Title
Computational Bases of Two Types of Developmental Dyslexia

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 18(0)

Authors
Harm, Michael W.
Seidenberg, Mark S.

Publication Date
1996

Peer reviewed

Computational Bases of Two Types of Developmental Dyslexia

Michael W. Harm and Mark S. Seidenberg
University of Southern California
Neuroscience Program
Los Angeles, CA 90089-2520
{mharm,marks}@gizmo.usc.edu

Abstract

The bases of developmental dyslexia were explored using connectionist models. The behavioral literature suggests that there are two dyslexic subtypes: "phonological" dyslexia involves impairments in phonological knowledge whereas in "surface " dyslexia phonological knowledge is apparently intact and the deficit may instead reflect a more general developmental delay. We examined possible computational bases for these impairments within connectionist models of the mapping from spelling to sound. Phonological dyslexia was simulated by reducing the capacity of the models to represent this type of information. The surface pattern was simulated by reducing the number of hidden units. Performance of the models captured the major behavioral phenomena that distinguish the two subtypes. Phonological impairment has a greater impact on generalization (reading nonwords such as NUST); the hidden unit limitation has a greater impact on learning exception words such as PINT. More severe impairments produce mixed cases in which both nonwords and exceptions are impaired. Thus, the simulations capture the effects of different types and degrees of impairment within a major component of the reading system.

1. Introduction

One of the attractions of the connectionist or parallel distributed processing approach is that it can be used to develop unified accounts of normal and disordered behavior. Effects of brain injury or developmental anomaly can be simulated by "damaging" components of a neural network model of normal performance. A prominent example of this approach is provided by research on reading and dyslexia. Becoming a skilled reader involves mastering the correspondences between spelling and pronunciation. Sejnowski and Rosenberg (1987) developed a neural network model of this process, and Seidenberg and McClelland (1989; hereafter SM89) used a similar model to account for detailed aspects of behavior. Dyslexia--failures to acquire age-appropriate reading skills despite normal intelligence and adequate opportunity to learn--is often associated with impairments mapping from spelling to sound (Castles & Coltheart, 1993). Our goal was to see if the behavioral impairments associated with dyslexia could be explained in terms of damage to a model of skilled reading.

There is an emerging consensus that there are two prominent subtypes of developmental dyslexia (Castles & Coltheart, 1993; Murphy & Pollatsek, 1994; Manis et al., 1996). The reading impairment observed in phonological dyslexia is apparently secondary to impaired processing of spoken language. Such children perform poorly on spoken language tasks such as counting the number of syllables in a word or deciding if two words rhyme (see Farmer & Klein, 1995, for review). In reading they are markedly impaired in their ability to use their knowledge of spelling-sound correspondences to pronounce novel letter strings (nonwords such as NUST). These children do not resemble younger children who are learning to read normally. The second subtype has been termed developmental surface dyslexia (Castles & Coltheart, 1993). Such children are also impaired in reading but their phonological processing capacities appear to be intact. They have particular difficulty learning to read words with irregular spelling-sound correspondences, such as GIVE and PINT. These children's performance closely resembles that of much younger normal readers; hence they exhibit a developmental delay.

There are two theoretical accounts of these phenomena, tied to models of normal word recognition. In the dual-route model (Coltheart et al., 1993), there are separate "lexical" and "nonlexical" mechanisms for pronouncing letter strings. The "lexical" mechanism involves knowledge associated with specific words; it provides the only way of pronouncing irregular words such as PINT and cannot be used to pronounce novel strings such as NUST. Surface dyslexia is thought to involve an impairment in acquiring this mechanism. The "nonlexical" mechanism consists of rules governing correspondences between graphemes and phonemes; it can be used to pronounce novel letter strings but not irregular words. Phonological dyslexia is thought to involve an impairment in acquiring the pronunciation rules. Note that Pinker's (1991) theory of the past tense is also a dual-route model, with a rule component distinct from a word-specific component.

A number of behavioral phenomena related to normal performance and effects of brain injury on reading present difficulties for the dual-route theory (Seidenberg, 1995; Plaut et al., 1996). Several aspects of developmental dyslexia present further challenges for this approach. The idea that phonological dyslexia involves an impairment in acquiring grapheme-phoneme correspondence rules misses the fact that these children have broader phonological impairments that are manifested in tasks other than reading. The idea that an impairment in the lexical mechanism underlies the surface pattern fails to explain the fact that such children tend to exhibit a broad developmental delay that affects all aspects of reading (Manis et al., 1996), not just exception words. Finally, it is an embarrassment for the dual-route theory that selective impairments in the two processing subsystems are rarely if ever observed. Most dyslexics are impaired in reading both exception words and nonwords; the dual-route theory can only explain this by assuming that both routes happen to be impaired in most cases, but there is no independent evidence that this is so (Manis et al., 1996).

An alternative to the dual-route account is provided by connectionist models in which there is a single, homogeneous mechanism for mapping between spellings and pronunciations. Such models provide a good account of a broad range of phenomena concerned skilled performance and breakdown following brain injury (SM89; Plaut et al., 1996). The different patterns of developmental dyslexia might be explained within such models in terms of different *types* of damage to a single underlying mechanism, rather than damage to different pronunciation *mechanisms*. For example, Manis et al. (1996) suggest that the phonological subtype could result from impairments in phonological representation in an SM89 style model. Such degraded representations would make it harder to acquire spelling-sound correspondences and also interfere with performance on other tasks involving phonological information. Similarly, the surface form could derive from a limitation on the capacity of the network to encode information--for example, limiting the number of hidden units. This would affect the learning of exception words (see SM89 for details) but a severe enough impairment would affect regular words and generalization as well.

The purpose of the present research was to assess the adequacy of the connectionist account by seeing if we could account for major aspects of the distinct dyslexic subtypes. We implemented a version of an SM89-style model of the mapping from spelling to sound and then ran versions with either phonological or capacity limitations. The models were assessed in terms of their capacities to learn words with regular and irregular pronunciations and to pronounce novel items.

2. Model Architecture

The phonological representation used in the simulations consisted of 6 slots, each slot corresponding to a phoneme in a monosyllabic word, and consisting in turn of 11 phonetic features: sonorant, consonantal, voiced, nasal, degree, labial, palatal, pharyngeal, lower_lip, tongue and radical. These features could take on a continuum of values ranging from between -1 and +1. The slot arrangement was vowel centered, and could encode syllables of CCVVCC format. A word could have at most two consonants before the vowel, and two after. Normal vowels were encoded as a single vowel phoneme and a second empty slot; diphthongs were encoded as pairs of vowel slots. The orthographic representations consisted of 8 slots, representing letter positions. Letters were encoded using a localist representation, with 26 units per position, and were also vowel centered. Up to 3 consonants could be represented before the initial, centered vowel, and up to 4 letters (consonants or vowels) after the vowel.

The 66 phonological units were fully connected to one another with initially random weights ranging from -0.001 to 0.001. An additional set of 20 cleanup units were added, with initially random weights going from each of the phonological units to each cleanup unit, and back from the cleanup units to the phonological units (see Figure 1). These units are analogous to the cleanup units used in semantic attractor networks (e.g., Plaut & Shallice, 1993). The direct connections between featural units within a phoneme were able to encode intraphonemic constraints; those between slots encoded constraints related to the sequence of phonemes. The cleanup units allowed higher order dependencies among features to be represented.

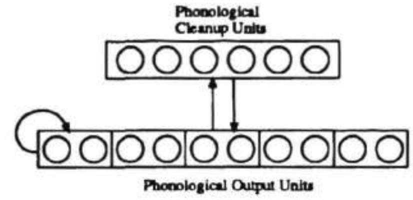

Figure 1: The Phonological Component

The schematic architecture of the "normal" reading model is shown in Figure 2. Orthographic units projected onto a set of 100 hidden units, which in turn projected onto the phonological units of the phonological component. The task of the reading model was to map orthographic representations of words onto the correct phonological units.

We then modified this architecture to examine the effects of phonological impairments on reading acquisition. Two conditions were used which imposed different limitations on the extent to which phonological information could be encoded. The first model was identical to the normal unimpaired model except that the weights in the phonological network were subject to weight decay (see Hinton, 1989). The effect of this decay is to apply pressure to the network to avoid large values on the weights. The network can still encode higher order relationships between the units, but the strength of these encodings is curtailed. A weight decay constant of 0.00005 was used.

In the second, more severely impaired simulation the cleanup units were deleted from the phonological attractor network. By removing the cleanup units, we disabled the network's ability to encode higher order relationships among the phonological units. This impaired phonological component had only direct connections between the phonological units, and hence was limited in the complexity of computations it could perform (see Minsky & Papert, 1969). Both of these simulations had 100 hidden units, the same number as the normal model.

In each condition we first trained the phonological component on a set of phonological word forms. The weights that resulted from this pretraining were used when each phonological component was incorporated in a model that learned to pronounce written words. The goal was to determine how reductions in the capacity to represent phonological information would affect performance on the spelling-sound mapping task.

To assess the effect of reducing the computational resources available to the reading task while preserving phonological knowledge an additional pair of simulations were run. They were identical in architecture to the normal model, except that one models had only 35 hidden units, and a second one had only 20. These models had the same phonological representation as used in the normal condition.

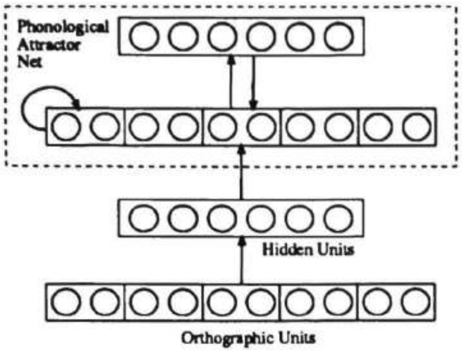

Figure 2: The reading model, with phonological output network in place

3. Training The Models

Phonological Pretraining

The phonological component was trained on a corpus of 3123 monosyllabic words using backpropagation through time (Rumelhart, Hinton & Williams, 1986). The weights from each phonological unit to itself were set to 0.75 and frozen. This gave the phonological units a tendency to hold onto their current value, but decay slowly to zero. Other weights were initially randomized. Each word was given a probability of presentation according to its estimated frequency of occurrence using a sample of 30 million words from the Wall Street Journal. The probability for each word was set to the logarithm of that word's frequency in the WSJ divided by the logarithm of the frequency of the most frequent word ("the"). Training proceeded as follows. A word was probabilistically chosen from the training set. For tick 0, the phonological units were set to the values corresponding to its phonological form. The network was allowed to run for 5 ticks, with all units unclamped for the last 4 ticks. The output of the network during ticks 2 through 4 was compared to the original phonological form of the word. Error was injected into the network based on the difference between output values and the targets, and the weights were adjusted so as to reduce this error. Then another word was chosen randomly and the process repeated. The overall effect of this training regime is to force the weights to encode statistical relationships between the phonological units. Training was halted after 1 million training trials. At the conclusion of training, the baseline network's mean sum squared error was 0.05, the network with weight decay on the phonological weights ended with a mean error of 1.8, and the network with no cleanup connections ended with a mean error of 0.8. These error scores are the average summed error over 66 output units, so the average deviance from unit output to target for the 3 simulations was, respectively, 0.001, 0.027 and 0.012.

To further assess the quality of the phonological representations, a simple pattern completion task was devised. In this task, for each of the words in the training set with an initial consonant cluster, the features of the second consonant slot was left unspecified while the remainder of the word form was clamped to that word's phonological form. The network was then run for 6 ticks, and the value which the unspecified phoneme was drawn into was assessed. The word was scored correct if the segment that was produced (evaluated by the nearest neighbor criterion, described in section 4) was a legal segment for English in that environment. For instance, for the word /b/l/u/, if the net output /b/r/u/ that would also be scored as acceptable. The normal network was able to insert a legal phoneme into the slot 75% of the time. The decay network could only produce a legal phoneme 46% of the time, and the network without cleanup units could only produce a legal output 10% of the time.

This test is not meant to be a full test of the networks' phonological competency, but rather a gross measure of the quality of the phonological attractor basins that the different conditions represent. The network with weight decay is impaired relative to the normal network due to the downward pressure on the magnitude of the weights during training. The network without cleanup units is even more impaired, because this task, like the XOR task, relies on the conjunctive use of other features in the word's environment, and not simply direct relationships between features.

Training on Reading Task

The pretrained phonological components were then used in models that were trained on the reading task. In both the normal model and the reduced resource models the standard phonological component training method was used. The phonologically impaired conditions utilized representations trained with weight decay or cleanup unit deletion, as described above. In each case, weights were initialized to the final values from the relevant pretraining phase. The remaining weights in the network (orthographic to hidden; hidden to phonological) were initialized to small random values. The model was then trained on the same corpus of 3123 words, again using log frequency to determine probabilities of being selected for training. For each word chosen, the orthographic units were clamped with the appropriate values for ticks 0-6. At tick 6, the phonological output was compared with the phonological target, error was injected into the network, and the weights were updated.

Four replications were run for each condition (normal, the 2 reduced resource conditions, and the two phonologically impaired conditions). For each simulation run, a different random number seed was used, resulting in different distributions of initial random weights, and a different ordering of the presentation of words.

4. Results

Two scoring methods were used. In the nearest-neighbor method, the phonological output of each 11 units within a phoneme slot is compared to the representations for each of the phonemes that exist in the training set. The phoneme that is closest in euclidean distance to the output is the one that is taken to be the output. A second, more stringent threshold method was also used, and unless otherwise noted will be the one reported below. For this measure, each feature of a phoneme had to be within a specified distance of the target for the phoneme to be counted as correct. A threshold value of 0.5 was used, covering 25% of the units' activation range of -1 to 1. In both cases, a word was scored as correct only if all of its phonemes were correct. To evaluate the networks' performance on words, we used a set of frequency 93 regular items such as BACK, and 92 exceptions, such as COMB, taken from the "surface list" developed by Patterson & Hodges (1992).[1] For nonwords we constructed a set of 367 items (e.g., GOMB, SOAD, FAIJE) taken from items used by McCann & Besner (1987), Glushko (1979) and Seidenberg et al. (1994). Regular words follow the putative spelling-sound correspondence rules of the language, and exceptions violate them. Nonwords assess the ability to generalize to untrained forms.

All models were evaluated after running for 8.5 million words. In almost all cases, learning had ceased long before this point (see Figures 3-6). Asymptotically, the normal models got an average of 98% of the training set correct when scored with the nearest-neighbor method and 83% of the nonwords[2]. Using the threshold method, the average results were 97% and 75%, respectively. Nonword performance is somewhat lower than levels reported for people, particularly with the threshold method. This measure is quite conservative, however; for example, some small deviations from target values that are scored as incorrect would not be perceivable in humans. Also, we have made no attempt to improve nonword performance using various techniques known to facilitate generalization (e.g., pruning, noise). Plaut et al. (1996) discuss other factors that affect nonword generalization.

Phonological Impairments

Figures 3 and 4 show the developmental curves for the impaired phonological knowledge conditions compared with the normal condition. All plots show the average of four simulation runs. With mild levels of phonological impairment (i.e., weight decay, Figure 3), there are decrements on both the rate of acquisition and asymptotic performance on nonwords, but very little effect on regulars and exceptions (see Figure 7 for a summary of the asymptotic conditions). With the nocleanup net (Figure 4), the exceptions also begin to show a decrement in rate of acquisition relative to the normal network. Acquisition of the capacity to generalize is also being slowed, though less than in the decay condition. Mild phonological impairment has little effect on the rate of acquisition for regular and exception items. With the more extreme impairment, there is slower acquisition of exceptions in addition to poor performance on nonwords throughout development.

These simulations capture the basic characteristic of phonological dyslexia, that nonword generalization is impaired more than performance on vocabulary words. In the relatively pure cases of phonological dyslexia, subjects' performance on regular and exception words is close to normal, while nonword generalization is poor (Castles & Coltheart, 1993). This result is produced by the decay condition. Many other phonological dyslexics exhibit a "mixed" pattern in which performance on exception words begins to be affected as well. This outcome was observed in the nocleanup condition.

The Reduced Resource Conditions

Figures 5 and 6 summarize the time course of training in the reduced resource conditions relative to the normal baseline model. At the end of training, the reduction to 35 HUs had almost no effect on regular words or on nonword generalization (see Figure 7). However, for exceptions, the 35 HU case shows a drop from 91% to 83% correct in asymptotic performance, relative to the normal network. Decreasing the number of HUs slows learning for all types of items (Figures 4 and 5), but the effect is biggest for the exceptions. With only 20 hidden units, there is a bigger effect on exceptions, and the developmental curves for nonwords and regulars begins to be affected as well.

These simulations capture basic characteristics of the surface dyslexia pattern. In relatively pure cases, reading of regular words and nonwords is intact, but exception words are impaired. With more severe deficits, the regulars and nonwords start to be affected, with exceptions most vulnerable.

5. Discussion

The simulations show that deficits associated with two major patterns of developmental dyslexia can be produced by different types of impairments to a model of normal performance. The phonological pattern derives from impairments in the capacity to represent this type of information. This account can explain why phonological dyslexics are also impaired on spoken language tasks such as rhyme detection. The phonological representations in question are not specific to reading; they are also used in the perception of spoken language. This pattern of correlated reading and spoken language deficits is more difficult to explain within the dual-route model, which attributes phonological dyslexia to an impairment in learning grapheme-phoneme correspondence rules. Why this should also affect spoken language tasks is not clear.

We have derived the surface pattern from a resource limitation, which slows learning across the board. The model retains the capacity to encode the simple and consistent spelling-sound correspondences and eventually masters them with sufficient training; however, its capacity to encode irregular words is limited. This represents an alternative to the standard dual-route account, which holds that the surface pattern results from damage to a "lexical" processing mechanism that encodes the pronunciations of all words. This approach has difficulty accounting for the prevalence of the mixed pattern, in which performance is impaired on both words and nonwords. Thus, on our view the surface pattern represents a kind of general developmental delay that has broad effects on acquisition but especially on

---

[1] Some items from their list were excluded because they cannot be represented in our scheme.

[2] A full listing of the items, with network outputs, is available on a web page at http://maestro.usc.edu:8080/mwharm/cogsci96.html

learning exceptions. Although we have derived the surface pattern by manipulating the number of hidden units, other types of anomalies could be expected to produce similar effects. For example, a visual-perceptual impairment that had the effect of degrading the input orthographic patterns would also cause broad learning delays with the largest impact on the words with unusual spellings or pronunciations.

Why do the different types of anomalies have different effects on network behavior? Degrading the phonological capacity of the network by eliminating the cleanup units and interconnections between featural units forces the network to memorize the training set, yielding poor generalization. With the additional units, the network can encode aspects of the structure of phonological space independently of how this information relates to orthography. An analogous condition exists with the weight decay simulation: the phonological network is prevented from fully developing high quality representations. The normal network, having developed rich attractors in the phonological component, can be more sloppy in its conversion from orthography to phonology, which discourages overfitting of the training set. In effect, it is less likely to become a whole-word reader because it has the phonological safety net in place. Because the hidden units have a higher demand placed on them in the face of phonological impairment, exception words are secondarily impaired: the network has effectively fewer hidden units to learn exception items, because it needs to recruit more to produce an accurate phonological output. In contrast, the reduced resource simulations do not have the capacity to memorize the training set, and focuses instead on the redundant correspondences that characterize the "regular" or "rule-governed" words. At asymptote performance on regulars and nonwords is relatively spared, with an impairment on the exception words. With more severe restrictions on computational resources, the net's capacity to encode even the relatively simple and consistent spelling to sound correspondences would be impaired.

In conclusion, these simulations provide further insight into the nature and causes of the two dissociable forms of developmental dyslexia, while demonstrating the validity of a model of word recognition that employs a single pronunciation mechanism rather than the two "routes" of the dual-route model. There is strong independent evidence concerning the existence of phonological processing impairments in children who exhibit the behavioral profile termed phonological dyslexia (see Farmer & Klein, 1995), and therefore our simulation of this deficit pattern by degrading the phonological component has considerable face validity. The surface pattern is less common and there is no independent evidence whether it arises from a resource limitation, a visual processing deficit, or some other cause. The kinds of computational impairments that could give rise to that behavioral pattern can be seen in the present work, motivating further investigations of their bases in human development.

Acknowledgments

Research support by NIMH grant 47566 and a Research Scientist Development Award to M.S. Seidenberg. We thank Lori Altmann for assistance in developing the phonological representation.

References

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*(2), 149-180.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed- processing approaches. *Psychological Review, 100*(4), 589-608.

Farmer, M., & Klein, R. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review*, 4(2), 460-493

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 674-691.

Hinton, G.E. (1989). Connectionist learning procedures. *Artificial Intelligence, 40*, 185-234.

Manis, F., Seidenberg, M., Doi, L., McBride-Chang, C., & Peterson, A. (1996). On the basis of two subtypes of developmental dyslexia. *Cognition 58*, 157-195.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception & Performance, 13*(1), 14-24.

Minsky, M., & Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.

Murphy, L., & Pollatsek, A. (1994). Developmental dyslexia: Heterogeneity without discrete subgroups. *Annals of Dyslexia, 44*, 120-146.

Patterson, K., & Hodges, J. R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia, 30*(12), 1025-1040.

Pinker, S. (1991). Rules of language. *Science 253*, 530-534.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*(5), 377-500.

Plaut, D. C., McClelland, J. L., Seidenberg, M., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review 103*(1), 56-115.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing, vol. 1.*, . Cambridge, MA: MIT Press.

Seidenberg, M.-S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Seidenberg, M. S., Plaut, D. C., Petersen, A.-S., McClelland, J. L., & et al. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 20*(6), 1177-1196.

Seidenberg, M.S., 1995. Visual word recognition: an overview. In J. L. Miller & P. Eimas (Eds), *Speech, Language and Communication*, 1995, Academic Press, Inc.

Sejnowski, T.J., & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English Text. *Complex Systems, 1*, 145-168.

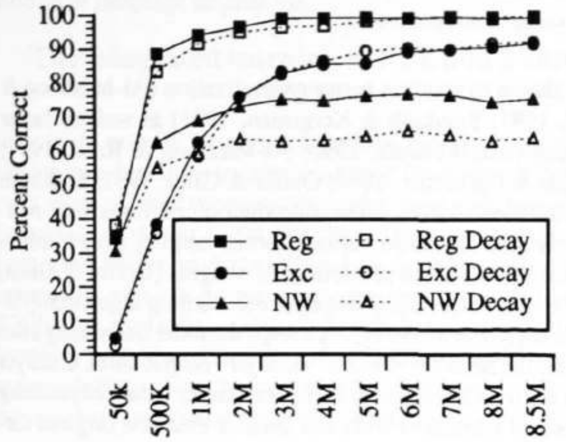

Figure 3: Normal network compared to phonological weight decay network. "Pure pattern" in which only nonwords are affected.

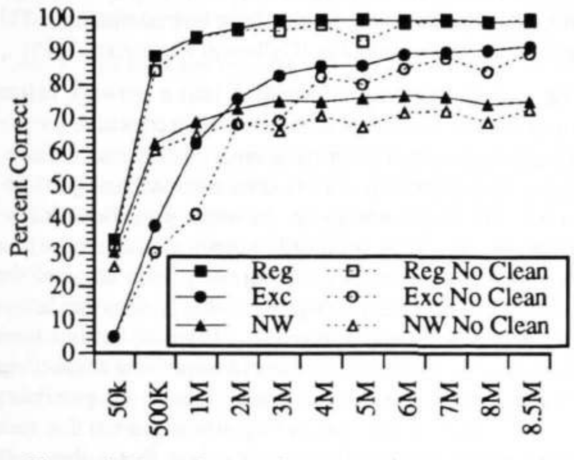

Figure 4: Normal network compared to network without phonological cleanup units. "Mixed" pattern in which exceptions are also affected.

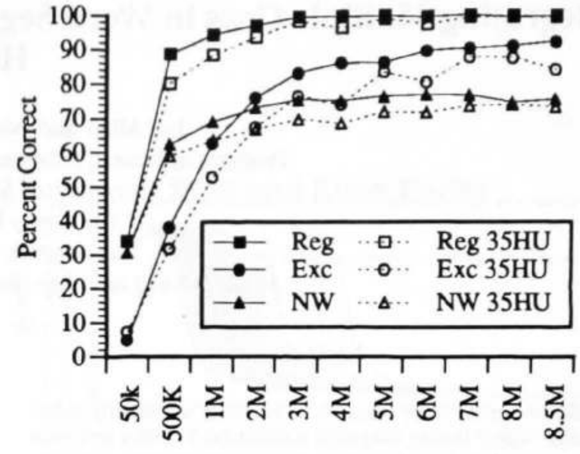

Figure 5: Normal network compared with 35 hidden unit network. Capacity to learn exceptions is impaired.

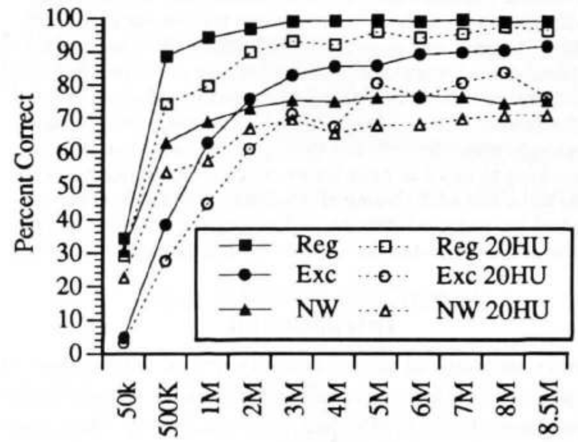

Figure 6: Normal network compared to 20 hidden unit network. Learning of regular correspondences starts to be impaired as well as exceptions.

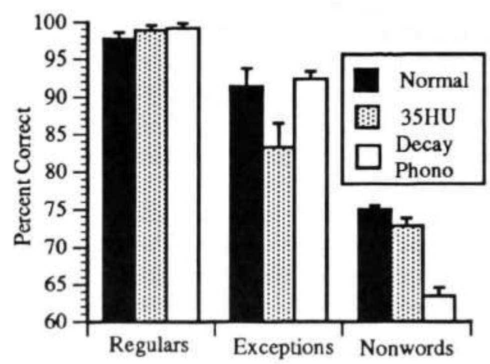

Figure 7: Asymptotic performance for crucial conditions illustrating double dissociation between exception learning and generalization.