

# A Model of Frame and Verb Compliance in Language Acquisition

Rutvik Desai\*

*Department of Computer Science and the Cognitive Science Program  
Indiana University, Bloomington, IN 47405  
rudesai@indiana.edu*

## Abstract

*Researchers studying word learning have discovered that the syntactic frame in which a word appears plays an important role in the interpretation of the word, and this importance diminishes gradually with the increase in age. The interpretation of the sentences based on the frame and the verb is known as frame and verb compliance respectively. Here, a connectionist model is presented that learns a miniature language by associating sentences with the corresponding “scenes.” In doing so, when the input to the network is changed to reflect the increasing linguistic experience of children, it exhibits a shift from frame to verb compliance. It is argued that these phenomena can be attributed to the increasingly combinatorial linguistic experience and representations that change with learning, and it is not necessary to invoke specialized mechanisms or principles.*

## 1. Introduction

Children learn new words rapidly. A common-sense explanation for vocabulary acquisition is that word meanings are learnt by observing real-world contingencies of their use. The meaning of *jump* is learnt from noticing that it occurs in the presence of jumping events. However, this simple explanation has several difficulties when attempting to account for acquisition of meaning of all words. Many of these problems are listed in [1]: (a) This theory fails to account for the fact that children with radically different exposure conditions (*e.g.*, the blind and the sighted) acquire much the same meanings, (b) many verbs are used for the same events and only provide a perspective on an event (*e.g.*, *chase* and *flee*), and (c) many verbs only differ in the level of specificity at which they describe single events (*e.g.*, *see*, *look*, *orient*).

In light of these problems, it has been suggested that children use another rich source of information, namely the syntactic context in which the words occur. This proposal is known as *syntactic bootstrapping* [1,2,3].

According to this hypothesis, children can use the knowledge of syntax to predict meanings of words. The learner observes the real world situations and also observes the language structures in which various words appear. If there is a correlation between meanings and a range of syntactic structures, the meaning (or some components of the meaning) of an unknown word can be predicted when it appears in a familiar structure.

### 1.1. Verb Compliance and Frame Compliance

One way to study the effect of syntax on the acquisition of word meaning is to use familiar words in a different or incorrect syntactic context and examine the effect on the interpretation of the word. For example, we can insert a transitive verb in an intransitive frame and examine how children interpret the sentence. If children are still learning about a verb, then they may more readily accept its occurrence in an incorrect frame. They are more likely to reject an incorrect frame when they have fully acquired the verb. If children interpret the sentence in accord with the frame, they are said to be *Frame Compliant*. If the interpretation fits more with the verb, they are *Verb Compliant*.

Frame and Verb Compliance are interesting for another theoretical reason. While children's verb use is overwhelmingly correct, a major exception to this appears somewhere around the age of 3. As reported by Bowerman [4,5], children sometimes use verbs in incorrect sentence frames, as in *\*Don't fall that on me* (to protest the impending dropping of an object by someone). Thus, children overgeneralize, *e.g.*, they use a verb transitively when only intransitive use is allowed, or vice versa. Children must learn eventually which uses are “licensed” for which verbs. For example, they must learn that *sink* can be used either transitively or intransitively, but *fall* and *go* allow only noncausal interpretation. How children overcome these overgeneralizations is a major question in language acquisition. This question is essentially the same as asking why children become Verb Compliant at some stage. When children show Verb

\*Currently at the Department of Neurology, Medical College of Wisconsin, Milwaukee, WI 53226.

Compliant behavior, they have sufficient confidence in their knowledge of verb meaning and syntax that they reject contradictory cues, which is exactly the requirement for overcoming the tendency to overgeneralize.

Now we look at some empirical evidence for compliance effects.

## 1.2. The Data

Naigles and colleagues [6,7] conducted experiments involving the approach described above. They asked 120 children, from 2.5 to 12 years of age, as well as adults, to enact grammatical and ungrammatical sentences using “Noah’s Ark” and wooden toy animals as props. Ungrammatical sentences were constructed by placing transitive verbs (*bring, take, push, put*) in intransitive frames (e.g., *\*The lion puts in the ark, \*The zebra brings*). Similarly, intransitive verbs (*come, go, fall, stay*) were inserted in transitive frames (e.g., *\*The elephant comes the giraffe*). The children’s enactment was deemed to be Frame Compliant if they modified the meaning of the verb to conform to the frame in which it was encountered (e.g., the elephant pushing or carrying the giraffe). It was considered Verb Compliant if they followed the restrictions of the verb (e.g., the elephant moving independently of giraffe).

Their overall results indicated that younger children, especially the 2-year-olds, were more Frame Compliant, enacting the ungrammatical sentences according to the demands of the frame and altering the meaning of the verb. They allowed the novel frames to influence the interpretation of the familiar verbs. Older children, and especially the adults, were more Verb Compliant, following the restrictions of the verb and repairing the sentence. Children at the intermediate ages were en route to the adult state, showing intermediate levels of Frame and Verb Compliance.

Similar experiments have been conducted with children with Down Syndrome (DS) [8]. The linguistic skills of children with DS are split in an interesting way. Relative to their syntactic knowledge (often measured by measured MLU or auxiliary use) their vocabulary growth is advanced. It was reported [8] that children with DS who had a “vocabulary age” of 6 years were syntactically like 3-year-olds. While children with DS were more Frame Compliant than their chronological-age mates, they also exhibit the move from Frame to Verb Compliance. Adolescents with DS show more Verb Compliance than gradeschoolers with DS. Thus, with the advance in syntactic knowledge, DS children also move toward Verb Compliance.

In this paper I present a connectionist model that attempts to explain the mechanisms by which this shift occurs. First, a network is presented that learns a miniature language by associating simple sentences to the

corresponding “scenes.” The network’s behavior with respect to the compliance effects is then examined. Then, various theories of compliance and the implications of network’s behavior are discussed. We end with a discussion of the nature of representations and input in the network.

## 2. The Network

The architecture of the network is shown in Figure 1. It contains recurrent connections in the hidden layer as in a Simple Recurrent Network [9] to handle temporal sequences of words. Recurrent connections on the output layer make it easier for it to remember what has been already learned from the earlier portion of the sentence.

Then input to the network consists of sentences or noun phrases (henceforth called “utterances”) describing one or two objects and optionally an action, generated by the grammar shown below:

S → NP | NP1 | NP is IV | NP1 are IV | NP is TV NP  
NP → DET N | DET SIZE N  
NP1 → NP and NP  
N → boy | girl | dog | mouse  
IV → jumping | dancing | running | walking  
SIZE → large | small  
TV → pushing | holding | hugging | kicking  
DET → a

One can divide the utterances generated by this grammar into five basic types: (a) N, (b) NN, (c) NV, (d) NNV, and (e) NVN. With optional adjectives describing the size, utterances such as *a girl and a big dog are jumping* or *a small dog and a big mouse* are obtained. These utterances are presented to the network sequentially, one word at each time step. Words are represented in a localist manner by turning on a single bit in the input layer. Also, *ing* is treated as a separate word, with the assumption that it can be discerned from the word stem as a separate unit. An end-of-utterance marker, STOP, is presented after the last word of each utterance, at which point all context units are reset.

On the output or the semantic end, the descriptions of scenes corresponding to the input utterance are presented as a 30-bit fixed-width vector. There are two slots for objects, and one for the action or the event taking place. Each object slot is divided into two slots of 4 and 6 units each, which represent the attribute large (1100) or small (0011) and type of object respectively. In the 10-bit event slot, the first 4 bits indicate whether the action is causal or non-causal (with activations 1100 and 0011 respectively), and the remaining 6 bits describe other features of the action. A distributed representation for each individual object and event is generated by turning on 3 randomly chosen bits in its slot. If each bit is viewed as representing

a feature, this creates representations with partially overlapping features. The slots for an attribute, object, or action not described in the scene are set to 0.

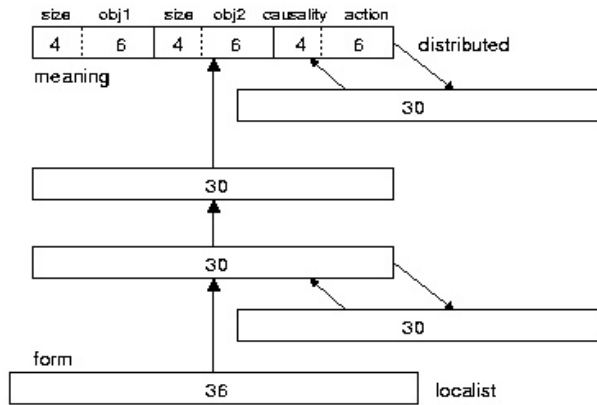


Figure 1. The network architecture.

### 3. Comprehension

We first test if the network is capable of performing the basic task of producing the correct scene corresponding to an input utterance. The set of all utterances generated by the grammar was probabilistically divided into two parts, one for training and the other for testing generalization. There is significant variability in the total number of different types of utterances generated by the above grammar: There are 12 utterances of type (a), 144 utterances of type (b), 48 utterances of type (c), and 576 utterances each of type (d) and (e). Hence, these utterances were included in the training set with differing probabilities: 1.0 for type (a), 0.4 for type (b), 0.7 for (c), and 0.2 for (d) and (e). To give more representation to utterances that have a lower frequency, and to ensure that they are not overwhelmed by other more frequent patterns, type (a) utterances were included thrice in the training set while type (c) utterances were included twice. The entire scene is presented as the target for every word. If there is a large difference in the type or token frequency between different types of utterances (e.g., many more transitives than intransitives), this can lead to local optima, since the network attempts prediction from incomplete information. The purpose of the chosen probabilities is only to avoid a severe imbalance in the type/token frequencies; other than that the values are not critical to the results.

The network was trained using backpropagation on the utterances in the training set. The initial weights were sampled from a uniformly random distribution between -0.2 and 0.2. The complete target was held constant for the duration of the entire utterance. This ensures that no words are given a special status and encourages the network to process the words as soon as they arrive. A

learning rate of 0.0005 and no momentum were used. The weights were updated at the end of each epoch. Training was continued till there was no significant improvement in the error.

To assess the performance of the network, if the activation of an output unit was less than 0.5, it was considered OFF, and it was taken as being ON otherwise. An utterance was declared to be processed correctly if, at the end of the utterance, all output units had the desired ON or OFF activations. With this criterion, in a average of 5 runs, 100% accuracy was achieved on a training set of 322 (different) utterances and 96% of utterances were processed correctly in the remaining 1034 utterances of the testing set, which the network had not seen during training. The network was, then, largely successful in this task of producing semantics given an utterance, or comprehension. Next, we look at the experiments regarding frame and verb compliant behavior in the network.

### 4. Frame and Verb Compliance in the Network

To qualitatively simulate the increasing vocabulary and linguistic experience of children, the network was trained in stages with increasing numbers of nouns, rather than with the entire vocabulary as in the basic comprehension task described in the previous section. Four transitive and four intransitive verbs were used in all stages. The number of nouns used at each stage was increased gradually from 1 to 7. No adjectives were used. The set of utterances at each was again divided probabilistically into training and testing sets. Utterances in the testing set were not part of the training set at any stage. The network started with the weights from the previous stage and was trained to near-perfect accuracy on the training set using a learning rate of 0.001.

Two types of ungrammatical utterances were generated to test the network. The first was a NVN transitive sentence with a known intransitive verb (e.g., *a dog is dancing a boy*) while the second was a NNV intransitive sentence with a known transitive verb (e.g., *a dog and a boy are holding*). Two transitive and two intransitive verbs were chosen and four sentences were generated with each verb using two nouns. Transitive verbs were inserted in intransitive sentences, and vice versa. We are interested in examining whether the network interprets these verbs in incompatible frames as depicting causal events or noncausal ones. Recall that there are four units in the network's output indicating causality, where the pattern 1100 stands for a causal meaning while 0011 stands for a noncausal interpretation. To assess the network's response, a variable  $\delta$  is defined as the mean activation of the first two units minus the mean activation of the last two units. A positive  $\delta$  indicates a transitive

response, while a negative  $\delta$  suggests an intransitive interpretation of the verb. The mean value of  $\delta$ , calculated from the eight transitive and intransitive sentences at the end of each stage, is shown in Figure 2.

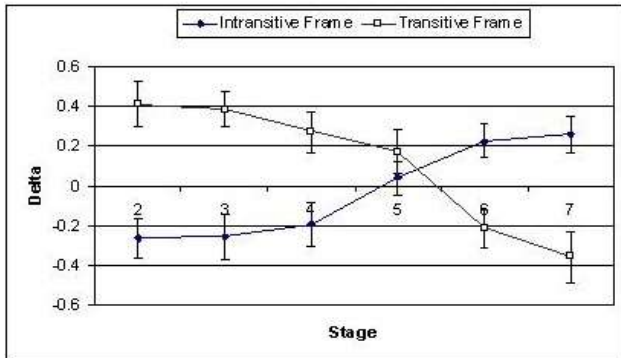


Figure 2. The value of the parameter  $\delta$  across different stages. A positive  $\delta$  implies a causal interpretation of the verb while a negative  $\delta$  indicates noncausal interpretation.

The conflict arising from the mismatch between the frame and the verb is indicated by the values of  $\delta$ . All four units receive some activation in most cases. However, in the initial stages the frame tends to “win,” as indicated by the higher activation of the units for causality (positive  $\delta$ ) in the case of transitive frame and units for noncausality (negative  $\delta$ ) in the case of intransitive frame. In other words, we get a Frame Compliant response in earlier stages with a small number of nouns. This behavior changes gradually with the increase in the number of nouns. With 7 nouns, there is still a conflict due to the mismatch, but now the response is more in accordance with the type of the verb. For the transitive verb, the output is closer to 1100 and for the intransitive verb it is closer to 0011, suggesting a Verb Compliant response.

## 5. Discussion

Frame and verb compliant behavior is closely related to the well-known overgeneralization errors made by children, and their recovery from those errors. We can consider some of the theories related to overgeneralization and frame and verb compliance, and ask what the implications of the model are.

### 5.1. Maturation

A maturation-based account is offered in [10]. Very briefly, verbs become organized into semantic subclasses known as *narrow range subclasses* as their representations are refined. The semantics of the verb

class determines whether the verbs allow alternations (e.g., causative and noncausative use in transitive and intransitive frames) or not. When the representation of a verb matches that of another verb in the same subclass that is known to alternate between causal and noncausal, the former is allowed to alternate as well. For example, motion verbs that encode path (e.g., *bring*, *take*, *go*) can be used either causally or noncausally, but not in both ways. On the other hand, motion verbs that encode manner, like *roll* and *bounce* can be used both transitively and intransitively. Overgeneralization occurs because a verb is used in the same manner as other semantically similar verbs in a subclass.

The shift to Verb Compliance may occur because the verb representations are elaborated to the extent that they have formed grammatically relevant narrow range subclasses. Some verbs no longer allow causal interpretation because they do not fit the semantic specification of the subclasses that are causal. At the time of puberty, those subclasses of verbs for which there has been no evidence of alternation become fixed or “closed.” After that, no new information about the verb is accepted, resulting in Verb Compliant behavior. For example, since *come* and *go* do not encode manner of motion, they do not match the specification of the alternating subclass of motion verbs (that includes *roll* and *bounce*). This subclass is closed at maturation, so *come* and *go* no longer allow causative interpretation.

As pointed out in [6,8], there are factors other than age that appear to affect compliance behavior and present a serious problem for this account. If the maturation-based account was correct, one would expect to observe an across-the-board shift from Frame to Verb Compliance, for all verbs and frames. But this is not the case. Some verbs elicit more Frame Compliance than others. For example, in the Naigles, Fowler, and Helm (1992) study, in the NVN frame, *come* and *go* elicited significantly more Verb Compliance than *stay* and *fall*. *Stay* and *fall* also differed from each other significantly. In the NV frame, *bring*, *take*, and *put* showed significantly more Frame Compliance than *push*. Secondly, some frames induce Frame Compliance to a later stage than others. For example, the shift to Verb Compliance for the NV frame is effectively complete at age 5. On the other hand, even 12-year-olds and adults continue to exhibit Frame Compliance for their NVNPN frame. Intransitive frames shifted earlier than transitive frames. Furthermore, the move towards Verb Compliance can occur at different ages, anywhere between 2 to 7 years of age.

In the study of children with DS [8], it was found that although these children were more Frame Compliant than their chronological age mates, they exhibit this shift also. Since their maturational progress is dissociated, one would expect that prepubescent children with DS will be no less Verb Compliant than their adolescent counterparts if the maturational account is correct. This appears not to

be the case and children with DS also move to become more Verb Compliant.

## 5.2. Mutual Exclusivity

Another proposal about recovery from overgeneralizations is termed The Mutual Exclusivity Principle (also called Contrast, Uniqueness, or Pre-emption) [5,11,12,13]. In brief, this principle is that children will allow only one lexical entry to occupy a semantic niche. When two words are determined to have similar meanings, one of them is pre-empted and removed from the lexicon. For example, causative *come* is basically equivalent to *bring*. Using Bowerman's (1982) example, during the period in which overgeneralized (causative) *come* is frequent in production, *bring* is practically nonexistent. When *bring* becomes more frequent eventually, the causative *come* declines. This can explain why some verbs elicit Verb Compliance. For example, transitive *bring* and *take* pre-empt causative uses of *come* and *go* respectively. We do not discuss all the details of Mutual Exclusivity here (see [5,6] for a more detailed discussion), but note that while Mutual Exclusivity may have some role to play in recovering from overgeneralizations, it does not account for all the effects found in the data. For example, it does not explain why intransitive *push* causes Verb Compliance earlier than intransitive *bring* or *take*. This principle also does not work for all the verbs, since for some verbs, it is difficult to find a similar meaning verb that can pre-empt its use in the right way.

## 5.3. Lexical Knowledge

A different account based on lexical knowledge is offered in [7]. This account relies on children's knowledge of individual verbs. Children's conjectures about verb meanings are refined by ongoing events as well as the structures in which they appear. At early stages of vocabulary acquisition, open-minded children assume that not all structures have as yet been heard and therefore certain properties of verbs (such as whether they encode causality) may be unknown to them. In this case they make use of the structural information provided by the frames. At some point, however, older children and adults feel warranted to believe that all the relevant information about the meaning has been obtained. Then they would perceive a novel structure as simply ill-formed, causing verb compliant behavior.

This theory can explain various effects in the data well. For example, the shift towards Verb Compliance is a function of individual verbs and individual frames because different amount of knowledge is accrued for them due to their differing frequencies in the input. While this account is supported by data, some important details remain unclear. One might ask where the so-called "open-

mindedness" in the initial stages and the confidence about the meaning at later stages come from. After hearing a verb in a certain number of contexts, exactly what makes a child more or less open-minded to accept new meanings? An answer is not offered in [7], but one possibility is to invoke some type of innate parameter or threshold that allows children to determine whether a certain amount of experience with a verb is enough to warrant confidence in the meaning of that verb.

## 5.4. Lexical Knowledge and Innate Principles

More recently, Lidz, Gleitman, and Gleitman [14] offer an explanation that involves both lexical knowledge and innate principles. It is best summarized by the following quote:

"The deduction of verb meaning based on an analysis of the surface structure is a learning heuristic. The learning device is asking itself, in effect: Assuming Principles [the *Theta Criterion* and the *Projection Principle*], what could be the meaning of the verb now heard, such that these principles projected this observed (surface) structure for it? Such a deductive procedure will be invoked only when the learner does not have secure knowledge of the verb in question." (p. 37)

Thus, when children know that they have secure knowledge of a verb, they assume that anyone who uses it otherwise has misspoken, resulting in Verb Compliance. Otherwise they invoke innate principles that state that participants in an event will line up one-to-one with noun-phrases in the clause [15], and make a decision based on that.

This account has a problem similar to that of the Lexical Knowledge account: Exactly how do children determine whether they have "secure knowledge" of the verb? An all-or-nothing decision about knowing a verb also seems to be involved in this account. A child either doesn't know the verb (and invokes the principles), or does (and rejects the frame). However, the knowledge of a verb is likely to be graded. Subjects even show different compliance effects for the same verb in different frames. The spontaneous remarks of subjects in [6] indicate that both children and adults are ambivalent about the sentences they are asked to act out. They have conflicting information and varying degrees of confidence in their knowledge, and hence it seems unlikely that are cleanly picking one path over the other.

## 5.5. Emergence

The network simulation suggests an alternative to the previous theories that does not rely on innate principles or overt determination of the knowledge of a verb. It can be

viewed as an extension of the Lexical Knowledge theory. The explanation of various effects, such as some verbs inducing more verb compliance than others and some frames continuing to elicit frame compliance till a later stage, is identical to the one offered by the Lexical Knowledge theory: They are input driven. The verbs that are experienced frequently and in multiple contexts, tend to elicit Verb Compliance early.

This account differs from the Lexical Knowledge theory the in *use* of the input. The network's Frame Compliance in the initial stages is a result of representations being more context-bound at those stages. The context-sensitivity of the representation is in turn a consequence of the memory, or the number of connections in the network and the number of patterns stored. With a certain amount of memory available, it is possible, and easier, to simply memorize entire syntactic-semantic patterns as wholes as long as there are relatively few patterns. This gives rise to context-bound representations in the hidden layer (see [16] for a more detailed discussion of context bound representations). Since frames and context have more weight in the representation, the network gravitates towards an interpretation based on the frame since the incompatibility of one word has a relatively small effect. The "open-mindedness" of the network, so to speak, in early stages is a result of the fact that the context plays an important role in early representations.

As more words are encountered in varied contexts, it is no longer feasible to store entire patterns individually because it entails excessive demands on memory. As a result, various components in an utterance are separated, and the words are gradually de-contextualized. As verbs (along with other words) attain their own separate representation, the effect of context is diminished and the relevant form/meaning mappings are strengthened. The words in various frames are encountered with many other words describing causal and noncausal events, and therefore do not exert significant influence with respect to the causality of an event. The network learns the remaining consistent correlations between groups of verbs and causal or noncausal events. This results in Verb Compliant behavior. Stated another way, the network exhibits open-mindedness in early stages and confidence about the meaning in later stages, but it is not a result of reasoning about the number of contexts in which words were encountered, or the confidence in the knowledge of verb meaning. Rather, Frame and Verb Compliant behavior is an "epiphenomenon," or an emergent consequence of the task, the input, the learning procedure, as well as the size and architecture of the network.

## 5.6. The Nature of Representations

The claim here is that the shift in compliance behavior is due to the diminishing role of context with increasing

linguistic experience. This may raise two questions: (1) What independent evidence is there that the network's early representations are context-bound and more become context-free later? (2) What is the independent evidence that context plays a large role in *children's* early representations that diminishes with age?

### Representations in the network

One way to gain some insight into the representations used by the network is to test its generalization ability at various stages. If the network has memorized or rote learned entire training patterns, it should perform poorly when the same words are encountered in a different context. On the other hand, if it has learned context-free word-level form/meaning mappings then they should be recognized regardless of the sentential context. The generalization performance for five networks on the respective testing sets at the end of different stages is shown in Figure 3.

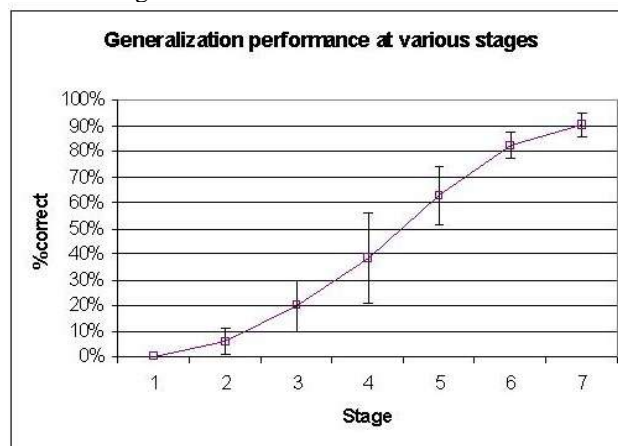


Figure 3. Performance on the testing set across stages.

The results suggest that in early stages, the network has not developed individuated representations of various components of the input utterances, and they develop after increasingly combinatorial input. St. John [16] found a similar pattern of results in his story-processing network. The network performed well at generalization only when it was exposed to highly combinatorial input, with many participants in the various slots in the representation of events.

Another method of examining the nature of network's representations involves probing the hidden unit activations directly as the utterances are processed. One way to examine the activation space is principal Components Analysis (PCA), which can measure the underlying dimensions of variation in the hidden unit activations. If two complex patterns are classified by a network with a simple strategy such as, for example, the presence or absence of one or two features, then one would expect few significant principal components in the hidden layer space because there would be few

underlying dimensions of variation. For a more complex decision-making process involving combinations of multiple features at different time-points, one would expect a high-dimensional space and more principal components would be needed to account for the variance.

Here, the set of utterances used in the first stage is passed through the network after each stage, and the 30 unit activations of the first hidden layer are recorded after each word. PCA is performed on this 30-dimensional set of vectors after each stage. The results of this analysis, as mean eigenvalues of five networks, are shown in Figure 4.

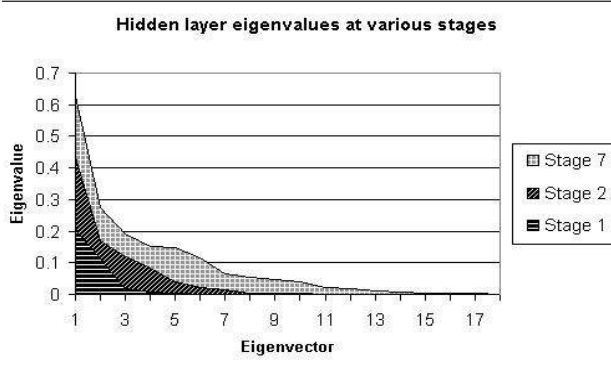


Figure 4. The mean eigenvalues of the hidden layer activation space at different stages, when processing the same set of utterances.

After the first stage, only 3 of the 30 eigenvalues are 0.01 or more on average. This number increases to 7 after stage 2, and to 13 after the last stage. The same set of utterances is processed differently, resulting in more eigenvectors of higher magnitude. During the processing of an utterance, in the early stages, once the network finds cues sufficient to distinguish one item or utterance from another, it does not need to analyze it further since it can predict the rest. The hidden layer activations remain relatively constant after that, resulting in few dimensions of variation overall. With context-free representations, activations change with each input word, resulting in more dimensions of variation [17].

### Representations in children

From studies of productions of children, there is evidence that children's early representations are context-bound too, and they gradually become de-contextualized. A compelling collection of evidence that children's early language is highly item-based, or based on specific linguistic items and expressions they comprehend and produce, is provided in [18,19,20]. There is also evidence that early productions are bound not only to linguistic context, but also to non-linguistic context such as actions, social routines, or salient events that occur frequently in the experiences of the young child [21].

### 5.7. The Nature of the Input

An important feature of the network is the changing input through different stages. The size of the training set, as well the vocabulary, is increased gradually<sup>1</sup>. However, one can view the training sets at different stages as being of the same size, since training involves cycling repeatedly over the same set. The stages are differentiated by increasing type frequency and decreasing token frequency, not by the number of utterances.

With respect to the increase in vocabulary, older children are more likely to have experienced higher number of words and word combinations simply by virtue of having more linguistic experience over a longer period of time, even if they experience the same linguistic environment. Although the vocabulary is explicitly changed here, it can be viewed as gradually sampling more utterances from a fixed set. The approach in [21] is noteworthy, where selective attention is used on a fixed input to model children's increasing experience with words to achieve the same effect as explicit addition of words to the vocabulary in a model of past-tense learning.

Secondly, there is evidence that the linguistic environment of children is not constant but changes with age. Child-directed speech (CDS) to younger children is syntactically and semantically simplified, is less diverse, and contains more high-frequency words [20,23,24,25,26]. Caretakers restrict their vocabularies when talking to young children [23,24], and the type-to-token ratio increases with age in CDS [23].

The simulation used in this work is admittedly small, and the utterances are simple. It should be noted that the sentences used in the experiments with 2- and 3-year-old children are also simple, and an attempt was made to use identical or very similar utterances in the simulations. It would also be desirable to use natural CDS from a database, rather than an artificial grammar. However, the use of raw CDS is more suitable in paradigms where no form/meaning associations are involved, and the task of the network is (typically) to predict the next word in the input. One can choose utterances from the CDS that are suitable for the task and relevant to the phenomenon under consideration, but that defeats the purpose of using natural CDS to some extent.

### 6. Conclusions

A connectionist network was presented that learns to comprehend utterances of a miniature language by associating them with the corresponding scenes. When the training set of the network is varied to reflect the increasing linguistic experience of children, the network

<sup>1</sup> Note that this is unlike early models of inflectional morphology, which were criticized for sudden changes in the input.

exhibits frame and verb compliance effects. The network's account of the shift from frame to verb compliance is similar to the lexical knowledge theory in that these effects are attributed to increasing experience with words in varied contexts. However, this account does not entail explicit rules or reasoning mechanisms. There is nothing in the network designed specifically to produce these effects; they emerge as a result of the network attempting to efficiently accomplish the task of associating utterances with scenes. The network's behavior is a consequence of various low-level parameters such as the number of weights, the number of units in each layer, the number of input patterns, and so forth. Children's compliance behavior, similarly, may change automatically when they have developed relatively context-free representations and sufficiently strong individual form-meaning mappings so that conflicting information is ignored, without making any explicit decision to do so. This work supports the view that specific mechanisms or behaviors can arise as a result of the nature of the task and the general characteristics of the tools employed to perform the task, without the presence of dedicated mechanisms.

## References

- [1] Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3-55.
- [2] Gleitman, L. & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science*, 1, 31-35.
- [3] Landau, B. & Gleitman, L. (1985). *Language and Experience: Evidence From the Blind Child*. Cambridge, MA: Harvard University Press.
- [4] Bowerman, M. (1977). The acquisition of rules governing "possible lexical items:" Evidence from spontaneous speech errors. In *Proceedings of Research in Child Language Development*, 13, 148-156.
- [5] Bowerman, M. (1982). Starting to talk worse: Clues to language acquisition from children's late speech errors. In S. Strauss (Ed.), *U-shaped Behavioral Growth*. New York: Academic.
- [6] Naigles, L., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb meanings. *Cognitive Development*, 7, 403-428.
- [7] Naigles, L., Gleitman, H., & Gleitman, L. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and Cognition: A Developmental Perspective*. Norwood, NJ: Ablex.
- [8] Naigles, L., Fowler, A., and Helm, A. (1995). Syntactic bootstrapping from start to finish with special reference to Down Syndrome. In M. Tomasello, and W. Merriman (Eds.), *Beyond Names of Things: Young Children's Acquisition of Verbs*. Hillsdale, NJ: Erlbaum.
- [9] Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- [10] Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.
- [11] Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum.
- [12] Clark, E. V. (1991). Acquisitional principles in lexical development. In S. Gelman and J. Byrnes (Eds.), *Perspectives on Thought and Language: Interrelations in Development*. Cambridge: Cambridge University Press.
- [13] Markman, E. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*. Cambridge: Cambridge University Press.
- [14] Lidz, J. Gleitman, H., & Gleitman, L. (2001). *Kidz in the Hood: Syntactic Bootstrapping and the Mental Lexicon*. (Tech. Rep. IRCS-01-01). Philadelphia, PA: University of Pennsylvania, Institute of Research in Cognitive Science.
- [15] Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Foris: Dordrecht.
- [16] St. John, M. F. (1992). The story gestalt: A model of knowledge-intensive processes in text comprehension. *Cognitive Science*, 16, 271-306.
- [17] Desai, R. (2002). Item-based Language Learning in Children and Connectionist Networks. *38th Annual Conference of the Chicago Linguistic Society*.
- [18] Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4(4), 156-163.
- [19] Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74, 209-253.
- [20] Cameron-Faulkner, T., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27, 843-873.
- [21] Jackson, D., Constandse, R. M., and Cottrell, G. W. (1996). Selective attention in the acquisition of the past tense. In *Proceedings of the 18th Cognitive Science Conference*. San Diego, CA: Erlbaum.
- [22] Barrett, M.D. (1986). Early semantic representations and early word usage. In S.A. Kuczaj & M.D. Barrett (Eds.), *The Development of Word Meaning*. New York: Springer.
- [23] Chapman, R. S. (1981). Mother-child interaction in the second year of life: its role in language development. In R. L. Schiefelbusch and D. Bricker (Eds.), *Early Language: Acquisition and Intervention*. Baltimore: University Park Press.
- [24] Hu, Q. (1994). A study of some common features of mothers' vocabularies. In J. L. Sokolov and C. E. Snow (Eds.), *Handbook of Research in Language Development Using CHILDES*. Hillsdale, NJ: Erlbaum.
- [25] Desai, R. (2002). *Modeling emergence in language acquisition*. Ph.D. Thesis, Indiana University Computer Science Department and the Cognitive Science Program.
- [26] Pine, J. (1994). The language of primary caregivers. In C. Gallaway, and B. J. Richards (Eds.), *Input and Interaction in Language Acquisition*. Cambridge, UK: Cambridge University Press.