Title
Token Frequency and Phonological Predictability in a Pattern Association Network: Implications for Child Language Acquisition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 11(0)

Authors
Marchman, Virginia
Plunkett, Kim

Publication Date
1989

Peer reviewed

Token Frequency and Phonological Predictability in a Pattern Association Network: Implications for Child Language Acquisition

Virginia Marchman
Department of Psychology
University of California, San Diego

Kim Plunkett
Institute of Psychology
University of Aarhus, Denmark

ABSTRACT

The degree to which the behavior of PDP models of pattern associations (Rumelhart & McClelland, 1986; 1987) approximates children's acquisition of inflectional morphology has recently been highlighted in discussions of the applicability of PDP to the study of human cognition and language (Pinker & Mehler, 1988). In this paper, we attempt to eliminate many of the limitations of the R&M model, adopting an empirical approach to the analysis of learning (hit rate and error type) in two sets of simulations in which vocabulary structure (token frequency) and the presence of phonological subregularities are manipulated. A 3-layer back propagation network is used to implement a pattern association task with strings that are analogous to four types of present and past tense English verbs. We overview resulting "competitions" when strings are randomly assigned to verb classes, in particular, the conditions under which different overgeneralization errors (both "pure" and "blended") are produced. In a second set of simulations, identical type and token frequencies are used, but strings are assigned to the identity and vowel change classes on the basis of phonological shape of the stem. Phonological cues are exploited by the system leading to overall improved performance. However, overgeneralizations continue to be observed in similar conditions. Token frequency works together with phonological subregularities to determine patterns of learning, including the conditions under which "rule-like" behavior will and will not emerge. The results are discussed with reference to behavioral data on children's acquisition of the English past tense.

INTRODUCTION

Most current perspectives in child language acquisition frame language learning and production in terms of symbolic and categorically defined principles or *rules* (Chomsky, 1980; Pinker, 1984; see Derwing & Skousen, 1989). Systems of rules are seen as the indispensable format within which to explain how the language learner achieves a linguistic system that is abstract enough to produce grammatical utterances, but at the same time will allow her to go "beyond the data" and create novel combinations (e.g., Berko, 1958). The goals of the acquisitionist are to outline when children can be said to master the rules that represent linguistic structure and to uncover the mechanisms by which that organization is achieved.

In the past tense in English, irregular or "strong" verbs are not seen to form their past tenses through a suffixation rule (e.g., "add -ed"), but are "exceptions." The majority of the irregular verb stems can be grouped into three general categories (see Bybee & Slobin, 1982; Pinker & Prince, 1988 for more detailed classifications): (a) *identity mapping* (or no marking - doing nothing to the stem, e.g., hit —> hit); (b) *vowel change* (transforming the vowel, e.g., come —> came; (c) *arbitrary* (there is no obvious structural relationship between the present and past tense form, e.g., go —> went). Children will sometimes *overgeneralize* the regular rule and produce erroneous forms such as "goed" in which the regular "-ed" ending cooccurs with irregular stems (e.g., Bybee & Slobin, 1982). Of course, children eventually learn to produce both regular and irregular past tense forms correctly. The apparent regression and subsequent improvement in children's abilities suggests a stage-like *reorganization* of the child's rule system (Bowerman, 1982) and is an oft-cited example of "U-shaped" development. These phenomena are seen as among the most persuasive pieces of behavioral evidence that language learning involves achieving a system in which general rules and their exceptions must come to peacefully *coexist*.

Recently, Rumelhart & McClelland (1986, 1987) set out to capture several of the "facts" of the acquisition of the English past tense (i.e., that the "-ed" suffix is often overgeneralized to irregular verbs and that learning proceeds along a "U-shaped" course) using a 2-layer perceptron network. Their goal was to suggest how a model of language acquisition might be able to avoid rule-based mechanisms and discrete symbols, yet still capture what children "do" at various points in acquisition. The ability of networks of this sort to "behave" as children do is intended to challenge the view that acquisition is *necessarily* a process of organizing and reorganizing explicitly represented rules and their exceptions. Models such as this one characteristically utilize distributed representations and focus on elaborating the microstructure or sub-symbolic nature of cognition and language (Smolensky, 1988). These claims have undergone considerable scrutiny and have been met with resistance in some circles (e.g., Pinker & Mehler, 1988). Several have questioned the details of the R&M past tense simulation, nominating it as the 'test case' for evaluating the extent to which connectionism can offer a substantive *alternative* approach to understanding the nature of linguistic and cognitive systems (Fodor & Pylyshyn, 1988).

Clearly, the R&M model is limited by several of its assumptions about the structure of the input within which language learning takes place. First, children do not hear present and past tense forms side-by-side in the input in the absence of semantic or communicative information. Nor do children receive an explicit "teacher" signal as feedback. In addition, Pinker & Prince (1988) point out that the network's production of overgeneralizations coincided directly with the introduction of a greater number of regular verbs into the training set. Learnability considerations undermine the results of a model that introduces *discontinuity* into the input, i.e., the "learner" is exposed to only a subset of the available linguistic data early in learning. Further, exemplars (i.e., tokens) of particular verbs were presented with equal frequency to R&M's system. It is highly unlikely that children hear each verb token with equal frequency (Bever, in press). Finally, Pinker & Prince criticize the R&M model for not incorporating higher-order representations such as "word," "root," "regular verb" and "irregular verb" that allow system-internal differentiations between two classes of verbs and two learning mechanisms (rote and rule). Certain aspects of the past tense system in English are conducive to the mechanisms embodied in PDP models, i.e. the abstraction of family resemblance clusters of phonological similarity between and among verbs in the *irregular classes*. However, this mechanism is neither necessary nor appropriate for capturing the "default" nature (or form independence) of verbs in the *regular class*. The operation of the regular rule is assumed to involve higher-level representations that are manipulated regardless of lower-level phonological stem representations. R&M's failure to incorporate symbolic constructs that capture the phonological differences between regular and irregular verbs is interpreted as a significant and fatal shortcoming of the approach.

In this paper, we extend the R&M work by exploring learning in two sets of simulations that are required to master mappings that are analogous to present and past tense forms in English. However, unlike the R&M work, our simulations do not use Wickelfeature representations, and back propagation is employed in a three-layer network. We adopt an empirical, comparative approach in systematically varying token frequency and the presence/absence of phonological subregularities in the input set. At no point are discontinuities introduced into the input set in any simulation.

METHOD

All simulations use an artificial language consisting of randomly generated, possible English CVC, VCC and CCV strings. (See Plunkett & Marchman (1989) for a complete description of the vocabulary and phonological representation used). Each consonant and vowel is represented by a pattern of features distributed across 6 units, reflecting phonological contrasts such as voiced/unvoiced, front/middle/back, etc. The representation of the suffix (2 units) is not phonological, but 3 "allomorphs" are possible depending on the voicing of the final consonant of the stem (analogous to, for example, /-t/ following voiceless stop). Twenty units are used to encode each present and past tense form, and thus, the model is restricted to processing fixed length strings. Eleven English-like vowel change transformations are used (e.g., /i/ —> /A/, ring —> rang). Each vowel can be transformed to one, two or three possible new vowels in the output.

All networks must learn four types of mapping. Thus, the network, like the child, must learn to deal with several different classes of transformations simultaneously. In the *Parent* simulations, the network is at a slight disadvantage compared to the child in that strings are assigned to the different classes *randomly* (except that assignment to vowel change class is conditional upon possession of a vowel which can undergo a legal transformation). Here, no more phonological similarity exists between the members of a given class than between members of different classes. In the *Phonological* simulations, we partially mimic subregularities which characterize the vowel change and identity verbs in English, by assigning verbs to classes on the bases of stem final CV sequences.

The members of the 4 classes are assembled from a "language" of 700 legal strings. For each simulation, total vocabulary size (500 unique strings) and number of unique strings in each class (type frequency) is held constant: ARBITRARY: 2 types; REGULAR: 410 types; IDENTITY: 20 types; VOWEL CHANGE: 65 types. The number of repetitions of a unique string (token frequency) is manipulated across simulation so that the network experiences some items more frequently than others within a given sweep through the data. The token frequencies used in both the *Parent* and the *Phone* simulations are listed in Table 1. All simulations used 20 input, 20 output, and 20 hidden units, and were run on the "rlearn" simulator (Center for Research in Language, UCSD) using a back propagation learning algorithm. Performance is assessed in terms of the percentage of correct outputs for each verb class. For incorrect outputs, we compute the closest phonological representation in order to estimate the actual "verbal" output of the network, and generate categories of error types (Tables 2 and 3).

RESULTS AND DISCUSSION

Previous results from several series of simulations using this architecture and vocabulary (Plunkett & Marchman, 1989) revealed that both type (class size) and token frequency of a class determine rate of learning and final level of performance within that class. These parameters also influence the degree to which characteristics of one class of mappings will be adopted by the network when forming the past tense forms of verbs in other classes. In general, variations in token frequency were found to have a *greater* effect on performance than type frequency. However, effects can be observed in many directions depending on which strategy is dominant in that simulation. Dominance of a particular strategy is determined by the relative type and token frequencies of the competing classes, in interaction with the global characteristics of the total mapping function that the network is required to perform. A noteworthy characteristic of these networks is their inability to map a large number of arbitrary stems simultaneously.

These "network facts" are informative for understanding children's acquisition of language in only a limited sense -- for our purposes, to the degree that the particular input configuration used accurately represents input to children. It is extremely difficult to determine the exact frequency of English verbs to which children are exposed and/or that children find salient in processing. As a first approximation, we use a single type frequency configuration representing the relative class sizes of English, and vary token frequency parametrically across simulation.

Parent Simulations

Manipulations of token frequency in the *parent* simulations influence the performance of a given class and the extent to which overgeneralization errors were observed. The percent of items correctly output in each simulation are provided in Table 1. When type frequency (class size) of an irregular class is small, increasing its token frequency results in a high level of performance without any deleterious effects on the dominant form of mapping (e.g., suffixation). However, if type frequency is relatively large and backed up by a high token frequency, then the performance on the dominant form of mapping deteriorates dramatically. For example, successful learning of the arbitrary mappings only occurs when class size is small yet each exemplar is presented to the system fairly frequently (i.e., token frequency is greater than 20). Because of the initial biases of these networks to perform identity mapping, performance on the arbitraries is poor unless these type/token constraints are met. Interestingly, once the number of exemplars is sufficient for mastery in the arbitrary class, performance is generally unaffected by manipulations of token frequency for the other types of mappings in the network.

When natural languages incorporate arbitrary forms, they are generally highly frequent and constitute a relatively small class of items. For the child acquiring language, these typological characteristics undoubtedly contribute to the early learning of these forms. However, children will often later overgeneralize the regular mappings to the arbitrary class (*go —> goed*), but eventually return to using the correct form. We also observe this effect in many of the *parent* simulations. However, unlike the R&M simulation, this behavior cannot result from a discontinuity in the vocabulary to which the network is exposed. In these networks, overgeneralizations on arbitrary mappings arise from the need to satisfy a variety of constraints *within the framework of a single mechanism*. The network is forced to reorganize its weight matrix to meet the requirements of the dominant form of mapping. Once this is achieved, however, the network reestablishes correct performance so that arbitrary forms may peacefully co-exist with stems from the other classes.

In addition, we observe that the regular and vowel change mappings frequently compete with each other for network resources in such a manner that neither class can be completely mastered simultaneously. Competition effects result in the production of complex patterns of overgeneralization errors. (See distribution of error types in Table 2). Vowel change overgeneralizations to the regular class can occur at the same time as suffixation overgeneralizations occur to the identity class. Blended errors are also observed (i.e., the application of two mapping regularities in a single form). In studies of children's acquisition of the past tense in English, irregular sub-regularities sometimes give rise to their own patterns of overgeneralization, albeit less frequently than the standard "add -ed" overgeneralization (Bybee & Slobin, 1982; Marchman, 1988). That is, children will sometimes "overgeneralize" a vowel change or identity mapping to a regular or irregular stem, producing errors such as *pick —> pack*, or combining mapping types to produce blended responses, such as *ated* (Kuczaj, 1977). But, because irregular forms are not formed using a "default" rule, errors of this type are generally thought to result via *analogy* to the phonological shape of individual stems (Pinker & Prince, 1988; MacWhinney, 1987). In the *parent* simulations, none of the overgeneralization errors, of both the standard and irregular variety, can be attributed to the phonological structure of the input set given that verbs were randomly assigned to classes. However, none of the simulations succeeded in reaching "adult-like competence" in all classes simultaneously. The next set of simulations, the *phones*, explores whether the addition of phonological predictability into the input set will enable these networks to master the past tense.

Phonological Simulations

English irregular verbs possess phonological properties that can be said to characterize the class, however, they are nevertheless insufficient to reliably predict class membership, i.e., both regular and identity stems end with a dental consonant. As discussed by Pinker & Prince (1988), the lack of phonological similarity within the regular class is crucial to the hypothesis that different mechanisms of past tense formation operate on irregular and regular verbs. The regular rule is applied generally, without reference to the properties of the stem; whereas, irregular transformations take this phonological information into account in the production of a past tense form.

In the *phone* simulations, we impose the following constraints on class assignment: (a) All identity stems must end in a dental, and (b) All vowel change stems are restricted to eleven possible VC stem final sequences. We also ensure that the regular class contains stems possible irregular stems, e.g., some regulars end in a dental. Two questions are relevant: (1) Do the additional constraints aid in the identification of class membership and lead to improved performance? (2) Do patterns of competition and overgeneralization occur when phonological sub-regularities are available to the network that are similar to those when such information is not available? The *phone* simulations repeat the type and token frequencies of the corresponding *parent* simulations (see Table 1), and represent a second approximation to the task facing the young child learning the relationship between the present and past tense forms of English verbs.

In general, all phonological simulations exhibit a higher level of performance, across mapping types, compared to the *parent* simulations, except for the arbitrary mappings in a few simulations (See Table 1). However, note that the regulars perform minimally better under the *phone* condition. The greatest improvement tends to occur when the token frequency of the *vowel change* class is relatively high (simulations 5, 17, 18 and 27). Since there are no differences between conditions other than the sub-regularities in the identities and vowel changes, we can attribute the lower performance of the regulars in the *parent* condition to the absence of these subregularities. In the *phones*, the phonological subregularities conspire to protect the regulars from interference, despite the facts that (a) the regular class contains stems that resemble the vowel change and identity classes (similar vowel and final consonant), and (b) there are no explicit features marking the regular stems as "regular."

Table 3 presents the distribution of error types in the *phone* simulations. In general, these networks treat regulars which end in a dental as identities, however, many "dental final" regulars are mapped correctly and other regular stems are mapped as vowel changes or blended. Regular stems that conform to the characteristics of the vowel change class are often mapped as vowel changes, though again not all regular stems with vowel change characteristics are incorrectly mapped. There was a clear-cut advantage for the identity mappings in the *phones*. The network makes use of the phonological sub-regularities, however, it is also not indiscriminate in its categorization of verbs into classes on that basis (although they are a source of error). In several of the *phone* simulations, the identity class achieves optimal performance. However, *en route*, the mapping undergoes several reorganizations in which some identity stems are alternately treated as regular and vowel change stems *after* having been mapped correctly (see Plunkett & Marchman, 1989, for a discussion of "U-shaped" learning in these networks). Finally, there was a moderate advantage for vowel change mappings in the *Phone* condition, particularly apparent in simulations 17, 19 and 27. In simulations 17 and 27, both the vowel change class and the identity class have relatively large token frequencies. In the *parent* condition, the lack of phonological sub-regularities permits identity mapping to "spill over" into the vowel change class. However, in the *phone* condition, the regularity in the identity class restricts the application of identity mapping to items that possess these characteristics and hence reduces the level of interference with the vowel change class.

The provision of phonological constraints on class membership enables the competition effects between classes to diminish, improving overall performance in the *phones*. Nevertheless, *patterns* of learning are observed that are similar to those in the *parent* simulations even though many errors do bear the stamp of the phonological structure of the input set. The predominant error types for both sets of simulations are similar: Regulars, the most common error is identity mapping; Identities, the most common error is suffixation; Vowel changes, suffixation, identity mapping and blending, in that order. Similarly, blending errors in the identity class are absent in both the *parent* and *phone* simulations.

Clearly, the *phone* simulations map input stems in light of the phonological information concerning class membership. These networks seem to increasingly resemble a rule-governed, categorical system as the constraints on the network (represented here as *external* pattern constraints rather than *internal* architectural constraints) are tightened. The constraining effect of the phonological sub-regularities is particularly apparent in those simulations which otherwise give rise to substantial competition effects (compare *phones* 5, 17, 18 and 27 to those in the *parent* set). Phonological sub-regularities can, thus, serve to both *support and constrain* the observed frequency effects, both factors working together toward successful performance across all classes. Just as these networks can partition the arbitrary mappings so that they appear immune to various parameter manipulations, the introduction of phonological sub-regularities results in a system which is increasingly impervious to token manipulations of the input vocabulary. Frequency effects do not disappear, but instead are modulated by the internal structure of the sets of items that the network is required to process across learning.

CONCLUSIONS

This paper explored the "acquisition" of the English past tense by a 3 layer back-propagation network. Token frequency and phonological regularities crucially affect how well the network solved the problem, as well as the degree to which it "overgeneralized" identity markings and vowel changes in addition to making the standard suffixation error. In some simulations, it may be useful to describe performance (both correct and erroneous) as the result of a general strategy or "rule." However, within any given simulation, overgeneralization errors were rarely restricted to a single type and each verb class was susceptible to a relatively idiosyncratic set of error types. While the production of errors has been the focus here, several networks were indeed able to successfully "memorize" arbitrary forms at the same time that they were overgeneralizing regularities in the other three classes. Yet, the mechanism guiding this memorization was the "same" as that guiding the rule-like overgeneralization behavior.

The degree to which these results are analogous to the acquisition patterns of children is not as yet totally clear. Some analyses suggest that children's production repertoires reflect a variety of competing strategies which determine both correct and incorrect performance (Derwing & Baker, 1986; Marchman, 1988). Children, like these networks, are not likely to be exclusively suffix generalizers, or identity mappers, but will produce several different types of errors in generating past tense forms throughout acquisition. Rule-based models explicate this phenomenon via the competition between two (or more) discrete and explicitly represented hypotheses which, at various points in development, undergo changes in how and when they are likely to apply (see Pinker & Prince, 1988). In these networks, probabilistic differences between individual mapping strategies are a by-product of the learning process. Output fluctuations are the result of the implicit encoding of similarity relationships between the input stems in the weight matrix of the network. High token frequencies tended to "localize" the zones of interference of mapping types while high type frequencies tend to extend them. In addition, errors such as *ated* or *stooded* which blend two potential regularities in a single form were also observed. However, "blends" were relatively rare overall and predominated only in the vowel change class. Identity stems virtually never underwent blending. The introduction of phonological sub-regularities further restricted the occurrence of blending errors. These types of errors are also produced by children, yet they are much less frequent than the standard overgeneralization of "add -ed," and are likely to occur later in development. Further analysis is required to outline the developmental priority of "pure" over "blended" overgeneralizations in these networks.

Within certain limits, these networks sometimes behaved as if they were doing what we know children do when acquiring a morphological system. Classic patterns of overgeneralization were elicited (*without* introducing continuities into the learning set) by manipulating token frequency in networks in which phonological information *was and was not* available to define class membership. Clearly, our representation of the input conditions is far from adequate, and semantic information *must* play a role in the disambiguation of certain present tense/past tense mappings. However, the degree to which these systems behave in ways that are reminiscent of phenomena of acquisition reinforces the assumption that there is much to be gained from careful study of the nature and structure of input in the problem of language acquisition.

ACKNOWLEDGEMENTS

A more detailed description of this research is provided in Plunkett and Marchman (1989). The research was supported in part by grants from Center for Neurodevelopmental Studies at UCSD (NIH PHS NS22343), the Danish Humanities Research Council, and the Danish Technical Research committee. The authors sincerely thank the members of the Center for Research in Language and the PDP Natural Language Processing Discussion group at UCSD.

Table 1. Token frequencies and percent of items correctly produced after 50 sweeps.

SIM	ARBITRARY Token Freq.	ARBITRARY Par	ARBITRARY Phone	REGULAR Token Freq.	REGULAR Par	REGULAR Phone	IDENTITY Token Freq.	IDENTITY Par	IDENTITY Phone	VOWEL CHANGE Token Freq.	VOWEL CHANGE Par	VOWEL CHANGE Phone
1	1	0	0	1	96	94	1	15	50	1	0	3
2	5	0	0	1	92	89	1	10	40	1	5	1
3	10	100	0	1	96	93	1	0	65	1	0	0
4	15	50	100	1	96	90	1	15	40	1	3	1
5	15	50	100	1	50	69	1	20	35	5	82	88
6	20	100	100	1	93	92	1	5	60	1	0	6
7	40	100	100	1	90	92	1	15	65	1	3	6
8	145	100	100	1	87	90	1	10	60	1	3	10
9	15	50	100	1	77	76	1	40	40	2	22	31
10	15	100	50	1	84	88	5	75	95	1	6	13
11	15	100	50	1	82	86	14	100	100	1	6	6
16	15	100	50	1	74	83	16	100	100	1	6	4
17	15	50	50	1	65	71	5	70	100	3	28	62
18	20	100	100	1	59	67	1	30	40	5	72	88
19	20	100	50	1	82	79	1	25	35	2	18	44
20	20	100	100	1	88	87	5	80	100	1	5	7
21	20	100	100	1	82	88	14	100	100	1	2	4
26	20	100	50	1	76	83	16	100	100	1	3	7
27	20	100	100	1	60	73	5	70	100	3	37	57
Mean	--	79	68	--	80	83	--	46	70	--	16	23

Table 2: Distribution of Error Types in Parent Simulations

SIM	REGULAR Inap Suf	REGULAR Iden	REGULAR Inap Vow-S	REGULAR Blend	REGULAR Vow Chan	IDENTITY Suf	IDENTITY Inap Suf	IDENTITY Vow Chan	IDENTITY Inap Vow	VOWEL CHANGE Suf	VOWEL CHANGE Iden	VOWEL CHANGE Inap Vow-S	VOWEL CHANGE Inap Suf	VOWEL CHANGE Blend	VOWEL CHANGE Inap Vow
Par1	19	50	19	12		82	18			54	8	18	3	8	3
Par2	59	28		3		94		6		48	16	11	8	11	3
Par3	31	13	31	6		85	15			60	6	20	11	3	
Par4	27	20	27	13		74	21			63	10	14	3	8	2
Par5	18	61	17	2	12	63	6	6		55				18	9
Par6	41	11	30	4		95	5			63	9	13	2	8	5
Par7	50	28	11			82	18			51	10	25	3	5	2
Par8	58	25	6			100				48	10	22	6	6	
Par9	20	30	18	16	2	92		8		20	24	4	2	14	27
Par10	37	55	3	2		80	20			41	20	14	3	11	6
Par11	38	53	4							27	29	13	11	4	9
Par16	35	56	4	1						36	41	3	8	8	2
Par17	22	45	11	6	6	17	83			7	20	20		20	27
Par18	15	28	25	8	4	71			7		19			50	13
Par19	18	46	8	9		89	6	5		48	16	10	65	13	3
Par20	27	51	4	7		75	25			52	19	12	7	3	7
Par21	43	35	10	4						36	34	9	5	6	5
Par26	42	52	3		2					42	35	3	5	3	6
Par27	20	53	9	3	3	67	33			11	14	19		11	46
Mean	33	39	13	6	5	78	23	6	7	42	19	13	9	11	10
sd(σ)	13	15	9	4	3	19	20	1	0	16	10	6	15	10	12

Table 3: Distribution of Error Types in the Phone Simulations.

SIM	TYPE OF MAPPING: REGULAR — ERROR TYPE					IDENTITY — ERROR TYPE				VOWEL CHANGE — ERROR TYPE					
	Inap Suf	Iden	Inap Vow-S	Blend	Vow Chan	Suf	Inap Suf	Vow Chan	Inap Vow	Suf	Iden	Inap Vow-S	Inap Suf	Blend	Inap Vow
Phone1	27	50	9	9			100			33	14	12	9	20	3
Phone2	28	61	4	7		100				28	22	2	9	18	6
Phone3	11	44	15	7		100				30	18	12	6	14	6
Phone4	35	51	8	3		92			8	39	15	14	3	17	4
Phone5	13	33	17	6	8	33	8	17	25					33	50
Phone6	23	57	7	3		87			13	35	8	11	11	20	9
Phone7	13	59	13	6		100				34	25	11	8	8	9
Phone8	11	53	26	5		86	14			42	12	7	3	15	13
Phone9	32	46	8	4	2	92				9	27		7	24	24
Phone10	4	83	2	2		100				22	10	12	14	14	19
Phone11	7	54	15	7						15	18	14	12	23	8
Phone16	21	56	15	2						21	11	9	9	26	10
Phone17	16	45	16	6	4						18		5	23	45
Phone18	10	34	14	11	13	50		10	30		50		33	17	
Phone19	20	34	14	4	5	85	15			14	17	3	3	14	39
Phone20	11	70	4	4	6					48	16	6	3	19	3
Phone21	8	71	4	6	8					34	25	9	5	17	5
Phone26	13	81	3	1						29	32	5	15	5	8
Phone27	14	46	13	3	9					4	29	11		25	14
Mean	16	54	11	5	7	84	34	14	19	27	20	9	9	19	15
sd(σ)	9	14	6	3	3	22	44	5	10	12	10	4	7	6	15

The error categories presented in Tables 2 and 3 are to be interpreted as follows:

Regular Errors

Inap Suf The Stem is suffixized but with the wrong suffix.

Iden The stem is treated as an identity stem.

Inap Vow-S The stem is appropriately suffixized but undergoes an illegal vowel transformation.

Blend The stem is appropriately suffixized but undergoes a legal transformation.

Vow Chan The stem is treated as though it were a vowel change stem.

Identity Errors

Suf The stem is treated as a regular stem.

Inap Suf The stem is treated as a regular stem but inappropriately suffixized.

Vow Chan The stem is treated as though it were a vowel change stem.

Inap Vow The vowel change transformation is inappropriate though may or may not be "illegal"

Vowel Change Errors

Suf The stem is treated as a regular stem.

Iden The stem is treated as an identity stem.

Inap Vow-S The stem is transformed as though it were a vowel change but the vowel change is "illegal" The stem is appropriately suffixized.

Inap Suf The stem is treated as though it were a regular but inappropriately suffixized.

Blend The stem undergoes a legal vowel change but is also appropriately suffixized.

Inap Vow The vowel change transformation is inappropriate though may or may not be "illegal"

REFERENCES

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150-177.

Bever, T. G. (in press). The demons and the beast -- Modular and nodular kinds of knowledge. To appear in C. Georgopoulos & R. Ishihara (Eds.), *Interdisciplinary approaches to language: Essays in honor of S-Y. Kuroda.* Kluwer: Dordrecht.

Borer, H. & Wexler, K. (1987). The maturation of syntax. In T. Roeper and E. Williams (Eds.), *Parameter Setting.* Dordrecht, Holland: Reidel.

Bowerman, M. (1982). Reorganizational process in lexical and syntactic development. In E. Wanner & L. Gleitman (Eds.) *Language Acquisition: The State of the Art.* Cambridge: Cambridge University Press.

Bybee, J. & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language, 58*, 265-289.

Chomsky, N. (1980). Rules and Representations. *Behavior and Brain Sciences, 3*, 1-61.

Derwing, B. L. & Baker, W. J. (1986). Assessing morphological development. In P. Fletcher & M. Garman (Eds.), *Language acquisition: Studies in first language development.* Second edition. Cambridge: Cambridge University Press.

Derwing, B. L., & Skousen, R. (1989). Real-time morphology: Symbolic rules or analogical networks. Proceedings of the 15th Annual meeting of the Berkeley Linguistics Society, Berkeley, CA. February.

Fodor, J. & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3-71.

Kuczaj, S. (1977). The acquisition of regular and irregular past tense verbs. *Journal of Verbal Learning and Verbal Behavior, 16*, 589-600.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, N.J.: Erlbaum.

Marchman, V. (1984). Learning not to overgeneralize. *Papers and reports on child language development, 24*, 69-74.

Marchman, V. (1988). Rules and regularities in the acquisition of the English past tense. *Center for Research on Language Newsletter, 2(4)*, April, 1988.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. & Mehler, J. (Eds.) (1988). *Connections and Symbols.* Cambridge, MA: MIT Press.

Pinker, S. & Prince, A. (1988). On Language and Connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 59-108.

Plunkett, K., & Marchman, V. (1989). Pattern association in a back propagation network: Implications for Child Language Acquisition. Technical Report #8902. Center for Research in Language, University of California, San Diego.

Rumelhart, D. E. & McClelland, J.L. (1986). On learning the past tense of English verbs. In D.E. Rumelhart & J.L. McClelland & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2.* Cambridge, MA: Bradford Books.

Rumelhart, D. E. & McClelland, J. L. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processing. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition.* Hillsdale, N.J.: Erlbaum.

Smolensky, P. (1988). The constituent structure of connectionist mental states: A reply to Fodor & Pylshyn. Technical Report, Department of Computer Science, University of Colorado, Boulder, CO.