

A Developmental Model of Syntax Acquisition in the Construction Grammar Framework with Cross-Linguistic Validation in English and Japanese

Peter Ford Dominey

Sequential Cognition and Language Group
Institut des Sciences Cognitives, CNRS
69675 Bron CEDES, France
dominey@isc.cnrs.fr

Toshio Inui

Graduate School of Informatics,
Kyoto University,
Yoshida-honmachi, Sakyo-ku, 606-8501,
Kyoto, Japan
inui@kyoto-u.ac.jp

Abstract

The current research demonstrates a system inspired by cognitive neuroscience and developmental psychology that learns to construct mappings between the grammatical structure of sentences and the structure of their meaning representations. Sentence to meaning mappings are learned and stored as grammatical constructions. These are stored and retrieved from a construction inventory based on the constellation of closed class items uniquely identifying each construction. These learned mappings allow the system to process natural language sentences in order to reconstruct complex internal representations of the meanings these sentences describe. The system demonstrates error free performance and systematic generalization for a rich subset of English constructions that includes complex hierarchical grammatical structure, and generalizes systematically to new sentences of the learned construction categories. Further testing demonstrates (1) the capability to accommodate a significantly extended set of constructions, and (2) extension to Japanese, a free word order language that is structurally quite different from English, thus demonstrating the extensibility of the structure mapping model.

1 Introduction

The nativist perspective on the problem of language acquisition holds that the <sentence, meaning> data to which the child is exposed is highly indeterminate, and underspecifies the mapping to be learned. This “poverty of the stimulus” is a central argument for the existence of a genetically specified universal grammar, such that language acquisition consists of configuring the UG for the appropriate target language (Chomsky 1995). In this framework, once a given parameter is set, its use should apply to new

constructions in a generalized, generative manner.

An alternative functionalist perspective holds that learning plays a much more central role in language acquisition. The infant develops an inventory of grammatical constructions as mappings from form to meaning (Goldberg 1995). These constructions are initially rather fixed and specific, and later become generalized into a more abstract compositional form employed by the adult (Tomasello 1999, 2003). In this context, construction of the relation between perceptual and cognitive representations and grammatical form plays a central role in learning language (e.g. Feldman et al. 1990, 1996; Langacker 1991; Mandler 1999; Talmy 1998).

These issues of learnability and innateness have provided a rich motivation for simulation studies that have taken a number of different forms. Elman (1990) demonstrated that recurrent networks are sensitive to predictable structure in grammatical sequences. Subsequent studies of grammar induction demonstrate how syntactic structure can be recovered from sentences (e.g. Stolcke & Omohundro 1994). From the “grounding of language in meaning” perspective (e.g. Feldman et al. 1990, 1996; Langacker 1991; Goldberg 1995) Chang & Maia (2001) exploited the relations between action representation and simple verb frames in a construction grammar approach. In effort to consider more complex grammatical forms, Miikkulainen (1996) demonstrated a system that learned the mapping between relative phrase constructions and multiple event representations, based on the use of a stack for maintaining state information during the processing of the next embedded clause in a recursive manner.

In a more generalized approach, Dominey (2000) exploited the regularity that sentence to meaning mapping is encoded in all languages by word order and grammatical marking (bound or free) (Bates et al. 1982). That model was based on

the functional neurophysiology of cognitive sequence and language processing and an associated neural network model that has been demonstrated to simulate interesting aspects of infant (Dominey & Ramus 2000) and adult language processing (Dominey et al. 2003).

2 Structure mapping for language learning

The mapping of sentence form onto meaning (Goldberg 1995) takes place at two distinct levels in the current model: Words are associated with individual components of event descriptions, and grammatical structure is associated with functional roles within scene events. The first level has been addressed by Siskind (1996), Roy & Pentland (2002) and Steels (2001) and we treat it here in a relatively simple but effective manner. Our principle interest lies more in the second level of mapping between scene and sentence structure.

Equations 1-7 implement the model depicted in Figure 1, and are derived from a neurophysiologically motivated model of sensorimotor sequence learning (Dominey et al. 2003).

2.1 Word Meaning

Equation (1) describes the associative memory, WordToReferent, that links word vectors in the OpenClassArray (OCA) with their referent vectors in the SceneEventArray (SEA)¹. In the initial learning phases there is no influence of syntactic knowledge and the word-referent associations are stored in the WordToReferent matrix (Eqn 1) by associating every word with every referent in the current scene ($\alpha = 1$), exploiting the cross-situational regularity (Siskind 1996) that a given word will have a higher coincidence with referent to which it refers than with other referents. This initial word learning contributes to learning the mapping between sentence and scene structure (Eqn. 4, 5 & 6 below). Then, knowledge of the syntactic structure, encoded in SentenceToScene can be used to identify the appropriate referent (in the SEA) for a given word (in the OCA), corresponding to a zero value of α in Eqn. 1. In this “syntactic bootstrapping” for the new word “gugle,” for example, syntactic knowledge of Agent-Event-Object structure of the sentence “John pushed the gugle” can be used to assign

¹ In Eqn 1, the index $k = 1$ to 6, corresponding to the maximum number of words in the open class array (OCA). Index $m = 1$ to 6, corresponding to the maximum number of elements in the scene event array (SEA). Indices i and $j = 1$ to 25, corresponding to the word and scene item vector sizes, respectively.

“gugle” to the object of push.

$$\begin{aligned} \text{WordToReferent}(i,j) &= \text{WordToReferent}(i,j) + \\ &\text{OCA}(k,i) * \text{SEA}(m,j) * \\ &\max(\alpha, \text{SentenceToScene}(m,k)) \end{aligned} \quad (1)$$

2.2 Open vs Closed Class Word Categories

Our approach is based on the cross-linguistic observation that open class words (e.g. nouns, verbs, adjectives and adverbs) are assigned to their thematic roles based on word order and/or grammatical function words or morphemes (Bates et al. 1982). Newborn infants are sensitive to the perceptual properties that distinguish these two categories (Shi et al. 1999), and in adults, these categories are processed by dissociable neurophysiological systems (Brown et al. 1999). Similarly, artificial neural networks can also learn to make this function/content distinction (Morgan et al. 1996). Thus, for the speech input that is provided to the learning model open and closed class words are directed to separate processing streams that preserve their order and identity, as indicated in Figure 2.

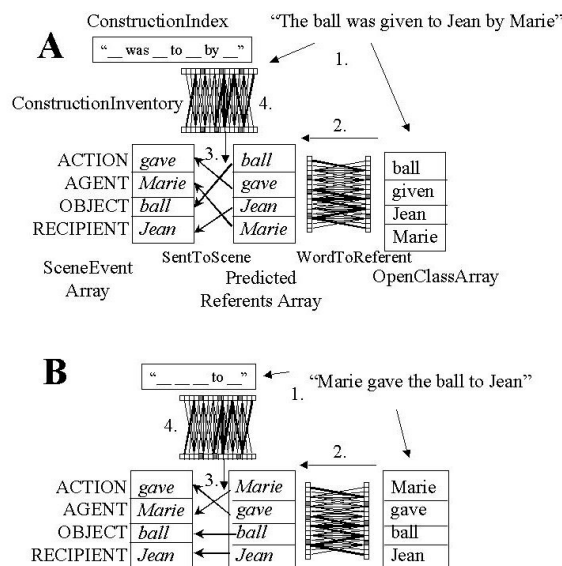


Figure 1. Structure-Mapping Architecture. 1. Lexical categorization. 2. Open class words in Open Class Array are translated to Predicted Referents in the PRA via the WordtoReferent mapping. 3. PRA elements are mapped onto their roles in the SceneEventArray by the SentenceToScene mapping, specific to each sentence type. 4. This mapping is retrieved from Construction Inventory, via the ConstructionIndex that encodes the closed class words that characterize each grammatical construction type.

2.3 Mapping Sentence to Meaning

Meanings are encoded in an event predicate, argument representation corresponding to the SceneEventArray in Figure 1 (e.g. push(Block, triangle) for “The triangle pushed the block”). There, the sentence to meaning mapping can be

characterized in the following successive steps. First, words in the Open Class Array are decoded into their corresponding scene referents (via the WordToReferent mapping) to yield the Predicted Referents Array that contains the translated words while preserving their original order from the OCA (Eqn 2) ².

$$PRA(k,j) = \sum_{i=1}^n OCA(k,i) * WordToReferent(i,j) \quad (2)$$

Next, each sentence type will correspond to a specific *form to meaning* mapping between the PRA and the SEA. encoded in the SentenceToScene array. The problem will be to retrieve for each sentence type, the appropriate corresponding SentenceToScene mapping. To solve this problem, we recall that each sentence type will have a unique constellation of closed class words and/or bound morphemes (Bates et al. 1982) that can be coded in a ConstructionIndex (Eqn.3) that forms a unique identifier for each sentence type.

The ConstructionIndex is a 25 element vector. Each function word is encoded as a single bit in a 25 element FunctionWord vector. When a function word is encountered during sentence processing, the current contents of ConstructionIndex are shifted (with wrap-around) by n + m bits where n corresponds to the bit that is on in the FunctionWord, and m corresponds to the number of open class words that have been encountered since the previous function word (or the beginning of the sentence). Finally, a vector addition is performed on this result and the FunctionWord vector. Thus, the appropriate SentenceToScene mapping for each sentence type can be indexed in ConstructionInventory by its corresponding ConstructionIndex.

$$ConstructionIndex = f_{\text{circularShift}}(ConstructionIndex, FunctionWord) \quad (3)$$

The link between the ConstructionIndex and the corresponding SentenceToScene mapping is established as follows. As each new sentence is processed, we first reconstruct the specific SentenceToScene mapping for that sentence (Eqn 4)³, by mapping words to referents (in PRA) and

² Index k = 1 to 6, corresponding to the maximum number of scene items in the predicted referents array (PRA). Indices i and j = 1 to 25, corresponding to the word and scene item vector sizes, respectively.

³ Index m = 1 to 6, corresponding to the maximum number of elements in the scene event array (SEA). Index k = 1 to 6, corresponding to the maximum number of words in the predicted

referents to scene elements (in SEA). The resulting, SentenceToSceneCurrent encodes the correspondence between word order (that is preserved in the PRA Eqn 2) and thematic roles in the SEA. Note that the quality of SentenceToSceneCurrent will depend on the quality of acquired word meanings in WordToReferent. Thus, syntactic learning requires a minimum baseline of semantic knowledge.

$$SentenceToSceneCurrent(m,k) = \sum_{i=1}^n PRA(k,i) * SEA(m,i) \quad (4)$$

Given the SentenceToSceneCurrent mapping for the current sentence, we can now associate it in the ConstructionInventory with the corresponding function word configuration or ConstructionIndex for that sentence, expressed in (Eqn 5)⁴.

$$ConstructionInventory(i,j) = ConstructionInventory(i,j) + ConstructionIndex(i) * SentenceToScene-Current(j) \quad (5)$$

Finally, once this learning has occurred, for new sentences we can now extract the SentenceToScene mapping from the learned ConstructionInventory by using the ConstructionIndex as an index into this associative memory, illustrated in Eqn. 6⁵.

$$SentenceToScene(i) = \sum_{j=1}^n ConstructionInventory(i,j) * ConstructinIndex(j) \quad (6)$$

To accommodate the dual scenes for complex events Eqns. 4-7 are instantiated twice each, to represent the two components of the dual scene. In the case of simple scenes, the second component of the dual scene representation is null.

We evaluate performance by using the WordToReferent and SentenceToScene knowledge to construct for a given input sentence the “predicted scene”. That is, the model will

references array (PRA). Index i = 1 to 25, corresponding to the word and scene item vector sizes.

⁴ Note that we have linearized SentenceToSceneCurrent from 2 to 1 dimensions to make the matrix multiplication more transparent. Thus index j varies from 1 to 36 corresponding to the 6x6 dimensions of SentenceToSceneCurrent.

⁵ Again to simplify the matrix multiplication, SentenceToScene has been linearized to one dimension, based on the original 6x6 matrix. Thus, index i = 1 to 36, and index j = 1 to 25 corresponding to the dimension of the ConstructionIndex.

construct an internal representation of the scene that should correspond to the input sentence. This is achieved by first converting the Open-Class-Array into its corresponding scene items in the Predicted-Referents-Array as specified in Eqn. 2. The referents are then re-ordered into the proper scene representation via application of the SentenceToScene transformation as described in Eqn. 7⁶.

$$PSA(m,i) = PRA(k,i) * SentenceToScene(m,k) \quad (7)$$

When learning has proceeded correctly, the predicted scene array (PSA) contents should match those of the scene event array (SEA) that is directly derived from input to the model. We then quantify performance error in terms of the number of mismatches between PSA and SEA.

3 Learning Experiments

Three sets of results will be presented. First the demonstration of the model sentence to meaning mapping for a reduced set of constructions is presented as a proof of concept. This will be followed by a test of generalization to a new extended set of grammatical constructions. Finally, in order to validate the cross-linguistic validity of the underlying principals, the model is tested with Japanese, a free word-order language that is qualitatively quite distinct from English.

3.1 Proof of Concept with Two Constructions

3.1.1 Initial Learning of Active Forms for Simple Event Meanings

The first experiment examined learning with sentence, meaning pairs with sentences only in the active voice, corresponding to the grammatical forms 1 and 2.

1. Active: The block pushed the triangle.
2. Dative: The block gave the triangle to the moon.

For this experiment, the model was trained on 544 <sentence, meaning> pairs. Again, meaning is coded in a predicate-argument format, e.g. push(block, triangle) for sentence 1. During the first 200 trials (scene/sentence pairs), value α in Eqn. 1 was 1 and thereafter it was 0. This was necessary in order to avoid the effect of erroneous

(random) syntactic knowledge on semantic learning in the initial learning stages. Evaluation of the performance of the model after this training indicated that for all sentences, there was error-free performance. That is, the PredictedScene generated from each sentence corresponded to the actual scene paired with that sentence. An important test of language learning is the ability to generalize to new sentences that have not previously been tested. Generalization in this form also yielded error free performance. In this experiment, only 2 grammatical constructions were learned, and the lexical mapping of words to their scene referents was learned. Word meaning provides the basis for extracting more complex syntactic structure. Thus, these word meanings are fixed and used for the subsequent experiments.

3.1.2 Passive forms

The second experiment examined learning with the introduction of passive grammatical forms, thus employing grammatical forms 1-4.

3. Passive: The triangle was pushed by the block.
4. Dative Passive: The moon was given to the triangle by the block.

A new set of <sentence, scene> pairs was generated that employed grammatical constructions, with two- and three- arguments, and active and passive grammatical forms for the narration. Word meanings learned in Experiment 1 were used, so only the structural mapping from grammatical to scene structure was learned. With exposure to less than 100 <sentence, scene>, error free performance was achieved. Note that only the WordToReferent mappings were retained from Experiment 1. Thus, the 4 grammatical forms were learned from the initial naive state. This means that the ConstructionIndex and ConstructionInventory mechanism correctly discriminates and learns the mappings for the different grammatical constructions. In the generalization test, the learned values were fixed, and the model demonstrated error-free performance on new sentences for all four grammatical forms that had not been used during the training.

3.1.3 Relative forms for Complex Events

The complexity of the scenes/meanings and corresponding grammatical forms in the previous experiments were quite limited. Here we consider complex <sentence, scene> mappings that involve relativised sentences and dual event scenes. A

⁶ In Eqn 7, index $i = 1$ to 25 corresponding to the size of the scene and word vectors. Indices m and $k = 1$ to 6, corresponding to the dimension of the predicted scene array, and the predicted references array, respectively.

small corpus of complex <sentence, scene> pairs were generated corresponding to the grammatical construction types 5-10

5. The block that pushed the triangle touched the moon.
6. The block pushed the triangle that touched the moon.
7. The block that pushed the triangle was touched by the moon.
8. The block pushed the triangle that was touched the moon.
9. The block that was pushed by the triangle touched the moon.
10. The block was pushed by the triangle that touched the moon.

After exposure to less than 100 sentences generated from these relativised constructions, the model performed without error for these 6 construction types. In the generalization test, the learned values were fixed, and the model demonstrated error-free performance on new sentences for all six grammatical forms that had not been used during the training.

3.1.4 Combined Test

The objective of the final experiment was to verify that the model was capable of learning the 10 grammatical forms together in a single learning session. Training material from the previous experiments were employed that exercised the ensemble of 10 grammatical forms. After exposure to less than 150 <sentence, scene> pairs, the model performed without error. Likewise, in the generalization test the learned values were fixed, and the model demonstrated error-free performance on new sentences for all ten grammatical forms that had not been used during the training.

This set of experiments in ideal conditions demonstrates a proof of concept for the system, though several open questions can be posed based on these results. First, while the demonstration with 10 grammatical constructions is interesting, we can ask if the model will generalize to an extended set of constructions. Second, we know that the English language is quite restricted with respect to its word order, and thus we can ask whether the theoretical framework of the model will generalize to free word order languages such as Japanese. These questions are addressed in the following three sections.

3.2 Generalization to Extended Construction Set

As illustrated above the model can accommodate 10 distinct form-meaning mappings or grammatical constructions, including constructions involving "dual" events in the meaning representation that correspond to relative clauses. Still, this is a relatively limited size for the construction inventory. The current experiment demonstrates how the model generalizes to a number of new and different relative phrases, as well as additional sentence types including: conjoined (John took the key and opened the door), reflexive (The boy said that the dog was chased by the cat), and reflexive pronoun (The block said that it pushed the cylinder) sentence types, for a total of 38 distinct abstract grammatical constructions. The consideration of these sentence types requires us to address how their meanings are represented. Conjoined sentences are represented by the two corresponding events, e.g. *took(John, key)*, *open(John, door)* for the conjoined example above. Reflexives are represented, for example, as *said(boy)*, *chased(cat, dog)*. This assumes indeed, for reflexive verbs (e.g. said, saw), that the meaning representation includes the second event as an argument to the first. Finally, for the reflexive pronoun types, in the meaning representation the pronoun's referent is explicit, as in *said(block)*, *push(block, cylinder)* for "The block said that it pushed the cylinder."

For this testing, the ConstructionInventory is implemented as a lookup table in which the ConstructionIndex is paired with the corresponding SentenceToScene mapping during a single learning trial. Based on the tenets of the construction grammar framework (Goldberg 1995), if a sentence is encountered that has a form (i.e. ConstructionIndex) that does not have a corresponding entry in the ConstructionInventory, then a new construction is defined. Thus, one exposure to a sentence of a new construction type allows the model to generalize to any new sentence of that type. In this sense, developing the capacity to handle a simple initial set of constructions leads to a highly extensible system. Using the training procedures as described above, with a pre-learned lexicon (WordToReferent), the model successfully learned all of the constructions, and demonstrated generalization to new sentences that it was not trained on.

That the model can accommodate these 38 different grammatical constructions with no modifications indicates its capability to generalize. This translates to a (partial) validation of the hypothesis that across languages, thematic role assignment is encoded by a limited set of

parameters including word order and grammatical marking, and that distinct grammatical constructions will have distinct and identifying ensembles of these parameters. However, these results have been obtained with English that is a relatively fixed word-order language, and a more rigorous test of this hypothesis would involve testing with a free word-order language such as Japanese.

3.3 Generalization to Japanese

The current experiment will test the model with sentences in Japanese. Unlike English, Japanese allows extensive liberty in the ordering of words, with grammatical roles explicitly marked by postpositional function words *-ga*, *-ni*, *-wo*, *-yotte*. This word-order flexibility of Japanese with respect to English is illustrated here with the English active and passive di-transitive forms that each can be expressed in 4 different common manners in Japanese:

1. The block gave the circle to the triangle.
 - 1.1 Block-ga triangle-ni circle-wo watashita .
 - 1.2 Block-ga circle-wo triangle-ni watashita .
 - 1.3 Triangle-ni block-ga circle-wo watashita .
 - 1.4 Circle-wo block-ga triangle-ni watashita .
2. The circle was given to the triangle by the block.
 - 2.1 Circle-ga block-ni-yotte triangle-ni watasareta.
 - 2.2 Block-ni-yotte circle-ga triangle-ni watasareta .
 - 2.3 Block-ni-yotte triangle-ni circle-ga watasareta .
 - 2.4 Triangle-ni circle-ga block-ni-yotte watasareta .

In the “active” Japanese sentences, the postpositional function words *-ga*, *-ni* and *-wo* explicitly mark agent, recipient and, object whereas in the passive, these are marked respectively by *-ni-yotte*, *-ga*, and *-ni*. For both the active and passive forms, there are four different legal word-order permutations that preserve and rely on this marking. Japanese thus provides an interesting test of the model’s ability to accommodate such freedom in word order.

Employing the same method as described in the previous experiment, we thus expose the model to <sentence, meaning> pairs generated from 26 Japanese constructions that employ the equivalent of active, passive, relative forms and their permutations. We predicted that by processing the *-ga*, *-ni*, *-yotte* and *-wo* markers as closed class elements, the model would be able to discriminate and identify the distinct grammatical constructions and learn the corresponding mappings. Indeed, the model successfully discriminates between all of the construction types based on the ConstructionIndex

unique to each construction type, and associates the correct SentenceToScene mapping with each of them. As for the English constructions, once learned, a given construction could generalize to new untrained sentences.

This demonstration with Japanese is an important validation that at least for this subset of constructions, the construction-based model is applicable both to fixed word order languages such as English, as well as free word order languages such as Japanese. This also provides further validation for the proposal of Bates and MacWhinney (et al. 1982) that thematic roles are indicated by a constellation of cues including grammatical markers and word order.

3.4 Effects of Noise

The model relies on lexical categorization of open vs. closed class words both for learning lexical semantics, and for building the ConstructionIndex for phrasal semantics. While we can cite strong evidence that this capability is expressed early in development (Shi et al. 1999) it is still likely that there will be errors in lexical categorization. The performance of the model for learning lexical and phrasal semantics for active transitive and ditransitive structures is thus examined under different conditions of lexical categorization errors. A lexical categorization error consists of a given word being assigned to the wrong category and processed as such (e.g. an open class word being processed as a closed class word, or vice-versa). Figure 2 illustrates the performance of the model with random errors of this type introduced at levels of 0 to 20 percent errors.

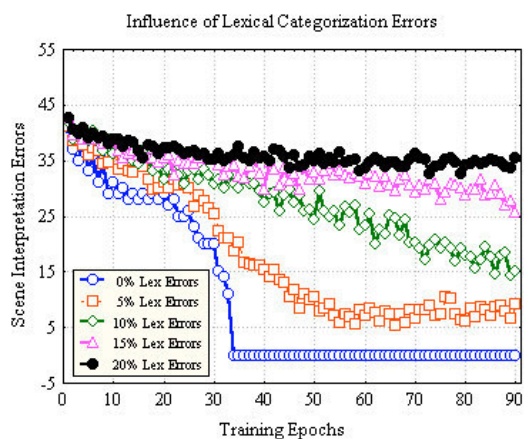


Figure 2. The effects of Lexical Categorization Errors (mis-categorization of an open-class word as a closed-class word or vice-versa) on performance (Scene Interpretation Errors) over Training Epochs. The 0% trace indicates performance in the absence of noise, with a rapid elimination of errors. The successive introduction of categorization errors yields a corresponding progressive impairment in learning. While sensitive to the errors, the system demonstrates a desired graceful degradation.

We can observe that there is a graceful degradation, with interpretation errors progressively increasing as categorization errors rise to 20 percent. In order to further assess the learning that was able to occur in the presence of noise, after training with noise, we then tested performance on noise-free input. The interpretation error values in these conditions were 0.0, 0.4, 2.3, 20.7 and 33.6 out of a maximum of 44 for training with 0, 5, 10, 15 and 20 percent lexical categorization errors, respectively. This indicates that up to 10 percent input lexical categorization errors allows almost error free learning. At 15 percent input errors the model has still significantly improved with respect to the random behavior (~45 interpretation errors per epoch). Other than reducing the lexical and phrasal learning rates, no efforts were made to optimize the performance for these degraded conditions, thus there remains a certain degree of freedom for improvement. The main point is that the model does not demonstrate a catastrophic failure in the presence of lexical categorization errors.

4 Discussion

The research demonstrates an implementation of a model of sentence-to-meaning mapping in the developmental and neuropsychologically inspired construction grammar framework. The strength of the model is that with relatively simple “innate” learning mechanisms, it can acquire a variety of grammatical constructions in English and Japanese based on exposure to <sentence, meaning> pairs, with only the lexical categories of open vs. closed class being prespecified. This lexical categorization can be provided by frequency analysis, and/or acoustic properties specific to the two classes (Blanc et al. 2003; Shi et al. 1999). The model learns grammatical constructions, and generalizes in a systematic manner to new sentences within the class of learned constructions. This demonstrates the cross-linguistic validity of our implementation of the construction grammar approach (Goldberg 1995, Tomasello 2003) and of the “cue competition” model for coding of grammatical structure (Bates et al. 1982). The point of the Japanese study was to demonstrate this cross-linguistic validity – i.e. that nothing extra was needed, just the identification of constructions based on lexical category information. Of course a better model for Japanese and Hungarian etc. that exploits the explicit marking of grammatical roles of NPs would have been interesting – but it wouldn’t have worked for English!

The obvious weakness is that it does not go further. That is, it cannot accommodate new construction types without first being exposed to a training example of a well formed <sentence, meaning> pair. Interestingly, however, this appears to reflect a characteristic stage of human development, in which the infant relies on the use of constructions that she has previously heard (see Tomasello 2003). Further on in development, however, as pattern finding mechanisms operate on statistically relevant samples of this data, the child begins to recognize structural patterns, corresponding for example to noun phrases (rather than solitary nouns) in relative clauses. When this is achieved, these phrasal units can then be inserted into existing constructions, thus providing the basis for “on the fly” processing of novel relativised constructions. This suggests how the abstract construction model can be extended to a more generalized compositional capability. We are currently addressing this issue in an extension of the proposed model, in which recognition of linguistic markers (e.g. “that”, and directly successive NPs) are learned to signal embedded relative phrases (see Miikkulainen 1996).

Future work will address the impact of ambiguous input. The classical example “John saw the girl with the telescope” implies that a given grammatical form can map onto multiple meaning structures. In order to avoid this violation of the one to one mapping, we must concede that form is influenced by context. Thus, the model will fail in the same way that humans do, and should be able to succeed in the same way that humans do. That is, when context is available to disambiguate then ambiguity can be resolved. This will require maintenance of the recent discourse context, and the influence of this on grammatical construction selection to reduce ambiguity.

5 Acknowledgements

This work was supported by the ACI Computational Neuroscience Project, The Eurocores OMLL project and the HFSP Organization.

References

- Bates E, McNew S, MacWhinney B, Devescovi A, Smith S (1982) Functional constraints on sentence processing: A cross linguistic study, *Cognition* (11) 245-299.
- Blanc JM, Dodane C, Dominey P (2003) Temporal processing for syntax acquisition. Proc. 25th Ann. Mtg. Cog. Science Soc. Boston
- Brown CM, Hagoort P, ter Keurs M (1999) Electrophysiological signatures of visual lexical

- processing: Open- and closed-class words. *Journal of Cognitive Neuroscience*. 11 :3, 261-281
- Chang NC, Maia TV (2001) Grounded learning of grammatical constructions, *AAAI Spring Symp. On Learning Grounded Representations*, Stanford CA.
- Chomsky N. (1995) The Minimalist Program. MIT
- Crangle C. & Suppes P. (1994) Language and Learning for Robots, CSLI lecture notes: no. 41, Stanford.
- Dominey PF, Ramus F (2000) Neural network processing of natural language: I. Sensitivity to serial, temporal and abstract structure of language in the infant. *Lang. and Cognitive Processes*, 15(1) 87-127
- Dominey PF (2000) Conceptual Grounding in Simulation Studies of Language Acquisition, *Evolution of Communication*, 4(1), 57-85.
- Dominey PF, Hoen M, Lelekov T, Blanc JM (2003) Neurological basis of language and sequential cognition: Evidence from simulation, aphasia and ERP studies, *Brain and Language*, 86, 207-225
- Elman J (1990) Finding structure in time. *Cognitive Science*, 14:179-211.
- Feldman JA, Lakoff G, Stolcke A, Weber SH (1990) Miniature language acquisition: A touchstone for cognitive science. In *Proceedings of the 12th Ann Conf. Cog. Sci. Soc.* 686-693, MIT, Cambridge MA
- Feldman J., G. Lakoff, D. Bailey, S. Narayanan, T. Regier, A. Stolcke (1996). L0: The First Five Years. *Artificial Intelligence Review*, v10 103-129.
- Goldberg A (1995) *Constructions*. U Chicago Press, Chicago and London.
- Hirsh-Pasek K, Golinkof RM (1996) *The origins of grammar: evidence from early language comprehension*. MIT Press, Boston.
- Kotovskiy L, Baillargeon R, The development of calibration-based reasoning about collision events in young infants. 1998, *Cognition*, 67, 311-351
- Langacker, R. (1991). *Foundations of Cognitive Grammar. Practical Applications, Volume 2*. Stanford University Press, Stanford.
- Mandler J (1999) Preverbal representations and language, in P. Bloom, MA Peterson, L Nadel and MF Garrett (Eds) *Language and Space*, MIT Press, 365-384
- Miikkulainen R (1996) Subsymbolic case-role analysis of sentences with embedded clauses. *Cognitive Science*, 20:47-73.
- Morgan JL, Shi R, Allopenna P (1996) Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping, pp 263-286, in Morgan JL, Demuth K (Eds) *Signal to syntax*, Lawrence Erlbaum, Mahwah NJ, USA.
- Pollack JB (1990) Recursive distributed representations. *Artificial Intelligence*, 46:77-105.
- Roy D, Pentland A (2002). Learning Words from Sights and Sounds: A Computational Model. *Cognitive Science*, 26(1), 113-146.
- Shi R., Werker J.F., Morgan J.L. (1999) Newborn infants' sensitivity to perceptual cues to lexical and grammatical words, *Cognition*, Volume 72, Issue 2, B11-B21.
- Siskind JM (1996) A computational study of cross-situational techniques for learning word-to-meaning mappings, *Cognition* (61) 39-91.
- Siskind JM (2001) Grounding the lexical semantics of verbs in visual perception using force dynamics and event logic. *Journal of AI Research* (15) 31-90
- Steels, L. (2001) Language Games for Autonomous Robots. *IEEE Intelligent Systems*, vol. 16, nr. 5, pp. 16-22, New York: IEEE Press.
- Stolcke A, Omohundro SM (1994) Inducing probabilistic grammars by Bayesian model merging/ In *Grammatical Inference and Applications: Proc. 2nd Intl. Colloq. On Grammatical Inference*, Springer Verlag.
- Talmy L (1988) Force dynamics in language and cognition. *Cognitive Science*, 10(2) 117-149.
- Tomasello M (1999) The item-based nature of children's early syntactic development, *Trends in Cognitive Science*, 4(4):156-163
- Tomasello, M. (2003) *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press, Cambridge.