Constructions emerging : a usage-based model of the acquisition of grammar

Beekhuizen, B.F.

Author: Beekhuizen, Barend
Title: Constructions emerging : a usage-based model of the acquisition of grammar
Issue Date: 2015-09-22

CHAPTER 8

Concluding remarks

Understanding how children acquire the language of their community within a limited amount of time is a central question in linguistics. The usage-based constructivist approach to language acquisition holds that children do so by using domain-general learning mechanisms such as social cognition and pattern recognizing mechanisms. Computational modeling, that is: simulating a child's behavior by formalizing and implementing important pieces of our favorite hypotheses as software, is becoming an increasingly important method in the field of language acquisition. I hope to have contributed to both the field of language acquisition and computational cognitive modeling with this dissertation by addressing four major points I presented at the outset:

- Achieving greater comprehensiveness of computational cognitive models
- Achieving greater naturalism in the computational modeling of the acquisition of meaning
- A reappreciation of the starting-small hypothesis within the usage-based framework
- A reassessment of proposed learning mechanisms (cognitive) and algorithms (computational).

I believe I have done so with the Syntagmatic-Paradigmatic Learner (SPL), a computational model of the acquisition of linguistic representations that aims to implement various aspects of a usage-based theory of language acquisition. Crucially, SPL starts off with no linguistic-representational content, and learns to comprehend as well as produce utterances. SPL processes utterances in a context of situations (the properties of which were derived from an empirical study presented in chapter 4), and in doing so, gradually builds up a constructicon, an inventory of both lexical and grammatical constructions. The 'learning mechanisms' involved in the learning process are best thought of as mere traces of processing operations, rather than actual hypothesis-testing operations (which is the metaphor, grounded in deductivist thought, that is often used to describe the acquistion of linguistic representations).

8.1 Recapitulating SPL

Let us briefly go over the main properties of SPL once more. The model uses the representational format of the construction, a pairing of signifying elements and a signified conceptualization. Starting with no representations, it tries to parse novel input items, pairings of an utterance and a set of situations to which the utterance possibly refers.

The set of situations was generated by the input generation procedure of Alishahi & Stevenson (2010). I modified this procedure to reflect the actual properties of the situational contexts of linguistic used events, as studied in chapter 4. In that chapter, we found that the levels for noise (the absence of some conceptual target from experience) and uncertainty (the overwhelming presence of conceptual non-targets in experience) typically used in computational modeling studies are low compared to the ones we find in actual caregiver-child interaction. I studied the latter by looking at a corpus of videotaped caregiver-child interaction and annotated the corpus for all conceptual elements reasonably thought to be present in the situation around the speech situation. Another insight following from this study was that chains of events are highly dependent on each other: if the mother engages in an action with a ball, it is very likely that she will engage in another action with the ball afterwards, or perhaps in the same action with another object. Given the tediousness of hand-coding the data, this method did not prove scalable to the demands of a computational model. The study of these properties of interaction 'in the wild', however, did lead to an adaptation of Alishahi & Stevenson's (2010) input generation procedure. In this adapted procedure, we generate pairs of an utterance and the situational context in which the utterance occurs, with the latter consisting of a set of situations, one of which is the target situation, unless the target situation is absent. Notably, the similarity of the situations within the situational context, and between subsequent situational contexts to each other is given by the similarity we found in the caregiver-child interaction. Furthermore, the setting of the parameters for noise and uncertainty was derived from the video data as well.

For every processed input item, the model arrives at an optimal analysis, and does so without engaging in utterance-wide optimization. That is: SPL processes the utterance linearly and while keeping track of only the most likely analysis up to that point. The best analysis constitutes the input for SPL's learning mechanisms. Through a set of learning mechanisms, SPL gradually builds up an inventory of constructions allowing it to comprehend and produce utterances. The learning mechanisms constitute the central innovation of the model in the aim to stay close to the usage-based approach as set out by Langacker (1988). I believe this aim has been fulfilled in the design of the model in several ways. Crucially, all of the learning mechanisms, with perhaps the exception of cross-situational learning, are online mechanisms. That is: they do not constitute post-hoc operations on the constructicon (the inventory of constructions), but rather reflect the traces left by the processing of the input item. These traces are found at several levels.

First, a trace of the most concrete representations of the utterances the processes is left in the representational system of SPL through the use of most-concrete constructions. This operation has the effect that highly concrete representations, if they are reinforced often enough, can become stronger over time. We can interpret this as the formation of category prototypes: the well-reinforced, highly-concrete representations are readily available to the model in analyzing and generating utterances.

Second, the mechanism of reinforcing the most-concrete used constructions, i.e. the most-concrete constructions, allows the model to accrue reinforcement mass for those constructions that are used frequently. The effect of this operation is that abstract constructions may obtain reinforcement if they are used to analyze utterances. Because the model only reinforces the most-concrete used construction, the reinforcement operation rewards patterns that are actually used. The usefulness of a construction is therefore determined by its frequency of use. Notably, this design feature implements Bybee's (2006) notion of type frequency. An abstract construction will typically only be reinforced once for each unique usage event for which it is used in an analysis. If the same usage event is encountered again, it is very likely that the more concrete construction blocks the use of the more abstract one. Routinization through high token frequency follows from the same learning operation: if a construction is used frequently, it is more readily available for subsequent analyses. If this construction happens to be a highly concrete one (i.e., one with many constituents lexically specified) the model will acquire such a construction as a routine.

Third, the model builds up increasingly long constructions through the use of the syntagmatization operation. Syntagmatization is the trace left by the processing of multiple, smaller, constructions for which the model has found no analysis in which they are connected to each other with a grammatical construction. These smaller constructions then form the constituents of a novel, wider, construction. Syntagmatization is the primary means through which SPL builds up grammatical constructions.

Finally, paradigmatization allows the model its potential to generalize to unseen usage events. By taking the joint structure of any two constructions that have been reinforced, the paradigmatization 'extracts' abstractions from more concrete constructions. These abstractions, however, are only extracted in the implementational sense: as no selection over them takes place, they can be considered immanent in the more concrete constructions from which they are abstracted, by simply restating their overlap. However, through the reinforcement of the most-concrete used construction, they can be reinforced themselves, in a way akin to Langacker's (2009) description of how abstractions may obtain unit status without the more concrete patterns doing so. This way, selection of 'good' or 'useful' abstractions takes place, but without any selection mechanism performing a global evaluation of the usefulness of a novel abstraction.

The model gets off the ground by the cross-situational learning mechanism, which compares recent usage events and extracts any reliable overlap as initial lexical constructions. Another way of obtaining lexical constructions is through the bootstrap operation. Bootstrapping is a property of the utterance analysis mechanism that fills a non-phonologically-specified slot of a construction with a substring of the utterance, by assuming that substring is an actual word filling that slot.

Both cross-situational learning and bootstrapping allow for the extraction of chunks: lexical constructions that are larger than a single word in the 'adult' language. These chunks, unlike what many within the usage-based framework assume, are not broken down by the paradigmatization operation. This would require the model to engage in a post-hoc re-analysis of the chunks, which was an operation I wanted to avoid, as it makes learning more than a mere by-product of processing.

8.2 The behavior of SPL

I evaluated SPL's behavior both in a comprehension (chapter 5) and a production (chapter 7) experiment. In the comprehension experiment, I looked at the performance of the model in identifying the correct situation out of all possible situations the utterance could refer to, as well as the coverage of the utterance and the situation with the best analysis. On all three measures, SPL gradually becomes a more competent language user over time. Similarly, for production, SPL was tested by having it generate utterances on the basis of a situation and its constructicon at that point in time. The generated utterances become longer over time, and increasingly capture the linguistic material found in the utterance that would have been produced by the input generation procedure. Interestingly, the model displayed high scores of precision, or correctness, from the outset: whatever it produced was mostly correct. This is in line with the finding that children mainly make errors of omission (leaving out elements present in adults' speech), but few errors of comission (producing linguistic elements an adult would not produce).

Next, I looked at the robustness of the model. Recall that we set the parameters for the similarity of the situations in the situational context, as well as the noise and uncertainty of the situational context on the basis of the empirical study of caregiver-child interaction. We may, however, ask how the model performs given different values for these parameters. I found that if the situations are similar to each other, the model is relatively robust to higher levels of noise and uncertainty (on the measures discussed above). Generating each situation independently of the previous one creates a situational context in which the situations are more dissimilar from each other, and in that condition, noise and uncertainty do affect the model's performance negatively. This suggest that the coherence of the situational contexts in which children have their early linguistic experiences plays an important role in bootstrapping a linguistic system: even if the child misidentifies the precise situation, the erroneously identified situation likely contains many elements that are correct.

It is, however, at a more detailed level that the interesting behavioral patterns can be seen, and especially from the failure of the model to behave as we expect, we learn important things about how the mechanisms work. In the two experimental chapters, I studied several behavioral patterns of the model in qualitative detail, to try to understand why the model behaves in certain ways.

In the production experiments, we observed that the number of expressed arguments grew over time as an effect of an increasing number of syntagmatized and subsequently paradigmatized constructions being acquired. I was not able to simulate the prevalence of subject omissions, but argued that this is likely due to a lack of pragmatics and of a right-edge processing bias, as, for instance, MOSAIC (Freudenthal et al. 2010) incorporates. What I did find was that the omission of early arguments was not only a matter of a small vocabulary: for many aspects of the situation the model had to express, it had a lexical construction available, but it simply did not have a grammatical construction ready to fit the lexical construction in. With this analysis, I provided a usage-based analysis of Berk & Lillo-Martin's (2012) finding that older children who have been deprived of linguistic input but are otherwise normally functioning, go through a two-word stage while having a far more extensive vocabulary than a eighteen-month old. An important caveat here is that the higher frequency of subject omissions over other argument omissions was not predicted by the model. Here, the model is somewhat more remote from reality. I argued that the most likely reason for this phenomenon is the information structure of discourse and the salience of the participants: if subjects typically denote less salient and discourse-given participants, we can expect them to be learned (through comprehension) and produced less frequently. An interesting extension of the current model would be to include a discourse model. This seems a relatively small step, since the current input generation procedure already involves chains of events and utterances, on the basis of which we can change the salience of certain referents and words.

A central question in language acquisition is why children sometimes overgeneralize argument-structure (and other) constructions and how they retreat from this overgeneralization. The overgeneralization of argument structure constructions and the subsequent retreat were modeled in chapter 7. The answer of SPL to these two questions is that it quickly builds up an inventory of abstract, generalizable, grammatical constructions (which it, however, hardly uses in comprehension) that it combines with verbs that cannot occur in these constructions (e.g., *you fall ball*). The presence of an alternative construction pre-empts this overgeneralization after a phase of overgeneralization. I argued that pre-emption works in two ways. First, the more entrenched this alternative construction is, the quicker the model retreats from overgeneralization. Second, we find an entrenchment effect of the 'correct' construction: when the model experiences more cases of *ball fall* with a causative meaning (someone dropping a ball), the constructions underlying such utterances are reinforced more, and because of this, highly general constructions allowing for the overgeneralization become less entrenched. I argued that, rather than describing this as entrenchment per se, we could better regard this effect as 'latent pre-emption', that is: as a pre-emption effect that is not seen in the behavior (the model does not produce *ball fall*, as it is less expressive than *you drop ball*), but that does block the use of a novel, erroneous, combination of an abstract construction and a verb.

8.3 The representations acquired by SPL

One interesting property of computational models is that we can study their representations independently of the model's behavior. I did so in chapter 6. A first finding reported there is that, even though all learning mechanisms are available over time, their use varies over time. For the acquisition of lexical constructions we found that cross-situational learning, the naïve method by means of which the model extracts similarities across linguistic usage events, is only used for the first few hundreds of input items. Afterwards, the model has built up an inventory of semi-open and open grammatical constructions that it can use to bootstrap the meaning of words it has not seen. The paradigmatization operation, secondly, displays interesting 'bursts' of activity over time, meaning that the model does not arrive at abstractions gradually, but encounters exemplars that 'unlock' new subspaces of the design space of linguistic representations.

The abstractions learned by SPL display the interesting property that they are not directly obvious from the behavior of the model in comprehension and production. If we would not have looked under the hood of the model, we might have arrived at the erroneous conclusion that its representational system is very concrete. This is a false line of reasoning: given the usage-based tenet that language users prefer the use of more concrete constructions over more abstract ones (as implemented in the probability model of SPL), we expect the highly concrete constructions to show up most of the time. However, representationally, the model has great potential for making generalizations. In fact, generalizations are found rather early, and the model spends the later iterations mainly by adding more relatively concrete constructions to the abstract ones that pre-empt the latter. This is not strange, given the overgeneralization behavior we observe in both children and SPL: once abstraction is available, the model will use it for expressivity's sake, unless it has something more concrete that is equally expressive.

An interesting feature of the abstractions found in the model is that they clearly reflect the type frequencies of the items occurring in them (cf. Bybee 2006): the transitive construction is strongly reinforced as a non-verb-specific construction, because many verbs occur in it, whereas the caused-motion construction is only seen with two verbs, and hence reinforced in verb-island-like constructions rather than as constructions that abstract over verbs.

Reversing the perspective, we furthermore saw how certain words are more readily learned as independent lexical constructions whereas others are primarily learned as the constituents of grammatical constructions. Notably, words referring to entities ('nouns'), are typically learned as independent entities. For the other kinds of words, there was more variation, both between the words and between simulations. Pronouns are used in a lot of different contexts, hence boosting the likelihood of their independent acquisition, but they are also used frequently *within* particular constructions. What we find for pronouns, as well as for prepositions and verbs displaying similar distributions, is that they are acquired independently in some simulations, but as 'bound' elements of constructions in others. I identified three possible factors that determined a word's independence. First, the more different elements occur in a slot, the more likely it is that the abstraction over them will be used in comprehension and production, and the more likely it is that the filler word will be acquired independently. Second, the frequency of the word in the slot: the higher this value is, the more likely it is that it will not be acquired independently, as it will be reinforced as part of a grammatical construction often. Finally, the word's 'promiscuity' matters: if a word occurs across the slots of many grammatical constructions, it is more likely that it will be acquired independently.

On several aspects of the representations, we found high degrees of 'individual' variation between the simulations: the abstraction of the representations as well as the relative independence of various words varied between simulations. This is interesting, as the various simulations display grossly the same behavior – they perform equally well on the global tasks in comprehension and production. I will return to this issue in section 8.5.

8.4 Desiderata and explananda

In chapter 2, I set out a list of theoretical desiderata and empirical explananda the model has to satisfy. Previous models have made important contributions by focussing on parts of this list and my aim was to bring all insights together. I believe SPL reasonably succeeds in doing so: table 8.1 displays the list and

desideratum/explanandum	(Chang 2008)	(Alishahi & Stevenson 2008)	(Kwiatkowski 2011)	(Beekhuizen & Bod 2014)	(Freudenthal et al. 2010)	(McCauley & Christiansen 2014*a*)	SPL
D1 (explicitness)	+	+	+	+	+	+	+
D2 (comprehensiveness)	◊	−	◊	◊	−	−	+
D3 (simultaneity)	◊	−	+	+	−	−	+
D4 (representational realism)							
D4-1 (qualitative grounding)	+	+	−	+	+	+	+
D4-2 (quantitative grounding)	+	+	+	+	+	+	+
D4-3 (immanence)	+	+	−	+	+	−	+
D5 (processing realism)							
D5-1 (heterogeneous structure building)	−	−	−	−	+	−	+
D5-2 (linear processing)	−	−	−	−	+	+	+
D6 (ontogenetic realism)							
D6-1 (cumulative complexity)	◊	−	−	−	+	+	+
D6-2 (learning-by-processing)	−	+	+	+	+	+	+
D6-3 (parts-to-whole and v.v.)	+	−	−	−	+	−	+
D6-4 (developmental continuity)	+	+	+	+	+	+	
D7 (explanatory insight)	+	+/−	+	+/−	+/−	+/−	+
D3-1 (unification)	−	+	−	−	+	−	+
E1 (decreasing argument omission)	◊	−	−	−	+	−	+
E2 (prevalence of subject omission)	◊	−	−	−	+	−	−
E3 (co-varying complexity)	−	−	−	−	−	−	−
E4 (overgeneralization and retreat)	◊	+	−	−	−	−	+
E5 (mechanisms overgeneralization)	−	−	−	−	−	−	+

Table 8.1: A comparison of SPL to the various learners discussed in section 2.5.

whether or not SPL satisfies each particular desideratum or explanandum.

To the best of my knowledge, SPL constitutes the first usage-based computational model that is able to parse and generate utterances while starting with no representational content (D2 and D3). Furthermore, I believe it most closely instantiates the full set of ideas put forward within the usage based perspective: the representations are both qualitatively and quantitatively grounded in the linguistic usage events through their reinforcement in analyzing the usage event. Any learned abstractions are furthermore immanent: they merely restate commonalities across more concrete constructions. In making the analyses, SPL reasonably satisfies the constraints on the realism of processing. Although this was not the focus of this dissertation, it satisfies the baseline conditions that processing is incremental over the utterance and does not involve the search for an optimal analysis over the full utterance.

Obviously, SPL is not a complete model: no model ever is, which is why we call it a model. Several design features of SPL function as 'stubs' in the model to make it work.[1] These stubs are well grounded in our knowledge of pragmatic reasoning, linguistic processing, and learning theory, but I do see room for improvement over the current formulations: a more gradient application of them, over the discrete 'constraints' that have been formulated for the model, is definitely a locus of such improvement.

On the empirical side, more evaluation of the model to experimental data is needed. My reason to focus on 'naturalistic' comprehension and production is that the natural situation of linguistic interaction forms a baseline: if we cannot explain that, the fact that we do understand behavior in artificial settings is to my mind a worthless one. After all, this is the context in which languages are culturally evolved and where the cognitive mechanisms involved in linguistic behavior are geared towards (whether developmentally or biologically). However, once we understand the naturalistic case (to some extent), going back and forth between the evaluation on naturalistic behavior and experimentally elicited behavior vastly enriches our knowledge of the cognitive mechanisms. I hope to contribute to this evaluation in future work.

Crucially, however, SPL satisfies the developmental desiderata: it obeys to the cognitive law of cumulative complexity by gradually building up more complex representations (both in length and abstraction) from simpler ones. All learning in the model can be seen as the traces left by the processing of the usage event: there are no reorganization operations on the constructicon as a whole, nor does the model 'allow' constructions 'in' or not on the basis of how useful they are: many representations are extracted from the usage events, but only a few get reinforced in subsequent usage events. The issue of parts-to-whole and whole-to-parts learning is interesting. SPL does parts-to-whole learning by means of the syntagmatization operator, but does not break down larger units into its components (e.g., when chunks are acquired). In chapter 6, I argued that this kind of offline blame assignment may be at odds with

[1] I owe this way of regarding aspects of the model to Suzanne Stevenson (p.c.).

the idea that learning is a by-product of processing. It requires the learner to re-analyze her previous experiences in terms of a novel conception. Perhaps this is not impossible, but I believe this aspect of the starting-big conception is in want of some more elaboration. Interestingly, SPL does display kinds of whole-to-parts learning, for instance through the bootstrapping operator, whereby a novel word is learned on the basis of a larger linguistic gestalt. Finally, the learning operators are available to the model throughout time, although, as I discussed earlier, their frequencies vary.

I believe SPL provides a good example of narrowing the gap between a theoretical conception and a computational model (D7). Most aspects of the model are readily interpretable as aspects of the usage-based perspective, as I have argued in chapter 3. SPL furthermore provides some unifying explanations: effects of type frequency, token frequency, overgeneralization and the retreat from overgeneralizatio all emerge simply from the reinforcement procedure of the model by means of which the representational potential changes over time.

Looking at the explananda, finally, we see that SPL meets explananda E1, E4, and E5. I did not discuss explanandum E3 anywhere in this dissertation and have not attempted to model it myself, but I do believe it to be a crucial empirical observation that future studies should address. Explanandum E2 is not met by the model: as I argued in chapter 7 it requires either an implemented notion of discourse salience or a right-edge bias. Perhaps adding either of those to SPL may help satisfy this explanandum.

8.5 Suggestions for the usage-based conception

All of the observations discussed in sections 8.2 and 8.3 are effects found within the computational model. As such, we may easily dismiss them as artefacts of the model. I believe, however, that in many cases this is not the best thing to do. SPL instantiates a rather close implementation of a usage-based conception of language acquisition, and as such constitutes a way of studying the various aspects of a usage-based account in interaction, something not possible in the lab or from the armchair. The interpretation can lead to two kinds of conclusions: either SPL is right about some aspect of the theory, or SPL is wrong, but then a better implementation of a particular cognitive mechanism has to be proposed in order to replace the proposal made in SPL.

A first aspect of the model I would like to draw attention to is the notion of a competence-performance distinction it embodies. Looking at the behavior of the model, it seems that it only has acquired highly concrete constructions. However, these are the constructions it uses most frequently (which is why they are stored at that level of concreteness in the first place), and for rarer events, the model quickly arrives at a high level of abstraction in its representational potential. With the back-and-forths between adherents of the early-abstraction and lexical conservatism perspectives, it is hard to find empirical data that are not contradicted by other data. The point I want to make with SPL, however, is that the usage-based perspective is not at odds with an early-abstraction view. Given the close implementation of an immanent abstraction procedure, SPL quickly arrives at abstraction. Perhaps children do so as well.

Secondly, SPL supports the view that, despite their linguistic behavior being roughly the same, different language users may have different representations from one another. We have seen this in chapter 6 for several phenomena: the number and abstraction of the constructions varies across simulations, and whereas some simulations operate on the basis of pronoun frames like '*you* X *it*', others have independent pronouns. Nevertheless, in all simulations, the model arrives at a very similar performance on the comprehension and production tasks.

In the discussion of the independence of items, a factor was found that, to my knowledge, has not been studied well within the usage-based framework. Earlier in this conclusion, I coined it 'promiscuity': the ease with which a word is used in the slots of various constructions. This may be a factor, besides type frequency of a constructional slot and the token frequency of a word in that slot and it would be interesting to study its effects on processing, both through corpus studies and experimental work.

8.6 Suggestions for cognitive modeling

A first central contribution of this dissertation is the empirical grounding of the situational context in empirical findings on the actual situational contexts in which children experience linguistic usage events. Although not scalable by itself to function as input to a computational model, the method did provide us with valuable insights in the situational contexts in which children acquire language. When studying the acquisition of meaningful units, I believe, one cannot simply make up reasonable estimations of the uncertainty and noise present in the situation, rather, an empirical grounding of these estimations is required.

Nonetheless, the way most computational models approach conceptualization is still far from perfect. Meaning is hard.[2] A future direction I would like to suggest is the combination of continuous representations with naturalistic settings. The use of resources like WordNet is simply not suited to capture the subtleties of constructional meaning, and, more importantly, displays a cultural bias. The induction of universal semantic maps and subsequent acquisition of categories within this map forms an interesting way forward (cf. Beekhuizen, Fazly & Stevenson 2014).

In computational modeling, the shadow of scalability is always looming. The case is not different for SPL, I believe. I presented the performance of

[2]Or, as Hugo Brandt Corstius, famously, and intranslatably said "Wat je ook doet, de semantiek gooit roet" (lit. 'whatever you do, semantics throws soot', 'whatever you do, semantics is a spoilsport').

the model given an empirically grounded toy setting: the model processes certain words, but not others, given a limited representation of the meaning. Nonetheless, the distribution of the words, as per Alishahi & Stevenson's (2010) input generation procedure, as well as the parameters of the situational noise and uncertainty, are grounded in empirical work. What I have attempted is to trade-off the often conflicting notions of the faithfulness to a theoretical perspective, achieving realism in the simplifying assumptions, maximizing a model's empirical coverage, and maximizing the number of things a model can do (comprehensiveness). I have mainly focussed on the faithfullness and comprehensiveness, perhaps slightly at the expense of attempting to find more empirical coverage. There is a time for everything and only through a methodologically heteredoxical approach to computational modeling can we use its full potential.

Another issue of scalability that I believe needs to be addressed in usage-based frameworks, is that of 'hard' constructions. Within the generative paradigm, several constructions have been proposed to be unlearnable from the input data alone. Two approaches are typically pursued within the usage-based framework to counter these claims, namely the reconceptualization of the construction (e.g., Verhagen (2005) for long-distance *Wh*-questions, or van Hoek (1997) for pronominal binding), and corpus-driven work, possibly involving computational models that show that one can arrive at, at least, representations leading to the correct outputs (Smets 2010, Bod & Smets 2012). However, computational models doing so typicaly do not take meaning into account. It would be interesting to see if such 'hard' constructions can be acquired in a framework such as SPL.

Finally, I believe that cognitive modeling should be used more as a tool for theory formation than only as a hypothesis testing device. Not that that latter should be done *less*, but I believe that we, as a community of linguists and cognitive scientists have not yet understood the full potential of the method, which goes well beyond the mere empirical evaluation of a theory. At all levels of the scientific process, modeling provides a tool for shaping our endeavors: as a discovery procedure, a helping hand (but also a constraint) in formulating and scrutinizing theories, a means of giving existence proofs, and a means of both making as well as evaluating predictions. I hope to have presented a case where many of these possible applications of computational modeling come together, and furthermore hope that more research along similar lines will be done.