Improving Norman's Reading Ability: Computational Simulation of Word Recognition with the Logogen Model

Stuart J. Soltysiak,
Department of Computer Science
University of Essex
Colchester CO4 3SQ

Jacqueline Masterson,
Department of Psychology
University of Essex
Colchester CO4 3SQ

November 1993

Abstract

A computer simulation of the revised logogen model (Morton 1979, 1980) of word recognition is presented. Evaluation of initial performance found reading accuracy was 50%. Assumptions made during the development of the simulation adversely affected performance, and these have been addressed in the work presented here. The simulation is now capable of word recognition at a level comparable to human skilled reading performance.

Word recognition models generally date from the 1960s to present. The logogen model (Morton, 1969, 1970) was one of the first and most influential models of word recognition to be developed. Its comprehensive coverage of spoken and visual word recognition provided a new and powerful vocabulary with which word recognition could be investigated. Subsequent experimental observations with skilled readers on facilitation effects in speeded naming tasks (e.g. Winnick and Daniel, 1970; Clark and Morton, 1983) and investigations aimed at discovering whether words or morphemes are the appropriate perceptual unit (e.g. Murrell and Morton, 1974) indicated that some predictions from the logogen model were not supported. A revised version of the logogen model was therefore put forward (Morton, 1979, 1980; Morton and Patterson, 1980). Computer simulations have become widespread during the last decade to the point that the *de facto* standard is that current theories are embodied in computer simulations. The use of simulations enables the specification, verification and validation of word recognition theories to be carried out in a rigorous manner. Unfortunately, due to a lack of specification of a number of processes the logogen model has not, in the past, been able to support simulations. Nevertheless, many recent simulations have been influenced by the logogen model, such as the interactive activation model of McClelland and Rumelhart (1981) or the DRC model of Coltheart, Curtis and Atkins (in press).

This paper briefly describes an implementation of the revised logogen model as a computer simulation. After reviewing initial performance of the simulation three experiments are reported which address the limitations of the original implementation. The paper closes with a discussion of the future directions in which the simulation will be applied.

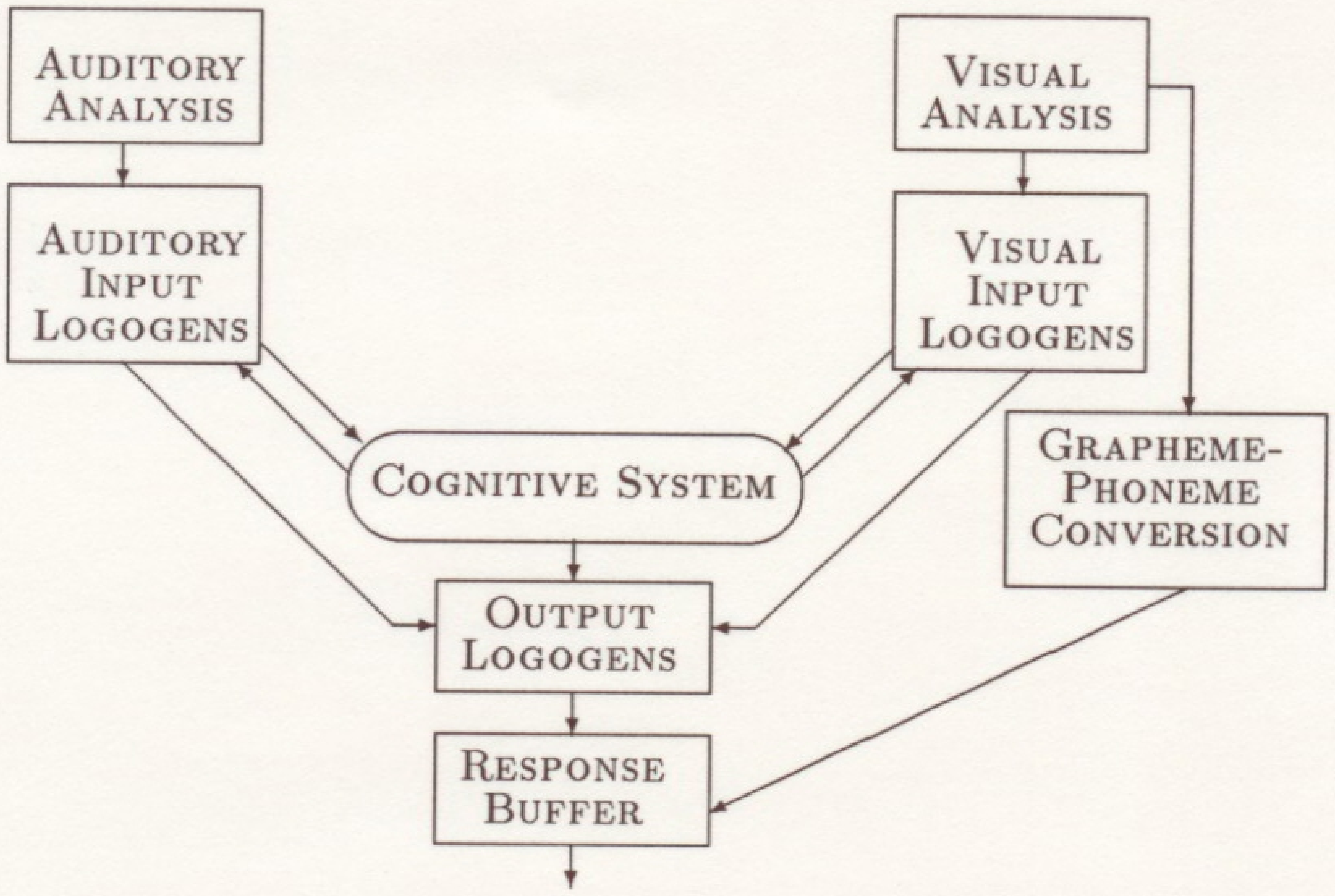

Figure 1: The revised logogen model (from Morton, 1980).

The Revised Logogen Model

The logogen model (Morton, 1969, 1970) was the first detailed attempt at modelling the processes involved in skilled reading[1]. The revised logogen model addressed limitations of the original version. The primary effect of the revision was to split the central component of the original model (the *logogen system*) into three separate logogen systems. Figure 1 shows the arrangement of the revised logogen model. Throughout this paper only the visual word recognition system will be discussed. However, many of the issues raised are relevant to spoken word recognition.

The fundamental principles of the logogen model are that words are represented by dictionary units (logogens) within the mental lexicon (cognitive and logogen systems), and that word recognition arises as a result of the activation of dictionary units by an input stimulus. Logogens are essentially word detectors, with one logogen for each word in the lexicon. Each logogen has a threshold which specifies the amount of activation required for an input to be identified as the word represented by the logogen, and furthermore this threshold is inversely related to the frequency of the word — high frequency words have lower thresholds than low frequency words. This accounts for the *word frequency effect* whereby high frequency words are recognised faster than low frequency words by skilled readers (e.g. Forster and Chambers, 1973).

Word recognition in the revised logogen model is achieved as follows. Presentation of an input stimulus initiates a preliminary "visual analysis" which results in the extraction of a set of features. These features are then used to activate any visual input logogens where there is a match with the features required by the logogen. Activation of a visual input logogen to threshold results in the logogen firing, denoting that the input has been visually identified. Two things then occur. First, the logogen's threshold is reduced by a small amount, thereby accounting for the *identity priming effect* — a word is recognised faster if it is preceded, or *primed*, by itself (e.g. Scarborough, Cortese and Scarborough, 1977). The second result is the transmission of a code to the cognitive system. The receipt of this code by the cognitive system results in access to the semantic attributes of the identified word. These attributes are subsequently transmitted to the output logogen system, where a second activation process commences in a similar fashion

---

[1] Skilled reading is an individual's ability to read competently and fluently.

to that for visual input logogens. The firing of an output logogen signifies full recognition of the word, and a response is made available as a set of instructions to operate the speech articulators. The response is temporarily stored in the response buffer before being output by the speech production system as the pronunciation of the recognised word. The input stimulus has been recognised by the access of lexical entries (logogens) within the lexicon (cognitive and logogen systems).

The *lexical route* described above is one of two processes in the revised logogen model by which skilled readers can process printed letter strings. Use of the lexical procedure means that only those words contained within the lexicon can be recognised. Skilled readers are also able to read nonwords such as TWIMP. To account for this ability the revised logogen model postulates the existence of a *nonlexical* route, operating as follows. Presentation of an input stimulus initiates parsing of the stimulus into graphemes. A grapheme is a letter (or letter cluster) whose pronunciation is represented by a single phoneme (e.g. C → /k/ as in CAT, or GH → /f/ as in TOUGH). Identified graphemes are then substituted by the corresponding phonemes by application of grapheme-phoneme correspondence (GPC) rules. The pronunciation of the input is then assembled from the phonemes for the individual graphemes.

This brief description of visual word recognition in the revised logogen model does not do it full justice. In particular its application to the explanation of several forms of acquired dyslexia by selective impairment of components (Morton and Patterson, 1980) has not been discussed. Nor indeed has the spoken word recognition system, although the underlying principles are identical for both visual and spoken word recognition. The implementation of a computer simulation of the model as described above is now presented.

All About Norman

Norman[2] is a computer simulation of the revised logogen model written in Common Lisp within the Allegro Common Lisp environment. During the development of Norman, described in full in Soltysiak (1993), several limitations in the descriptions of the logogen model were identified, namely inadequate specification of the various representations and processes contained within the model. Since the development of a computer simulation requires explicit specification of all aspects of a model, it was found necessary to make a number of important assumptions in the course of developing the simulation:

Input Representation Letter and feature representations are based on those employed by McClelland and Rumelhart (1981) in their interactive activation model.

Feature Analysis Features are extracted from the input stimulus in a pseudo-parallel manner. The word as a whole is analysed for each feature, rather than by a serial letter-by-letter approach.

Logogen Activation Every feature that is extracted results in all logogens which detect that feature in the specified position experiencing activation.

Logogen Selection The first logogen activated to threshold fires, preventing all other logogens from firing.

---

[2] Named as a result of the large volume of incorrect responses of "norm" during initial evaluation of spoken word recognition performance.

Semantic Representation Word meanings are represented by syntactic and morphological information plus a numeric identifier unique to each word.

Output Representation The model makes available a list of phonetic features which specify articulation details (those of Chomsky and Halle, 1968, are used). These are converted into phonemes and output as the response.

The Nonlexical Route The simulation has a restricted set of GPC rules available, covering single letter-sound correspondences only. Graphemic parsing simply consists of the identification of letters in the input.

These wide-ranging assumptions were necessary given the lack of direction from descriptions of the logogen model. Several opaquely described aspects of the model have not been implemented in the current simulation (existence of contextual input from cognitive to visual input logogen systems; the direct connection of the input and output logogen systems; and the output of a response as a silent rehearsal "feedback loop" rather than speech output). These omissions do not, however, affect the fundamental operating principles of the logogen model. Norman was provided with a lexicon of 898 four-letter monosyllabic words, obtained from the Oxford Psycholinguistic Database (Quinlan, 1992). Only those words deemed to have 'standard' word status (i.e. those that are not archaic, foreign, colloquial, and so on) were used.

Evaluation of the Simulation

Initial evaluation of the simulation was aimed at identifying the validity of the assumptions made and, where necessary, formulating alternative approaches to consider. Full details of the evaluation are contained in Soltysiak (1993); a brief overview only is given here.

Four evaluation experiments were conducted to investigate (1) initial performance of the simulation; (2) effects of random internal organisation of logogen systems; (3) effects of threshold reduction; and (4) an alternative method of feature extraction and logogen activation increase. Each experiment involved the presentation of all known word stimuli once[3]. Norman's performance in the experiments is given in Table 1. The results showed that there was a negligible effect of threshold reduction (0.11% increase in performance), with minor effects of random logogen ordering (1% decrease in performance) and feature extraction and activation increase (1.33% decrease in performance). Three *types*[4] of error were made by Norman:

Experiment	*Number Correct*	*Number Incorrect*
1	460 (51.22%)	438 (48.78%)
2	451 (50.22%)	447 (49.78%)
3	461 (51.33%)	437 (48.67%)
4	448 (49.89%)	450 (50.11%)
Average	455 (50.67%)	443 (49.33%)

Table 1: Overall reading performance across all four evaluation experiments. Experiment 1 served as the datum against which subsequent results were compared.

---

[3] A known word here refers to a word possessing a lexical entry (logogen) in the simulation's lexicon.

[4] An error *type* defines the location of an error in the simulation. An error *class* (see later) represents a comparison of error types between experiments.

Identification Errors The input stimulus is mis-identified by the visual input logogen system, caused by the wrong visual input logogen firing.

Response Errors The input stimulus is correctly identified, but the wrong output logogen fires, causing the output of an incorrect response.

Identification + Response Errors The input stimulus is subject to error at both visual input and output logogen systems.

	Error Type		
Experiment	*Identification*	*Response*	*Identification + Response*
1	81.05%	11.19%	7.76%
2	84.56%	13.87%	1.57%
3	81.69%	1.21%	7.78%
4	82.44%	10.00%	7.56%
Average	82.44%	11.57%	6.17%

Table 2: Overall distribution of error types across all four evaluation experiments.

The distribution of error types is given in Table 2. Errors made by Norman were located in two areas of the simulation. Identification errors are made within the visual analysis/visual input logogen stages; while response errors are made within the output logogen system. Two main causes of error were identified: (1) the process used to increase logogen activation according to extracted features; and (2) the process used to select a logogen to fire.

Word recognition by the nonlexical (GPC) route was poor, with an error rate of 81.29%. This was simply a consequence of the limited GPC rules implemented, which do not cover a sufficient number of English spelling-sound correspondences to permit reasonable performance. Norman's poor nonlexical reading performance caused by an under-developed nonlexical route, together with a highly error-prone lexical route, showed some similarity to the reading performance of some developmental dyslexics. For example, the developmental phonological dyslexic SB, reported by Masterson, Hazan & Wijayatilake (in press), makes 10% errors in reading words of mixed frequency of four letters in length. In reading nonwords SB performs very poorly (52% error with four-letter nonwords) and shows use of only simple letter-sound correspondences. SB has achieved a high level of literacy in spite of his dyslexia, and other individuals with this form of dyslexia have higher error rates, resembling the pattern of performance shown by Norman even more closely. However, while the behaviour exhibited by Norman appeared to resemble that of some developmental phonological dyslexics, the simulation was intended as a computational implementation of the logogen model of *skilled reading*. The pattern of performance represented a side-effect of the implementation rather than an overt attempt to model the behaviour of developmental phonological dyslexics.

Having briefly covered the implementation and evaluation of the computer simulation, the motivations for the present work were the modification of the simulation according to recommendations made in Soltysiak (1993) for the improvement of reading performance. It is these that are now presented.

Improving Norman's Reading Performance

The evaluation experiments in Soltysiak (1993) indicated that Norman's reading performance could be improved by the implementation of alternative methods for the increase of logogen activation levels and for the selection of a logogen to fire. Improvements in nonlexical reading would be expected to arise from expansion of the GPC rule set, which is briefly discussed later. The experiments reported here concentrate upon Norman's lexical reading performance. The procedures for increasing logogen activation and selecting a logogen to fire are somewhat interlinked in this simulation. Logogens experience activation whenever a feature extracted from the stimulus is detected by them, this process continuing until a logogen has reached threshold and fires. The improvements suggested here revolve around a clearer distinction between these procedures. The proposed approach is to adopt a mechanism analogous to the deadline procedure suggested by Coltheart, Davelaar, Jonasson and Besner (1977). The present proposal suggests alternative procedures for (1) increasing logogen activation levels; and (2) selecting a logogen to fire. Each of the experiments reported below involved the presentation of all of the 898 words in Norman's lexicon once.

Experiment One — The Deadline

Coltheart *et al.* (1977) proposed the use of a deadline to permit the logogen model to make lexical decisions about letter strings presented on a VDU screen (i.e. to decide whether letter strings were real words or not). Their deadline operated by specifying a limited amount of time for logogens to become activated to threshold: if no logogen has reached threshold after the deadline has passed, assume that none will, and therefore a "no" response should be made. The deadline was made flexible such that it would be determined as a function of the amount of activity within the logogen systems — the more activity there is the shorter the deadline should be since it is very likely that a logogen will reach threshold.

Incorporation of the deadline mechanism in the present simulation involves the following. Visual analysis of an input and the activation of logogens takes place for a limited amount of time. When the deadline has expired, a logogen is selected to fire (using the 'first-to-threshold' procedure employed in Soltysiak, 1993). At present the deadline is sufficiently long to permit complete analysis of the input and for all logogens to receive the maximum amount of activation possible from the input (in the original implementation feature analysis and activation of visual input logogens proceeded until a logogen reached threshold and fired; now analysis and activation is carried out for a predetermined length of time before any logogen is considered for firing). The deadline procedure has been applied to both visual input and output logogen systems. All other aspects of the simulation remain as before. Table 3 shows the reading performance with the deadline mechanism incorporated.

Response type	*Number of responses*
Correct	545 (60.69%)
Incorrect	353 (39.31%)

Table 3: Overall reading performance, Experiment One.

Discussion

The deadline mechanism has improved reading performance by 10%, and has eliminated response errors (and combined identification + response errors). The response errors in the original implementation were a result of the semantic representation used: the semantic features for some words were a subset of the features for other words. Since activation was incremented until a logogen reached threshold, logogens with a subset of features were likely to reach threshold before those logogens with the full set of features, and thus fire before the correct logogen could. By using the deadline mechanism all logogens receive maximum activation, and hence the correct logogen can now fire. There is, however, a question about the assumption that the input stimulus is perfectly analysed[5]. Complete analysis furnishes visual input logogens with the definitive set of features present in the stimulus, therefore providing the clearest possible evidence for the identity of the stimulus. However, if the amount of time available for visual analysis were to be restricted, fewer features could be extracted and the potential for confusion of the stimulus arises (or as a worst case Norman would fail to identify the stimulus). The magnitude of potential performance degradation depends upon the deadline setting (i.e. the amount of time allowed for visual analysis). Krueger and Shapiro (1979) investigated the word superiority effect using the Rapid Serial Visual Presentation (RSVP) technique. Word superiority describes the effect obtained with skilled readers whereby letters in words are more accurately perceived than letters in pseudowords, which in turn are more accurately perceived than letters in nonwords. RSVP involves the presentation of stimuli one at a time at a fixed rate and at the same location on a display. The rate of presentation can be controlled by the experimenter. In Krueger and Shapiro's experiments subjects were asked to respond "yes" if a target letter appeared in a presented stimulus and "no" otherwise. Stimuli were presented at different rates, ranging from 1 to 30 stimuli per second (equivalent to 60–1800 words per minute; typical reading speed is 300 words per minute). All stimuli comprised six uppercase letters. They found a word advantage over nonwords (of 2–4%) for all subjects. The best advantage levels were obtained with presentation rates of 3–5 stimuli per second, with no advantage observed at all for presentation rates in excess of 10 stimuli per second. From these results they concluded that "feature extraction seems to be largely accomplished within the first 100 msec following stimulus onset" (Krueger and Shapiro, 1979, p.665); and that word context does not aid the extraction of features but rather their interpretation (since there were no effects of word context when stimuli were presented for 100 msec or less). Juola, Ward and McNamara (1982) also reported a word superiority effect over presentation durations of 100–300 msec. It appears, therefore, that feature extraction takes place in a very short period of time, and if the amount of time available for feature extraction is reduced, recognition performance is degraded. The operation of this simulation may be in line with experimental observations should the deadline for visual analysis be reduced.

Experiment Two — The Activation Mechanism

In the simulation as it is currently set up, visual input logogens experience activation whenever they detect a feature which has been extracted from the stimulus. This results in many logogens experiencing activation, some of which may share few features in common with the stimulus. For example, suppose the feature *a* is present in letters 1, 2 and 4 of a word. In the present approach each visual input logogen with the *(a, 1)* combination is activated, then each logogen with the

---

[5] The interactive activation model of McClelland and Rumelhart (1981) also assumes perfect extraction of features. The input to their model is a list of features (present and absent) for each letter of the input. Their model makes no attempt to perform the extraction process, but rather concentrates upon the process of interpreting the features via its feature/letter/word detectors. The present simulation differs from the McClelland and Rumelhart model in that it extracts features from an input.

(a, 2) combination and then every logogen with the *(a, 4)* combination is activated. Continuing this example it is easy to see the large number of logogens which would experience activation for the feature a in letters 1, 2 and 4:

$$\{\{(a,1)\},\{(a,2)\},\{(a,4)\},\{(a,1),(a,2)\},\{(a,1),(a,4)\},\{(a,2),(a,4)\},\{(a,1),(a,2),(a,4)\}\}$$

If, however, the logogens which experience activation were restricted to only those capable of detecting the feature a in letters 1, 2 *and* 4, there would be a substantial reduction in the number of logogens activated: $\{(a,1),(a,2),(a,4)\}$. The reduction of the number of logogens experiencing activation entails a reduction in the number of logogens that can be activated to threshold and thus compete to fire. Therefore word recognition performance should improve. The present simulation has been modified to incorporate this in conjunction with the deadline mechanism described in Experiment One. The adopted approach is almost identical to that previously employed in Experiment 4 of Soltysiak (1993). Table 4 shows the performance of the simulation.

Response type	*Number of responses*
Correct	818 (91.09%)
Incorrect	80 (8.91%)

Table 4: Overall reading performance, Experiment Two.

An analysis of error classes was made using the following criteria:

Identical Errors Where the error in the current experiment is identical in every respect to that of the datum experiment.

Different Errors Where the error has been made in both current and datum experiments, but where the response made in the current experiment is different to that of the datum experiment.

New Errors Where the error in the current experiment was correctly responded to in the datum experiment.

Experiment One served as the datum for comparison of these results. According to these criteria, error classes in this experiment are distributed as shown in Table 5. As discussed in the last section the only error type now made in the modified simulation is the identification error.

Error class	*Number of occurrences*
Identical	63 (78.75%)
Different	17 (21.25%)
New	0

Table 5: Distribution of error classes, Experiment Two.

Discussion

Quite clearly the process by which logogens experience activation has a major effect upon reading performance. This experiment has achieved a 30% improvement in performance over Experiment

Response type	*Number of responses*
Correct	474 (83.18%)
Incorrect	151 (16.82%)

Table 6: Overall reading performance, additional experiment for Experiment Two.

One. These results pose two questions. First, how are differences between the results of this experiment and those of Experiment 4 of Soltysiak (1993) explained? This question is important given that the approach used is, to all intents and purposes, identical for the two experiments. There are two possible answers: Either the operation of the deadline mechanism in conjunction with this new procedure facilitated the improvement, or a very small modification of the activation procedure over that used in Experiment 4 in Soltysiak (1993) led to the improvement[6]. The suspicion is that it is the latter rather than the former which is responsible for the improvement. To isolate the real cause of improvement in performance, an additional experiment was conducted which transplanted the new activation procedure into the version of Norman that was described in the original implementation (Soltysiak, 1993). The results (Table 6) clearly show improved performance. The original version which gave the results for Experiment 4 of Soltysiak (1993), shown in Table 1, was inaccurate and therefore produced incorrect results. Errors made by Norman in this additional experiment covered all three types (identification, response and identification + response). The extra experiment allows the following conclusions to be drawn: (1) the greatest contribution to improved performance comes from the alternative logogen activation increase procedure; and (2) the addition of the deadline mechanism permits an additional 10% improvement.

The second question addresses the issue of how words are identified. Currently a wholistic-style approach is used — words are identified through their component features. This contrasts with the widely held view that visual word identification is hierarchical, that is words are detected from letters and letters are detected from features which are detected in the stimulus (e.g. Johnston and McClelland, 1980; McClelland and Rumelhart, 1981; Coltheart *et al.*, in press). This issue will be discussed at greater length in the general discussion, suffice it to say that hierarchical identification does not appear to be a necessary process for reasonably accurate word recognition.

Experiment Three — The Selection Process

Logogens are currently selected on a 'first-to-threshold, first-to-fire' basis. This clearly invites error since no attempt is made to consider alternative logogens, one of which may be the correct one. This has been addressed in this experiment: Once the deadline has expired, visual input logogens are scanned to produce a subset of logogens which have all been activated to (and beyond) threshold. This subset is then searched to identify which logogen has exceeded its threshold by the greatest amount (previously the first logogen encountered after the deadline had expired that had reached threshold was selected as the response). This logogen is then selected and processing in the simulation continues as before. The selection by greatest excess activation

---

[6] The modification was made to correct an oversight in the previous use of the procedure: the wrong values had been given to the activation increase parameter. The activation received by a logogen should be identical whichever approach is used since all features need to be identified: a logogen capable of detecting the feature a in positions (1, 2, 4) of a word requires exactly the same amount of activation whether it detects the feature in each position independently (the original approach) or simultaneously (the new approach). The previous attempt (Experiment 4, Soltysiak, 1993) provided only a fraction of the required activation, an error which was corrected in the current use of the approach. This was purely error correction and does not represent parameter manipulation to fit the data.

permits the logogen that fires to be the one which most closely matches the input stimulus, and therefore it is most likely to be the correct logogen. This has been implemented in Norman in conjunction with the deadline and alternative activation increase procedure. Performance of the simulation is given in Table 7.

Response type	*Number of responses*
Correct	859 (95.66%)
Incorrect	39 (4.39%)

Table 7: Overall reading performance, Experiment Three.

Experiment Two served as the datum for comparison of these results. Errors were classified using the criteria described above, with the distribution of classes given in Table 8. For all classes errors were solely identification errors.

Error class	*Number of occurrences*
Identical	22 (56.41%)
Different	4 (10.26%)
New	13 (33.33%)

Table 8: Distribution of error classes, Experiment Three.

Discussion

Selecting logogens to fire on the basis of greatest excess activation improves recognition performance by 4.5%. Analysis of the distribution of error classes provided interesting observations. Most of the errors were either identical to those of Experiment Two, or they remained errors but produced different responses. This suggests an additional source of error beyond the selection of logogens. Investigation of the large number of new errors can shed light upon this. New errors comprised one-third of the total number of errors made. Since all errors are identification errors, the problem must be located at the visual analysis and visual input logogen system stages of the model. Analysis of new errors revealed that for eleven of the errors, the responded word and target word differ by only one letter. The other two errors differ by two letters. Moreover, the substituted letters have very similar representations in the typeface used (e.g. M/N, B/D, O/U, R/P), many of them differ by just one feature (e.g. M/N, B/D, R/P). Therefore, very close visual similarity plays a role in the mis-identification of input stimuli. However, this explanation only covers part of the issue: it does not explain how the wrong visual input logogen can fire in preference to the correct logogen given the new selection procedure. Further analysis of the errors revealed that the frequency of a word influences the selection of a logogen to fire. Since higher frequency words have lower thresholds, they require fewer features to activate the logogen to threshold. Therefore, if the target word (e.g. TOME) differs from a visually similar word by just one feature (e.g. TONE) and the target is of lower frequency (as in this example) the logogen for the visually similar word is likely to reach threshold before that of the target. Since logogen selection is based upon the greatest amount of activation in excess of threshold, the visually similar word is likely to have a larger excess because of its lower threshold, and is thus more likely to be selected to fire. At present visual input logogens are only informed of the presence of a feature, and receive activation accordingly. Logogens are ignorant in the sense that they do not know whether other features are absent or are present but obscured in the input. If logogens were to be explicitly told of both the presence and absence of features, logogen activation could be altered accordingly. If a feature is present, the activation of the logogens which detect it can be increased as described above. If, however, the feature is absent, all logogens which require the feature to be present should receive inhibition in the form of a reduction in activation. This should make the difference between visually similar words much smaller and therefore selection by greatest activation beyond threshold should return the correct logogen.

One prediction that is made by the simulation operating with this selection mechanism is that there should be no effect of logogen ordering upon performance, simply because order of logogens is not a factor in selecting logogens to fire. The effects observed in the initial evaluation (Table 1, Experiment 2) were small (a 1% decrease in performance), being a result of the serial search used to select a logogen to fire. To test this prediction Experiment Three was repeated using the same randomly ordered logogen systems as had been used in the initial evaluation. The results for the two experiments were identical except for one response. There are two logogens for the homograph LIVE, one for the verb "to live" (/l i v/), and one for the adjective "live" (/l a i v/). In Experiment Three the logogen corresponding to /l a i v/ was selected to fire; in the additional experiment the logogen corresponding to /l i v/ was selected. Although the response was different, the logogen which fired on both occasions did, nevertheless, detect the same features (since /l i v/ and /l a i v/ have exactly the same orthographic representations[7]). Since this was a very small (1 in 39) occurrence it is possible to say that ordering of logogens does not affect performance.

General Discussion

The experiments reported here have permitted Norman's reading performance to be substantially improved. The modifications have addressed limitations of the original implementation, and now represent a more acceptable level of performance for a model of skilled reading. Evaluation of the original simulation suggested Norman exhibited the behavioural characteristics of developmental phonological dyslexics, with overall reading performance well below that for skilled readers. The modified simulation is now capable of behaviour which is very similar to that obtained from skilled readers. Several recent experiments (Table 9) have found that skilled readers typically make 2–3% errors in reading single words. Norman currently makes 4.4% errors, a figure not too far from the experiments reported in Table 9. There remain, however, aspects of the simulation in need of further elaboration.

Experiment	*Error Rate*
Taraban & McClelland (1987)	3.02%
Monsell, Doyle & Haggard (1989)	2.2%
Andrews (1989)	1.76%
Andrews (1992)	2.08%

Table 9: Error rates for skilled readers from four recent naming experiments. All studies involved words of mixed frequencies presented one at a time on a VDU using a randomised list.

---

[7] This raises the question of whether there should be two input logogens detecting the same set of features. If only one input logogen were to exist, different pronunciations could be obtained by more than one output logogen. This is a more plausible approach, since each word has only one detector, but different interpretations of it (accessed from the cognitive system) would permit alternative pronunciations to be produced by activation of more than one output logogen.

Firstly, Norman has received no further development of his nonlexical (GPC) route. Consequently his nonlexical reading performance remains at the 81% error rate observed during the initial evaluation. This may become an important factor when the operation and interaction of lexical and nonlexical routes is considered. Some interpretations of behavioural data suggest that the lexical and nonlexical routes operate in parallel, with the nonlexical route being slower (e.g. Patterson and Morton, 1985; Seidenberg, Waters, Barnes and Tanenhaus 1984). The simultaneous operation of the lexical and nonlexical routes may serve to correct potential errors. Norman does not enjoy such simultaneous operation of his lexical and nonlexical reading routes, and therefore any error correction this may offer is not present. Norman can read lexically and nonlexically, but not both in parallel[8]. Expansion of the GPC route is planned for future work, which is likely to require the development of both the GPC rule set and the graphemic parser. Expanding the GPC rules available to Norman requires the coverage of a larger set of English spelling-sound correspondences. Two approaches have been used, either exhaustive hand generation or the use of a learning mechanism (in a fashion similar to the DRC model of Coltheart *et al.*, in press). Improving the graphemic parser is a necessity with a larger GPC rule base since rules will exist for graphemes of more than a single letter. These graphemes must be identified in the letter string, and thus the current parser requires modification. The approach to be adopted for this has yet to be confirmed.

The experiments performed so far with the simulation have evaluated the validity of several assumptions made about the model rather than determining the success with which Norman simulates the behaviour of skilled reading. This latter point is the natural progression. Behavioural experiments include manipulations of psycholinguistic variables, such as word frequency, and use a variety of paradigms (such as speeded naming, lexical decisions, identity priming, and so on). At present it is anticipated that several important experimental findings will be investigated with Norman:

Word frequency effects (e.g. Forster and Chambers, 1973). High frequency words are responded to faster than low frequency words.

Identity priming effects (e.g. Scarborough *et al.*, 1977). Recognition of a word is faster when primed by itself.

Lexicality effects (e.g. Rubenstein, Garfield and Milikan, 1970). Words have an advantage over nonwords (nonwords take longer to reject than words take to be accepted in lexical decision experiments).

Legality effects (e.g. Gough and Cosky, 1977). Globally illegal sequences (e.g. SJMF) are rejected faster than partially or fully legal sequences. Of partially legal sequences, early illegality

---

[8] Indeed, how to achieve this parallelism in a language such as Lisp is not readily apparent. This problem is not restricted to Norman however, since other computer simulations use conventional programming languages. McClelland and Rumelhart's (1981) interactive activation model makes two important assumptions about parallel processing: that letters in words are processed in parallel and moreover that the three levels in the model's hierarchy — features, letters and words — operate in parallel also. The computer simulation of their model (written in C) does not implement true parallelism. Processing is assumed to operate in a series of discrete time steps. At each time step a strictly sequential process is performed to update the activation of each level of detectors and the effects of inter- and intra-level excitation and inhibition. Parallelism is achieved only in the sense that each level in the hierarchy is updated during every time step and that all levels reach their final "state" during the same time step. We suspect that similar problems may be present in the DRC model of Coltheart *et al.* (in press), which is a computational dual-route model of reading identical in principle to the revised logogen model as described above. Coltheart *et al.* employ the interactive activation model for visual word identification (replacing visual analysis and visual input logogens of the revised logogen model), together with a nonlexical GPC route. Their equivalent of the response buffer of the revised logogen model consists of an interactive activation network in which the outputs of both the lexical and nonlexical routes compete to arrive at the response to make. How the nonlexical route operates in parallel with the lexical route is not made clear.

(e.g. SJIP) causes faster rejection than late illegality (e.g. SAJE).

Orthographic neighbourhood effects (e.g. Andrews, 1989). Low frequency words with a small orthographic neighbourhood (the set of words that can be formed by changing just one letter in the word at a time) take longer to accept than low frequency words with a large orthographic neighbourhood.

RSVP effects (e.g. Krueger and Shapiro, 1979). A word advantage over nonwords is observed for letter strings presented up to 10 stimuli per second. (see discussion of Experiment One).

Regularity effects (e.g. Glushko, 1979). Words involving the use of regular GPC rules (e.g. CAT, HORSE, QUICK) are recognised faster than words involving exceptions to the rules (e.g. YACHT).

The final issue to be discussed is the process of visual analysis in word recognition, raised in Experiment Two. The current wholistic approach to word recognition is one which is not widely supported in other models, the belief being that words are visually identified by a hierarchical system (McClelland, 1976; Adams, 1979; Johnston and McClelland, 1980; McClelland and Rumelhart, 1981; Coltheart *et al.*, in press). The logogen model's vague exposition of visual analysis prompted the assumptions embodied in the current approach. However, a hierarchical approach could be incorporated into the current simulation. To achieve this the visual analysis component would require expansion to include letter as well as feature detectors, with the identified letters serving as the evidence for the activation of visual input logogens (the same principle Coltheart *et al.*, in press, have adopted for their DRC model). However, Howard (1987) reports studies of TM, a deep dyslexic patient who finds distorted words (e.g. d o g, d o g, d o g or aLtErNaTiNg cAsE) difficult to read (contradicting reported findings of no significant format distortion effects, e.g. McClelland, 1976; Adams, 1979). TM is almost entirely unable to match letters across case. Howard argues that "such a patient presents problems for the theories of word recognition that suppose that word recognition units are addressed by abstract letter codes alone"(Howard, 1987, p.33). Howard's observation that TM possesses no abstract representations for the component letters of words leads him to conclude that TM does read by recognition of words without an intermediate stage of activation of their component letters. This clearly goes against hierarchical claims for visual word recognition. The exact operation of visual analysis in word recognition is thus somewhat confusing given that both wholistic and hierarchical theories are supported by experimental evidence. However, hierarchical approaches have been embodied in several recent computer simulations of word recognition (McClelland and Rumelhart, 1981; Coltheart *et al.*, in press), and the use of hierarchical visual word identification within the logogen model should be seriously considered.

In summary, Norman represents a computer simulation of the revised logogen model which is capable of reading performance comparable to that of skilled readers. Development of the simulation continues, with emphasis on expansion of the nonlexical route and elaboration of the process of visual word identification. In addition the simulation is to be applied to behavioural data to determine to what extent a number of experimental manipulations affect Norman in the way they affect the reading performance of skilled human readers.

References

ADAMS, M.J. *Models of Word Recognition.* Cognitive Psychology, **11**, 133-176, 1979.

ANDREWS, S. *Frequency and Neighborhood Effects on Lexical Access: Activation or Search?.* Journal of Experimental Psychology: Learning, Memory and Cognition, **15**, 802–814, 1989.

ANDREWS, S. *Frequency and Neighborhood Effects on Lexical Access: Lexical Similarity or Orthographic Redundancy?.* Journal of Experimental Psychology: Learning, Memory and Cognition, **18**, 234–254, 1992.

CHOMSKY, N. & HALLE, M. *The Sound Pattern of English.* Harper and Row, 1968.

CLARKE, R. & MORTON, J. *Cross-modality Facilitation in Tachistoscope Word Recognition.* Quarterly Journal of Experimental Psychology, **35A**, 79–96, 1983.

COLTHEART, M., *Lexical Access in Simple Reading Tasks.* In G. Underwood (ed) *Strategies of information Processing.* North Holland, 1978.

COLTHEART, M., DAVELAAR, C., JONASSON, J.T. & BESNER, D. *Access to the Internal Lexicon.* In S. Dornic (ed) *Attention and Performance VI*, Academic Press, 1977.

COLTHEART, M., CURTIS, B. & ATKINS, P. *Models of Reading Aloud: Dual-Route and Parallel Distributed Processing Approaches.* Psychological Review, in press.

FORSTER, K.I. & CHAMBERS, S.M. *Lexical Access and Naming Time.* Journal of Verbal Learning and Verbal Behaviour, **12**, 627–635, 1973.

GLUSHKO, R.J. *The Organisation and Activation of Orthographic Knowledge in Reading Aloud.* Journal of Experimental Psychology: Human Perception and Performance, **5**, 674–691, 1979.

GOUGH, P.B. & COSKY, M.J. *One Second of Reading Again.* In N.J. Castellan, D.B. Pisani & G.R. Potts (eds) *Cognitive Theory,* II, Erlbaum, 1977.

HOWARD, D. *Reading Without Letters?.* In M. Coltheart, G. Sartori & R. Job (eds) *The Cognitive Neuropsychology of Language*, Lawrence Erlbaum Associates, 1987.

JOHNSTON, J.C. & McCLELLAND, J.L. *Experimental Tests of a Hierarchical Model of Word Identification.* Journal of Verbal Learning and Verbal Behaviour, **19**, 503–524, 1980.

JUOLA, J.F., WARD, N.J., & McNAMARA, T. *Visual Search and Reading of Rapid Serial Visual Presentations of Letter Strings, Words, and Text.* Journal of Experimental Psychology: General, **111**, 208–227, 1982.

KRUEGER, L.E., & SHAPIRO, R.G. *Letter Detection With Rapid Serial Visual Presentation: Evidence Against Word Superiority at Feature Extraction.* Journal of Experimental Psychology: Human Perception and Performance, **5**, 657–673, 1979.

MASTERSON, J., HAZAN, V. & WIJAYATILAKE, L. *Phonemic Processing Problems in Developmental Phonological Dyslexia.* Cognitive Neuropsychology, in press.

McCLELLAND, J.L. *Preliminary Letter Identification in the Presentation of Words and Nonwords.* Journal of Experimental Psychology: Human Perception and Performance, **2**, 80-91, 1976.

McCLELLAND, J.L. & RUMELHART, D.E. *An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings.* Psychological Review, **88**, 5, 1981.

MONSELL, S., DOYLE, M.C. & HAGGARD, P.N. *Effects of Frequency on Visual Word Recognition Tasks: Where are They?.* Journal of Experimental Psychology: General, **118**, 43–71.

MORTON, J. *The Interaction of Information in Word Recognition.* Psychological Review, **76**, 2, 165–78, 1969.

MORTON, J. *A Functional Model for Memory.* In D.A. Norman (ed) *Models of Human Memory*, Academic Press, 1970.

MORTON, J. *Word recognition.* In J. Morton & J.C. Marshall (eds) *Psycholinguistics Series 2: Structures and Processes*, Paul Elek, 1979

MORTON, J. *The Logogen Model and Orthographic Structure.* In U. Frith (ed) *Cognitive processes in Spelling*, Academic Press, 1980.

MORTON, J. & PATTERSON, K.E. *A New Attempt at an Interpretation or an Attempt at a New Interpretation.* In M. Coltheart, K. Patterson & J.C. Marshall (eds) *Deep Dyslexia* Routledge & Kegan-Paul, 1980.

MURRELL, G.A. & MORTON, J. *Word Recognition and Morphemic Structure*. Journal of Experimental Psychology, **102**, 936–968, 1974.

PATTERSON, K.E. & MORTON, J *From Orthography to Phonology: An Attempt at an Old Interpretation.* In K.E. Patterson, J.C. Marshall & M. Coltheart, *Surface Dyslexia*, Lawrence Erlbaum Associates, 1985.

QUINLAN, P.T. *Oxford Psycholinguistic Database*. Oxford University Press, 1992.

RUBENSTEIN, H., GARFIELD, & MILIKAN *Homographic Entries in the Internal Lexicon.* Journal of Verbal Learning and Verbal Behaviour, **9**, 487–494, 1970.

SCARBOROUGH, D.L., CORTESE, C., & SCARBOROUGH, H.S. *Frequency and Repetition Effects in Lexical Reading*. Journal of Experimental Psychology: Human Perception and Performance, **3**, 1–17, 1977. b

SEIDENBERG, M.S., WATERS, G.S., BARNES, M.A. & TANENHAUS, M.K. *When Does Irregular Spelling of Pronunciation Influence Word Recognition?*. Journal of Experimental Psychology, **84**, 74–81, 1984.

SOLTYSIAK, S.J. *Word Recognition and the Logogen Model: A Computational Investigation*. Unpublished MSc. Dissertation, Department of Computer Science, University of Essex, 1993.

TARABAN, R., & McCLELLAND, J.L. *Conspiracy Effects in Word Pronunciation.* Journal of Memory and Language, **26**, 608–631, 1987.

WINNICK, W.A. & DANIEL, S.A. *Two Kinds of Response Priming in Tachistoscopic Recognition*. Journal of Experimental Psychology, **84**, 74–81, 1970.