CONSTRAINING INTERACTIVITY: EVIDENCE FROM ACQUIRED DYSLEXIA[1]

Gordon D. A. Brown

Department of Language and Linguistics
University of Essex

ABSTRACT

It is sometimes claimed that interactive-activation models are too powerful, and that it is difficult to constrain them adequately. I illustrate this problem by showing that the basic interactive-activation architecture has several different possible sources for effects of spelling-to-sound regularity on word naming. I then show how data can constrain the architecture. New data lead to a rather different and more constrained version of the interactive-activation model to account for spelling-to-sound conversion. Analysis of the errors made by patients suffering from acquired surface dyslexia confirms the predictions of the constrained model. It is concluded that the traditional interactive-activation framework must be considerably constrained to account for normal and disturbed word naming.

INTRODUCTION

An early version of the interactive-activation (IA) model (McClelland & Rumelhart 1981; Rumelhart & McClelland, 1982) successfully accounted for contextual effects on letter perception. Since then, the IA framework has been used to account for human performance in a wide variety of domains.

One reason for the popularity of the IA framework is that it provides a general and powerful mechanism for building cognitive models. Some researchers have worried that the resulting models may even be too powerful, and difficult to constrain. In this paper I show that this worry is sometimes justified, for a number of different IA architectures can

1 This work was supported by the Economic and Social Research Council (U.K.), reference number C08250011. Reprint requests to: Department of Language and Linguistics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England.

predict the "basic findings" in the psychology of spelling-to-sound conversion. Nevertheless, new findings can constrain IA models in this domain, and an appropriately constrained model makes novel predictions that are testable.

A second reason for the popularity of the IA framework is the compatibility of IA models with neural-level modeling techniques. It is not always plausible to interpret IA models as neural nets directly (McClelland, 1985); IA modelers would not always claim that there is just one neuron per node in their IA model. Nevertheless, it is typically assumed that an IA model could easily be cashed out in terms of a more distributed neural network (see Smolensky, 1986). So current IA models often come in between neural modeling and the functionalist approach: much IA modeling is not neural-level because it is not distributed, and it is not functionalist because it is not hardware-independent and because it involves sub-symbolic processing.

The fact that IA models are intended to be cashed out in neural terms means that they should make predictions about the behavior of patients suffering from neurological impairment. That is, IA models and their distributed implementations should not only be able to account for graceful degradation of performance under damage; they should also account for those cases involving severe brain injury where degradation is *not* graceful and leads to quite specific symptom complexes. Progress has already been made in this area, using both local and distributed models (e.g. Cottrell, 1985; Hinton & Sejnowski, 1986; McClelland & Rumelhart, 1986). One aim of the present paper is to present further evidence that an IA model can make novel predictions about the nature of these impairments, and to show that these predictions are upheld. The data can in turn constrain the architecture of the IA model.

BACKGROUND

In this paper I will be concerned with one procedure: the conversion of orthographic representations to phonological representations. This provides us with a classic computational-level mapping problem (Marr, 1982), in which one set of representations (of printed words) must be mapped into another set of representations (of word pronunciations). This particular mapping problem is a difficult one, because the pronunciation of an English word cannot reliably be predicted from its orthography. McClelland and Rumelhart (1981) mention spelling-sound translation as a suitable domain of application for their IA model, and indeed refer to the work of Glushko (1979) as a source of inspiration.

Humans can derive the correct pronunciations of words, even though some words have pronunciations that are not predictable from their spelling. Words will be called *exceptional* or *irregular* here when they contain orthographic segments of at least two letters that are pronounced differently in several other words (see Henderson, 1985, for a discussion of terminology). For example, the word *PINT* has an irregular or exceptional pronunciation compared with its orthographic neighbors such as *MINT*, *HINT*, *TINT* etc.[2] So the exception word *PINT* may be contrasted with *PILL*, which has a pronunciation that is regular and consistent (cf. *MILL*, *HILL*, *TILL* etc.).

The basic experimental finding is that it takes longer to prepare pronunciations of exception words like *PINT* than to prepare pronunciations of consistent words like *PILL* (Glushko, 1979). This exception-word effect is more likely to be obtained when the words are low in frequency and when subjects process the words more slowly (Seidenberg, Waters, Barnes & Tanenhaus, 1984; Seidenberg, 1985a). So subjects do sometimes make use of spelling-to-sound correspondence information in word naming, and this is more likely to happen when processing is slow and there is more time for phonological information to become activated (Seidenberg et al., 1984; Seidenberg, 1985a).

A variety of more detailed findings has been obtained, and many different models have been put forward to account for the findings (for recent reviews, see Humphreys & Evett, 1985; Kay, 1985). Many of the models are basically IA in orientation, although as most of them have not been implemented it is not always clear exactly what predictions they make. The model that accounts for the widest range of data is that of Seidenberg and his colleagues (Seidenberg et al., 1984; Seidenberg, 1985a; 1985b; in press; Waters and Seidenberg, 1985). This model can account for the basic effects of spelling-to-sound characteristics on word naming and lexical decision time, and the interactions of such effects with word frequency and subject speed, within the IA modeling tradition. Seidenberg (in press) has developed and extended this model to account for effects of morphological and syllabic structure on lexical processing. Sejnowski and Rosenberg (1986) have implemented a connectionist system, *NETtalk*, which exhibits great success in learning the spelling-to-sound constraints in English using the back-propagation algorithm described in Rumelhart, Hinton and Williams (1986) (see also Rosenberg & Sejnowski, 1986). So the basic IA framework is apparently very successful in accounting for a wide variety of sophisticated experimental data and task performance. But it may be that this is because the framework is insufficiently constrained in certain respects, as we see below.

How do interactive-activation models predict the exception word effect, whereby words like *PINT* with exceptional pronunciations take longer to pronounce than matched words like *PILL* with regular consistent pronunciations? Figure One is similar to the full version of the IA model set out in McClelland and Rumelhart (1981), although it differs in that it contains separate lexical levels for orthography and phonology and does not include a feature level. Any such a model can easily be extended to include intermediate levels between words and letters, representing sub-lexical letter and phoneme clusters (Brown,

2 For discussion purposes, we consider just the pronunciation of terminal trigrams in four-letter words. Of course, some letter clusters in an exception word will be pronounced regularly; the model to be discussed takes account of this.

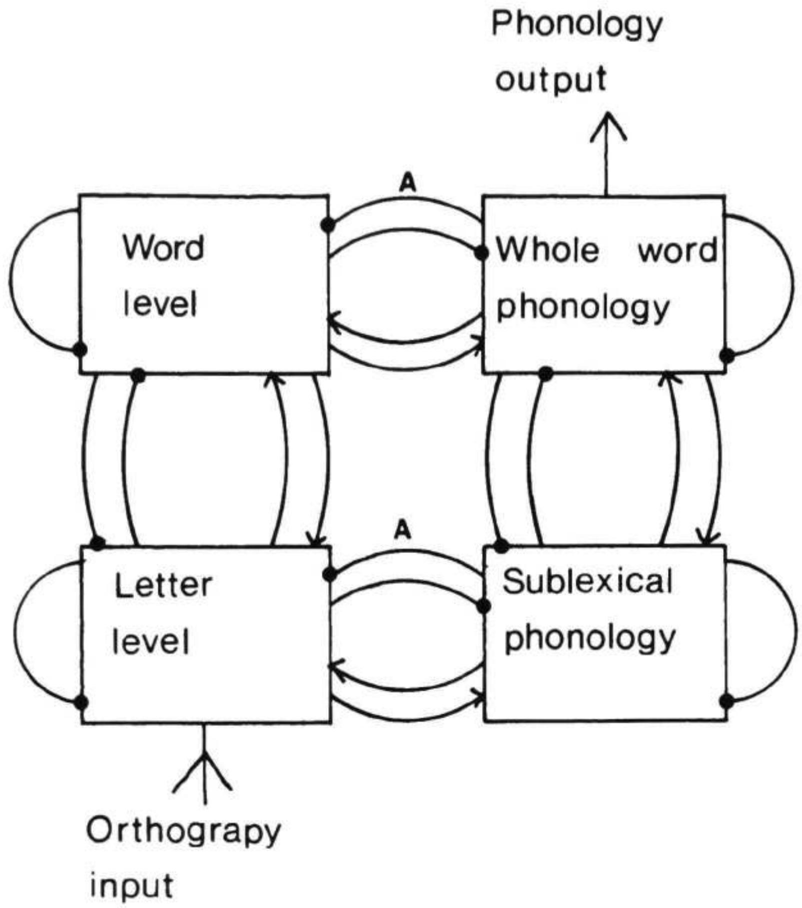

Figure One

1987; Seidenberg, in press). Note that the model has inhibition flowing from words to letters, letters to words, phonology to orthography and orthography to phonology, as well as mutual inhibition within each level[3]. If an IA model contains both orthographic and phonological levels, with connections between them, there are at least two ways in which exception-word effects will be predicted. One way is by inhibition flowing from the phonological to the orthographic levels (pathway A in Figure One). This would operate in the following way. A node for an orthographic segment with more than one possible pronunciation will activate more than one node in the corresponding phonological level. For example, the intermediate-level orthographic node for *-INT* (as in *PINT* and *MINT*) will activate phonology nodes

3 Although the letter-to-letter inhibition and the word-to-letter inhibition were both set at zero in the simulation reported by McClelland and Rumelhart (1981). Furthermore, McClelland (1985) reports a problem with letter-to-word inhibition, which is that if three competing letter nodes are all equally active, two of them will produce enough inhibition to cancel out excitation from the remaining candidate.

corresponding to pronunciations /aɪnt/ and /ɪnt/ (not shown on Figure One). These sub-lexical phonology nodes will then cause activation to spread to lexical-level phonology nodes such as /paɪnt/, /mɪnt/, /lɪnt/ etc. If these lexical-level phonology nodes can inhibit lexical-level orthography nodes (pathway A), the result will be that lexical-level nodes for exception words like *PINT* will build up activation more slowly than nodes for consistently-pronounced words like *PILL*. This is mainly because *PINT*'s orthographic neighbors cause inconsistent phonology to be activated, and this inconsistent phonology indirectly inhibits activity in *PINT*'s lexical-orthographic node, via phonology-to-orthography feedback.

But IA models will normally predict exception effects even if there is no feedback from phonological to orthographic levels (Brown, 1987; Seidenberg, in press). A model will still predict exception-word effects if there is *mutual inhibition* within the phonological levels, even if pathway A does not exist, because there will be slower activation of phonology whenever inconsistent phonology is active. An example is the case above where the two different phonological nodes activated at the same level are /aɪnt/ and /ɪnt/. Because of mutual inhibition, both of these will become activated more slowly than they would have on their own. This is because of the reciprocal nature of the inhibition: the most highly activated member in a "winner take all" network will always win eventually (Feldman & Ballard, 1982), but when inhibition is mutual the winner will win more slowly if it has more competition. Just how much more slowly it wins will of course depend on the precise nature of the mutual inhibition function. So, IA models will predict delayed pronunciation of exceptionally-pronounced words even when there is no feedback from phonological to orthographic levels.

Yet another interactive-activation architecture has been put forward to account for exception effects (e.g. Glushko, 1979; Kay & Marcel, 1981; Marcel, 1980). This allows phonology to be activated only as a result of activation within the lexical-level orthographic levels. If there is downward inhibition from phonology to orthography, or if there is mutual inhibition within the phonology levels, exception effects will be predicted. The way this works will depend on whether only whole-word or sublexical phonology is represented, but either case will lead to inhibition for words with orthographic neighbors pronounced differently. Some implementations of this possibility will be equivalent to the architecture in Figure One (see Marcel, 1980 for a detailed discussion).

CONSTRAINING THE ARCHITECTURE

The previous section demonstrated that interactive activation models could account for the exception-word effect in a number of different ways. This suggests that the architecture is underdetermined by the data. However, Brown (1987) has claimed that the reason an exception word like *PINT* takes longer to pronounce than a consistent word like *PILL* is not in fact due to interference coming from the activation of inconsistent phonology associated with *PINT*'s orthographic neighbors. Rather, it is because of the low *frequency* of the spelling-sound correspondence *-INT* -> /aɪnt/ compared with the high-frequency correspondence -*ILL* -> /ɪl/. The contrast can easily be seen by considering a word like *SOAP*. This word is not inconsistently pronounced, because there are no orthographic neighbors pronounced differently. *SOAP* is the only four-letter English word ending in -*OAP* (remember, we are considering just the pronunciations of the last three letters in four-letter words). And it turns out that words like *SOAP* are delayed in pronunciation just as much as exception words (*PINT*) compared to consistent words like *PILL*, even though *SOAP* does not have differently-pronounced orthographic neighbors to cause interference. What this strongly suggests is that the *frequency* of, not the *number of exceptions* to, a spelling-to-sound correspondence in a word determines the speed with which that word is pronounced by normal adults. Brown (1987) therefore suggested that the strength of a link between an orthographic node and corresponding phonology nodes will depend on the frequency of that spelling-sound correspondence in the language (Seidenberg, in press, makes the same suggestion). The implemented version of the model also contains spelling-sound correspondences at many different levels (letters, bigrams, trigrams etc).

What I want to do now is to outline the implications of these data for the architecture of interactive activation models. As discussed above, the standard IA model has between-level inhibition, and within-level mutual inhibition, both of which predict that there should be effects of a word's spelling to sound regularity on the time taken to pronounce that word. Yet no such effect exists: effects that have previously been attributed to this variable are in fact due to the frequency of spelling-to-sound correspondence. The inhibitory mechanisms that predict the effect must therefore be removed. (I ignore the unattractive alternative possibility that the inhibition is simply too small to be detectable experimentally.) It is a simple matter to remove feedback from phonological to orthographic levels. This involves removing the relevant excitatory and inhibitory connections between phonological and orthographic levels from the full version of the model set out by McClelland and Rumelhart (1981). But the second mechanism that predicts exception-word effects is mutual inhibition within the phonological levels. Many previous researchers have attributed exception effects to this source. It seems undesirable to remove this mutual inhibition, because of the possible saturation if too many nodes at a given level can be active at the same time.

However, there are several ways to preserve the desirable inhibition within a level without the undesirable side-effect of predicting non-existing experimental findings. In a typical IA model a node on a given level is connected to all the other nodes on the same level. A node i with activation a_i will sum the inhibitory evidence reaching it, and (ignoring decay and incoming activation) its activation at time $(t+\delta t)$ will be given by something of the form:

$$a_i(t+\delta t) = a_i(t)\ (1-n_i(t))$$

where n_i is the summed incoming inhibition (assumed less than 1.0) from other nodes in the same layer, and the node has a resting level of zero. This has the effect that even the node with the highest activation level will be inhibited to some extent by its neighbors, i.e., mutual inhibition. What is needed, however, is a case where the highest-activated node receives no inhibition itself, but inhibits all the other nodes in the level (non-mutual inhibition). Shastri and Feldman (1984) discuss suitable types of unit, with which it is possible for each participating node to receive inhibition dependent on the highest activation of any participating node. This kind of scheme has two quite independent advantages. The first advantage is that the number of connections needed is drastically reduced. In a non-distributed layer with N nodes, the number of connections needed to serve a mutual inhibition process will be a quadratic function of N. But using the Shastri and Feldman "max-calculator" units, the number of connections needed will be only a linear function of N, because there is a master node which receives activation from each node in the layer, and sends the maximum activation it receives back to each node as inhibition. Furthermore, the existence of a separate master node provides a means of control over the within-layer inhibition. This feature is useful for strategic purposes (see Cottrell, 1985).

In our implemented version of such a system, the effect of inhibition is given by the following form of equation, which gives the new activation of a participating node after a cycle of inhibition:

$$a_i(t+\delta t) = a_i(t)\ (1 + \beta[a_i(t) - M(t)])$$

where $M(t)$ is the maximum activation at time t of any node participating in the WTA system, and β is a constant.

In other words, each node is inhibited to an extent that depends on the difference between its activation and the activation of the most active node in the layer. For the most active node itself, of course, this difference is zero, and so there will be no inhibition. This is then a WTA network *par excellence* (because the rich get richer without paying tax on the way).

In terms of the model of phonological processing, this is what is necessary. The within-level inhibition prevents saturation of the network, without slowing down in any way the activation of the winning node. In the currently implemented version of the model, the inhibition works in this way (although the only within-level inhibition is at the lexical-level orthographic and phonological levels in the implemented version of the model, see Brown, 1987). So, according to the constrained model, the reason that *PINT* is named more slowly than *PILL* is because of the low frequency of the spelling-sound correspondence in *PINT*, and not because of interference from *PINT*'s differently-pronounced orthographic neighbors.

We therefore have a resolution in which a different inhibition scheme, which may be independently preferred on the grounds that it requires fewer nodes and allows the possibility of strategic control over inhibition, also provides a better account of the data.

The IA model is therefore considerably constrained. The data suggest that there is no feedback from phonological to orthographic levels, and that the within-level inhibition is not mutual inhibition.

SURFACE DYSLEXIA

We have claimed that the powerful IA architecture needs to be constrained to account for empirical data from normal subjects. But a crucial test of the newly-constrained model is its ability to make novel predictions. We now examine the predictions made by the constrained IA model for the performance of patients suffering from various forms of acquired dyslexia. The syndrome most relevant to the present model is that of surface dyslexia.

Surface dyslexic patients are able to synthesize pronunciations of non-words, but have difficulty in pronouncing many words with exceptional pronunciations. Furthermore, these patients have difficulty in defining homophones (see accounts in Marshall & Newcombe, 1973; Shallice & Warrington, 1980; Coltheart, Masterson, Byng, Prior & Riddoch, 1983; Shallice, Warrington & McCarthy, 1983; Kay & Lesser, 1985; Patterson, Marshall & Coltheart, 1985). These symptoms lead naturally to the suggestion that surface dyslexics are making use of sub-lexical spelling-to-sound correspondence information, and that their access to a semantic lexicon is often via a phonological representation. Most surface dyslexics can pronounce some exception words, especially when they are high-frequency (Bub, Cancelliere & Kertesz, 1985) suggesting that some lexical-level correspondences are preserved. Also they show lexicality effects (Marcel, 1980), suggesting lexical-level involvement (although not all patients show lexicality effects: Shallice et al., 1983; Kay & Lesser, 1985).

Shallice et al. (1983) show that their patient, HTR, is affected by "degrees of irregularity", and can pronounce many "mildly irregular" words correctly. Mildly irregular words are defined as words that contain a spelling-to-sound correspondence that is the second most frequent in the language. Note that although this is a measure of the *relative* frequency of a spelling-sound correspondence, it is likely to be correlated with the *absolute* frequency of that correspondence. It is therefore difficult to tell which of the two factors is causing the effects. The data lead Shallice et al. to conclude that surface dyslexics fall on a continuum according to the size of orthographic units they can translate to phonology. In general, the consensus view is that surface dyslexics are impaired on "irregular" words, where regularity is in some way defined in terms of other words containing the same orthographic segment pronounced differently. In terms of an interactive activation model with spelling-sound links at many different levels, it is reasonable to conclude that surface dyslexics have preserved low-level correspondences but have lost most high-level spelling-to-sound correspondences. Indeed, Shallice et al. give an account very similar to this, and point out that if higher frequency or early-acquired correspondences were more likely to be preserved, surface dyslexia could result. This is because level of correspondence is confounded with frequency of correspondence, because high-level correspondences will occur in fewer words.

What I want to do now is examine the predictions of the newly-constrained IA model for the nature of the errors made by these surface dyslexic patients. The main relevant properties of the constrained version of the model are that (*a*) connections between orthographic and phonological nodes are weighted according to the frequency of that spelling-sound correspondence in the language; (*b*) there is no feedback from phonological to orthographic levels, and (*c*) within-level inhibition is not mutual inhibition.

When higher-level correspondences are abolished, it is reasonable to suppose that the highest frequency correspondences within that level are most likely to be disrupted. Therefore, words that contain low-frequency spelling-to-sound correspondences are most likely to be pronounced incorrectly or not at all, because the correct pronunciation of these words relies on the use of high-level (i.e. lexical or trigram level) correspondences. Most previous models have assumed in contrast that words with exceptional or irregular pronunciations will be susceptible to disruption.

The prediction made by the constrained IA model is, then, that surface/semantic dyslexics will be more likely to make errors on word containing unusual spelling-to-sound correspondences at high levels. The regularity of the word, where regularity is defined (as it normally is) in terms of the number of a word's neighbors that are spelt similarly but pronounced differently, should have no effect.

THE ANALYSIS

Several researchers have examined the prediction that surface dyslexics should make more errors on irregular words. Most have used the lists of regular and irregular words published by Coltheart et al. (1979), and many have published full listings of the words that their patient pronounced wrongly. It is therefore possible to re-analyse these data to determine whether it is in fact the number of differently-pronounced but similarly-spelled words that impairs performance, or whether it is in fact the frequency of the spelling-to-sound correspondence in the word. There are six complete published corpora of errors on the Coltheart et al. words. These are found in Coltheart et al. (1983); Shallice et al. (1983), Kay and Lesser (1985), and Saffran (1985). A number of other papers do include corpora, but these either contain only a subset of the errors made errors (e.g. Margolin, Marcel & Carlson, 1985; Masterson, Coltheart & Meara, 1985) or are based on different sets of regular and irregular words (e.g. Newcombe & Marshall, 1985). We therefore analysed the six complete corpora, although two are from one patient. This patient was tested on two occasions two months apart (Saffran, 1985); on the first occasion 31 errors were made, on the second occasion 17 of those 31 words were misread along with 10 other words. This is similar to the normal overlap between two different patients.[4] One of the corpora came from a "developmental surface dyslexic" ("C.D." in Coltheart et al., 1983); five from acquired surface dyslexics.

For each four-letter and five-letter monosyllabic word in the Coltheart et al. (1979) lists (N=54) we calculated a measure of the frequency of the spelling-to-sound correspondence, and the exceptionality of the spelling-to-sound correspondence. These were calculated by looking at the phonology asociated with each of the trigrams in the word. The exceptionality of each trigram within a word was calculated as the cumulative Kucera & Francis (1967) frequency of all same-length words containing the same trigram in the same position but pronounced differently. The exceptionality of each word was the sum of its trigram exceptionalities. The spelling-sound frequency of each trigram was calculated as the cumulative frequency of all same-length words (including the word in question) containing the same trigram pronounced the same way. The spelling-sound frequency of a word was then obtained by summing the spelling-sound frequencies of the trigrams within that word.

There are therefore two measures for each word, one relating to the frequency of the spelling-sound correspondences contained in the word, and the other relating to the exceptionality of the word (i.e. the number of other words containing the same orthographic segment pronounced differently). And the prediction is that the number of errors made by surface dyslexics will be related to spelling-sound frequency, in contrast to previous claims that the relevant factor will be spelling-sound irregularity.

The 27 of our words classified as irregular by Coltheart et al. had a median exceptionality of 231, and a median spelling-sound frequency of 203. The words classified as regular by Coltheart et al. had a median exceptionality of 54, and a median spelling-sound frequency of 498.

Overall, ignoring the Coltheart et al. classification, there was a clear negative correlation between spelling-sound frequency and error rate: Spearman's Rho = -0.39, t=3.0, p<.01. In contrast, there was no correlation between error rate and exceptionality: Rho = 0.19, t=1.4, p>.10. This clearly supports the prediction made by the constrained IA model discussed above; errors are more likely to be made on words containing infrequent spelling-sound correspondences rather than on words with irregular spelling-sound correspondences.

It could be argued that these correlations result from our own definition of exceptionality, which is based only on high-level spelling-sound correspondences. In fact this is unlikely, because the higher-level correspondences are apparently more susceptible to damage in surface dyslexics. Nevertheless, it could also be argued that four-letter and five-letter words should be analysed separately, in case word length has an independent influence on error rate (over and above the tendency for longer words to contain less frequent spelling-sound correspondences). Therefore, further analysis was carried on on the 12 four-letter monosyllabic words and the 15 five-

4 The figures here are based on Saffran's corpus rather than on the figures in the accompanying text.

letter monosyllabic words classified by Coltheart et al. as irregular. The results were as follows.

For five-letter irregular words, spelling-sound frequency correlated significantly and negatively with error rate: Rho = -0.59, t=2.7, p<.02. Exceptionality did not correlate significantly with error rate: Rho = -0.17, t=0.6, p>.20. A similar pattern of correlation was observed for the four-letter words, although the correlation between spelling-sound frequency and error rate failed to reach significance: Rho = -0.32, t=1.1, p>.20. Exceptionality again failed to correlate with error rate: Rho = 0.05, t=0.2, p>.20.

In combination, these results clearly suggest that surface dyslexics tend to make more errors on words containing infrequent spelling-to-sound correspondences, rather than on words with exceptional spelling-sound patterns. Apparent effects of exceptionality have in fact been due to spelling-sound frequency. This is exactly the pattern of results predicted by the constrained version of the IA model.

It should be noted that we have not controlled for the frequency of purely orthographic regularity in our analysis; this is impossible to do for the words for which error corpora have been reported. This is unlikely to be a major problem, for Brown (1987) found effects of spelling-sound frequency when orthographic regularity was controlled for, and other authors have obtained spelling-sound effects that are not due to orthography. And in reaction-time experiments it is unlikely that orthographic frequency and spelling-sound regularity effects would exactly cancel each other out across a wide range of word frequency and subject decoding speed.

It should also be noted that the spelling-sound frequency of a word is related to the frequency of occurrence of that word, because the frequency of a word contributes to the frequency of the spelling-sound correspondences contained within it. Again, a number of experiments have found effects of spelling-sound characteristics when word frequency is controlled. For example it is clear that surface dyslexics do make more errors on the Coltheart-irregular than on the Coltheart-regular words even though the two sets of words are matched for word frequency. The claim here is just that the effects are really due to the confounded factor of spelling-sound frequency. Indeed, our model interprets effects of word frequency on word naming time as being due to the frequency of spelling-sound correspondences in that word, including the lexical-level spelling-sound correspondence (the strength of which will depend directly on word frequency).

CONCLUSION

We have shown that the full interactive activation framework when applied to the domain of spelling-to-sound conversion is in some respects too powerful, because many different inhibition mechanisms could give rise to delayed processing of words with exceptional pronunciations. Because effects which have previously been seen as exceptionality effects are in fact simple spelling-to-sound frequency effects, the model needs to be constrained. A more constrained IA model is discussed, which has no feedback from phonological to orthographic levels and which also uses a different mechanism for within-level inhibition. This has the dual advantages of giving a better account of the data and requiring fewer within-level inhibitory connections. The constrained model gives rise to novel predictions about the errors made by surface dyslexic patients, and these predictions are confirmed. Thus it is both possible and necessary to constrain models within the interactive-activation framework.

BIBLIOGRAPHY

BROWN, G. D. A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language*, 26, 1-23.

BUB, D., CANCELLIERE, A., & KERTESZ, A. (1985). Whole-word and analytic translation of spelling to sound in a non-semantic reader. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

COLTHEART, M., BESNER, D., JONASSON, J. T. & DAVELAAR, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology*, 31, 489-507.

COLTHEART, M., MASTERSON, J., BYNG, S., PRIOR, M., & RIDDOCH, J. (1985). Surface Dyslexia. *Quarterly Journal of Experimental Psychology*, 35A, 469-495.

COTTRELL, G. (1985). *A connectionist approach to word sense disambiguation.* PhD. thesis, TR 154, University of Rochester.

FELDMAN, J. A., & BALLARD, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.

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.

HENDERSON, L. (1985). Issues in the modelling of pronunciation assembly in normal reading. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

HINTON, G. E., & SEJNOWSKI, T. J. (1986). Learning and relearning in Boltzmann machines. In D. E. Rumelhart and J. L. McClelland (Eds.), *Parallel Distributed Processing Volume 1: Foundations.* Cambridge, Mass: MIT Press.

HUMPHREYS, G. W., & EVETT, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route hypothesis. *The Behavioral and Brain Sciences*, 8, 689-740.

KAY, J. (1985). Mechanisms of oral reading: A critical appraisal of cognitive models. In A. W. Ellis, (Ed.), *Progress in the psychology of language*, Vol. 2, London: Erlbaum.

KAY, J., & LESSER, R. (1985). The nature of phonological processing in oral reading: Evidence from surface dyslexia. *Quarterly Journal of Experimental Psychology*, 37A, 39-81.

KAY, J., & MARCEL, A. J. (1981). One process, not two, in reading aloud: Lexical analogies do the work of non-lexical rules. *Quarterly Journal of Experimental Psychology*, 33A, 397-414.

KUCERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*, Brown University Press.

MARCEL, A. J. (1980). Surface dyslexia and beginning reading: A revised hypothesis of the pronunciation of print and its impairments. In M. Coltheart, K. E. Patterson & J. C. Marshall, (Eds.), (1980). *Deep Dyslexia*. London: RKP.

MARGOLIN, D. I., MARCEL, A. J., & CARLSON, N. R. (1985). Common mechanisms in dysnomia and post-semantic surface dyslexia: processing deficits and selective attention. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

MARR, D. (1982). *Vision*. San Francisco: W. H. Freeman & Co.

MARSHALL, J. C., & NEWCOMBE, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research*, 2, 175-199.

MASTERSON, J., COLTHEART, M., & MEARA, P. (1985). Surface dyslexia in a language without irregularly spelled words. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

McCLELLAND, J. L. (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitive Science*, 9, 113-146.

McCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.

McCLELLAND, J. L., & RUMELHART, D. E. (1986). Amnesia and distributed memory. In J. L. McClelland and D. E. Rumelhart (Eds.), *Parallel Distributed Processing Volume 2: Psychological and Biological Models.* Cambridge, Mass: MIT Press.

NEWCOMBE, F., & MARSHALL, J. C. (1985). Reading and writing by letter sounds. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

PATTERSON, K. E., MARSHALL, J. C., & COLTHEART, M. (1985). (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

ROSENBERG, C. R., & SEJNOWSKI, T. J. (1986). The spacing effect on NETtalk, a massively-parallel network. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society,* 72-89.

RUMELHART, D. E., HINTON, G. E., & WILLIAMS, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart and J. L. McClelland (Eds.), *Parallel Distributed Processing Volume 1: Foundations.* Cambridge, Mass: MIT Press.

RUMELHART, D. E. & McCLELLAND, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review,* 89, 60-94.

SAFFRAN, E. M. (1985). Lexicalisation and reading performance in surface dyslexia. In K. E. Patterson, J. C. Marshall and M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

SEIDENBERG, M. S. (1985a). The time course of phonological code activation in two writing systems. *Cognition,* 19, 1-30.

SEIDENBERG, M. S. (1985b). Constraining models of word recognition. *Cognition,* 20, 169-190.

SEIDENBERG, M. S. (in press). Reading complex words.

SEIDENBERG, M. S., WATERS, G. S., BARNES, M. A., & TANENHAUS, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior,* 23, 383-404.

SEJNOWSKI, T. J., & ROSENBERG, C. R. (1986). NETtalk: A parallel network that learns to read aloud. The John Hopkins University, Technical Report JHU/EECS-86/01.

SHALLICE, T., & WARRINGTON, E. K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. E. Patterson & J. C. Marshall, (Eds.), (1980). *Deep Dyslexia.* London: RKP.

SHALLICE, T., WARRINGTON, E. K., & McCARTHY, R. (1983). Reading without semantics. *Quarterly Journal of Experimental Psychology,* 35A, 111-138.

SHASTRI, L., & FELDMAN, J. A. (1984). Semantic networks and neural nets. TR131, University of Rochester.

SMOLENSKY, P. (1986). Neural and conceptual interpretations of PDP models. In J. L. McClelland and D. E. Rumelhart (Eds.), *Parallel Distributed Processing Volume 2: Psychological and Biological Models.* Cambridge, Mass: MIT Press.

WATERS, G. S., & SEIDENBERG, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory & Cognition,* 13, 557-572.