



# Prospects for usage-based computational models of grammatical development: argument structure and semantic roles

Stewart M. McCauley and Morten H. Christiansen\*

The computational modeling of language development has enabled researchers to make impressive strides toward achieving a comprehensive psychological account of the processes and mechanisms whereby children acquire their mother tongues. Nevertheless, the field's primary focus on distributional information has led to little progress in elucidating the processes by which children learn to compute meanings beyond the level of single words. This lack of psychologically motivated computational work on semantics poses an important challenge for usage-based computational accounts of acquisition in particular, which hold that grammatical development is closely tied to meaning. In the present review, we trace some initial steps toward answering this challenge through a survey of existing computational models of grammatical development that incorporate semantic information to learn to assign thematic roles and acquire argument structure. We argue that the time is ripe for usage-based computational accounts of grammatical development to move beyond purely distributional features of the input, and to incorporate information about the objects and actions observable in the learning environment. To conclude, we sketch possible avenues for extending previous approaches to modeling the role of semantics in grammatical development. © 2014 John Wiley & Sons, Ltd.

## How to cite this article:

*WIREs Cogn Sci* 2014, 5:489–499. doi: 10.1002/wcs.1295

## INTRODUCTION

In recent decades, cognitive science has increasingly relied upon computational modeling for existence proofs, hypothesis testing, and as a source of predictions on which to base empirical research. Nowhere is this trend more apparent than in developmental psycholinguistics, where, for over three decades (Ref 1), computational models have increasingly contributed to the long-standing debate over the nature of syntax acquisition. Computational modeling—as

a methodology—promises to provide a rigorous, explicit account of the psychological mechanisms whereby children acquire grammatical knowledge, as they move from a limited understanding of the surrounding social context to a seemingly unbounded capacity for communicating novel information. In recent years, computational models have been used extensively—though certainly not exclusively—to develop usage-based approaches to grammatical development. In particular, models have served to provide existence proofs, demonstrating that specific types of linguistic knowledge can, in principle, be learned from the input.

Usage-based computational accounts of grammatical development have primarily focused on what can be learned from distributional information. This

\*Correspondence to: christiansen@cornell.edu

Department of Psychology, Cornell University, Ithaca, NY, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.

approach has met with considerable success, illuminating the learning of syntactic categories,<sup>2</sup> specific developmental patterns,<sup>3</sup> and the acquisition of construction-like units,<sup>4</sup> in addition to illustrating the emerging complexity of children's grammatical knowledge more generally.<sup>5,6</sup> Yet, distributional approaches are unlikely to provide a complete account of children's *language use*. Crucially, distributional models have contributed little to our understanding of how the child computes *meaning*; to become a fully productive language user, the child must learn to compute meanings based on previously unencountered utterances, and to generate novel utterances conveying meanings they themselves wish to communicate.

The relative lack of semantic information in computational accounts of grammatical development stems in part from the difficult challenge of simulating naturalistic semantic representations that children may use. Moreover, the disciplinary segregation within developmental psycholinguistics further exacerbates the problem: separate subfields have typically focused on largely distinct areas, along traditional boundaries, such as those dividing phonology from word learning, word learning from syntax acquisition, and syntax acquisition from semantic development. As a result, much of the computational work on grammatical development has focused on structural considerations and this presents a serious challenge for usage-based approaches to acquisition, which hold that grammatical learning and development are tied to form-meaning mappings.<sup>7,8</sup> While incorporating semantics is therefore a pressing challenge for usage-based accounts in particular, the importance of meaning for grammatical development has also been emphasized in generativist approaches (e.g., Refs 9 and 10). Existing theoretical positions form a broad spectrum regarding the extent to which semantics is relied upon but many converge on the idea that grammatical development involves learning from meaning in context to at least some degree.

A comprehensive usage-based computational account is therefore faced with the considerable task of approximating learning from naturalistic semantic input while capturing the interplay between form and meaning in grammatical development. This challenge is made all the more daunting when one considers the full range of what semantic learning involves, from tense and aspect to anaphora to quantifiers and interrogatives. To make matters worse, comprehension and production involve rich conceptual representations that extend beyond what can be represented by current formalisms such as truth-conditional representations or first-order logic (e.g., Ref 11).

While accounting for such aspects of semantics presents a major challenge for computational models, initial progress has been made in a key area tied to early grammatical development: verb-argument structure and semantic role assignment. In what follows, we review existing computational models that instantiate usage-based principles to capture such linguistic development. The success of these models, we argue, is encouraging with respect not only to better understanding the psychological mechanisms involved in acquiring syntax but also with respect to the prospect of future work in modeling semantics-driven grammatical development more broadly.

To move toward a more complete usage-based account of grammatical development, we propose that computational models should aspire to meet a few basic challenges concerning the role of semantic information in model input, the linguistic tasks performed, and the ways in which performance is evaluated:

- (1) *Models should aim to capture aspects of language use.* Computational models of grammatical development should attempt to simulate the processes whereby children learn to interpret meanings during comprehension and to produce utterances that convey specific intended meanings (this requires that models incorporate approximations of learning from meaning in the contexts in which utterances are encountered, rather than from purely distributional information). This offers the advantage that models can be evaluated on their ability to capture relevant developmental psycholinguistic data, which necessarily involves tasks related to comprehension and/or production. Without the ability to model developmental data, it is uncertain whether the linguistic knowledge acquired by a model is actually necessary or sufficient to give rise to children's linguistic behavior.
- (2) *Models should make simplifying assumptions clear and explicit, motivating them with developmental data.* Computational accounts of language acquisition must make simplifying assumptions not only about the psychological mechanisms they seek to capture but also the nature of the input. This is especially true of semantic/perceptual input to models, given the challenge of creating naturalistic semantic representations. If possible, researchers should aim to motivate their decisions by appealing to psychological data (e.g., the decision to supply a predefined set of categories of some sort to a model could be supported with evidence that

children acquire those categories prelinguistically). As a corollary to this, models should only make simplifying assumptions where necessary and, where possible, employ naturalistic input (such as corpora of child-directed speech). When a model makes unnecessary or unmotivated simplifying assumptions, it becomes more difficult to assess how much of the model's performance is due to what it is capable of learning versus what is already built in.

- (3) *Models should adhere to psychologically plausible processing constraints.* Models intended as mechanistic accounts should aim to process input in an *incremental* fashion, rather than performing batch learning (e.g., by processing an entire corpus in a single step). This allows the model to approximate developmental trends when the trajectory of learning is examined, increasing the range of developmental data available for evaluating the model (e.g., longitudinal data or data from children in different age groups). Models should also aim to employ computations that are in principle capable of processing input *online*, in accordance with psycholinguistic evidence for the incremental nature of sentence processing in both children and adults.<sup>12,13</sup> Incorporating psychologically implausible processes means the model may be less likely to scale up to deal with more naturalistic data. Aside from the limitations this places on the model's ability to illuminate our understanding of the psychological mechanisms involved in grammatical development, it curtails its chances of contributing to the future of the field more broadly by serving as the basis for the construction of more comprehensive models.

In what follows, we provide an overview of existing usage-based computational models of verb argument structure learning that incorporate semantic information. We cover models of semantic role assignment, verb-argument structure construction learning, and models that learn about semantic roles and argument structure in the service of simulating comprehension and production processes more directly. Throughout, we evaluate models according to the challenges outlined above. We conclude by offering potential directions for extending existing computational approaches and for incorporating more naturalistic approximations of semantic input using readily available resources and techniques.

## MODELS OF SEMANTIC ROLE ASSIGNMENT

The notion of semantic roles (also referred to as *thematic roles*), such as *agent* and *patient*, was initially proposed by linguists working toward alternatives to early approaches to formal semantics,<sup>14,15</sup> but now enjoys widespread acceptance in theoretical linguistics. In the domain of formal approaches to syntax, semantic roles have been incorporated to varying degrees in argument structure analyses (e.g., Refs 16 and 17). Semantic roles are also widely accepted in psycholinguistics, where empirical work has built support for their psychological reality through evidence for adults' use of role information during online sentence comprehension (e.g., Refs 12, 18, and 19).

Thus, it is unsurprising that among the earliest computational models of language development to incorporate semantic information were those which learned to assign semantic roles to sentence constituents, providing an initial step toward capturing argument structure in comprehension processes. An early, representative example is the connectionist model of McClelland and Kawamoto,<sup>20</sup> a nonrecurrent network featuring a single layer of trainable weights. The model receives input in the form of static representations of sentences (consisting of a single verb and up to three noun phrases), in which words are represented in a distributed fashion by lists of semantic microfeatures (e.g., SOFTNESS, VOLUME, BREAKABILITY). The model is then trained to activate the semantic representations of the correct words filling up to four fixed semantic roles: AGENT, PATIENT, INSTRUMENT, and MODIFIER. The authors therefore characterize the key problem faced in learning argument structure as one of assigning a fixed set of (possibly prelinguistic) semantic roles to constituents where little to no ambiguity exists in the environment. The model successfully learns the role assignment task, generalizes to novel words, and is capable of disambiguating meanings based on sentence context. Nevertheless, the model's limitations are substantial: the static nature of the input representations severely limits the complexity of the sentence types the model can learn from, while the use of four fixed semantic roles and a lack of function words in the input further restricts what can be learned by the model.

This approach was later extended by St. John and McClelland,<sup>21</sup> who present a model that builds interpretations as sentence constituents are processed incrementally. The model employs separate input and output components: the input architecture is a simple recurrent network (SRN; Ref 22), while the

output side of the model is trained to respond to queries about sentences and their meanings. The SRN learns from sequences of sentence constituents (verbs, simple noun phrases, and prepositional phrases) to incrementally revise its predictions about the entire event described. The output component of the model is trained through back-propagation to respond with the appropriate semantic role when probed with a sentence constituent, and vice versa. As with McClelland and Kawamoto,<sup>20</sup> the authors characterize the problem facing the learner as one of assigning a predefined set of roles to sentence constituents. The model successfully learns to predict meanings incrementally, for both active and passive sentences, and generalizes to novel sentences and structures. Nonetheless, the model shares a number of limitations with that of McClelland and Kawamoto, including the use of a small number of fixed semantic roles. Despite the limitations of the model and its predecessor, subsequent models have successfully extended the basic framework to more comprehensive accounts, demonstrating that the general approach can scale up to more complex grammars (e.g., Ref 23; discussed below). Both models serve as valuable initial steps toward incorporating meaning into usage-based models, successfully demonstrating that a statistical approach based on thematic roles can in principle bootstrap basic aspects of grammar and achieve semantic and syntactic learning simultaneously.

Moving connectionist approaches closer to a more complete account of argument structure, Allen<sup>24</sup> describes a further model of semantic role assignment which introduces proto-role units in addition to each thematic role. An additional improvement is in the sentences used as input to the model, which are drawn from the CHILDES database<sup>25</sup> rather than generated by an artificial grammar (as is the case with most connectionist models of language learning). During exposure, verb and preposition input units are held constant while arguments are presented sequentially. Through these fixed frames, Allen implicitly characterizes the problem facing the learner as akin to one of learning argument structure constructions (e.g., Ref 7), although the task facing the model involves assigning roles to constituents (as in the previous approaches). The model exhibits syntactic bootstrapping, capturing role interpretations and semantic features for novel verbs. Despite this, the model is limited by the use of unambiguous feedback about correct semantic roles, and a built-in special status afforded to verbs in the linguistic input. While the model has a small vocabulary and input corpus, it is likely that the model would scale up to deal with more representative input (as suggested by subsequent connectionist

models discussed below). Allen and Seidenberg<sup>26</sup> extend this model, using it to propose a theory of grammaticality judgment. Furthermore, the fixed frame approach has been successfully applied in subsequent models with broader coverage (e.g., Ref 27; discussed below).

A further connectionist model of role assignment is presented by Morris et al.<sup>28</sup> Words are presented sequentially to a SRN that learns to map constituents to a small set of semantic roles, similar to previous models. A number of different sentence types are used as input to the model, featuring both experiential and action verbs. While the authors view the problem facing the learner in much the same way as McClelland and Kawamoto,<sup>20</sup> for instance, they go further in demonstrating that such an approach can both make contact with developmental data and yield unique insights into grammatical development. The model exhibits a pattern of generalization and undergeneralization for specific sentence types that approximates developmental psycholinguistic findings (e.g., Ref 29). Crucially, the authors use an analysis of the network's hidden layer representations to trace the emergence of an implicit 'subject' category, which is acquired entirely through the model's semantic processing, in the absence of any syntactic architectural features. Despite these successes, the model's coverage is limited by its impoverished semantics: semantic information is tied entirely to feedback about semantic roles. In addition to the model's particular limitations, it shares a number of limitations with previous approaches: for instance, constituents are mapped to a small set of predefined semantic roles. Although the input corpus—and resulting vocabulary size—is restricted due to computational considerations, the model appears capable of scaling up to deal with more naturalistic input (as suggested by subsequent, similar SRN models discussed below, such as Ref 23).

Recent statistical models of semantic role assignment have moved beyond the computational limitations of neural networks, successfully scaling up to deal with naturalistic input in the form of corpora of child-directed speech. The model of Connor et al.,<sup>30</sup> for instance, takes a subsection of the CHILDES database<sup>25</sup> as input. While the model instantiates the 'structure mapping' account of syntactic bootstrapping,<sup>31</sup> and is therefore at odds with usage-based theory on a conceptual level, its ability to scale up to more naturalistic input and learn from ambiguous semantic role information is useful in thinking about usage-based models. In the 'structure mapping' approach, children are innately biased to align each of the nouns in a sentence with a verb

argument. This allows the number of nouns appearing in a sentence to guide comprehension in the absence of knowledge of verb meanings. The model of Connor et al. captures this general notion by learning to assign a predefined set of semantic roles to arguments using a classifier, scaffolded by intermediate structural representations. An unsupervised hidden Markov model (HMM) is employed to cluster unlabeled words into part-of-speech categories using sequential distributional information, with a preclustering procedure used to create an initial division of function and content words. A 'seed list' of nouns is used to identify HMM states as potential argument states. The algorithm then chooses the word most likely to be the predicate for a sentence, based on the HMM state most likely to appear with the number of argument states identified in the sentence. With this amount of structural knowledge in place, the model is able to deal with ambiguous semantic feedback in the form of an unordered superset of semantic roles for each sentence (with the constraint that at least one of the roles truly exists). Importantly, feedback from the semantic role labeling task is used to refine the model's intermediate structural representations. The model successfully learns useful abstract features for semantic role assignment, and generalizes to sentences featuring novel verbs.

In addition to learning from naturalistic linguistic input, the Connor et al. model goes beyond previous models of role labeling in its ability to learn from ambiguous semantic feedback. While it represents the state of the art in psycholinguistic computational approaches to role labeling, the model is not without limitations. As with previous models, a fixed set of pre-defined semantic roles is used, and role information provides the only semantic input to the model; there is no further approximation of learning from a scene or event. Furthermore, the structure mapping approach necessitates learning from static representations of entire utterances rather than processing utterances in an incremental, online manner.

The above models of semantic role labeling provide an important step toward a more comprehensive computational account of grammatical development. The early connectionist approaches successfully demonstrate that aspects of argument structure can be learned through idealized semantic representations, suggesting a number of avenues for expanding the input to models to include meaning in context. Such models can trace potential routes for the emergence of abstract grammatical knowledge through purely semantic processing (e.g., Ref 28).<sup>a</sup>

## MODELS OF VERB-ARGUMENT CONSTRUCTION LEARNING

A number of more recent computational models have moved beyond the role labeling task, approaching the problem of acquiring verb-argument structure as one of learning grammatical constructions (e.g., Ref 7). Although not explicitly construction-oriented, Niyogi<sup>32</sup> provided an important precursor to models of argument structure learning, through a Bayesian approach to learning the semantic and syntactic properties of verbs. Niyogi's model learns from utterance-scene pairs consisting of sentences and accompanying semantic representations, made up by small sets of hand-coded features. The model is robust to noise, capable of learning from a small number of verb exposures, and exhibits both syntactic and semantic bootstrapping effects, successfully using syntactic frames to learn verb meanings and verb meanings to learn the syntactic frames in which a verb can be used. Despite these successes, the model is trained on a small language with a severely restricted vocabulary and range of sentence types. The model additionally relies on a considerable amount of built-in knowledge, including the structure of its hypothesis space and the prior probabilities over hypotheses.

More directly invoking construction grammar approaches, Dominey<sup>27</sup> presents a model of construction learning that is trained on a small artificial language, but uses simple processing mechanisms to learn from utterances paired with video data. Input to the model is derived from videos of an experimenter enacting and narrating scenes involving three distinct objects (e.g., a red cylinder). The narration is processed by a speech-to-text system, while the video is analyzed by an automated system tracking the contact that occurs between the items in the scene. Scene events are then encoded by such elements as duration of contact and object displacement, with the object exhibiting greater relative velocity encoded as the agent. This leads to such event representations as Touch(AGENT, OBJECT) and Push(AGENT, OBJECT, SOURCE). The model employs a modular architecture, acquiring initial word meanings through cross-situational learning. Utterances are then processed such that open- and closed-class words are automatically routed to separate processing streams. The model then uses the arrangement of closed class words in the input sentences to identify unique sentence types, which are then used to build up an inventory of constructions. Through these design choices, Dominey represents the problem facing the learner as one of learning partially abstract constructions based on item-based frames, rooted in previously acquired knowledge of the open-class/closed-class distinction.

The model is evaluated according to its ability to construct an internal scene representation corresponding to an input sentence, which it was able to do for a number of both active and passive constructions, in addition to relative clause constructions, with generalization to novel sentences. Despite this ability, the model is quite limited in scale, with a vocabulary of less than 25 words and an inventory of just 10 constructions. Because of its reliance on a predefined set of closed-class items to identify sentence structures that are assumed to be unique and nonoverlapping, it is unclear whether the model would successfully scale up to more naturalistic input in the form of corpora of child-directed speech. However, the general framework has been extended to cover Japanese<sup>33</sup> and French,<sup>34</sup> with a somewhat expanded vocabulary and inventory of learned constructions.

More recent accounts of verb-argument construction learning have attempted to deal with more naturalistic input. One such model is that of Chang,<sup>35</sup> who applies the Embodied Construction Grammar approach (ECG; see Ref 36 for a review) to the problem of learning item-based constructions in grammatical development. ECG is highly compatible with core principles of construction grammar, but places a strong emphasis on the importance of sensorimotor data and embodiment for determining the semantic content of constructions, invoking the notion of image schemas (e.g., Ref 37). Input to the model consists of utterances from a corpus of child-directed speech, accompanied by information about intonation, discourse properties (e.g., speaker, addressee, activity type, focus of joint attention), and an idealized representation of the visual scene. The model is initialized with a set of predefined schemas (corresponding to actions, objects, and agents) and a set of lexical constructions for individual words. The model acquires new constructions by forming relational maps to explain form-meaning mappings that the current grammar cannot account for, or by merging constructions into more general constructions.

While successfully acquiring useful verb-based constructions, Chang's approach has been applied to a limited range of constructions and requires a significant amount of hand encoding; it is not clear that it would scale up to a broader coverage. Furthermore, the learning mechanisms (involving minimum description length calculations, Bayesian statistics, etc.) involved in Chang's modeling approach may not be compatible with an incremental, on-line account of learning. Nevertheless, Chang's approach is encouraging for the prospect of a semantics-driven approach to grammatical development, and may be the currently best computational instantiation of the core

principles of various theoretical approaches emerging from cognitive linguistics (e.g., Refs 7 and 11).

Perhaps the most comprehensive model of argument structure construction learning is that of Alishahi and Stevenson,<sup>38</sup> based on incremental Bayesian clustering. Like the model of Connor et al.,<sup>30</sup> this model does not assume access to the correct semantic roles for arguments. However, unlike the model of Connor et al., the model of Alishahi and Stevenson does not have access to a fixed set of predefined roles, but instead learns a probability distribution over the semantic properties of arguments, capturing the development of verb-argument structure and semantic roles themselves, simultaneously. To approximate the semantics of nouns, lexical properties are extracted from WordNet.<sup>39</sup> This yields a list ranging from specific to more general properties (e.g., *cake*: {baked goods, food, solid, substance, matter, entity}), with considerable overlap among the more general properties across nouns. Input to the model consists in incrementally presented *argument structure frames*, each of which corresponds to an utterance and includes: the semantic properties for each argument; a set of hand-constructed semantic primitives for the verb (e.g., *eat*: {act, consume}); a set of hand-constructed event-based properties for each argument (e.g., {volitional, affecting, animate...}); the number of arguments; and the relative positions of the verb, arguments, and function words in the corresponding utterance. The authors add ambiguity to the input in the form of missing features. The frames are incrementally submitted to a Bayesian clustering process that groups similar frames into argument structure 'constructions'. In line with usage-based approaches, the model captures verb-specific semantic profiles for argument positions early in training. With continued exposure to the input corpus, these item-based roles gradually develop into more abstract representations, capturing the semantic properties of arguments across a range of verbs. The model is additionally capable of successfully capturing the meanings of novel verbs in ambiguous contexts.

Despite moving beyond previous approaches, the model of Alishahi and Stevenson is not without limitations. While the use of WordNet allows for automated creation of semantic properties for nouns, the use of hand-coded semantic primitives for verbs and event-based argument properties offers a crude approximation of learning from actual events, and restricts the input to frequent verbs. The use of static input representations means a lack of incremental sentence processing, and a considerable amount of built-in knowledge is provided, such as pre-existing knowledge of noun and verb categories.

Perfors et al.<sup>40</sup> present a further model of argument structure construction learning, which bears some similarities to that of Alishahi and Stevenson<sup>38</sup> while serving to underscore the importance of considering the distributional and semantic dimensions of the task simultaneously. The authors describe a hierarchical Bayesian approach primarily concerned with the distributional properties of verbs appearing in the dative alternation (e.g., Ref 41). Input to the model is extracted from CHILDES<sup>25</sup> and divided into epochs, allowing the model to approximate an incremental trajectory while learning in batch. A purely distributional version of the model learns from both positive and (indirect) negative evidence and successfully forms appropriate alternating and nonalternating verb classes, but overgeneralizes lower frequency verbs beyond the constructions in which they appear. In a subsequent version of the model, however, the inclusion of a single semantic feature (with three possible values corresponding to three classes of verb) leads to more child-like performance (e.g., Ref 41), with less overgeneralization. The model serves to underscore the potential power of distributional information as a basis for learning about argument structure while also demonstrating what can be gained by the introduction of even highly-idealized semantic information. Despite the insights provided by the model, it has a number of limitations: the model possesses prior knowledge about the uniformity and distribution of constructions in the input, and, as a result, it is unclear as to how heavily model's performance depends on its prespecified knowledge and whether it could serve as the basis for a more fully empiricist approach. Furthermore, the model focuses on a very restricted domain (the dative alternation); the authors note that it remains uncertain whether their approach would scale up to deal with a more complex dataset featuring a greater number of verbs and constructions.

## LEARNING ARGUMENT STRUCTURE THROUGH COMPREHENSION AND PRODUCTION

A number of models have successfully captured aspects of argument structure by learning to comprehend and produce utterances in an incremental, online fashion. Among the earliest and most comprehensive models in this vein is the Connectionist Sentence Comprehension and Production (CSCP) model of Rohde,<sup>23</sup> a large-scale SRN which is trained on a more complex subset of English than used with previous models, including features such as multiple verb tenses, relative clauses, and sentential complements. The semantic component of the model

consists of meanings encoded in distributed featural representations, and is trained using a query network (as in Ref 21). Comprehension in the model consists in learning to output an appropriate sentence meaning, given an incrementally presented sequence of words; as part of this process, the model learns to predict the next word in a sequence. Production involves learning to predict a series of words, given a static representation of sentence meaning (the most strongly predicted word is selected as the start of the utterance, and so forth). Thus, comprehension and production are tightly interwoven in the model. The model achieves strong performance on a number of tasks, successfully processing a wide range of sentence types, including sentences featuring multiple clauses. Importantly, the model also captures a number of psycholinguistic effects related to verb-argument structure, including structural priming, argument structure preference, and sensitivity to structural frequency.

The CSCP demonstrates that the general approach adopted by previous connectionist accounts of semantic role labeling can scale up to approximate online comprehension and production in an integrated model, with more complex input. Furthermore, the model acquires knowledge of argument structure through its attempts to comprehend and produce utterances, consistent with usage-based theory. Despite its comprehensive coverage, the model leaves something to be desired in the training of its semantic system: it remains unclear what psychological processes or mechanisms the model's fill-in-the-blank style query network would correspond to. Nevertheless, Rohde's model is perhaps the most comprehensive connectionist approach to language learning.

A similar—and somewhat more developmentally focused—model of acquisition through comprehension and production is provided by Chang et al.,<sup>42</sup> who use the Dual-path Model of Chang<sup>43</sup> to capture aspects of grammatical development within a connectionist framework. The Dual-path Model uses two distinct sets of connection weights: the first set captures the 'sequencing' of linguistic material, and is connected to the second set of weights which captures mappings between word forms, lexical semantics, event properties, and semantic roles (the 'message' component of the model). As with the above-discussed models of semantic role labeling, the Dual-path model simplifies the problem facing the learner considerably by assuming the correct mapping between semantic roles and lexical-semantic representations (via fast-changing weights). However, semantic roles in the model (five in total) do not instantiate traditional thematic roles (such as AGENT or PATIENT), but instead correspond to general

properties of a visual scene. For instance, a single role represents patients, themes, experiencers, and figures, while another role corresponds to goals, locations, ground, recipients, and so forth. The model is tasked with learning to correctly produce the words of a sentence when presented with a corresponding meaning representation (a task which can, in principle, be reversed to evaluate the model's comprehension performance). The Dual-path Model can successfully capture infant preferential-looking data<sup>44</sup> as well as data from elicited child productions.<sup>45</sup> It has also been used to successfully simulate structural priming effects.<sup>46</sup>

Like the model of Rohde,<sup>23</sup> the Dual-path Model is among the most comprehensive computational accounts of grammatical development to incorporate an active role for semantics, simulating online comprehension and production processes while making contact with a range of psycholinguistic data. While the model operates over a variety of hand-constructed sentence types (and has been successfully extended to cover word order biases in English and Japanese,<sup>47</sup> in addition to the acquisition of relative clauses<sup>48</sup>), the input to the model is nevertheless limited in scope, relative to models that learn from full corpora of child-directed speech. However, computational demands aside, it is likely that the general approach could scale up to deal with a more realistic set of input data. The model is further limited by its automatic alignment of lexical-semantic representations with the appropriate semantic roles, which are predefined and fixed, and thus does not capture the emergence of abstract roles or the ambiguity inherent in semantic feedback.

A further online, incremental approach to grammatical development is that of Mayberry et al.,<sup>49</sup> who present a recurrent network model of comprehension that incorporates a number of desirable features. Rather than simply learning to map linguistic input onto semantic roles, input to the model features representations of actions and entities in a scene (featuring two events), which remain active as the corresponding utterance unfolds incrementally. The model learns to output a meaning representation capturing the relationship between the particular action and entities described by the input sentence; this is done incrementally, in that the model's interpretation changes as each utterance unfolds. The model also captures anticipatory processing through prediction of likely utterance continuations. The model's selection of the appropriate scene is modulated by an utterance-driven attention mechanism, in the form of a gating vector. In addition to its general psycholinguistic features, the model's performance provides a qualitative fit to eye-tracking data from previous studies using

the visual world paradigm (e.g., Ref 50). Like other connectionist approaches, the grammar generating the linguistic input to the model is quite simple, and the model's vocabulary size is severely limited. However, given the effectiveness of the model's attention mechanism in processing semantic representations inspired by the visual world paradigm, it is likely that the model would successfully scale up to more representative input.

The models reviewed in this section successfully acquire argument structure by taking usage-based theory to its natural conclusion: by modeling language learning as *language use*, rather than relying on traditional notions of grammar induction as a separate process. A key challenge for the future will be to move this general approach beyond the computational restrictions inherent in connectionist techniques, by implementing usage-driven learning in higher-level statistical models capable of scaling up to deal with input in the form of entire corpora of child-directed speech (e.g., Refs 51 and 52) as well as more complex, multilayered semantic representations.

## EVALUATING AND EXTENDING EXISTING MODELS

Despite their limitations, existing models' ability to acquire aspects of verb-argument structure by approximating learning from meaning in context is encouraging for the prospect of more fully comprehensive usage-based models of grammatical development. In order to move toward models that better illuminate the psychological processes and mechanisms driving acquisition, the simplifying assumptions made by these and other models must continue to be examined and updated in the context of developmental data. For instance, the vast majority of the models discussed here rely on semantic role information in some capacity, based on a fixed set of predefined semantic roles. Developmental psycholinguistic work suggests that knowledge of abstract roles such as AGENT and PATIENT emerges gradually in development and is scaffolded by linguistic experience,<sup>53</sup> in line with the view that children acquire semantic roles gradually from the input. Despite the widespread acceptance of semantic roles, there has been little agreement on what semantic roles consist in or what part they play in language use; researchers have argued for a variety of approaches, with granularity ranging from verb-specific roles (e.g., Ref 54) to broad proto-roles (e.g., Ref 55).<sup>b</sup> A more fully comprehensive model of language development will need to address the nature and acquisition of semantic roles themselves (as in Ref 38), which represents an important step

toward understanding the ways in which linguistic and conceptual knowledge interact with and reinforce one another in learning argument structure.

Usage-based models will eventually need to move beyond argument structure and other aspects of so-called basic syntax to explore a broader range of grammatical phenomena. Given the success of idealized semantic information in helping to capture aspects of argument structure, it may prove that usage-based models will be better equipped to learn more difficult aspects of grammar after taking semantics into account: rather than involving purely structural considerations, meaning may also be central to learning complex grammatical phenomena, such as subject-auxiliary inversion (cf., Ref 56). Thus, in order to expand the grammatical coverage of models, researchers may need to expand the range of nonlinguistic information available as input (e.g., the above-cited account of subject-auxiliary inversion involves knowledge of tense), while also taking steps to ensure that the inclusion of highly idealized semantic input is not tantamount to building grammatical knowledge itself into the model. This will likely involve moving beyond the currently available tools. While existing resources such as FrameNet,<sup>57</sup> VerbNet,<sup>58</sup> and WordNet<sup>39</sup> constitute potentially rich sources of information for guiding the construction of features that can be combined with other tools (e.g., semantic shallow parsers) to automate the construction of idealized scenes for input to models concerned with argument structure, they are clearly insufficient for moving closer to the broader goal of modeling semantics more generally.

Researchers must also consider the amount of ambiguity present in the nonlinguistic information used as input to models. Simply randomizing the presence or absence of idealized referents may not yield representative input; for instance, Matusevych et al.<sup>59</sup> analyze the differences between contextual information generated based on child-directed speech itself versus hand-tagging of child–adult interaction videos, concluding that utterance-based meaning representations greatly oversimplify the task facing the learner. Matusevych et al., however, offer an automated technique for generating paired linguistic and idealized visual information that reflects the statistical properties of hand-tagged video data.

Finally, it must be recalled that meaning also involves social knowledge. To deal with more naturalistic semantic input and plausible degrees of ambiguity, models may need to incorporate learning from social information, including social feedback (e.g., Ref 60), reflecting the semi-supervised nature of the learning task. Previous models of word learning have successfully incorporated idealized social cues (e.g., Ref 61), and Chang<sup>35</sup> provides an initial step toward extending such an approach to grammatical development.

## CONCLUSION

We have provided a brief overview of the prospects and challenges of incorporating learning from semantic information into usage-based models of grammatical development, focusing on initial successes in modeling argument structure. Importantly, though, most of these challenges, if not all, are not unique to usage-based models but apply to varying degrees to *all* models that seek to understand the role of meaning in syntactic acquisition (e.g., as exemplified by the Connor et al.<sup>30</sup> model of thematic role assignment). We see, as a key goal for future work, the extension of these models to deal with increasingly naturalistic input and to cover the role of semantics in acquiring a broader range of grammatical knowledge. More generally, we expect that the lessons learned from the approaches surveyed here—as initial steps toward developing more comprehensive usage-based computational accounts of acquisition—are likely to have broad applications to both the modeling and theoretical understanding of grammatical development.

## NOTES

<sup>a</sup> Moreover, using a slightly simplified version of the Morris et al.<sup>28</sup> SRN model, Reali and Christiansen<sup>62</sup> demonstrated how network limitations on mapping from words to thematic roles can drive cultural evolution of a consistent word order from an initial state with no constraints on the order of words.

<sup>b</sup> We thank an anonymous reviewer for reminding us of this.

---

## ACKNOWLEDGMENTS

We would like to thank Laura Wagner and two anonymous reviewers for helpful comments and suggestions. This work was partially supported by BSF grant number 2011107 awarded to MHC.

## REFERENCES

1. Pinker S. Formal models of language learning. *Cognition* 1979, 7:217–283.
2. Redington M, Chater N, Finch S. Distributional information: a powerful cue for acquiring syntactic categories. *Cogn Sci* 1998, 22:425–469.
3. Freudenthal D, Pine JM, Gobet F. Understanding the developmental dynamics of subject omission: the role of processing limitations in learning. *J Child Lang* 2007, 34:83–110.
4. Solan Z, Horn D, Ruppin E, Edelman S. Unsupervised learning of natural languages. *Proc Natl Acad Sci USA* 2005, 102:11629–11634.
5. Bannard C, Lieven E, Tomasello M. Modeling children's early grammatical knowledge. *Proc Natl Acad Sci USA* 2009, 106:17284–17289.
6. Bornsztajn G, Zuidema W, Bod R. Children's grammars grow more abstract with age: evidence from an automatic procedure for identifying the productive units of language. *TopICS* 2009, 1:175–188.
7. Goldberg AE. *Constructions at Work*. New York: Oxford University Press; 2006.
8. Tomasello M. *Constructing a Language*. Cambridge: Harvard University Press; 2003.
9. Culicover PW, Jackendoff R. *Simpler Syntax*. Oxford: Oxford University Press; 2005.
10. Culicover PW, Nowak A. *Dynamical Grammar*, vol. 2. Oxford: Oxford University Press; 2003.
11. Langacker RW. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press; 2008.
12. Altmann G, Kamide Y. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 1999, 73:247–264.
13. Borovsky A, Elman JL, Fernald A. Knowing a lot for one's age: vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *J Exp Child Psychol* 2012, 112:417–436.
14. Fillmore C. The case for case. In: Back E, Harms RJ, eds. *Universals in Linguistic Theory*. London: Holt, Rinehard, and Winston; 1968, 1–88.
15. Jackendoff R. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press; 1972.
16. Bresnan J. *Lexical-Functional Syntax*. Oxford: Blackwell; 2001.
17. Chomsky N. *Lectures on Government and Binding*. Berlin: Mouton de Gruyter; 1981.
18. Carlson G, Tanenhaus M. Thematic roles and language comprehension. In: Wilkins W, ed. *Syntax and Semantics: Vol. 21. Thematic Relations*. San Diego: Academic Press; 1988, 263–291.
19. Trueswell JC, Tanenhaus MK, Garnsey SM. Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. *J Mem Lang* 1994, 33:285–318.
20. McClelland JL, Kawamoto AH. Mechanisms of sentence processing: assigning roles to constituents of sentences. In: McClelland JL, Rumelhart DE, eds. *Parallel Distributed Processing*, vol. 2. Cambridge, MA: MIT Press; 1986, 318–362.
21. St. John MF, McClelland JL. Learning and applying contextual constraints in sentence comprehension. *Artif Intell* 1990, 46:217–257.
22. Elman JL. Finding structure in time. *Cogn Sci* 1990, 14:179–211.
23. Rohde DL. A connectionist model of sentence comprehension and production. Unpublished Doctoral Dissertation, *Carnegie Mellon University*; 2002.
24. Allen J. Probabilistic constraints in acquisition. In: Sorace A, Heycock C, Shillcock R, eds. *Proceedings of the GALA '97 Conference on Language Acquisition*. Edinburgh: University of Edinburgh Human Communications Research Center; 1997, 300–305.
25. MacWhinney B. *The CHILDES Project: Tools For Analyzing Talk*, vol. 1. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
26. Allen J, Seidenberg MS. The emergence of grammaticality in connectionist networks. In: MacWhinney B, ed. *The Emergence of Language*. Mahwah, NJ: Lawrence Erlbaum Associates; 1999, 115–151.
27. Domney PF. Learning grammatical constructions in a miniature language from narrated video events. In: *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates; 2003, 354–359.
28. Morris WC, Cottrell GW, Elman J. A connectionist simulation of the empirical acquisition of grammatical relations. In: Wermter S, Sun R, eds. *Hybrid Neural Symbolic Integration*. Berlin: Springer; 2000, 175–193.
29. Maratsos M, Fox DE, Becker J, Chalkley MA. Semantic restrictions on children's passives. *Cognition* 1985, 19:167–191.
30. Connor M, Fisher C, Roth D. Online latent structure training for language acquisition. In: Walsh T, ed. *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press; 2011, 1782–1787.
31. Fisher C. Structural limits on verb mapping: the role of analogy in children's interpretations of sentences. *Cogn Psychol* 1996, 31:41–81.
32. Niyogi S. Bayesian learning at the syntax-semantics interface. In: *Proceedings of the 24th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates; 2002, 697–702.
33. Dominey PF, Inui T. A developmental model of syntax acquisition in the construction grammar framework with cross-linguistic validation in English and

- Japanese. In: *Proceedings of the First Workshop on Psycho-computational Models of Language Acquisition*. Stroudsburg, PA: Association for Computational Linguistics; 2004, 33–40.
34. Dominey PF, Hoen M, Inui T. A neurolinguistic model of grammatical construction processing. *J Cogn Neurosci* 2006, 18:2088–2107.
  35. Chang NCL. Putting meaning into grammar learning. In: *Proceedings of the First Workshop on Psycho-computational Models of Language Acquisition*. Stroudsburg, PA: Association for Computational Linguistics; 2004, 17–24.
  36. Bergen B, Chang NCL. Embodied construction grammar. In: Hoffmann T, Trousdale G, eds. *Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press; 2013, 168–190.
  37. Lakoff G. *Women, Fire, and Dangerous Things*. Chicago, IL: University of Chicago Press; 1987.
  38. Alishahi A, Stevenson S. A computational model of learning semantic roles from child-directed speech. *Lang Cogn Process* 2010, 25:50–93.
  39. Miller GA. Nouns in WordNet: a lexical inheritance system. *Int J Lexicog* 1990, 3:245–264.
  40. Perfors A, Tenenbaum JB, Wonnacott E. Variability, negative evidence, and the acquisition of verb argument constructions. *J Child Lang* 2010, 37:607–642.
  41. Gropen J, Pinker S, Hollander M, Goldberg R, Wilson R. The learnability and acquisition of the dative alternation in English. *Language* 1989, 65:203–257.
  42. Chang F, Dell GS, Bock K. Becoming syntactic. *Psychol Rev* 2006, 113:234–272.
  43. Chang F. Symbolically speaking: A connectionist model of sentence production. *Cogn Sci* 2002, 26:609–651.
  44. Hirsh-Pasek K, Golinkoff RM. The intermodal preferential looking paradigm: a window onto emerging language comprehension. In: McDaniel D, McKee C, Cairns HS, eds. *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press; 1996, 105–124.
  45. Tomasello M. Do young children have adult syntactic competence? *Cognition* 2000, 74:209–253.
  46. Bock JK. Meaning, sound, and syntax: lexical priming in sentence production. *J Exp Psychol* 1986, 12:575–586.
  47. Chang F. Learning to order words: a connectionist model of heavy NP shift and accessibility effects in Japanese and English. *J Mem Lang* 2009, 61:374–397.
  48. Fitz H, Chang F, Christiansen MH. A connectionist account of the acquisition and processing of relative clauses. In: Kidd E, ed. *The Acquisition of Relative Clauses: Processing, Typology and Function* (TILAR Series). Amsterdam: John Benjamins; 2011, 39–60.
  49. Mayberry MR, Crocker MW, Knoeferle P. Learning to attend: a connectionist model of situated language comprehension. *Cogn Sci* 2009, 33:449–496.
  50. Knoeferle P, Crocker MW. The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cogn Sci* 2006, 30:481–529.
  51. McCauley SM, Christiansen MH. Learning simple statistics for language comprehension and production: the CAPPUCINO model. In: Carlson L, Hölscher C, Shipley T, eds. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society; 2011, 1619–1624.
  52. McCauley SM, Christiansen MH. *Language Learning as Language Use: A Computational Model of Children's Language Comprehension and Production*. Manuscript in preparation. Ithaca, NY: Cornell University; 2014.
  53. Shayan, S. Emergence of roles in English canonical transitive construction. Unpublished doctoral dissertation, *Indiana University*; 2008.
  54. McRae K, Ferretti TR, Amyote L. Thematic roles as verb-specific concepts. *Lang Cogn Proc* 1997, 12:137–176.
  55. Dowty D. Thematic proto-roles and argument selection. *Language* 1991, 67:547–619.
  56. Bouchard D. Solving the UG problem. *Bioling* 2012, 6:1–31.
  57. Baker CF, Fillmore CJ, Lowe JB. The Berkeley Framenet project. In: *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*. Stroudsburg, PA: Association for Computational Linguistics; 1998, 86–90.
  58. Schuler, KK. VerbNet: a broad-coverage, comprehensive verb lexicon. Unpublished doctoral dissertation, *University of Pennsylvania*; 2005.
  59. Matussevych Y, Alishahi A, Vogt P. Automatic generation of naturalistic child–adult interaction data. In: Knauff M, Pauen M, Sebanz N, Wachsmuth I, eds. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society; 2013, 2996–3001.
  60. Goldstein MH, Schwade JA. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol Sci* 2008, 19:515–522.
  61. Yu C, Ballard DH. A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing* 2007, 70:2149–2165.
  62. Reali F, Christiansen MH. Sequential learning and the interaction between biological and linguistic adaptation in language evolution. *Interact Stud* 2009, 10:5–30.