The Acquisition of Lexical and Grammatical Aspect in a Self-Organizing Feature-Map Model

Ping Li
Ping@Cogsci.Richmond.Edu
Department of Psychology
University of Richmond
Richmond, VA 23173, USA

Abstract

This study uses self-organizing feature maps to model the acquisition of lexical and grammatical aspect. Previous research has identified a strong association between lexical aspect and grammatical aspect in child language, on the basis of which some researchers proposed innate semantic categories (Bickerton, 1984) or prelinguistic semantic space (Slobin, 1985). Our simulations indicate that this association can be modeled by self-organization and Hebbian learning principles in a feature-map model, without making particular assumptions about the structure of innate knowledge. In line with results from Li (1999), our study further attests to the utility of self-organizing neural networks in the study of language acquisition.

Introduction

Most linguistic theories of tense and aspect recognize two kinds of aspect: lexical aspect refers to the inherent temporal meanings of a verb, whereas grammatical aspect refers to a particular viewpoint toward the described situation. For example, whether the verb characterizes a situation as having a temporal boundary or an end result is a matter of lexical aspect, whereas whether the sentence presents a situation as ongoing (progressive/imperfective) or completed (perfective) is a matter of grammatical aspect. In English as well as in many other languages, lexical aspect is typically encoded by verb semantics, whereas grammatical aspect is encoded by morphological markers (e.g., English suffixes *–ing* and *–ed*).

Linguists have developed several systems to capture lexical aspect and grammatical aspect (see Comrie, 1976; Smith, 1997). For lexical aspect, the best-known system is Vendler's (1957) four-way classification of verbs into activities, accomplishments, achievements, and states: (1) activity verbs like *walk* and *run* encode situations as consisting of successive phases over time with no inherent endpoint; (2) accomplishment verbs like *build a house* also characterize situations as having successive phases, but differ from activities in that they encode an inherent endpoint (e.g., house-building has a terminal point and a result); (3) achievement verbs encode situations as punctual and instantaneous, e.g., *recognize a friend* and *cross the border*, and (4) state verbs encode situations as involving homogeneous states with no inherent endpoint, e.g., *know*, *want*, and *possess*. 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). With respect to grammatical aspect, there are two major categories, according to Comrie (1976): imperfective and perfective. Imperfective aspect presents a situation from an internal point of view, often as ongoing (progressive) or enduring (continuous), whereas perfective aspect presents a situation from an external perspective, often as completed. In English, the imperfective-perfective contrast is realized in the difference between the progressive *–ing* and the past-perfective *–ed*.[1]

Studies of language acquisition have long documented the interaction between lexical aspect and grammatical aspect in child language and in adult second language learning (for a comprehensive review, see Li & Shirai, 2000). In particular, researchers have found that young children initially tend to restrict tense-aspect morphology to specific categories of lexical aspect. This restricted or "undergeneralized" use is found in diverse languages such as Chinese, English, French, Italian, Japanese, and Turkish (see Li & Shirai, 2000 for a review). For example, English-speaking children tend to associate the use of the progressive marker *–ing* only with atelic, activity verbs, whereas they associate the past-perfective marker *–ed* only with telic verbs (accomplishments and achievements).[2] This strong association weakens over time, and eventually children develop adult-like competence in using both the progressive and the perfective suffixes with different lexical aspect categories.

Capitalizing on this strong association in early child language, some researchers hypothesized that children have innate semantic categories that roughly correspond to the lexical aspect distinctions of verbs. In particular, Bickerton (1984) argued that the semantic distinctions between *punctual* (e.g., *jump*) and *nonpunctual* (e.g., *walk*), and between *state* (e.g., *want*) and *process* (e.g. *walk*) are biologically programmed as part of a Language Bioprogram. Bickerton's initial claim for the proposed bioprogram was based on evidence from creole languages, but he also drew on the following evidence from early child language: (1) children treat achievement verbs (punctual) differently from activity verbs (nonpunctual) in their use of grammatical morphology; (2) children treat state verbs differently from activity (process) verbs, in that they use *–ing* only with process verbs and never with state verbs. These patterns prompted Bickerton

---

[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 languages for tense and aspect.

[2] Some studies also report a third association between the habitual *–s* and state verbs, e.g., Clark (1996).

that children use tense-aspect morphology early on only to mark the bioprogrammed semantic distinctions. In a similar proposal with a somewhat different perspective, Slobin (1985) proposed that children come to the language acquisition task with a pre-structured semantic space in the Basic Child Grammar. This semantic space contains a universal, uniform set of prelinguistic semantic notions, initially independent of the child's linguistic experience, and they act like magnets to strongly attract the mapping of grammatical forms of the input language. Two contrasting categories, *process* and *result*, are in this space, and thus children would tend to map the progressive *–ing* to the process (atelic) verbs and the past-perfective *–ed* to the result (telic) verbs early on.

In this study, we entertain the same empirical results with an alternative proposal that rejects the strong version of the nativist argument on innate semantic categories.[3] In previous empirical studies (Li & Bowerman, 1998; Li & 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 linguistic input, rather than innate biases. In parental speech, there are probabilistic associations between progressive markers and atelic verbs, and between perfective markers and telic verbs (see Li & Shirai 2000 for a review); children's initial undergeneralizations (restricted uses of morphology) might reflect their analyses of these probabilities. This study is a detailed implementation of this idea in a connectionist model. In previous connectionist work (Li, 1999), we explored the use of self-organizing neural networks, in particular, the self-organizing feature maps as a model of language acquisition. Our model was applied to overgeneralization and recovery phenomena in the acquisition of English reversive prefixes (*un-* and *dis-*), in connection with the acquisition of structured semantic representations (the cryptotypes of verbs). In this study, we extend this line of research to examine the undergeneralization of aspectual suffixes (*-ing*, *-s*, and *–ed*), in connection with the acquisition of semantic categories of lexical aspect. More important, we attempt to show (1) how a multiple feature-map 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 speech from the CHILDES database (MacWhinney, 1995). Evidence from such a study could shed light on the processes of lexical and morphological development in child language.

Several important properties of self-organizing feature maps make them particularly well suited to the study of lexical and morphological acquisition (see Li, 1999, for a discussion). First, they belong to the class of unsupervised learning networks that require no explicit teacher; learning is achieved entirely by the system's organization in response to the input. These networks provide computationally more relevant models for language acquisition: one could argue that child language acquisition in the natural setting (especially the organization and reorganization of the lexicon) is largely a self-organizing process that proceeds without explicit teaching (MacWhinney, 1998). Second, self-organization in these networks allow for the gradual formation of structures as activity bubbles on 2-D maps, as a result of extracting an efficient representation of the complex statistical regularities from the high-dimension input space (Kohonen, 1989). This property allows us to model the emergence of semantic categories as a gradual process of lexical development. Self-organizing feature maps are also biologically plausible models: one could conceive of the human cerebral cortex as essentially a self-organizing map (or multiple maps) that compresses information on a 2-D space (Spitzer, 1999). Third, several self-organizing maps can be connected via *Hebbian learning,* a co-occurrence learning mechanism, according to which the associative strength between two neurons is increased if the neurons are both active at the same time (Hebb, 1949). In a multiple feature-map model (Miikkulainen, 1997), initially, all units on one map could be associated with all units on the other map; as self-organization takes place, the associations become more focused, so that eventually only the maximally active units on the two (or more) maps are associated. This procedure allows us to model one-to-many or many-to-many associations between forms and meanings on the basis of how often they co-occur and how strongly they co-activate in the representation. In short, self-organization and Hebbian learning are two important computational principles that aid us in the understanding of lexical representation and morphological generalization in language acquisition.

Method

Network Architecture

DISLEX is a multiple feature-map model of the lexicon that relies on self-organization and Hebbian learning principles (Miikkulainen, 1997). In this study, we use the basic architecture of DISLEX to model the acquisition of lexical and grammatical aspect. In this model, different feature maps dedicated to different types of linguistic information (orthography, phonology, or semantics) are connected through associative links via Hebbian learning. During learning, an input pattern activates a unit or a group of units on one of the input maps, and the resulting bubble of activity 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. Figure 1 presents a schematic diagram of the architecture of the model.

---

[3] Note that it is fundamental to the Language Bioprogram hypothesis that the semantic categories are biologically hardwired, whereas this is left more open in the Basic Child Grammar hypothesis.
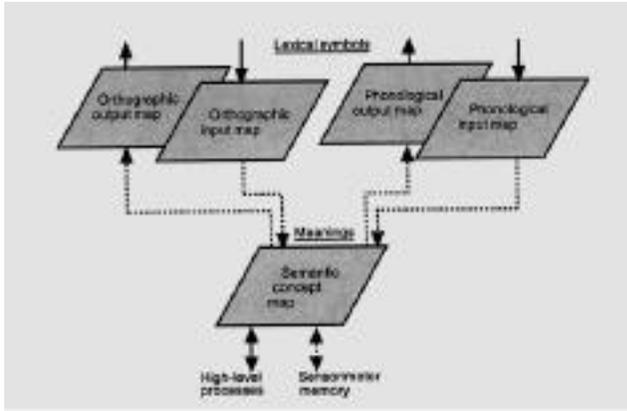
Figure 1: A multiple feature-map model of the lexicon (Miikkulainen, 1997)

In this study, we used no orthographic maps since we were modeling lexical and morphological acquisition in young children who are preliterate. We constructed two self-organizing maps, each of the size of 50 x 50 units, one for the organization of phonological input (henceforth the phonological map), and the other for the organization of semantic input (the semantic map). All simulations were run on a SUN Ultra workstation, using the DISLEX codes configured by Miikkulainen (1999).

Input Representations

In order 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 caregivers' speech in the CHILDES database (MacWhinney, 1995). We extracted all the utterances of the parents, caregivers, and experimenters from the CHILDES database in over 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 from this corpus was chosen as an input to the network if it occurred in the parental or caregivers' speech for five or more times in a given age period (see below). With this criterion we selected a total of 562 words (types) as input to our network. They were inputted to the network in four stages, according to the age groups at which they occurred (see below).

Previous connectionist models of language acquisition have often relied 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. Representations of linguistic information in this way are often subject to the criticism that the network works precisely because of the use of certain linguistic features. To overcome potential limitations associated with this approach, we used more realistic input data to simulate the acquisition of aspect. We represented our inputs as follows.

Phonological representations to our network were based on a syllabic template coding developed by MacWhinney and Leinbach (1991). Instead of a simple phonemic representation, this representation reflects current autosegmental 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 (vowel). 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 trisyllabic 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.

Semantic representations to our network were based on the lexical co-occurrence analyses in the Hyperspace Analogue to Language (HAL) model of 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. We used 100 dimensions for the encoding of each vector.

Task and Procedure

Upon training of the network, a phonological input representation of the verb was inputted to the network, and simultaneously, the semantic representation of the same input was also presented to the network. By way of 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. The phonological representations of the corresponding aspectual suffixes were also co-activated with the phonological and semantic representations of the verb, depending on whether the verb co-occurs with –ing –ed, or –s in the parental speech in the CHILDES database. 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. (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 the period when the child is 12 months old. A total of 186 verbs fit

our selection criteria for the period when the child is between 13-18 months old, out of which 34 (18%) occurred with –*ing*, 9 (5%) with –*ed*, and 9 (5%) with –*s*. (2) *Input Age 2;0* (19-24 months). 324 verbs were selected, which include the 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 verbs 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 verbs 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.

Results and Discussion

We focus here on three levels of analysis of our modeling results: the role of input, the emergence of lexical aspect categories, and the formation and relaxation of strong associations between lexical and grammatical aspect.

The Role of Input

One important rationale behind the current modeling effort is the understanding of the role of the linguistic input in guiding children's acquisition of lexical and grammatical aspect. Earlier we emphasized the relationship between patterns observed in children's speech and those in adult speech with respect to the interaction between lexical and grammatical aspect. But a simple correlation between children's and adults' 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 if a connectionist network, endowed with self-organization and Hebbian learning principles, is able to display learning patterns as the child does. If so, we can conclude that self-organization and Hebbian learning could provide the necessary kinds of mechanisms that drive the formation of patterns in children's acquisition. In this way, our modeling enterprise provides insights into the mechanisms that underlie the learning process.

Table 1 presents a summary of the major patterns of the network's learning according to the tense-aspect suffixes it produced at the different learning stages. It shows the results of the network's production of three suffixes,–*ing*, *-ed,* and –*s* with three types of verbs, atelic, telic, and stative.[4] The results are based on the analyses of the activation of units on the phonological map that each verb in the semantic map activated, after the network had been trained for 200 epochs at each stage. The table does not include instances in which the network produced multiple suffixes with a given verb (see Table 3 for these instances).

The results as shown in Table 1 are highly consistent with empirical patterns observed in early child language: the use of the progressive aspect is closely associated with atelic

---

[4] Atelic verbs correspond to Vendler's activities, telic to accomplishments and achievements, and stative to state.

verbs that indicate ongoing processes, while that of perfective aspect is closely associated with telic verbs that indicate actions with endpoints or end results. In particular, in early child English, –*ing* is restricted to activity verbs, the perfective/past marker –*ed* restricted to telic verbs, and habitual –*s* restricted to stative verbs (see Introduction). Our network, having taken in input patterns based on realistic adult speech, behaved in the same way as children do. For example, at Input Age 1;6, the network produced –*ing* predominantly with atelic verbs (75%), –*ed* overwhelmingly with telic verbs (82%), and –*s* exclusively with stative verbs (100%). Such associations remained strong at Age 2, but gradually became weaker at later stages (the association between –*s* and stative verbs remained strong throughout).

Table 1: Network's production of grammatical suffixes with lexical aspect categories*

	TENSE-ASPECT SUFFIXES					
	Age 1;6			Age 2;0		
VERB	*-ing*	*-ed*	*-s*	*-ing*	*-ed*	*-s*
Atelic	75	18	0	66	16	0
Telic	25	82	0	28	84	0
Stative	0	0	100	0	0	100
	Age 2;6			Age 3;0		
	-ing	*-ed*	*-s*	*-ing*	*-ed*	*-s*
Atelic	64	26	0	52	9	10
Telic	31	74	0	44	77	10
Stative	0	0	100	4	14	80

* Values represent percentages of verbs that occurred with the given suffix. Note that the percentages within a given column does not always add to 100%, reflecting the fact that some verbs could not be easily classified into one or the other category. This is also true for other tables.

Interestingly, when we analyzed the actual input to our network, we found similar patterns. Recall that the input to our network was based on the adult speech from the CHILDES database. Table 2 presents the percentages of use of suffixes with different verb types in the input data.

Table 2: Percentage of use of grammatical suffixes with lexical aspect categories in the input data

	TENSE-ASPECT SUFFIXES					
	Age 1;6			Age 2;0		
VERB	*-ing*	*-ed*	*-s*	*-ing*	*-ed*	*-s*
Atelic	69	22	0	74	15	17
Telic	28	77	33	24	77	0
Stative	3	0	67	2	8	83
	Age 2;6			Age 3;0		
	-ing	*-ed*	*-s*	*-ing*	*-ed*	*-s*
Atelic	67	23	20	67	23	23
Telic	25	69	20	31	65	8
Stative	8	8	60	2	12	69

This high degree of correlation between the network production and the input shows that our network was able to

learn on the basis of the information of the co-occurrences between lexical aspect (verb types) and grammatical aspect (use of suffixes). This learning ability was due to the network's use of Hebbian learning in computing and registering (a) when the semantic, phonological, and morphological properties of a verb co-occur and (b) how often they do so.

Note that the patterns of the two tables are consistent and similar, but not identical. This is important because if the learner, child and network alike, simply mimicked what's in the input by recording each individual word and suffix and their co-occurrence, the learner would have no productive control of the relevant linguistic device and would simply produce the patterns verbatim. Our results suggest that the associations between verb types and suffixes are stronger in the network's production than in the input to the network. Our network, like the child, behaved more restrictively than what is in the input with respect to the use of tense-aspect suffixes (see Li & Shirai, 2000, for details on this point).

The Emergence of Lexical Aspect Categories

As discussed earlier, a distinct property of feature maps is that the structures in the network's representation can be clearly visualized as activity bubbles or patterns of activity on a 2-D map. Figure 2 presents a snapshot of the network's self-organization of the semantic representations of 186 verbs at the end of the Input Age 1;6.
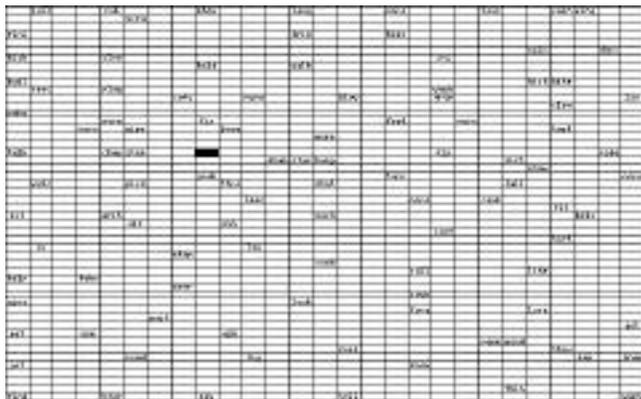

Figure 2: Network representations of verbs in the semantic map. Only the left portion of the complete map is shown due to space limit. Words longer than four letters are truncated.

An examination of this map shows that the network has clearly developed structured semantic representations that correspond to categories of lexical aspect such as telic verbs, atelic verbs, and stative verbs. Our network formed clear clusters of verbs by mapping similar verbs onto nearby regions of the map. For example, towards the lower right-hand corner, stative verbs like *feel, know, think, remember, wonder, love,* and *like* are mapped to the same region of the map. A second cluster of verbs occurs towards the lower left-hand corner, including verbs like *see, read, hear, say, ask,* and *tell*, all being verbs of visual or auditory activities. Still a third chain of verbs can be found in the middle-to-left portion of the map, including verbs like *catch, fix, break, knock, grab, and throw*, all of which are telic verbs indicat-

ing actions that lead to clear end results. Finally, a cluster of verbs can be found spanning the upper end of the map, including (from left to right) *rub, scrub, sleep, shout, laugh, drink, walk, kiss, cry, swim,* and *dance*, all of which are atelic activity verbs, and many of them co-occur with *–ing* early on. In contrast to this layer of verbs, the left-most columns feature primarily telic verbs, such as *finish, hide, build, reach, make, go, give, get,* and *find*. Of course, the network's representation at this point is still incomplete, as self-organization is still moving continuously from diffuse to more focused patterns of activity.

Crucially, on the one hand, these clusters form concentrated patterns of activity, providing the basis for semantic categories, and on the other hand, they also form focused associative pathways to the phonological and morphological representations of verbs on the other feature maps. When concentrated activities occur both horizontally (within a 2-D map) and vertically (across the maps), the semantic categories of lexical aspect will behave like magnets for the mapping of grammatical morphemes. Thus, when new verbs share enough similarities with verbs of a lexical aspect and fall within these clusters, their mapping to corresponding grammatical aspect will be readily assimilated through the existing associative pathways going from verb semantics to suffixes. This explanation provides an alternative account of the Basic Child Grammar, according to which the initial semantic categories that strongly attract grammatical mappings are privileged and pre-linguistic.

The results from our modeling offer a new way of thinking about the acquisition of categories of lexical aspect. Verbs in a lexical aspect category form complex relationships, in that they vary in (a) how many semantic features are relevant to the category, (b) how strongly each feature is activated in the representation of that category, and (c) how features overlap with each other across category members. Traditional analytical methods 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 featural overlapping and weighted featural composition (see Li & MacWhinney, 1996 for a discussion).

From Strong Associations to Diverse Mappings

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. 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.

This co-occurrence-and-constraint process is modeled in our network by Hebbian learning of the associative connections between forms and meanings. Hebbian learning can account for the relaxation of the associations as well as the

strong associations. Table 3 presents the same simulation results as Table 1, except here we added the multiple suffixation patterns -- a given verb was counted for multiple number of times in the table depending on the number of suffixes with which it co-occurred.

Table 3: Network's production of grammatical suffixes with lexical aspect categories (multiple suffixations)

	TENSE-ASPECT SUFFIXES					
	Age 1;6			Age 2;0		
VERB	-ing	-ed	-s	-ing	-ed	-s
Atelic	75	16	0	62	29	6
Telic	28	75	0	32	66	31
Stative	0	8	100	0	4	63
	Age 2;6			Age 3;0		
	-ing	-ed	-s	-ing	-ed	-s
Atelic	64	40	44	52	38	30
Telic	32	60	12	43	53	26
Stative	0	0	44	5	9	44

A comparison of this table with Table 1 reveals that for the early stages (1;6 and 2;0) the two tables are very similar; for the later stages, however, they become different, mainly with respect to the uses of *–ed* and *–s*. Detailed analyses show that over 50% of all suffixed verbs had more than one suffixes at Input Age 3;0, compared to only 5% at Input Age 1;6. These results suggest that multiple suffixation might be a driving force for the learner to break from the strong associations to more diverse associations between lexical and grammatical aspect. There was relatively little change with the *–ing* verbs, because the majority of the verbs early on were atelic verbs that took *–ing*. These same patterns were also found to be true of the parental input in the CHILDES database (see Li & Shirai, 2000, for detailed discussion).

Conclusion

In this paper I showed that self-organizing neural networks can be used successfully to model language acquisition, following up on Li (1999). Self-organization and Hebbian learning in such networks are two important computational principles that can account for the psycholinguistic processes in the acquisition of lexical and grammatical aspect. Focused associative pathways between forms and meanings lead to particularly strong associations between lexical aspect and grammatical aspect, thereby to undergeneralized patterns of grammatical morphology as observed in early child language. Increasing associative links along with incremental vocabulary growth lead to diverse mappings. Finally, self-organization of the semantic structure of verbs leads to the formation of lexical aspect categories, on the basis of the network's analysis of the complex relationships in a high-dimensional space of language use. Our results lend insights into the mechanisms and processes of lexical-morphological acquisition, and may also generate interests in further empirical studies against which we can compare detailed patterns of modeling results.

References

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences, 7*, 173-188.

Burgess, C. & Lund, K. (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes, 12,* 1-34.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge, England: Cambridge University Press.

Hebb, D. (1949). *The organization of behavior: A neuropsychological theory.* New York, NY: Wiley.

Kohonen, T. (1989). *Self-organization and associative memory.* Heidelberg: Springer-Verlag.

Li, P. (1999). Generalization, representation, and recovery in a self-organizing feature-map model of language acquisition. In M. Hahn & S.C. Stoness (eds.), *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp.308-313). Mahwah, NJ: Lawrence Erlbaum.

Li, P., & Bowerman, M. (1998). The acquisition of lexical and grammatical aspect in Chinese. *First Language,18*, 311-350.

Li, P., & MacWhinney, B. (1996). Cryptotype, overgeneralization, and competition: A connectionist model of the learning of English reversive prefixes. *Connection Science, 8*, 3-30.

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect.* Berlin and New York: Mouton de Gruyter.

MacWhinney, B. (1995). *The CHILDES project: Tools for analyzing talk* (2nd Ed). Hillsdale, NJ: Erlbaum.

MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology, 49,* 199-227.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40,* 121-157.

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

Slobin, D. (1985). Crosslinguistic evidence for the Language-Making Capacity. In D. Slobin (ed.), *The crosslinguistic study of language acquisition.* Vol.2. Hillsdale, NJ: Erlbaum.

Smith, C. (1991). *The parameter of aspect.* Dordrecht: Kluwer.

Spitzer, M. (1999). *The mind within the net.* MIT Press.

Vendler, Z. (1957). Verbs and times. *Philosophical Review, 66*, 143-160.