issue. One approach is the lexicalist account (e.g. Pinker (1989)) in which it is assumed that there exists a list of universal and innate ‘semantic roles’ (also called ‘thematic’ or ‘theta’ roles). In the lexicon it is then specified how many arguments a particular verb takes and which semantic roles they play. For example, the verb *push* (as in *Jack pushes a block*) is listed as a two-place predicate which assigns the roles ‘agent’ and ‘patient’ to its arguments. These roles are then ‘projected’ onto the syntactic structure of the sentence through a limited (and usually universal) set of linking rules. Differences in syntactic structures are taken as indicators for differences in the semantic role list of a verb.

Recently, however, the lexicalist approach has come under serious criticism. Goldberg (1995, p. 9–14) points to the fact that lexicalists are obliged to posit implausible verb senses in the lexicon. For example, a sentence like *she sneezed the napkin of the table* would count as evidence that the verb *sneeze* is not only an intransitive verb as in *she sneezed*, but that it also has a three-argument sense ‘X causes Y to move to Z’ and that it assigns the roles ‘agent’, ‘patient’ and ‘goal’ to its arguments. The lexicalist approach also fails to explain coherent semantic interpretations in creative language use and coercion effects, for example in A gruff ‘police monk’ *barks them back to work* (Michaelis, 2003, p. 261).

As an alternative, Goldberg (1995) proposes a constructionist account which we will adopt in this paper. Here, a verb's lexical entry contains its verb-specific 'participant-roles' rather than a set of abstract semantic roles. To take *push* as an example again, two participant-roles are listed: the 'pusher' and the 'pushed'. These participant-roles have to be "semantically *fused*" with semantic roles, which Goldberg calls 'argument roles' (p. 50) and which are slots in argument-structure constructions. Constructions are like the linking rules of the lexicalist approach in the sense that they are a mapping between meaning and form, but the difference is that they carry meaning *themselves* and that they add this meaning to the sentence. So instead of positing different senses for the verb to accommodate sentences such as *he pushed a block* and *he pushed him a block*, parts of the meaning are added by the verb and other parts are contributed by the constructions. For example, in *he pushed him a block* the 'recipient'-role is added by the ditransitive construction which maps the meaning 'X causes Y to receive Z' to a syntactic pattern.

In the constructionist account, semantic roles are no longer treated as universal nor as atomic categories. This is supported by empirical evidence from both cross-linguistic studies as from research on individual languages (Croft, 2001). Even for a specific category such as the English dative, the "relation between form and meaning is rather indirect and multi-layered" (Davidse, 1996). Moreover, it is shown that there is a gradient evolution from lexical items to become more grammaticalized (Hopper, 1987), which leads more and more linguists to the conclusion that pre-existing categories don't exist (Haspelmath, 2007).

The constructionist account is more plausible from an empirical point of view, but so far it leaves two questions unanswered: where do semantic roles come from and how exactly does 'fusion' work? This paper addresses both issues through experiments on artificial language evolution. It first proposes a fully operational implementation of the constructionist approach using the computational formalism Fluid Construction Grammar (Steels & De Beule, 2006, FCG). Next, the experiment itself is described. Since the experiment deals with artificial languages, the examples in this paper should not be confused with actual grammar descriptions, but rather as indicators of the minimal requirements for explaining semantic roles.

## 2. Semantic Roles and Fusion in Fluid Construction Grammar

In FCG, a language user's linguistic inventory is organized as a network of rules which is dynamically updated through language use. Figure 1 illustrates the relevant part of a speaker's network for the utterance *Jack pushes a block*. There are three lexical rules on the left for *jack*, *push*, and *block*, which introduce the individual meanings of these words. In a logic-based representation, the complete meaning can be represented as  $\{\exists v, w, x, y, z: \text{jack}(v), \text{block}(w), \text{push}(x), \text{push-1}(x, y), \text{push-2}(x, z)\}$ . Note that the lexical rule for *push* contains two participant-roles and that these are represented as predicates themselves. Instead of the names 'pusher' and 'pushed', the more neutral labels 'push-1' and 'push-2' are used.

The careful reader will have noticed that there is still a problem with the meaning: the variables  $v$  and  $y$  are bound to the same object (jack) so they are coreferential. Similarly, the variables  $w$  and  $z$  are coreferential because they are bound to the same object (block). Expressing coreferentiality between variables introduced by different predicates is one of the most important functions of grammar and languages have developed various strategies for doing so (e.g. word order in English and case marking in German). As is standard practice in computational linguistics, coreferential linking is achieved by making the variables equal (Steels, 2005), which results in the following meaning for the sentence:  $\{\exists v, w, x: \text{jack}(v), \text{block}(w), \text{push}(x), \text{push-1}(x, v), \text{push-2}(x, w)\}$ .

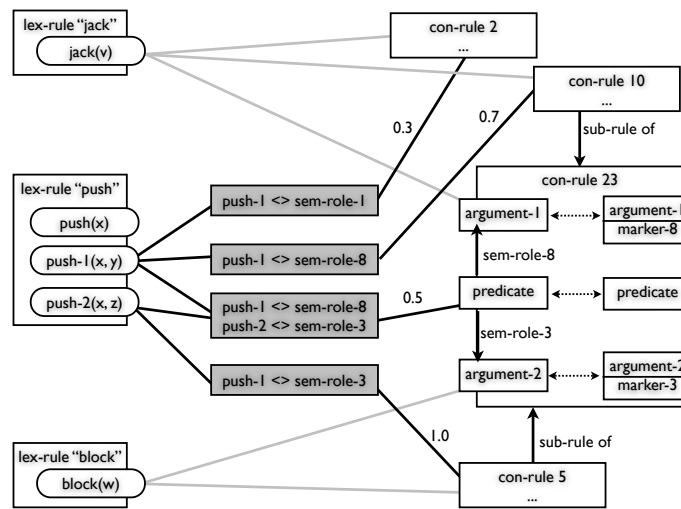


Figure 1. The fusion of an event's participant-roles and a construction's semantic roles is achieved through fusion links which are dynamically updated through language use.

In the FCG implementation, the composition of meanings including the establishment of coreference is taken care of by con-rules which thus implement argument-structure constructions in construction grammar (Goldberg, 1995). The con-rules map a semantic frame (the left pole) to a syntactic pattern (the right pole). The semantic frame contains a set of semantic roles and the syntactic pattern includes simple 'case markers' that immediately follow the arguments of which they indicate the semantic role.<sup>a</sup> An example utterance could be *push jack-BO block-KA* where *BO* indicates that *jack* plays sem-role-8 (which fuses with

<sup>a</sup>The experiment only focuses on the emergence of semantic roles. It therefore assumes a one-to-one mapping of semantic roles to grammatical markers.

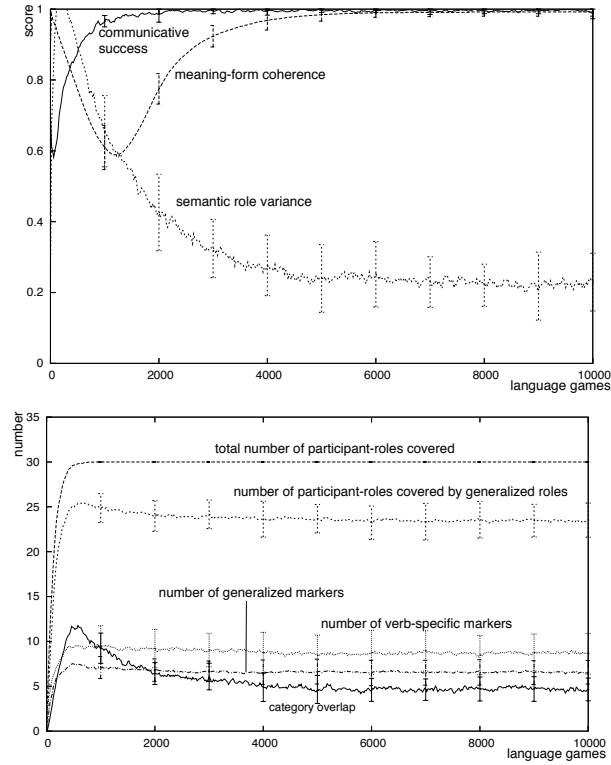


Figure 2. The top graph shows that the agents rapidly reach communicative success and that they converge on a coherent set of semantic roles after 5,500 language games. The semantic role variance reaches almost zero. The bottom graph gives more details on the roles themselves.

‘push-1’) and where *KA* indicates that *block* plays sem-role-3 (which fuses with ‘push-2’). There are also links between con-rule 23 and con-rule 5 and con-rule 10 which means that the latter two are sub-rules of con-rule 23. For convenience’s sake, these sub-rules are only illustrated as nodes in the network.

The fusion of the event-specific participant-roles and the semantic roles of a construction is specified in ‘fusion links’, which are the grey boxes in Figure 1. The fusion links represent all possible fusions known by an agent which can be extended if needed. Each of the links fuses a participant-role with a semantic role within a specific con-rule. This link has a ‘confidence score’ between 0 and 1 which indicates how successful this fusion has been in past communicative acts. For example, ‘push-1’ can be fused with ‘sem-role-8’ in con-rule 10 with a con-

confidence score of 0.7. There is a competing fusion link in which ‘push-1’ is fused with ‘sem-role-1’ in con-rule 2, but this link only has a confidence score of 0.3 so the other one is preferred. Finally, ‘push-1’ can also be fused with ‘sem-role-8’ in con-rule 23, which also contains the semantic role ‘sem-role-3’. In this case, the fusion has a confidence score of 0.5.

This fine-grained scoring mechanism allows speakers of a language to cope with the fuzzy edges of grammatical categories, which is necessary because grammar rules have to be applicable in a flexible manner. A network of rules, as opposed to a limited set of linking rules, is also an elegant way of capturing the complex and multilayered mapping between form and function in language.

### 3. Experiments on the Emergence of Semantic Roles

This paper hypothesizes (a) that the emergence of semantic roles is *triggered* by the need to reduce the cognitive effort of interpretation and to avoid misinterpretation, and (b) that generalizations and grammatical layers are *developed* as a side-effect of reusing existing linguistic structures in new situations. To test these hypotheses, the same experimental set-up was used as Steels and Baillie (2003).

The experiment involves a population of 5 artificial agents which play description games about dynamic real-world scenes. Equipped with a vision system and embodied through a pan-tilt camera, the agents are capable of extracting event descriptions from the scenes. During a game one agent describes an event in the scene to another agent. The game is a success if the hearer agrees with that description. In order to focus exclusively on the emergence of semantic roles, the agents are given a lexicon at the beginning of an experiment but no grammar.

The agents are autonomously capable of detecting when there might be communicative problems through self-monitoring (Steels, 2003). This enables the agent to detect whether variables are coreferential and thus whether there are missing links in the meaning of an utterance (Steels, 2005). If the speaker detects one missing link (but no more), he will try to repair this problem. The hearer’s learning strategy works in the same way, except that he has more uncertainty because he has no access to the speaker’s intended meaning. By comparing the parsed utterance to his world model, however, the hearer may exploit the situatedness of the communicative act to solve the missing link problem as well. Repairing a missing link can be done by classification or by combination.

Repair by classification occurs when the missing link involves a participant-role which the speaker encounters for the first time (e.g. push-1) which we will call the target-role. The agent will first check whether he already knows a semantic role for an analogous participant-role (source-role) that might be reused. Analogy works by (1) taking the event of the target-role and the event that was used to construct the source-role, (2) decomposing them into their event structures, and then (3) constructing a mapping between the two. For example, a ‘walk-to’-event can be decomposed into an event structure that starts with two

non-moving participants and then one participant approaching the other. Event structures themselves are represented as a series of micro-events. The algorithm takes all the participant-roles of the micro-events in which the target-role occurs and maps them onto the corresponding participant-roles in the source event structure. An analogous mapping is defined as when the filler of those corresponding roles is always the same. In case of multiple analogies, the source role which covers the most specific participant-roles is chosen. The source role will then be generalized so that it also covers the target-role. If no analogy could be found, the agent will create a new con-rule which maps the target-role to a newly invented marker. In both cases, fusion links are created and updated for later usage.

Repair by combining existing rules occurs when the speaker wants to express a two- or three-place predicate and already has separate rules that link some of the coreferential variables, but not all of them. The agent will then try to combine these existing rules into a new con-rule. New fusion links are created and family links (sub- and super-rules) are kept between the new con-rule and the rules that were used for creating it. In this way, a network of rules as seen in Figure 1 gradually emerges which improves linguistic processing.

Given the population dynamics of the experiment, several semantic roles may be created and generalized in local language games and then start to propagate among the agents. This automatically creates conflicting solutions, however, so the roles start competing with each other for survival and for covering as much participant-roles as possible. Language thus becomes a complex adaptive system in its own right, very much like a complex ecosystem. There are two types of selectionist forces at work: functional (i.e. some roles are more analogous and therefore better suited for covering a participant-role) and frequency-based.

To be able to align their grammars with each other, agents consolidate their linguistic inventory after each game by updating the scores of the fusion links. Since each construction has its own place in the grammar, fusion links are needed for each specific construction (see Figure 1). However, there is a danger of lingering incoherence if the scores of the fusion links are updated independently of each other. For example, the fusion link between ‘push-1’ and ‘sem-role-1’ may win the competition for single-argument utterances whereas the fusion with ‘sem-role-8’ may win for two-argument utterances. This is incompatible with observations in natural languages which develop a coherent system for argument-structure constructions. In order to solve this problem, the agents apply a consolidation strategy of multi-level selection. Instead of updating only the fusion links that were actually used during processing, all the compatible fusion links are updated as well. Compatible fusion links are links that are related to sub- or super-rules of the applied con-rule. These scores are increased if the game was a success while all the competing links are decreased by lateral inhibition. The scores are lowered if the game was a failure. The exact algorithm and experiments on multi-level selection are reported in more detail in Steels, van Trijp, and Wellens (2007).

#### 4. Results and Discussion

The results show that the agents succeed in developing a coherent system of semantic roles. The top graph in Figure 2 shows that the agents rapidly reach communicative success and that they learn all the case markers after 2,000 language games. It takes them another 3,500 games before they reach total meaning-form coherence. Meaning-form coherence is measured by taking the most frequent form to cover a participant-role and divide this by the total number of forms circulating in the population. Inversely, the semantic role variance – which measures the distance between the semantic role sets of the agents – reaches almost zero which means that the agents have aligned their semantic roles.

The bottom graph of Figure 2 gives more details about the roles themselves. The semantic role overlap indicates that there is still competition going on for 5 participant-roles. The graph also shows that there are 9 verb-specific markers whereas 7 have already become more generalized. These 7 markers cover 24 of the 30 participant-roles in the experiment. Figure 3 gives a snapshot of the evolution of case markers in one agent. It shows that there is a gradual continuum between more lexical, verb-specific markers and more grammaticalized markers which cover up to 8 participant-roles. Similar observations have been made in natural languages by grammaticalization studies (Hopper, 1987).

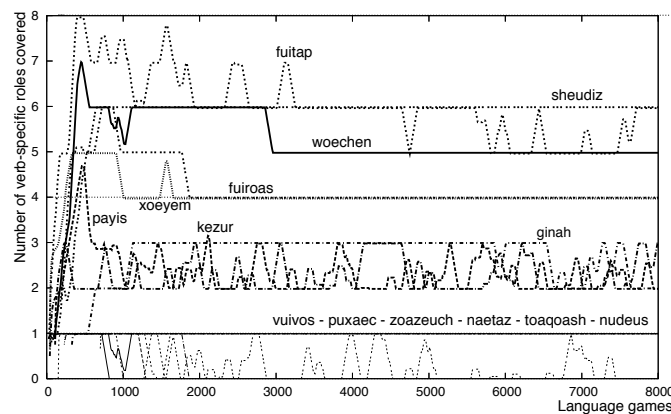


Figure 3. The evolution of case markers in one agent. For example “fuitap” covers 8 specific roles after 600 games, but is in conflict with other markers and in the end covers 6 roles. The graph shows the continuum between more specific and more generalized semantic roles.

#### 5. Conclusion

This paper showed that experiments on artificial language evolution can be highly relevant for linguistic theories. It proposed a fully operational implementation of

the constructionist account to predicate-argument structure in Fluid Construction Grammar. By embedding this approach in experiments with embodied artificial agents, a coherent explanation was presented on the emergence of semantic roles. The results of the experiments showed that semantic roles can emerge as a way to avoid misinterpretation and to reduce the cognitive effort needed during parsing, and that they are further grammaticalized by reuse through analogy.

**Acknowledgement** This research was funded by the EU FET-ECAgents Project 1940. The FCG formalism is freely available at [www.emergent-languages.org](http://www.emergent-languages.org). I am greatly indebted to Luc Steels (who implemented the first case experiment in 2001), director of the Sony Computer Science Laboratory Paris and the Artificial Intelligence Laboratory at the Vrije Universiteit Brussel, the members of both labs, and Walter Daelemans, director of the CNTS at the University of Antwerp.

## References

- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford UP.
- Davidse, K. (1996). Functional dimensions of the dative in english. In W. Van Belle & W. Van Langendonck (Eds.), *The dative. volume 1: Descriptive studies* (pp. 289 – 338). Amsterdam: John Benjamins.
- Goldberg, A. E. (1995). *A construction grammar approach to argument structure*. Chicago: Chicago UP.
- Haspelmath, M. (2007). Pre-established categories don't exist. *Linguistic Typology*, 11(1), 119–132.
- Hopper, P. (1987). Emergent grammar. *BLC*, 13, 139–157.
- Michaelis, L. A. (2003). Headless constructions and coercion by construction. In E. Francis & L. Michaelis (Eds.), *Mismatch: Form-function incongruity and the architecture of grammar* (pp. 259–310). Stanford: CSLI Publications.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge: Cambridge UP.
- Steels, L. (2003). Language re-entrance and the 'inner voice'. *Journal of Consciousness Studies*, 10(4-5), 173–185.
- Steels, L. (2005). What triggers the emergence of grammar? In *Aisb'05: Proceedings of eelc'05* (pp. 143–150). Hatfield: AISB.
- Steels, L., & Baillie, J.-C. (2003). Shared grounding of event descriptions by autonomous robots. *Robotics and Autonomous Systems*, 43(2-3), 163–173.
- Steels, L., & De Beule, J. (2006). Unify and merge in fluid construction grammar. In P. Vogt, Y. Sugita, E. Tuci, & C. Nehaniv (Eds.), *Symbol grounding and beyond*. (pp. 197–223). Berlin: Springer.
- Steels, L., van Trijp, R., & Wellens, P. (2007). Multi-level selection in the emergence of language systematicity. In F. Almeida e Costa, L. M. Rocha, E. Costa, & I. Harvey (Eds.), *Proceedings of the 9th ecal*. Berlin: Springer.