Acquisition of aspect in self-organizing connectionist models*

XIAOWEI ZHAO AND PING LI

Abstract

Two connectionist networks, DISLEX and DevLex-II, were used in this study to model the acquisition of lexical and grammatical aspect. Both models use multi-layered self-organizing feature maps, connected by associative links trained according to the Hebbian learning rule. Previous empirical research has identified a strong association between lexical aspect and grammatical aspect in child language, on the basis of which some researchers argue for innate semantic categories or prelinguistic predispositions. Our simulations indicate that such an association can emerge from dynamic self-organization and Hebbian learning in connectionist networks, without the need of a priori assumptions about the structure of innate knowledge. Our modeling results further attest to the utility of self-organizing neural networks in the study of language acquisition.

1. Introduction

Aspect is one of the key linguistic categories for expressing temporal concepts in many languages. In contrast to another important category of temporality, tense, which is often used to locate the relationship between time of event and time of speech, aspect typically characterizes how a speaker views the temporal contour of a situation described, for example, the beginning, continuation, or completion of a situation. Aspect is also one of the earliest devices acquired by children, and as such the scientific study of it provides significant insights into not only how young children acquire temporal notions, but also what psycholinguistic mechanisms underlie the general acquisition processes.

Linguistics 47–5 (2009), 1075–1112 0024–3949/09/0047–1075
DOI 10.1515/LING.2009.038

1.1. *Two kinds of aspect*

Linguists generally distinguish between two kinds of aspect, grammatical aspect and lexical aspect (under various names; see Li and Shirai 2000 for a review). Grammatical aspect is related to aspectual distinctions which are often marked explicitly by linguistic devices, such as the inflectional suffixes and auxiliaries in English. It is also known as the viewpoint aspect (Smith 1997) which refers to a particular viewpoint toward the situation being talked about. According to Comrie (1976), there are two major categories of grammatical aspect: imperfective and perfective. Imperfective aspect presents a situation with an internal point of view, often as ongoing (progressive) or enduring (continuous), whereas perfective aspect presents a situation with an external perspective, often as completed. In English, the imperfective-perfective contrast is realized in the difference between the progressive *be V-ing* and the past-perfective *-ed*.[1]

Lexical aspect, on the other hand, refers to the characteristics inherent in the temporal meanings of a verb, for example, whether the verb encodes an inherent end point of a situation, or whether the verb is inherently stative (i.e., continuous and homogeneous) or punctual (i.e., momentary and instantaneous). Most researchers adopt Vendler's (1957) classification as the standard treatment of inherent semantics of verbs, which involves four categories: activities, accomplishments, achievements, and states. Activity verbs like *walk*, *run* and *swim* encode situations as consisting of successive phases over time with no inherent end point. Accomplishment verbs like *build a house* also characterize situations as having successive phases, but unlike activities they encode an inherent endpoint (e.g., house-building has a terminal point and a result). Like accomplishments, achievement verbs also encode a natural endpoint, but unlike accomplishments and activities they encode events as punctual and instantaneous, that is, as having no duration, such as in *fall*, *recognize a friend* and *cross the border*. Finally, state verbs encode situations as homogeneous, with no successive phase or endpoints, involving no dynamicity, such as *know*, *want* and *love*. In addition, on the basis of whether the verb encodes endpoints, linguists also call activity and state verbs "atelic" (no endpoint), and accomplishment and achievement verbs "telic" (with endpoint).

In English, grammatical aspect and lexical aspect often interact with each other in complex fashions. Uses of the inflectional suffixes, *-ing*, *-ed* and *-s* are in many cases constrained. For example, progressive aspect *-ing* does not occur often with state verbs; thus while "John knows the boy" is good, "John is knowing the boy" sounds odd (Smith 1983). There are also combinatorial constraints between *-ing* and event verbs; for example, "The book is falling off the shelf" is odd when used to refer to the actual falling down, but is good when used to mean a preliminary stage (i.e., prior to actual falling; Smith 1997). These kinds of constraints may reflect the intricate relationships between language use and characteristics of the described event. For example, as pointed out by Brown (1973), many events with an end result last for such a short period of time that any description of them is unlikely to occur during the period, such as the actions of *fall*, *drop*, and *break*. Thus it is rare for speakers to describe the "ongoing-ness" of such events with *-ing* but more natural for them to describe the "completeness" using past-perfective forms.

1.2. *Acquisition of aspect in children*

The study of grammatical aspect, lexical aspect, and their interactions has attracted research attention from many perspectives and in many languages in the last decade (e.g., Klein et al. 2000; Li and Shirai 2000; Shirai 1998; Shirai and Andersen 1995, to name a few). Accompanying this interest in aspect is the crosslinguistic study of aspect in child language. How do children acquire the two kinds of aspect and their interactions in different languages? In many acquisition studies researchers have found a recurring pattern of association between lexical aspect and grammatical aspect: children initially tend to restrict tense-aspect morphology to specific categories of lexical aspect. For example, English-speaking children initially tend to use progressive marker *-ing* only with atelic, activity verbs, whereas past-perfective marker *-ed* only with telic verbs (accomplishment and achievements) at an early stage of development (Harner 1981; McShane and Whittaker 1988; Shirai and Andersen 1995). This restricted or "undergeneralized" pattern of use has led to much intense debate with respect to various theoretical frameworks (see Li and Shirai 2000 for review). An early suggestion from Bickerton (1981, 1984) was that children have innate semantic categories that roughly correspond to the lexical aspect distinctions of verbs (e.g., punctual-nonpunctual, state-process distinctions), and these categories are biologically programmed as part of a Language Bioprogram. Bickerton relied on both data from creole languages and child language acquisition to support his proposal that children's acquisition of tense-aspect morphology has a biological basis. Subsequent crosslinguistic studies, however, have provided counter evidence to this hypothesis (e.g., Li and Bowerman 1998; Mapstone and Harris 1985; Shirai 1994; Shirai and Andersen 1995), and led researchers to propose a variety of input-driven hypotheses about how children acquire tense-aspect morphology and lexical semantics of verbs (see Li and Shirai [2000] for a review).

The goal of the current study is not to provide support for one or the other hypothesis. Rather, our goal here is to give some mechanistic accounts for how empirically observed patterns could emerge out of simple computational principles. In previous empirical studies (Li and Bowerman 1998; Li and Shirai 2000), we proposed that the initial lexical-morphological associations could arise as a result of the learner's analyses of the verb-morphology co-occurrence probabilities in the input environment of the language learner. In parental speech, there are probabilistic associations between progressive markers and atelic verbs, and between perfective markers and telic verbs (Shirai and Andersen 1995); children's initial undergeneralizations (restricted uses of morphology) might reflect their analyses of these probabilities. In this study, we specifically simulate how connectionist networks can analyze these probabilities and arrive at patterns that resemble children's patterns of acquisition.

1.3. *Modeling the acquisition of aspect with self-organizing neural networks*

One of the first tasks that networks as well as children need to handle is to arrive at semantic representations of verbs for the acquisition of lexical aspect. For children, they can go about acquiring semantic representations in roughly three ways. First, they may analyze the semantic properties of verbs through a verb's co-occurrence with situational contexts. For example, de Lemos (1981) showed how parents intentionally draw infants' attention to event structures, such as difference between result and process, through carefully modeled input in parental speech. Yu (2006) and Yu and Smith (2006) also provided general evidence with respect to how the child can exploit word-to-world relationships to extract meanings of nouns and verbs. Second, children may extract semantic representations through a verb's co-occurrences with other words (e.g. *break* typically occurs with *glasses*, while *tear* with *pieces of paper*). In previous research, Li et al. (2000) showed how global lexical co-occurrence information in sentences can be used by the learner to derive accurate lexical semantic representations. There is also a large literature on the syntactic bootstrapping of word meanings that is related to this type of co-occurrence analysis (Gleitman 1990; Naigles 1990). Finally, children may compute the co-occurrences of particular grammatical morphemes (e.g., -*ed* and -*ing*) with sets of semantic features that turn up repeatedly across lexical items (Behrend 1995; Behrend et al. 1995; Maratsos and Chalkley

1980). While all three ways of semantic representations can be simulated in connectionist networks, our models presented here focus on the latter two. In a previous work, Farkas and Li (2002) successfully constructed a connectionist model called WCD (word co-occurrence detector) to extract semantic representations, in close resemblance to the second method of semantic acquisition.

In the last few years Li and colleagues (Li 2003, 2006; Li et al. 2004; Li et al. 2007; Zhao and Li 2005) have explored self-organizing neural networks as candidates of cognitively and neurally plausible models of language acquisition. Their model has been applied to the study of a number of problems, including the acquisition of lexical categories (Li et al. 2004), vocabulary spurt (Li et al. 2007), bilingual lexical processing (Li and Farkas 2002), and AOA (age of acquisition) effects (Li et al. 2004; Zhao and Li 2005). In this study, we build on this line of research to examine the acquisitions of grammatical aspect (*-ing*, *-s* and *-ed*), in connection with the acquisition of semantic categories of lexical aspect. In particular, we attempt to show (1) how a multi-layer self-organizing neural network model is able to capture the processes of semantic organization that leads to distinct lexical aspect categories that have been claimed to be innate or otherwise prelinguistic, and (2) how the model could derive child-like semantic-morphological associations on the basis of analyzing patterns in parental input speech based on the CHILDES database (Li et al. 2000; MacWhinney 2000). Evidence from our study could also shed light on the mechanisms of lexical and morphological development in child language in general.

Compared to other developmental connectionist models, most of which rely on supervised learning algorithms (see models reviewed in Elman et al. 1996), self-organizing neural networks, especially the so-called self-organizing maps (SOM), have several important properties that make them particularly well suited to the study of lexical and morphological acquisition (see Li 2003, 2006 for discussion). First, they belong to the class of unsupervised learning networks that require no explicit teacher; learning is achieved by the system's organization in response to the input. Such networks provide computationally more relevant models for language acquisition, given that in real language learning children do not receive constant feedback about what is incorrect in their speech, or the kind of error corrections provided by supervised learning algorithms (see Li 2003; MacWhinney 1998, 2001; Shultz 2003 for discussion). Second, self-organization in these networks allow for the gradual formation of structures as changes of activity bubbles on 2-D maps, as a result of extracting an efficient representation of the complex statistical regularities inherent in the high-dimension input space[2] (Kohonen 2001). In particular, the network organizes information first in large areas of the map and gradually zeros in on to smaller areas; this zeroing-in is a process from diffuse to focused patterns, as a function of the network's continuous adaptation to the input characteristics. Third, the self-organizing map can fall into a topography-preserving state, which means nearby areas in the map respond to inputs with similar features. This property allows us to model the emergence of semantic categories as a gradual process of lexical learning. Finally, several self-organizing maps can be connected via Hebbian learning, a well-established biologically plausible learning principle, according to which the association strength between two neurons is increased if the neurons are both active at the same time (Hebb 1949). Strong physiological evidence of Hebbian learning exists in the form of the so-called *long-term potentiation* (LTP) in the hippocampus, a very important area for learning or formation of long-term memory in the brain (Haykin 1999). Although Hebbian learning itself is not an inherent property of the self-organizing algorithm, when incorporated, the SOM model would have strong implications for language acquisition: it can account for the process of how the learner establishes relationships between word forms, lexical semantics, and grammatical morphology, on the basis of how often they co-occur and how strongly they are co-activated in the representation.

In sum, models based on the above properties can: (1) allow us to track the development of the lexicon clearly as an emergent property in the network's self-organization (from diffuse to focused patterns or from incomplete to complete associative links), (2) allow us to model one-to-many or many-to-many associations between forms and meanings in the development of the lexicon and morphology, and (3) provide us with a set of biologically plausible and psychologically relevant computational principles to study language acquisition. They are biologically plausible because the human cerebral cortex can be considered as essentially a self-organizing map (or multiple maps) with topography-preserving ability (Kohonen 2001; Spitzer 1999); they are psychologically relevant because child language acquisition is essentially a self-organizing process (Li 2003; MacWhinney 2001). In fact, a number of such models have already been developed in recent studies of language processing and acquisition, such as DISLEX (Miikkulainen 1997), DevLex (Li et al. 2004), DevLex-II (Li et al. 2007), and SEMANT (Silberman et al. 2007). Building on these models, our networks discussed below rely on self-organization and Hebbian learning principles to acquire lexical aspect, grammatical aspect, and their associations and interactions. We report here two modeling studies that have been designed to simulate aspect acquisition, based on the DISLEX model and the DevLex-II model, respectively.[3]

2. Method

2.1. *Study I: modeling aspect acquisition with DISLEX*

2.1.1. *A sketch of the DISLEX model.* DISLEX, a multiple SOM model of the lexicon, was first introduced by Miikkulainen (1993, 1997). In this model, different self-organizing maps are connected through associative links via Hebbian learning. Each map is dedicated to a specific type of linguistic information (e.g., orthography, phonology, and semantics), and is trained as a standard SOM. A SOM works roughly as follows (see Kohonen 2001 for details). A two-dimensional topographic map is constructed for the organization of input representations, where each node (or "neuron") is a location on the map that has input connections to receive external stimulus patterns. On the map, a neuron k has a vector $\vec{m}_k$ associated with it to represent the weights of the input connections to it. At each training step of SOM, an external input pattern (e.g., the phonological or semantic information of a word in our study) is randomly chosen and presented to all the nodes on the map; this activates many nodes on the map, according to how similar by chance the input pattern is to the weight vectors of the nodes, and the node that has the highest activation is declared the winner (the Best Matching Unit or BMU). Once a node becomes active in response to a given input, the weight vectors of that node and its neighboring nodes are adjusted, so that they become more similar to the input and the nodes will respond to the same or similar inputs more strongly the next time. In this way, every time an input is presented, an area of nodes will become activated on the map (the "activity bubbles") and the maximally active nodes are taken to represent the input. Initially activation occurs in large areas of the map, that is, large neighborhoods, but gradually learning becomes focused and the size of the neighborhoods reduces. This process continues until all the inputs have found some maximally responding nodes as their BMUs. As a result of this self-organizing process, the statistical structures implicit in the input are represented as topographical structures on the 2-D space. In this new representation, similar inputs will end up activating the same nodes in nearby regions, yielding meaningful activity bubbles that can be visualized on the map.

In standard SOM, the identification of winners on the map uses the following rule to update the weights of nodes around a winner or BMU:

$$(1)\quad \vec{m}_k(t+1) = \vec{m}_k(t) + \alpha(t) \cdot [\vec{x} - \vec{m}_k(t)] \quad \text{for all } k \in N_c$$

Here, $\alpha(t)$ is the learning rate for the map, which changes with time t. $\vec{x}$ is the input stimulus of current training step. N_c indicates the set of nodes in the neighborhood of the winner c. $\vec{m}_k$ is the input weight vector of a node k on the map. If the node k belongs to the nodes in the neighborhood of the winner c, its weight should be adjusted according to this equation; otherwise, it remains unchanged.

In DISLEX, an input pattern activates a node or a group of nodes on one of the input maps, and the resulting activity bubble propagates through the associative links and causes an activity bubble to form in the other map. If the direction of the associative propagation is from phonology or orthography to semantics, comprehension is modeled; production is modeled if it goes from semantics to phonology or orthography. The activation of co-occurring lexical and semantic representations leads to continuous organization in these maps, and to adaptive formations of associative connections between the maps. The weights of the associative links between the features maps are updated according to the Hebbian learning rule (Hebb 1949), as in (2):

$$(2) \quad \Delta w_{kl} = \beta \cdot \alpha_k^S \cdot \alpha_l^D$$

where w_{kl} is the unidirectional associative weight leading from node k in the source map to node l in the destination map, and α_k^S and α_l^D are the associated node activations. β is a constant learning rate. The associative weight vectors are then normalized according to (3), and normalization is carried out over all associative links of the source node.

$$(3) \quad w_{kl}(t+1) = \frac{w_{kl}(t) + \Delta w_{kl}}{\{\sum_l [w_{kl}(t) + \Delta w_{kl}]^2\}^{1/2}}$$

Using these basic features of the DISLEX model, in this study, we constructed two self-organizing maps, one for the organization of phonological input and one for the organization of semantic input. Each map consisted of 50×50 nodes. All simulations were run on a SUN Ultra workstation, using the DISLEX codes configured by Miikkulainen (1999).

2.1.2. *Input data and representations for DISLEX.* To model the role of linguistic input in children's acquisition of lexical and grammatical aspect, we selected as our input data the parental or caregiver's speech in the CHILDES database (MacWhinney 2000). We extracted utterances produced by parents, caregivers, and experimenters from the CHILDES database in about half of the English corpus (from Bates to Korman). Although not all of these utterances are child-directed, they form a representative sample of the speech that children are exposed to (e.g., dinner table talks, activities of free plays, and storytelling). A verb type from this corpus was chosen as an input to the network if it occurred in the total parental or caregivers' speech for five or more times in a given age period (a bare verb form and its inflected forms were calculated as different verb types). With this criterion we selected a total of 562 words (types) as input to our network. They were submitted to the network in four stages, according to the age groups at which they occurred (see 2.1.3. below for details)

Many previous connectionist models of language acquisition rely on the use of artificial input/output representations (e.g., randomly generated patterns of phonological or semantic representations), or representations that are constructed ad hoc by the modeler. Models based on such representations run the risk of modeling developmental patterns that have little to do with the actual learning task. To make direct contact with realistic learning, in this study, our model represents input words according to their realistic linguistic features, as follows.

To represent the phonology of the verbs, we used a syllable-based template coding developed by MacWhinney and Leinbach (1991) (see also Li and MacWhinney 2002). Instead of a simple phonemic representation, this representation reflects current auto-segmental approaches to phonology, according to which the phonology of a word is made up by combinations of syllables in a metrical grid, and the slots in each grid made up by bundles of features that correspond to phonemes, C's (consonants) and V's (vowels). The MacWhinney-Leinbach model used 12 C slots and 6 V-slots that allowed for representation of words up to three syllables. For example, the 18-slot template CCC VV CCC VV CCC VV CCC represents a full tri-syllabic structure in which each CCCVV is a syllable (the last CCC represents the consonant endings). Each C is represented by a set of 10 feature units and each V by a set of 8 feature units. The suffixes *-ing*, *-ed*, and *-s* are also represented by these feature units.

Semantic representations to our network were based on the lexical co-occurrence analyses in the Hyperspace Analogue to Language (HAL) model (Burgess and Lund 1997). HAL represents word meanings through multiple lexical co-occurrence constraints in large text corpora. In this representation, the meaning of a word is determined by the word's global lexical co-occurrences in a high-dimensional space: a word is anchored with reference not only to other words immediately preceding or following it, but also to words that are further away from it in a variable co-occurrence window, with each slot (occurrence of a word) in the window acting as a constraint dimension to define the meaning of the target word. Thus, a word is represented as a vector that encodes the entire contextual history of that word in a high-dimensional space of language use (see Li et al. 2000 for application of HAL to children's lexical acquisition). Here, we used 100 dimensions to encode each vector.

2.1.3. *Task and procedure.* Upon training of the network, a phonological input representation of the verb was presented to the network, and simultaneously, the semantic representation of the same input was also presented to the network. Through self-organization, the network formed an activity on the phonological map in response to the phonological input, and an activity on the semantic map in response to the semantic input. When an input verb was in its inflectional form, both its bare verb form and the suffix *-ing -ed*, or *-s* were presented to the network for training simultaneously. The phonological representations of the suffixes *-ing -ed*, or *-s* were presented only to the phonological map. Under this situation, the nodes that corresponded to the suffix's phonological representations would co-activate with those corresponding to the bare verb forms on the phonological map, and with those corresponding to the verbs on the semantic map. As the network received input and continued to self-organize, it simultaneously formed associations through Hebbian learning between the two maps for all the active units that responded to the input. The network's task was to create new representations in the corresponding maps for all input words and to be able to map the semantic properties of a verb to its phonological shape and its morphological pattern.

To observe effects of the interaction between lexical and grammatical aspect in the parental input on the network's learning and representation, we designed four stages to train the network, according to the different age groups of our input data. At each stage the network was trained for 200 epochs; that is, each verb type (a bare verb or a verb with suffix) was presented to the network 200 times.

1. Input Age 1;6 (13–18 months). Although parental/caregivers data in the CHILDES database are available from an age when the child is 6 months old, there are relatively few morphological markings prior to age one. A total of 186 verb types fit our selection criteria for the period when the child is between 13 and 18 months old, out of which 34 (18%) occurred with *-ing*, 9 (5%) with regular past form *-ed*,[4] 9 (5%) with *-s*, and the remaining verbs are verb stems in corpus without suffixes.
2. Input Age 2;0 (19–24 months). 324 verb types were selected, which include new verbs as well as verbs from the previous stage, among which 76 (23%) occurred with *-ing*, 23 (7%) with *-ed*, and 24 (7%) with *-s*.

3. Input Age 2;6 (25–30 months). 419 verb types were selected, among which 82 (20%) occurred with *-ing*, 35 (8%) with *-ed*, and 31 (7%) with *-s*.
4. Input Age 3 (31–36 months). 562 verb types were selected, among which 123 (22%) occurred with *-ing*, 70 (12%) with *-ed*, and 61 (11%) with *-s*.

These stages ensure an incremental growth of vocabulary and a coarse frequency coding: a verb or a suffix was presented to the network for the number of times it occurred across the four stages. Thus, during training, if a verb occurred at two different stages, the network would receive the verb as input twice, on both the phonological map and the semantic map; if it occurred at three stages, the network would receive it three times, and so on. For example, if a verb *talk* occurred with suffix *-ing* at stage 2, 3 and 4, but not at stage 1, then at the end the network would receive three times the phonological representation of *talk*, that of *-ing*, and the semantic representation of *talk*. In this way, words that occurred earlier would have higher frequency as the input to the network. In essence, this procedure modeled a coarse frequency of the verb-morphology association, even though the frequency variable was not explicitly manipulated in our simulations. The learning parameter (α) and the neighborhood size (*Nc*) systematically varied in our simulations, with α varying from 0.1 to 0.0 at each stage of training. *Nc* started large to cover the entire map and decreased to 1 at the end of each stage. At the beginning of each stage, *Nc* was slightly increased relative to that at the end epoch of the previous stage, to accommodate the increased number of input verbs at the new stage.

2.2. *Study II: modeling aspect acquisition with DevLex-II*

2.2.1. *A sketch of DevLex-II.* DevLex-II is a new multiple self-organizing neural network for modeling early lexical acquisition. It is based on and adapted from the DevLex model (Farkas and Li 2002; Li et al. 2004). It has been developed to account for some empirical phenomena in early lexical acquisition, such as "vocabulary spurt" and early word production errors (Zhao and Li 2005; Li et al. 2007). The basic structure of DevLex-II is similar to that of DISLEX and DevLex, with some differences noted below. Figure 1 shows the architecture of DevLex-II.

DevLex-II uses three layers of SOMs to process three basic levels of linguistic information: phonological content, semantic content, and

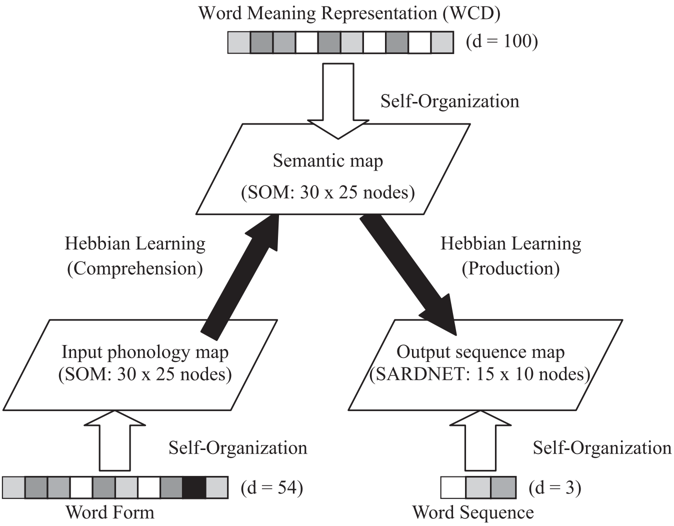

Figure 1. *The architecture of DevLex-II model. Each of the three self-organizing maps (SOM) takes input from the lexicon and organizes phonology, semantics, and phonemic sequence information of the vocabulary, respectively. The number of nodes in each map is indicated in parentheses. The dimension of input vector for each map is indicated by 'd =' in parentheses next to the input representation symbols. The maps are connected via associative links updated by Hebbian learning.*

output phonemic sequence. The addition of the phonemic sequence layer represents a step forward from the original DevLex model, and is inspired by models of word learning based on temporal sequence acquisition (e.g., Gupta and MacWhinney 1997). It is designed to simulate the challenge to young children when they need to develop better articulatory control of the phonemic sequences of words. Just as the learning of auditory sequences requires the mediation of memory systems, the learning of articulatory sequences requires support from the rehearsal in phonological working memory (Gathercole and Baddeley 1993; Gupta and MacWhinney 1997).

In our implementation of this idea, the activation pattern corresponding to the phonemic sequence information of a word is formed according to the algorithms of SARDNET (James and Miikkulainen 1995), which works slightly differently from the standard SOM algorithm. At each training step, phonemes are input into the sequence map one by one, according their order of occurrence in the word. The winning unit of a phoneme is found and the responses of nodes in its neighborhood are adjusted as shown in Equation 1. Once a unit is designated as the winner, it becomes ineligible to respond to the subsequent inputs in the sequence. In this way, same phonemes in different locations of a word will be mapped to different (but adjacent) nodes on the map as an result of the network's topography-preserving ability. When the output status of the current winner and its neighbors is adjusted according to Equation 1, the activation levels of the winners responding to phonemes before the current phoneme will be adjusted by a number γ^{d}, where γ is a constant and d is the distance between the location of the current phoneme and the previous phoneme that occurred in the word. This adjustment is intended to model the effect of phonological short-term memory during the learning of articulatory sequences; the activation of the current phoneme could be accompanied by some rehearsal of previous phonemes due to phonological memory, which deepens the network's or the child's impression of previous phonemes. Further details of DevLex-II are discussed in Li et al. (2007).

As in Study 1, the associative links between any two layers of maps are trained by Hebbian learning (see Equation [2]), such that the activation of a word on the form map can evoke the activation of a word on the meaning map via form-to-meaning links, thereby modeling word comprehension, and the activation of word meaning can cause the formation of word sequence via meaning-to-sequence links, thereby modeling word production. In DevLex-II, we say that a word has been learned in comprehension, when a node in the destination map (word meaning map) becomes consistently activated as the "winner" for a given input from the source map (word form map). We say that a word has been learned in production, when several nodes in the word sequence map become activated sequentially as winners that represent the word's phonemes.

2.2.2. *Input representations for DevLex-II.* As with DevLex, we used the PatPho system to construct the phonological patterns for word forms. PatPho is a generic phonological pattern generator for neural networks, which fits every word (up to tri-syllables) onto a template according to its vowel-consonant structure (Li and MacWhinney 2002). PatPho uses the same phonological method as in MacWhinney and Leinbach (1991) (see discussion in Section 2.1.2), but relies on articulatory features of phonemes (Ladefoged 1982) to represent the phonemes, *C*s and *V*s, and a phoneme-to-feature conversion process produces real-value or binary feature vectors for any word up to three syllables. In short, Patpho can code each input word in our simulation by the template *CCCVVCCCVVCCCVVCCC*, and then replace each phoneme with its appropriate representation using real-value or binary numbers.[5] For example, a verb *pick* with its progressive marker *-ing* would be encoded as p*CC*ɪ*V*k*CC*ɪVŋ*CCVVCCC*, and is represented as the following vector:

/pɪkɪŋ/: *1 0.45 0.733 0 0 0 0 0 0 0.1 0.1 0.185 0 0 0 1 0.921 0.733 0 0 0 0 0 0 0.1 0.1 0.185 0 0 0 0.75 0.921 0.644 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0*

In this representation, the first three units [1 0.45 0.733] indicate the phonetic features of phoneme /p/, the second and third sets of three units indicate that no more consonants follow /p/ in this word (hence zeros). The representation is left-justified, which means that in a given syllable, the representation of the phoneme is pushed toward the left side of the template (rather than the right side).

On the output sequence map, the phonemes of a word are processed one by one, so we need representations for each of the 38 English phonemes. Using the method of PatPho, we can represent these phonemes by three-dimensional real-value vectors. In particular, in the vector, the first dimension indicates whether the phoneme is a vowel or a consonant, and in the case of a consonant, whether it is voiced or voiceless. The second dimension indicates the position for vowels and manner of articulation for consonants, and the third dimension indicates the sonority for vowels and place of articulation for consonants (see Li and MacWhinney 2002).

With respect to the semantic representation of the input, in Study 1 we used representations based on the HAL method (Burgess and Lund, 1997). In Study 2, we used a special recurrent network called WCD (word co-occurrence detector) to generate HAL-like vectors. The main difference between HAL and WCD is that the former generates stationary vectors while the latter allows us to generate vectors that dynamically change with the learning history: lexical representations enrich over time as a function of learning the number of co-occurring words in the input sentences. WCD allows us to build semantic representations on the fly, incorporating more and more different words in a context, until the size of the lexicon reaches a given target level. Metaphorically, this learning scenario can be compared to filling the holes in a Swiss cheese: initially there may be more holes than cheese (*shallow* representations) but the holes get filled up quickly as the co-occurrence context expands with more words being acquired (*rich* representations).

Briefly, WCD works as follows (see Farkas and Li 2001, 2002; Li et al. 2004 for details). It reads through a stream of input sentences one word at a time, and learns the transitional probabilities between words which it represents as a matrix of weights. Given a total lexicon sized N, all word co-occurrences can be represented by an $N \times N$ contingency table, where the representation for the i_{th} word is formed by concatenation of i_{th} column vector and i_{th} row vector in the table. Hence, the two vectors correspond to the left and the right context, respectively; WCD transforms these probabilities into normalized vector representations for word meanings, which in turn are read by self-organizing maps (after Random Mapping, a procedure to achieve reduced uniform dimension of vectors; see Ritter and Kohonen 1989). As in study 1, different inflectional forms of the same verb are considered as different items, and therefore WCD will derive different (but also similar) representations for them. For example, *playing* and *played* will be represented distinctly, but since the co-occurrence contexts for these words will overlap significantly (e.g., the co-occurring words tend to be *ball*, *toys*, etc.), the representations for them will also tend to be similar. An example of a semantic representation generated by WCD is shown below (only part of the vector is shown, the full vector contains 2002 units). Here, every two units of the vector represent the normalized co-occurrence possibility between the verb *picking* and another word in the lexicon. The odd unit represents the possibility that *picking* happens before a given word, and the even unit represents the possibility that *picking* follows a given word in the context.

Picking: *0.000000 0.000000 0.006004 0.007211*
0.003548 0.017577 0.000000 0.000000 0.000000 0.000000
0.000000 0.000000 0.000000 0.001202 0.000000 0.000300
0.000000 0.000000 0.000000 0.000000......

As in Study 1, DevLex-II also uses as its input data the parental or caregivers' speech in the CHILDES database (MacWhinney 2000). Here we extracted all parental or caregivers' utterances from the complete English database (i.e., the available database in 2002). Due to the expanded size of the corpus (about 2.6 million word tokens), the criterion of verb selection was also modified: a verb type was chosen as input if it occurred in the parental speech for fifty or more times in a given age period. As in Study 1, the verbs were divided into four stages, according to the age groups (Age 1;6, 2:0, 2;6, 3;0) at which they occurred. To increase the accuracy of WCD representations, we also analyzed the selected verbs along with the nouns, adjectives, and close-class words from the MacArthur-Bates Communicative Development Inventories (the CDI, Toddler's List; Dale and Fenson 1996; homographs and homophones, word phrases, and onomatopoeias were excluded). These CDI words along with the verbs that fit our selection criterion (a total of 1001 words) served as the input contexts of WCD. We computed the semantic representations of the vocabulary at each of the four growth stages, resulting in

4 different data sets with increasing complexity in semantic representation. The four growth stages had the following vocabulary composition:

1. Input Age 1;6 (13–18 months): a total of 62 verb types fit our selection criteria for the period before age 1;6; 35 of these verbs occurred with *-ing*, 13 with *-ed*, 14 with *-s*.
2. Input Age 2;0 (19–24 months): 100 verb types were selected, which included the new words as well as words from the previous stage; 58 occurred with *-ing*, 19 with *-ed*, 23 with *-s*.
3. Input Age 2;6 (25–30 months): 154 verb types were selected, among which 86 occurred with *-ing*, 32 with *-ed*, 36 with *-s*.
4. Input Age 3 (31–36 months): A total of 184 verb types were selected, out of which 97 verbs occurred with *-ing*, 41 with *-ed*, and 46 with *-s*. This stage included all verbs that occurred in previous stages plus new ones.[6]

2.2.3. *Training parameters.* The size of the network was 30 $\times$ 25 nodes for the phonological map and the semantic map, and 15 $\times$ 10 nodes for the phonemic map. These numbers were chosen to be large enough to discriminate among the words and phonemes in the lexicon, while keeping the computation of the network tractable. The learning rate $\alpha(t)$ and neighborhood radius (*Nc*) were set similarly as in Study 1, where they changed according to different age periods. At Stage 1 (Age 1;6), for each layer, the learning rate α started from 0.4 (α*_start*), gradually decreased to 0.1 (α*_end*); the radius of a winner's neighborhood on the phonological or the semantic layer (*Nc_som*) gradually decreased from 13 to 0, and that on the phonemic output layer (*Nc_sard*) gradually decreased from 5 to 0. At Stage 2 (Age 2;0), the parameters were set to 0.4 (α*_start*), 0.1 (α*_end*), 8 (*Nc_som*), 5 (*Nc_sard*); at Stage 3 (Age 2;6), these parameters were 0.2, 0.05, 5, 3, respectively, and at Stage 4 (Age 3;0), they were 0.1, 0.05, 3, 3 respectively. The learning rate β for associative links between levels was kept constant at 0.1 during the entire training process. For each stage, the network was trained for 50 epochs, which means that each verb in a given stage was presented to each map 50 times.

3. Results and discussion

In this section we will focus on three levels of analysis for the two models' simulation results: the role of input, the emergence of structured semantic representations, and the role of Hebbian learning.

3.1. *Role of input*

An important rationale behind our simulations is for us to understand the role of linguistic input in guiding children's acquisition of lexical and grammatical aspect. Earlier we have emphasized the relationship between patterns observed in children's speech and those in adult speech with respect to the interaction between lexical aspect and grammatical aspect. But a simple correlation between children's and adult's patterns tells us only that the child is sensitive to the linguistic environment and is able to incorporate information from that environment into his or her own speech. It does not tell us how the child actually does this, or what mechanisms allow the child to do this. Thus we wanted to test whether connectionist networks, endowed with self-organization and Hebbian learning principles, are able to display learning patterns as the child does. Our networks receive phonological and semantic representations of input words from actual adult speech along with phonemic sequence (morphology) information of these words. If this kind of network is able to produce patterns like those we found in children's speech, on the basis of learning of the input, we can then conclude that self-organization and Hebbian learning provide the necessary kinds of mechanisms that drive the formation of patterns in children's acquisition. In this way, our modeling enterprise sheds light on the mechanisms that underlie the learning process.

Tables 1 and 2 provide summaries of the major patterns from the DISLEX and DevLex-II models, respectively, according to the tense-aspect

Table 1. *Percentage of use of tense-aspect suffixes with different verb types across input age groups in DISLEX's production and in parental input data (including verbs with multiple suffixes)*

VERBS	TENSE-ASPECT SUFFIXES											
	Age 1;6			Age 2;0			Age 2;6			Age 3;0		
	-ing	*-ed*	*-s*	*-ing*	*-ed*	*-s*	*-ing*	*-ed*	*-s*	*-ing*	*-ed*	*-s*
Network production												
Activity	72	16	0	62	29	6	64	40	44	52	38	30
Telic	28	75	0	32	66	31	32	60	12	43	53	26
Stative	0	8	100	0	4	63	0	0	44	5	9	44
Parental Input data												
Activity	65	22	33	66	30	29	62	40	42	60	40	43
Telic	32	77	33	32	65	25	32	54	26	33	44	24
Stative	3	0	33	3	4	46	6	6	32	7	16	33

Table 2. *Percentage of use of tense-aspect suffixes with different verb types across input age groups in DevLex-II's production and in parental input data (including verbs with multiple suffixes)*

VERBS	TENSE-ASPECT SUFFIXES											
	Age 1;6			Age 2;0			Age 2;6			Age 3;0		
	-ing	*-ed*	*-s*	*-ing*	*-ed*	*-s*	*-ing*	*-ed*	*-s*	*-ing*	*-ed*	*-s*
Network production												
Activity	73	0	29	69	27	33	61	24	35	62	30	37
Telic	27	75	14	21	53	28	32	62	27	31	62	26
Stative	0	25	57	10	20	39	7	14	38	7	8	37
Parental Input data												
Activity	63	23	29	62	26	26	63	22	33	60	29	35
Telic	31	62	29	31	58	26	29	66	25	32	59	24
Stative	6	15	43	7	16	48	8	12	42	8	12	41

suffixes the model produced at different learning stages. The two tables present the results of the networks' production of three suffixes, *-ing*, *-ed*, and *-s*, with three types of verbs, activity, telic and stative.[7] The results were based on the analysis of the networks' production ability; that is, how semantic representations induce activations on corresponding feature maps (phonological map in DISLEX and output phonemic sequence map in DevLex-II) through associative pathways. The analysis was done by inspecting the nodes that each verb on the semantic map activated, after the network had been trained for a specified number of epochs at each stage (200 epochs for DISLEX and 50 epochs for DevLex-II).

In particular, the testing of DISLEX's production ability of target verbs can be described as follows. At the end of each training stage, verb types in the lexicon of the current stage are presented to the semantic map one by one. For a verb type, its best matching unit or BMU on the semantic map is found, and in turn this node propagates its activation to the phonological map through the associative links. As a result of this propagation, some nodes on the phonological map become activated, and the network checks if the nodes were BMUs for the phonological representations of corresponding verb stems and suffixes; if they are, we say that the verb has been correctly produced. For example, when the verb *kicking* is shown to semantic map, two units on the phonological map may become activated, and if the two nodes are the BMUs for the phonological representations of *kick* and *-ing* respectively, we say that the verb *kicking* has been correctly produced by DISLEX. Running the test over the entire lexicon, we can calculate how many verbs in each suffix category or aspect category are correctly produced, as the results shown in Table 1.

The testing of DevLex-II's word production ability is similar to that of DISLEX, with a slight modification. This time, the winner node on the semantic map propagates its activation to the output sequence map rather than the phonological map, and several nodes in the sequence map become activated sequentially as winners that represent the word's phonemes. Then the network checks to see if every node is the BMU of a unique phoneme, according to the Euclidean distance between its input weight vector and the feature representation of every phoneme. If it is, the phoneme closest in Euclidean distance to the current winner becomes its retrieved phoneme; if it is not, the pronunciation of this phoneme has failed. Finally, the pattern of the retrieved phoneme sequence is treated as the output of word production. When the retrieved phonemic sequence matches up with the actual word's phonemic sequence, we say that the word has been correctly produced. For example, if the word *kicking* is shown to the semantic map, correct production occurs only when the consecutively activated nodes on the output phonemic map are the BMUs for /k/ /ɪ/ /k/ /ɪ/ /ŋ/ in a sequence.

The results of the two tables are quite similar and highly consistent with empirical patterns observed in early child language: the use of imperfective aspect is closely associated with activity verbs that indicate ongoing process, while the use of perfective aspect is closely associated with telic verbs that indicate actions with endpoints or end results. In particular, in early child English, the progressive marker *-ing* is highly restricted to activity verbs, the perfective/past marker *-ed* restricted to telic verbs, and the third person singular *-s* restricted to stative verbs (Bloom et al. 1980; Brown 1973; Clark 1996; Shirai 1991). Our networks, having taken in input patterns based on realistic parental speech, behaved in the same way as children do. For example, at Input Age 1;6, the networks produced *-ing* predominantly with activity verbs (72% for simulations based on DISLEX, 73% for those based on DevLex-II), *-ed* overwhelmingly with telic verbs (75% for both simulations), and *-s* with stative verbs (100% for DISLEX, 57% for DevLex-II; *-s* with stative verbs combinations were generally rare — 5–7 cases in our simulations — so the percentage discrepancy is not too meaningful). Such associations were strong at all four stages (especially for *-ing* and *-ed*), but they tended to become weaker over time.

Interestingly, when we analyzed the actual input to our networks (based on parental speech), we found similar patterns. Tables 1 and 2 also presented the percentages of the use of suffixes with different verb types in the input data for DISLEX and DevLex-II, respectively. An analysis of these tables indicate that in the input data there are also clear associations between -*ing* and activity verbs, -*ed* and telic verbs, and that these associations are strong throughout the four stages, as also found previously by Shirai (1991) and Olsen et al. (1998). The association between -*s* and stative verbs is not so obvious, especially in the input data for DISLEX. However, when we examined the larger corpus which was used for DevLex-II, we found that such association also exists. The degree to which the networks' production matches up with the input patterns indicates that the two networks in our study were able to learn on the basis of the information of the co-occurrences between lexical aspect (verb types) and grammatical aspect (verb morphology). This learning ability was due to the networks' use of Hebbian associative learning in computing if the semantic, phonological, and phonemic properties of a verb co-occur and how often they do so.

Note that in Tables 1 and 2, the patterns of the parental input and the network productions are consistent and similar, but not identical. It shows that the two networks' productions were not simply verbatim mimics of what's in the input by recording each individual word and suffix and their co-occurrence. This is important and shows that our networks have their own productive control of the relevant linguistic information. Our results indicate that the associations between verb types and suffixes are stronger in the networks' productions than they are in the input to the networks (at least for the early training stages). It shows that the two networks, like children, behave more restrictively than what is in the input with respect to the correlations between lexical aspect and grammatical aspect.

To see the data more clearly, we illustrate the patterns with Figure 2 to show the percentages of the use of suffixes with different verb categories in network's productions (Figure 2a) and in parental input data (Figure 2b) for DevLex-II at Input Age 1;6. Comparing Figures 2a and 2b, we can see that the network's production patterns are consistent with patterns in the parental input data, but the network has more restricted use of the suffixes.

3.2. *Emergence of structured semantic representations*

Elsewhere we have proposed an account of semantic development as an emergent process in which semantic features are connected in a system to support lexical categories, like in the formation of semantic cryptotypes (Li and MacWhinney 1996; Li 2003; Li et al. 2004; Hernandez et al. 2005; see also Rogers and McClelland 2004 for similar discussion).

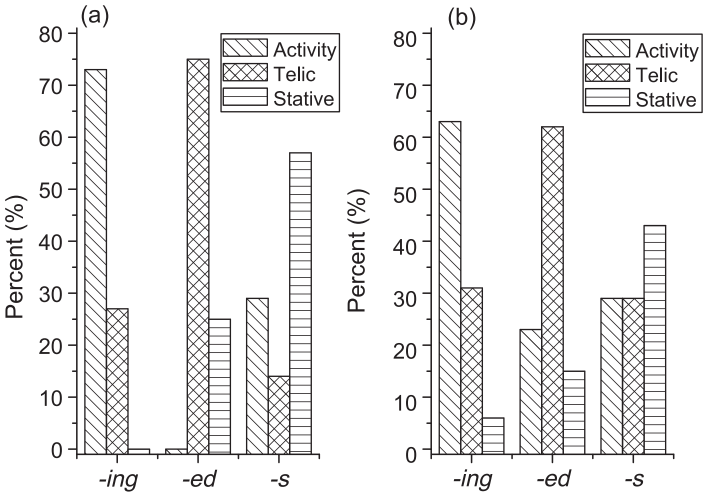

Figure 2. *Percentages of the use of suffixes with different verb types at Input Age 1;6 in (a) network production in DevLex-II, and (b) parental input data. Data are based on Table 2.*

The basic idea is that a given verb may be represented with multiple linguistic features, and the features themselves often co-occur and overlap in different verbs. For example, the verb *screw* may be viewed as having both a meaning of circular movement and a meaning of binding or locking, and the verb *zip* may be viewed as sharing both the "binding/locking" meaning and the "covering" meaning. Moreover, both *screw* and *zip* involve hand movements. Features may also vary in the strength with which they are represented in different verbs. For example, the verb *wrap* may be viewed as having the covering meaning. However, in some cases, the action of wrapping may also involve circular movements. Children may acquire such complex feature-to-verb relationships through statistical analyses of the three kinds of co-occurrences as discussed earlier (co-occurrences of verbs with situational contexts, with other words, and co-occurrences of particular grammatical morphemes with semantic features), leading to feature-based organization of verb categories. In the simulations here, we provided our networks with verb that are represented with multiple semantic features (lexical co-occurrence constraints, extracted by HAL or WCD), and we wanted to see how categories of lexical aspect could emerge from the self-organizing learning process.

As discussed earlier, a particularly useful property of self-organizing feature maps is that the statistical structures in the representations can be clearly visualized as activity bubbles or patterns of activity on a two dimensional map in a topography-preserving structure. Given that both DISLEX and DevLex-II represented semantic information from the high-dimensional space of verb usage in parental input, we hypothesize that verbs with similar aspectual properties should cluster together on the feature map. Figure 3 present a snapshot of DevLex-II's self-organization of the semantic representations of verbs (with suffixes) at the end of the learning process (i.e., Stage 4, Input Age 3;0).

An examination of this map shows that the network has clearly developed structured semantic representations that correspond to different lexical aspect categories. It formed clear clusters of verbs by mapping verbs with similar combination of semantic features onto nearby regions of the map, and several interesting observations can be made: (1) The most obvious structure of the map is that the words can be roughly divided into three main clusters according to the suffix that a verb stem takes, -*ing*, -*ed*, or -*s* (see Figure 3). (2) Within each area, there are also groups that correspond to categories of lexical aspect such as telic verbs, activity verbs, and stative verbs. For example, towards the lower left-hand corner of the larger part of -*s* area (the pale gray area), stative verbs, like *loves*, *knows*, *likes*, *wants*, and *needs* are mapped to the same region. Another example can be found in the -*ing* area (the area without shading): although most verbs clustered in this area are activity verbs such as *working*, *sitting*, *crawling*, *walking*, *sleeping*, *etc.*, there is also a cluster of telic verbs (at the middle-to-lower portion of the map) such as *wiping*, *fixing*, *hitting*, *putting*, *cutting*, *throwing*, *making*, and *getting.* (3) The distribution of lexical aspect is closely related to the distribution of grammatical aspect. Not only it is the case that the -*ed* area contained mostly telic verbs and the -*ing* area mostly activity verbs, but also telic verbs that take -*ing* were closer to the -*ed* area (e.g., *going*, *jumping*, *messing*, *picking* and *cleaning*, all bordering the -*ed* area). (4) Verbs with the same stem but different suffixes are also often mapped to regions not far away from one another, for example, *fixing* and *fixed*, *pushing* and *pushed*, *turns* and *turned* at the middle area of the map, and *playing* and *played* at the lower right corner of the map.

The emergence of structured semantic representations on our model can also be verified by a simple method called *k*-nearest neighbor (*k*-NN) algorithm (Duda et al. 2000). As a classical method in the field of pattern recognition for classifying objects into different classes, the basic idea of k-NN is to predict the class of a point in a dataset according to the most frequent class label of its *k* nearest neighbors. Implementing this method in our semantic map (see also Li et al. 2004), we can evaluate if a verb in our lexicon was mapped to a node close in Euclidean distance to other verbs belonging to the same class. This allows us to have a rough

Figure 3. *Emergent representations in the semantic map of DevLex-II after Input Age 3;0. Differently shaded regions indicate different aspect categories corresponding to different suffixes* -ed, -ing, *and* -s*. Within each category, verbs with the same lexical aspect are often grouped together.*

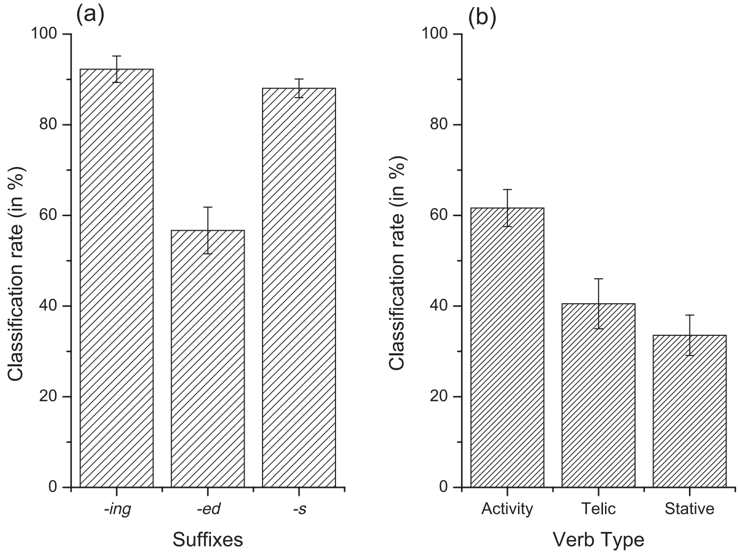

Figure 4. *Classification rates calculated by a 5-NN classifier according to lexical and grammatical representations of the verbs on the semantic map of DevLex-II. Classifications are based on: (a) the suffix that a verb stem takes:* -ing*,* -ed*, and* -s*; (b) the lexical aspect of the verb: activity, telic, and stative. The error bars indicate the standard deviations based on 5 trials.*

idea of the overall compactness of different lexical classes. Here, we conducted a 5-NN analysis of verb representations on the semantic map according to the suffix a verb stem takes, -*ing*, -*ed*, or -*s*. As shown in Figure 4a, the semantic map has developed clear clusters for different suffixes: for the category of -*ing*, the classification rate is about 92%, which means that 92 percents of verbs suffixed with -*ing* are located within a nearest neighborhood according to k-NN; for the categories of -*s* and -*ed*, the classification rates are 88% and 60%, respectively. We also conducted a 5-NN analysis of the verbs according to their lexical aspect properties: as shown in Figure 4b, the classification rates for activity verbs, telic verbs, and stative verbs are 61%, 41%, and 34%, respectively. The relatively low classification rates of the verb categories, compared with those of suffix categories, indicate that the organization of verb meanings according to lexical aspect is subordinate to the organization of verb suffixes on the map. In general, these results from the 5-NN analysis are consistent with our visual analyses of the semantic maps.

These observations lead us to conclude that the map has formed structured representations for grammatical aspect markers, such as -*ed*, -*ing*,

-*s*, and that the interaction between grammatical aspect and lexical aspect is reflected in the correlation between grammatical morphology and verb types, and in the categories of lexical aspect such as activity, telic, and stative. Of course, the network's representation structure on the semantic map is still not perfect, due to the complexity of semantic features of verbs and the difficulty in expressing the complex information from the higher dimensional inputs on the two dimensional map. In fact, the self-organization process can be considered as one that extracts the most important components of information from a dataset and expresses the outcome in compressed format, such as in Principal Component Analysis (PCA). During this process as the original information in a high-dimensional space is projected to a two dimensional space, some loss of information is inevitable, which explains why the two clusters of -*s* words never group together on the map and why not all verbs with the same stem are mapped together.

Finally, in Figure 5, we present sketches of the lexical aspect categories as they emerged on the semantic map at each training stage (from age 1;6 through age 3;0). These snapshots from Study 2 (similarly in Study 1; see Li and Shirai 2000) show us clearly the development of structures corresponding to different lexical categories on the map, and how they were gradually constructed based on linguistic input without a priori structure. At early stages of training, the structure of the map was relatively simple and easy to change, due to limited linguistic input at these stages. The overall representation was sparse, and the network was still in an unstable state. As new words were added into the lexicon, the semantic representation evolved into a more complex structure with stable basis. We can see that after Stage 2, the basic clusters of the map became consolidated; new words could not significantly change those distributions, but could only be added to the regions established by other words that shared similar features. For example, comparing the upper right corner of the map at the end of training stages 3 and 4, we can see that newly acquired words *looks* and *sounds* were filled into the area near the words *feels* and *tastes*.

The emergence and organization of categories on the semantic map could also be explicitly monitored by a map reorganization measure as shown in Figure 6 (see Li et al. 2004 for a description of the method). In particular, for any pair of two adjacent stages (e.g. stage of Age 2;0 and Age 2;6), we evaluated map reorganization as the Euclidean distance of the same word in the two maps, averaged over all words currently present in the map. For example, in the two maps for the stage of Age 2;0 (with 100 verb types) and the stage of Age 2;6 (with 154 verb types), we compared only the 100 verbs that were common to both maps. Results of this procedure, as shown in Figure 6, indicate that the map underwent signif-

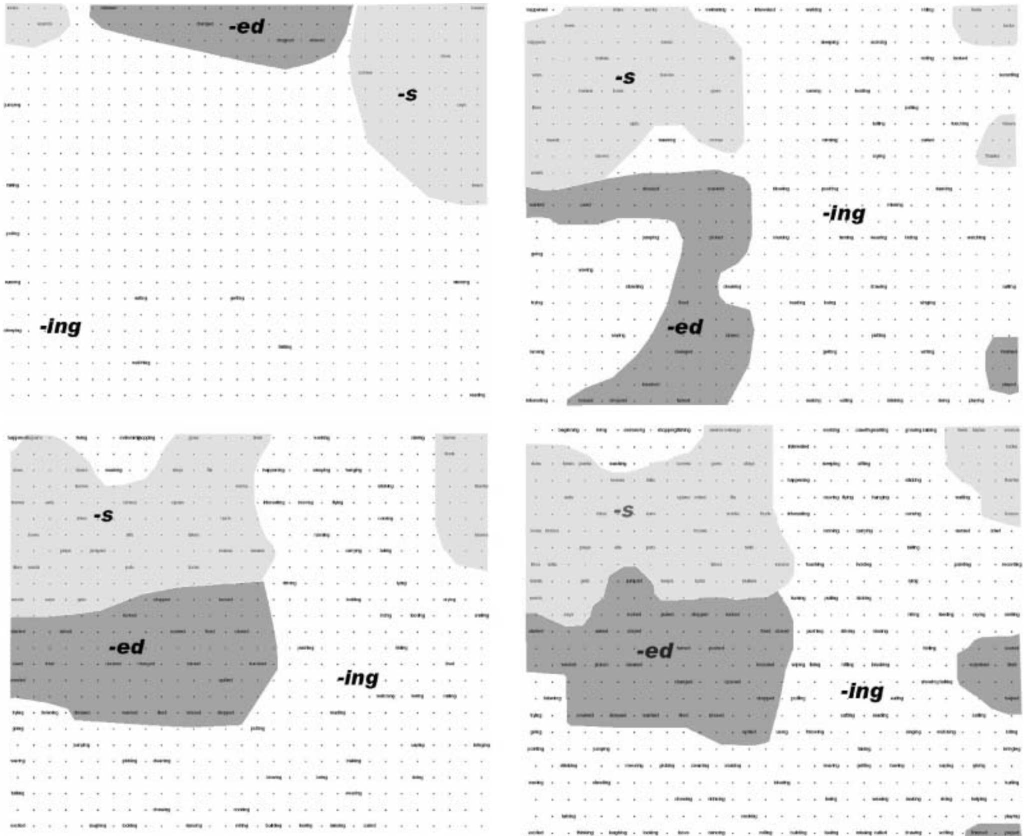

Figure 5. *Development of basic structure on the semantic layer of DevLex-II across the four stages. (a), (b), (c), and (d) indicate Input Ages 1;6, 2;0, 2;6, and 3;0, respectively.*

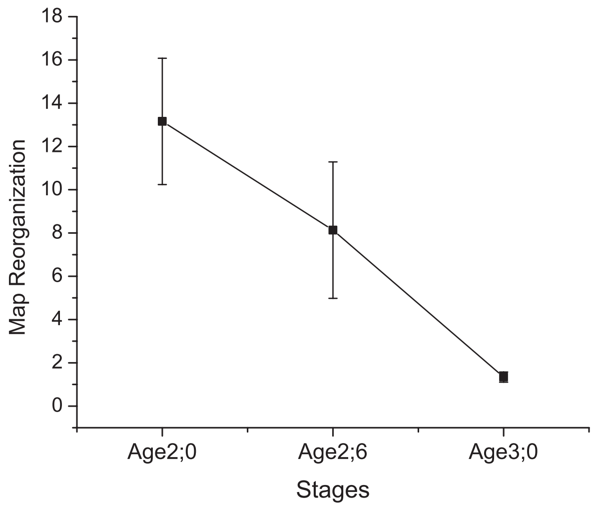

Figure 6. *Semantic maps' reorganization as a function of time (training stages), computed as the average amount of word shifts in the map's underlying grid. Each map was compared with its predecessor at the previous stage (the index of x-axis refers to the successor stage), by comparing the positions of the words common to both maps. The results are based on 5 trials.*

icant reorganization at initial stages of development, but reorganization gradually disappeared toward the later stages. The average amount of word shifts early on was large, but became smaller over time. Thus, the map's ability to radically restructure semantic representation decreased as the basic structure of meaning became consolidated, and this explains why later learned words could only be simply added to the existing structure of the network, as shown in Figure 5.

The results from our modeling offer a new way of thinking about the representation of lexical aspect and its interaction with grammatical aspect. Verbs in a lexical aspect category form complex relationships, in that they vary in (1) how many linguistic features are relevant to the category, (2) how strongly each feature is activated in the representation of the category, and (3) how features overlap with each other across category members. For example, *spill* may be viewed as indicating both a punctual and a resultative meaning; *close* may involve both a change of state and a completive meaning; and the feature "punctual" may be represented more strongly in *jump* than in *fall*: in a natural setting a single jump occurs instantaneously, whereas falling need not (e.g., we could still say that a leaf fell from a tree even if it drifted down slowly). For example, in Figure 3, we can find that the word *jump*, with its progressive marker *-ing*, is much closer than *falling* to the *-ed* area.[8] With varying degrees of connections from semantic features to verb forms, verbs can form clusters or categories that differ overall in lexical aspect. Traditional analytical methods from linguistics and psycholinguistics are much less effective, if not impossible, in dealing with these complex semantic relationships. By contrast, connectionist models that rely on distributed feature representations and nonlinear learning are ideally suited to accounting for the properties of feature overlapping and weighted feature composition. DISLEX and Devlex-II models that we discuss here provide clear examples of how we may solve semantic problems via weighted feature composition (see also Li and MacWhinney 1996; Li 2003, 2006; Li et al. 2004).

3.3. *Role of Hebbian learning*

The above results suggest that the learning of grammatical suffixes is not simply the learning of a rule such as adding *-ing* or *-ed* to a verb to mark the progressive aspect or the perfective aspect, but the accumulation of associative strengths that hold between a particular suffix and a complex set of semantic features distributed across verb forms (which support the emergence of a lexical aspect category). This learning process can be best described as a statistical, probabilistic process in which the learner implicitly tallies and registers the frequency of co-occurrences (strengthening what goes with what) and co-occurrence constraints (inhibiting what does not go with what) among the semantic features, lexical forms, and tense-aspect suffixes.

How do our networks accomplish this learning? The co-occurrence-and-constraint process is modeled in our networks by Hebbian learning of the associative connections between forms and meanings. For example, in DevLex-II, the phonological, semantic, and phonemic sequence representations of a verb are co-activated in three separate feature maps in our network, along with the corresponding suffixes with which the verb co-occurs in the input. Hebbian learning, according to which the associative strength is a function of the degree of co-activation of the two items in question, determines how strongly the connections form between the form, the meaning, and the suffix at any given point during the learning process.

Along with the self-organization of forms, meanings and morphologies on the corresponding feature maps, Hebbian learning provides the network with focused pathways from forms to meanings or from meanings to forms (including the suffixes). When concentrated patterns of activity have formed on the feature maps and on their associative connections, the network can readily "comprehend" new input (e.g., from phonological form to meaning) or "produce" the output (e.g., from meaning to phonemic sequence). This comprehension or production process can be a quick mapping, especially when the new input is sufficiently similar to members of an existing cluster and falls within the area of that cluster. For example, on the semantic map, most activity verbs are grouped together in the area corresponding to progressive aspect marker *-ing*, and the strong associative pathways are established between these verbs and the suffix *-ing* on the phonemic sequence map; so when a new activity verb is sent into the network, it is highly likely to be mapped to the suffix *-ing*, using the existing, learned pathway between the feature maps. In contrast, non-prototypical associations (e.g., between stative verbs and *-ing* or *-ed*) have formed much weaker associative pathways because such associations occur only infrequently in the input. Thus, a new stative verb will have fewer chances to be mapped to *-ed* or *-ing* initially without much learning. This process might also explain why our network produced stronger associations between particular verb categories and particular suffixes than the strength of the actual associations in the input (see Section 3.1), because the pathways between these categories and suffixes are prototypical and serve the "magnet" role for incoming words.

To make this view clearer, we can draw the strength of associative links from each node on the semantic map to the suffix *-ing* represented on the output phonemic sequence map at the end of training (Stage 4). The phonemic representation of suffix *-ing* is /ɪŋ/, so on the phonemic map, the suffix *-ing* is represented by the combination of nodes corresponding to phonemes /ɪ/ and /ŋ/. In Figure 7, we can see clearly that some areas on the semantic map have much stronger associations with *-ing* than other areas. (The strength of the associative links is represented by the gray scales on the map: the darker it is, the weaker the associative strength is). Comparing Figures 3 and 7, we can see that the light gray areas (indicating strong associations) match well with areas corresponding to *-ing* on the semantic map, and the charcoal grey areas (indicating weak associations) match with the areas corresponding to *-ed* and *-s*.

Thus, DISLEX and DevLex-II not only allow for the formation of categories of lexical aspect through self-organization in the feature maps, but also provide a mechanism for the formation of prototypical associations between lexical aspect and grammatical aspect via Hebbian learning. Moreover, as we will argue here, the same Hebbian learning mechanism also accounts for the dilution or relaxation of the strong associations.

As seen in Tables 1 and 2, our networks' production shows strong associations between lexical aspect and grammatical suffixes, and the strong

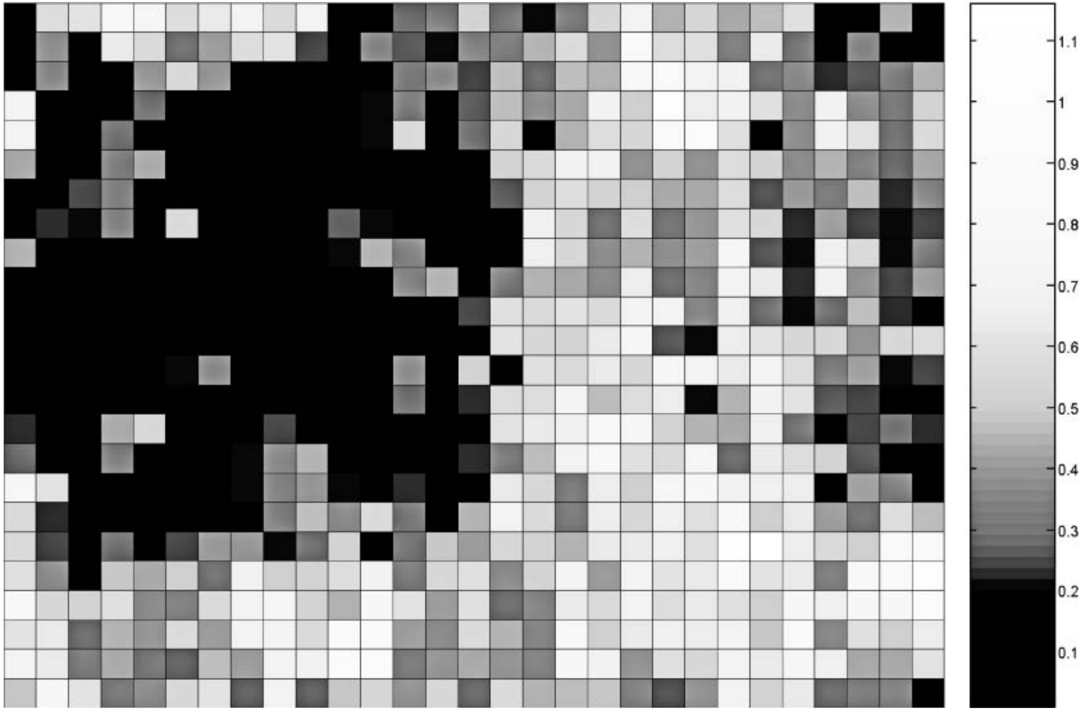

Figure 7. *Strength of associative links from each node on the semantic map to the grammatical suffix* -ing *on the phonemic sequence map (the combination of nodes responding to phonemes /ɪ/ and /ŋ/). The gray scales indicate the amount of strength of associative links. The darker the gray scale of a node is, the weaker the association between this node and the suffix* -ing.

Table 3. *Number and percentage verbs with multiple suffixes across input age groups in DevLex-II's production and in parental input data*

	Age 1;6	Age 2;0	Age 2;6	Age 3;0
Network production	0/22* (0%)	21/75 (28%)	50/124 (40%)	83/169 (49%)
Parental Input data	18/62 (29%)	35/100 (35%)	68/154 (44%)	97/184 (53%)

* a out of b verbs have more than one suffixes

associations weaken over time. This pattern is consistent with empirical results from child language (see detailed discussion in Chapter 4 in Li and Shirai 2000). How does the weakening of strong associations happen? To answer this question, we analyzed DevLex-II's production at different stages. First, we calculated the number of verbs with more than one suffixes (see Table 3), and found that early on, there were few multiple suffixations with the same verbs (zero at Input Age 1;6), whereas later on there were many more such cases (83 out of 169 at Input Age 3;0). This result suggests that multiple suffixations might be a driving force for the learner to break from the strong associations to more diverse associations between lexical aspect categories and grammatical suffixes. Second, verbs that took multiple suffixes were mostly telic and activity verbs but few stative verbs (among the 37 acquired verb stems with multiple suffixes at the end of Input age 3;0, 14 were telic verbs, 19 were activity verbs, but only 4 were stative verbs). This indicates that telic and activity verbs are more susceptible to changes away from the strong associations than are stative verbs. Third, there are four possibilities of multiple suffixation: a given verb stem might take both -*ing* and -*ed*, both -*ing* and -*s*, both -*ed* and -*s* or all three suffixes. In DevLex-II, multiple suffixations with -*ing* and -*s* or -*ing* and -*ed* occurred more frequently than with -*ed* and -*s* or all three suffixes (e.g., for Input age 3;0, the numbers were 19, 12, 1 and 5 verb stems, respectively). This shows that -*ed* and -*s* may be less compatible within a given verb than are -*ing* and -*ed* or -*ing* and -*s*.

These analyses demonstrate a dynamical picture of the learning system in the development from strong prototypical association between lexical aspects and grammatical aspects to more diverse associations. This picture is consistent with empirical data from child language (Shirai and Andersen 1995), although how the three patterns of multiple suffixation discussed above would map directly to patterns in child language needs to be further tested in empirical research. Clearly, Hebbian learning can account for the developmental process in the weakening of the strong prototypical associations. This development involves a transitional process in the restructuring of the mappings among phonological, semantic, and morphological patterns. The restructuring in our network is based on the network's ability to reconfigure associative pathways — in particular, to form new pathways between suffixes and verbs — and its ability to weaken or eliminate old pathways that were the basis of the prototypical associations. The adjustment of associative connections through Hebbian learning is proportional to how strongly the nodes in the corresponding maps are co-activated. For example, if given morphological nodes (e.g. nodes corresponding to /ɪ/ and /ŋ/, which together represent the suffix *-ing* on the phonemic sequence map) and a given semantic node have more chances to become co-activated, the strength of their associative links are correspondingly increased. Conversely, if a given morphological node and a given semantic node have fewer chances to become co-activated, the strengths of their associative links are correspondingly decreased. This is the reason why we see some areas on the semantic map having strong associative links with the suffix *-ing*, while other areas having very weak or no associations with it.

This learning procedure is basically input-driven, and therefore our default assumption is that the transition from strong prototypical associations to diverse mappings in the network is triggered by changes in the distributional properties in the input (but recall that our network does not simply mimic input verbatim, see discussion in 3.1). To verify this assumption, we also analyzed the parental speech in our CHILDES corpus that served as the input to our network. Our analyses indicated that almost all observations in the output of the network are also true in the input. These include (1) the number of multiple suffixed verbs increases with time (see Table 3); (2) telic and activity verbs tend to have multiple suffixes (for Input Age 3;0, 15 telic verb stems and 23 activity verb stems had more than one suffix, compared to 7 of stative verb stems); (3) *-ing* and *-s* or *-ing* and *-ed* are more likely to occur on same verb stems (for Input Age 3;0, 22 verb stems had both suffixes *-ing* and *-s*, 16 had both *-ing* and *-ed*, but only 4 had both *-s* and *-ed*, and 5 had all the three suffixes). The high correlation between patterns in the input and those in the output, and the network's development therein, suggest that the network has the ability to establish the necessary associative pathways for the mapping between verb meanings and grammatical morphology, through Hebbian learning.

As a final note, self-organization and Hebbian learning can also shed light on the overgeneralization issue in the acquisition of the English past tense. Our simulations point to two general conclusions with respect to past tense overgeneralization. First, given that children's early use of *-ed* is restricted to the aspectual meaning of telic verbs, overgeneralizations of *-ed* will not occur across the board for all types of verbs but will rather be restricted to telic verbs initially (see also Shirai 1991). Second, overgeneralizations of *-ed* are not only semantically restricted, but also sometimes semantically motivated. In our network models, semantic pathways formed via Hebbian learning can provide the basis for the production of overgeneralization errors. For example, in Figure 3, in the middle of the semantic map, suffixed telic words *locked* and *closed* were mapped on to nearby region since they share similar semantic properties. During learning, the phonology, the semantics, and the phonemic output representations (including the representation of *-ed*) of *locked* were co-activated, and similar co-activations also happened to *closed*. At the end of training, the BMU nodes of the words *locked*, *closed* and their neighboring nodes developed strong associations with the phonemic representation of *-ed*. Now, imagine what will happen when a telic irregular verb *shut* is input to the network: (1) the verb is highly likely to be mapped onto nearby regions of *locked* and *closed*, since all the three words share similar semantic features, and (2) the network could associate the semantics of *shut* with *-ed* because of the strong associations between the semantic region and the suffix, even though the model has only learned the associations for *lock/close* and *-ed*, and not *shut* and *-ed*. Although our main focus in our study is on the relationship between lexical aspect categories and aspect morphology, and as such in our simulations we did not actually test the irregular past tense forms, our hypothetical overgeneralization example of *shut* comes as a natural result of the structure of the network's semantic representations (which in turn is due to self-organization) and of the associative mappings of semantic features, lexical forms, and morphological devices (due to Hebbian learning) (see results for DISLEX discussed in Li and Shirai 2000: Ch. 7).

4. Conclusion

In this paper we presented two self-organizing neural networks, DISLEX and DevLex-II, to see how they can be used successfully to model the acquisition of lexical and grammatical aspect, and to provide insights into issues regarding the role of linguistic input, the representation of lexical categories of verbs, and the development of prototypical to nonprototypical associations. Self-organization and Hebbian learning in these networks are two important computational principles that can account for the psycholinguistic processes in the acquisition of lexical and grammatical aspect. Our simulations demonstrate: (1) the networks are able to display patterns of association as observed in empirical acquisition studies, on the basis of its analyses of input characteristics; (2) self-organization of the semantic structure of verbs leads to the formation of lexical aspect categories and grammatical aspect categories, on the basis of the network's analysis of the complex feature-to-verb, and verb-to-morphology relationships in language use; (3) focused associative pathways established by Hebbian learning between meanings and morphology lead to particularly strong associations between lexical aspect and grammatical aspect, thereby to undergeneralized patterns of grammatical morphology as observed in early child language; and (4) associative links between forms and meanings along with incremental vocabulary growth lead to diverse mappings, first with a relatively small number of words and morphemes and then spreading to others, which accounts for how the strong associations gradually weaken or dissolve in children's language.

As we mentioned earlier, the goal of our study is to determine whether simple but biologically plausible computational principles in self-organizing neural networks can account for empirically observed patterns in children's acquisition of lexical aspect and grammatical morphology. In particular, we wanted to see if our networks, without a priori stipulations about the structure of meaning or concept, can display the early strong associations between lexical aspect and grammatical aspect, and how they can also move away from the strong associations over time to approach adult patterns of aspect use. Our simulations here, along with other published studies (Farkas and Li 2001, 2002; Li 2003; Li et al. 2004; Li et al. 2007; Zhao and Li 2005), clearly serve to demonstrate the utility of self-organizing neural networks for unraveling mechanisms underlying lexical and grammatical acquisition, particularly with respect to the role of input and emergent lexical categories. Our study may also serve to generate interests in further empirical studies against which we can compare detailed patterns in our modeling results (e.g., the multiple suffixation patterns). Finally, one must note that in the actual learning situation the child has available a myriad of other types of information grounded in visual or other perceptual interactions with parents or caregivers, whereas our network does not have these types of information (see Yu and Smith [2006] for a recent analysis). What we have demonstrated here is the ability of artificial neural networks based on self-organization and Hebbian learning to mechanistically account for developmental patterns in language acquisition even without real-world visual or perceptual information.

Received 7 March 2006
Revised version received
29 October 2006

University of Richmond
Pennsylvania State University

Notes

* This research was supported by a grant from the National Science Foundation (BCS-0131829) to PL. Correspondence address: Ping Li, Department of Psychology, Pennsylvania State University, 614 Moore Building; University Park, PA 16802-3106, USA. E-mail: pul8@psu.edu.

1. Note that *-ed* marks both past tense and perfective aspect in English, just as *-s* marks both present tense and habitual aspect. Separate affixes are often used in other language for tense and aspect.
2. Inputs of SOM are often multivariate data with the form of high-dimensional vectors. These vectors can be considered as points distributed around a high-dimensional hyperspace. SOM can project these inputs onto nodes on a two-dimensional map. At first, an area of nodes ("activity bubble") will respond to an input, and then the bubbles will gradually reduce in size, until only one node is active. Details of SOM are further discussed in 2.1.1.
3. A preliminary version of the first model was reported in Li (2000) and Li and Shirai (2000: Ch. 7).
4. Although the irregular verbs were also included in the simulations based on DISLEX, we did not examine their relationship with regular past tense/perfective forms (but see our discussion in Section 3.3 on the issue of overgeneralization).
5. In our current simulations we used the real-value vectors.
6. Unlike in Study 1, bare verb forms were excluded from our simulations, as well as irregular past tense forms and nonverbs. Exclusion of these forms simplifies the simulation task and makes the analysis more tractable, given that DevLex-II involves a more complex network architecture than DISLEX. Moreover, our major goal here is to see whether the use of verbal suffixes is correlated with the lexical aspect of verbs, and as such our simulations are focused on suffixed verbs.
7. Our analyses below deviate slightly from a strict Vendlerian four-way classification, because accomplishment and achievement verbs are often difficult to separate without an extensive analysis of the sentence and speech context. Thus in what follows telic verbs include both accomplishments and achievements.
8. Jump can also be construed iteratively, so that *jumping* refers to a series of jumps, which is why we may see that children use *jumping* more frequently than *jumped* in actual speech.

References

Behrend, Douglas A. 1995. Processes involved in the initial mapping of verb meanings. In Michael Tomasello & William E. Merriman (eds.), *Beyond names for things: Children's acquisition of verbs*, 251–273. Hillsdale, NJ: Lawrence Erlbaum.

Behrend, Douglas A., L. Lynn Harris & Kelly B. Cartwright. 1995. Morphological cues to verb meaning: Inflections and the initial mapping of verb meanings. *Journal of Child Language* 22. 89–106.

Bickerton, Derek. 1981. *Roots of language.* Ann Arbor, MI: Karoma.

Bickerton, Derek. 1984. The language bioprogram hypothesis. *Behavioral and Brain Sciences* 7. 173–188.

Bloom, Lois, Karin Lifter & Jeremy Hafitz. 1980. Semantics of verbs and the development of verb inflection in child language. *Language* 56. 386–412.

Brown, Roger. 1973. *A first language.* Cambridge, MA: Harvard University Press.
Burgess, Curt & Kevin Lund. 1997. Modeling parsing constraints with high-dimensional context space. *Language and Cognitive Processes* 12. 1–34.
Clark, Eve V. 1996. Early verbs, event-types, and inflections. In Carolyn E. Johnson & John H. V. Gilbert (eds.), *Children's language*, vol. 9. Mahwah, NJ: Lawrence Erlbaum.
Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
Dale, Philip S. & Larry Fenson. 1996. Lexical development norms for young children. *Behavior Research Methods, Instruments & Computers* 28. 125–127.
Duda, Richard O., Peter E. Hart, & David G. Stork. 2000. *Pattern classification*, 2nd edn. New York, NY: John Wiley.
Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.
Farkas, Igor & Ping Li. 2001. A self-organizing neural network model of the acquisition of word meaning. In Erik M. Altmann, Axel Cleeremans, Christian D. Schunn & Wayne D. Gray (eds.), *Proceedings of the fourth international conference on cognitive modeling*, 67–72. Mahwah, NJ: Lawrence Erlbaum.
Farkas, Igor & Ping Li. 2002. Modeling the development of the lexicon with a growing self-organizing map. In H. John Caulfield et al. (eds.), *Proceedings of the Sixth Joint Conference on Information Science*, 553–556. Durham, NC: Association for Intelligent Machinery.
Gathercole, Susan E. & Alan D. Baddeley. 1993. *Working memory and language.* Hillsdale, NJ: Erlbaum.
Gleitman, Lila. 1990. The structural sources of verb meaning. *Language Acquisition* 1. 3–55.
Gupta, Prahlad & Brian MacWhinney. 1997. Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language* 59. 267–333.
Harner, Lorraine. 1981. Children talk about the time and aspect of actions. *Child Development* 52. 498–506.
Haykin, Simon. 1999. *Neural networks: A comprehensive foundation*, 2nd edn. Upper Saddle River, NJ: Prentice Hall.
Hebb, Donald O. 1949. *The organization of behavior: A neuropsychological theory*. New York, NY: John Wiley.
Hernandez, Arturo, Ping Li & Brian MacWhinney. 2005. The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences* 9. 220–225.
James, Daniel L. & Risto Miikkulainen. 1995. SARDNET: A self-organizing feature map for sequences. In Gerald Tesauro, David S. Touretzky & Todd K. Leen (eds.), *Advances in neural information processing* systems *(NIPS '94)*, 577–584. Cambridge, MA: MIT Press.
Klein, Wolfgang, Ping Li & Henriëtte Hendriks. 2000. Aspect and assertion in Mandarin Chinese. *Natural Language and Linguistic Theory* 18. 723–770.
Kohonen, Teuvo. 2001. *The self-organizing maps*, 3rd edn. Berlin: Springer.
Ladefoged, Peter. 1982. *A course in phonetics*, 2nd ed. San Diego, CA: Harcourt Brace.
de Lemos, Claudia. 1981. Interactional processes and the child's construction of language. In Werner Deutsch (ed.), *The child's construction of language*. London: Academic Press.
Li, Ping. 2000. The acquisition of tense-aspect morphology in a self-organizing feature map model. In Lila R. Gleitman & Aravind K. Joshi (eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*, 304–309. Mahwah, NJ: Lawrence Erlbaum.

Li, Ping. 2003. Language acquisition in a self-organizing neural network model. In Philip T. Quinlan (ed.), *Connectionist models of development: Developmental processes in real and artificial neural networks.* 115–149. Hove & Brighton: Psychology Press.

Li, Ping. 2006. In search of meaning: The acquisition of semantic structure and morphological systems. In June Luchjenbroers (ed.), *Cognitive linguistics investigations: A cross languages, fields and philosophical boundaries*, 109–137. Amsterdam: John Benjamins.

Li, Ping & Melissa Bowerman. 1998. The acquisition of lexical and grammatical aspect in Chinese. *First Language* 18, 311–350.

Li, Ping, Curt Burgess & Kevin Lund. 2000. The acquisition of word meaning through global lexical co-occurrences. In Eve V. Clark (ed.), *Proceedings of the thirtieth Stanford child language research forum*, 167–178. Stanford, CA: Center for the Study of Language and Information (CSLI).

Li, Ping & Igor Farkas. 2002. A self-organizing connectionist model of bilingual processing. In Roberto Heredia & Jeanette Altarriba (eds.), *Bilingual sentence processing*, 59–85. Amsterdam: North-Holland Elsevier.

Li, Ping, Igor Farkas & Brian MacWhinney. 2004. Early lexical development in a self-organizing neural networks. *Neural Networks* 17. 1345–1362.

Li, Ping & Brian MacWhinney. 1996. Cryptotype, overgeneralization, and competition: A connectionist model of the learning of English reversive prefixes. *Connection Science* 8. 3–30.

Li, Ping & Brian MacWhinney. 2002. PatPho: A phonological pattern generator for neural networks. *Behavior Research Methods, Instruments, and Computers* 34. 408–415.

Li, Ping & Yasuhiro Shirai. 2000. *The acquisition of lexical and grammatical aspect*. Berlin & New York: Mouton de Gruyter.

Li, Ping, Xiaowei Zhao & Brian MacWhinney. 2007. Dynamic self-organization and early lexical development in children. *Cognitive Science* 31. 581–612.

MacWhinney, Brian. 1998. Models of the emergence of language. *Annual Review of Psychology* 49. 199–227.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, Brian. 2001. Lexicalist connectionism. In Peter Broeder & Jaap M. Murre (eds.), *Models of language acquisition: Inductive and deductive approaches*, 9–32. Oxford: Oxford University Press.

MacWhinney, Brian & Jared Leinbach. 1991. Implementations are not conceptualizations: Revising the verb learning model. *Cognition* 40. 121–157.

Mapstone, Elizabeth R. & Paul L. Harris. 1985. Is the English present progressive unique? *Journal of Child Language* 12. 433–441.

Maratsos, Michael & Mary Anne Chalkley. 1980. The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In Keith E. Nelson (ed.), *Children's language*, vol. 2, 127–214. New York, NY: Gardner.

McShane, John & Stephen Whittaker. 1988. The encoding of tense and aspect by three- to five-year-old children. *Journal of Experimental Child Psychology* 45. 52–70.

Miikkulainen, Risto. 1993. *Subsymbolic natural language processing: An integrated model of scripts, lexicon, and memory.* Cambridge, MA: MIT Press.

Miikkulainen, Risto. 1997. Dyslexic and category-specific aphasic impairments in a self-organizing feature map model of the lexicon. *Brain and Language* 59. 334–366.

Miikkulainen, Risto. 1999. The DISLEX simulator. http://www.cs.utexas.edu/users/nn/pages/software/dislex.2.1.tar.Z.

Naigles, Letitia. 1990. Children use syntax to learn verb meaning. *Journal of Child Language* 17. 357–374.

Olsen, Mari B., Amy Weinberg, Jeffrey A. Lilly & John E. Drury. 1998. Mapping innate lexical features to grammatical categories: Acquisition of English -*ing* and -*ed*. In Morton Ann Gernsbacher & Sharon J. Derry (eds.), *Proceedings of the twentieth annual conference of the Cognitive Science Society*, 794–799. Mahwah, NJ: Lawrence Erlbaum.

Ritter, Helge & Teuvo Kohonen. 1989. Self-organizing semantic maps. *Biological Cybernetics* 61. 241–254.

Rogers, Timothy T. & James L. McClelland. 2004. *Semantic cognition: A parallel distributed processing approach.* Cambridge, MA: MIT Press.

Shirai, Yasuhiro. 1991. Primacy of aspect in language acquisition: Simplified input and prototype. Los Angeles, CA: University of California, Los Angeles dissertation.

Shirai, Yasuhiro. 1994. On the overgeneralization of progressive marking on stative verbs: Bioprogram or input? *First Language* 14. 67–82.

Shirai, Yasuhiro. 1998. The emergence of tense-aspect morphology in Japanese: Universal predisposition. *First Language* 18. 281–309.

Shirai, Yasuhiro & Roger W. Andersen. 1995. The acquisition of tense-aspect morphology: A prototype account. *Language* 71. 743–762.

Shultz, Thomas. 2003. Semantic boost on episodic associations: An empirically-based computational model. Cambridge, MA: MIT Press.

Silberman, Yaron, Shlomo Bentin & Risto Miikkulainen. 2007. Semantic boost on episodic associations: An empirically-based computational model. *Cognitive Science* 31(4). 645–671.

Smith, Carlota. 1983. A theory of aspectual choice. *Language* 59. 479–501.

Smith, Carlota. 1997. *The parameter of aspect.* Dordrecht: Kluwer.

Spitzer, Manfred. 1999. *The mind within the net: Models of learning, thinking, and acting.* Cambridge, MA: MIT Press.

Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66. 143–160.

Yu, Chen. 2006. Learning syntax-semantics mappings to bootstrap word learning. In Ron Sun (ed.), *Proceedings of the twenty-eighth annual conference of the cognitive science society*, 924–929. Mahwah, NJ: Lawrence Erlbaum.

Yu, Chen & Linda B. Smith. 2006. Statistical cross-situational learning to build word-to-world mappings. In Ron Sun (ed.), *Proceedings of the twenty-eighth annual conference of the cognitive sciences society*, 918–923. Mahwah, NJ: Lawrence Erlbaum.

Zhao, Xiaowei & Ping Li. 2005. A self-organizing connectionist model of early word production. In Bruno G. Bara, Lawrence Barsalou & Monica Bucciarelli (eds.), *Proceedings of the twenty-seventh annual conference of the cognitive science society*, 2434–2439. Mahwah, NJ: Lawrence Erlbaum.