Evolution of the Dual Route Cascaded Model

of Reading Aloud

Kevin Chang

University of Waterloo, Canada

Abstract

The time for skilled readers to name a non-word increases as the number of letters increase, and also increases when the stimulus is degraded. These effects are known as the *length effect* and *stimulus quality effect*, respectively. Besner and Roberts (2002) reported that the joint effect of these two factors on RT are *additive* in skilled readers. They also reported that the leading computational model of basic processes in reading, Coltheart, Rastle, Perry, Langdon & Ziegler's Dual Route Cascaded model (2001), produces an *under additive* interaction between these two factors. Besner and Roberts argued that this qualitative difference challenges DRC's fundamental assumption of cascaded processing. They proposed that thresholding early processing in the model would allow the model to simulate the human results. The present work implements such a threshold at the letter level in DRC. The new model successfully reproduces the joint effects of letter length and stimulus quality seen in skilled readers.

Introduction

The last two decades have seen a proliferation of computational models of reading aloud. Unlike the traditional "verbal models," where researchers describe the model's characteristics (Jacobs & Grainger, 1994), computational models are implemented as computer programs that can be executed and their internal workings analyzed. Computational models are increasing in popularity because they contain several desirable characteristics not shared by the traditional verbal models. First, ambiguity inherent in verbal model's specification is avoided because implementing a model requires completeness. That is, the researchers must explicitly spell out all components of a theory for the computer program to execute. Second, computational models can be assessed by comparing the model's performance with human data.

A number of computational models of skilled reading exist (e.g., McClelland & Rumelhart, 1981; Grainger & Jacobs, 1996; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, McClellnd, Seidenberg, & Patterson, 1996). To date, Coltheart et al.'s Dual Route Cascaded Model (DRC) is considered the most successful. DRC simulates 18 effects seen in the naming task and numerous effects in lexical decision. In addition, Coltheart et al. were able simulate various forms of acquired dyslexia. Coltheart et al. comment that:

"…… the set of phenomena that the DRC model can simulate is much larger than the set that any other current computational model of reading aloud can simulate; and, to the best of our knowledge, there is no effect seen in reading aloud that any of these other models can simulate but that DRC cannot."

We first describe the Dual Route Cascaded model. Second, we review evidence

reported by Besner and Roberts (2002) that appears problematic for the model. We then report and test a new implementation of part of the model, and show that it is able to simulate the data reported by Besner and Roberts (2002).

The Dual Route Cascaded model

As the name suggests, the Dual Route Cascaded model (DRC) has two core assumptions. First, processing throughout the model is cascaded. That is, any activation in earlier modules starts flowing to later modules immediately. Second, there are two routes for translating print into sound: a lexical route, which utilize word-specific knowledge, and a non-lexical Grapheme-to-Phoneme Conversion (GPC) route, which utilize a sub-lexical spelling-sound correspondence rule system. These routes can be seen in Fig 1.

Figure 1

The assumption of cascaded processing is derived from McClelland and Rumelhart's (1981) seminal work on the Interactive Activation model (IA) of context effect in letter perception. In fact, DRC is an extension of this model, in which the essentials of the feature and letter level processing modules (top part of Figure 1) are maintained.

Another feature of the IA model and most of DRC is that processing is done in parallel. For example, all features across the stimulus array are extracted in parallel. Similarly, all the letters units are activated in parallel. Indeed, processing occurs in parallel within all modules except the GPC module, where processing is serial. The

serial processing nature of the GPC module is explained more thoroughly in a later section.

The Two Routes

A second major assumption of DRC is that there are two routes underlying the process of converting print to sound. One is the lexical route and the other is the non-lexical GPC route. The lexical route translates the pronunciation of a word based on word specific knowledge. The route consists of three components: the semantic system, the orthographic lexicon, and the phonological lexicon, as seen in the left part of Figure 1. The semantic system computes the **meaning** of a word, whereas the lexicons compute the words' **orthographic** and **phonological** form. Currently, the semantic system is not implemented and will not be discussed further. Representations of a word in the orthographic lexicon and the phonological lexicon are linked so that activation in one leads to activation of the other. For instance, the letters "c," "a" and "t" will activate the orthographic representation of "cat," which will then activate its phonological representation of /k{t/. Frequency scaling is also applied to each orthographic and phonological lexicon. Thus, a high frequency word such as "the" will be named faster than a low frequency word such as "quench."

The non-lexical route differs from the lexical route in both the knowledge base and the type of processing it employs. The non-lexical route generates the pronunciation of letter string (be it a word or a non-word) via a set of sub-lexical spelling-sound correspondence rules. The set of rules is encapsulated in the GPC module. One important feature of the GPC module is that its processing is serial. The GPC module applies rules serially left to right to a letter. That is, letters activate phonemes in a serial, left to right fashion. Activation of the second phoneme does not start until a constant number of cycles after the start of activation of the first letter. For example, given a

non-word like "bant", the corresponding translation would be: B -> /b/, A -> /{/, N -> /n/, and T -> /t/. Coltheart et al. (2001) argue the non-word letter length effect produced by DRC is a direct consequence of serial processing in the GPC module. That is, because GPC processes letters serially, the time to name a non-word increases as the length of non-word increases. This phenomenon parallels the *letter length effect* in human performance (Weekes, 1997). As we shall see in the later sections, the letter length effect turns out to be a good vehicle for examining how DRC's non-lexical route operates.

The lexical route utilizes word-specific knowledge to determine the corresponding pronunciation, whereas the non-lexical route translates graphemes into phonemes via a set of sub-lexical spelling-sound correspondence rules. Thus, given a word that is known to the reader, the correct pronunciation is quickly generated by the lexical route. A non-word that cannot be found in the orthographic lexicon and hence cannot be read by the lexical route can be read by the non-lexical route. Although the set of sub-lexical spelling-sound correspondence rules can also be applied when naming known words, the resulting pronunciation will regularize the pronunciation of exception words (e.g. PINT is pronounced /pInt/). Together, an intact system of lexical and non-lexical routes is capable of pronouncing both words and non-words. Detailed discussion of this issue can be found in Coltheart et al. (2001). Here, we only consider how the model simulates the pronunciation of non-words. Therefore, only the operation of the non-lexical route is considered further.

GPC Route: An Example

A detailed walk through how GPC translate the non-word "bant" is provided below. This walk is useful because it provides insight into the actual implementation of the model.

Given the non-word "bant" to read, the model works as follows. On cycle 1, the stimulus is loaded into the model and the features making up each letter are set to 1 (present) or 0 (absent). On every subsequent cycle, activation is passed from the feature units to the letter units in parallel across all features and letter positions. Because the processing at the letter level is parallel and cascaded, all letter positions are activated at the same time and activation cascades to the orthographic level and GPC module immediately. Unlike the orthographic level, where activation occurs in parallel, the GPC module is constrained by its serial processing. Starting at cycle 10, the GPC module starts processing the first letter. The sub-lexical spelling-sound correspondence rule system is searched until a rule is matched to the first letter. The GPC module receives the same letter input until 17 cycles later, when the second letter is admitted to the GPC module. At cycle 27, the first two letters are fed into the GPC module. The rule system is then searched until a rule matched the first two letters. If such a rule cannot be found, the rule system will find a rule matching the first letter, and another rule matching the second letter. That is, the rule system will always try to match the longest grapheme. The translation process continues with the GPC module receiving an additional letter every 17 cycles, until all letters have been translated to phonemes. DRC is said to have named the stimulus when all phonemes receive activation of 0.43.

Besner and Roberts (2002)

Factorial experiments in which a factor that affects the rate of processing (e.g., stimulus quality) is varied in conjunction with another factor (e.g., word frequency) have been used for over a quarter of a century to evaluate different non-computational accounts of visual word recognition (e.g., Becker & Killion, 1977; Besner & Smith, 1992; Borowsky & Besner, 1993; Meyer, Schvaneveldt & Ruddy, 1975; Stanners, Jastrzembski & Westbrook, 1975; Stolz & Neely, 1995). To date, virtually none of the

published computational models of visual word recognition have explored whether such results can be successfully simulated. Besner and Roberts (2002) explored the impact of slowing the rate of early processing on the letter length effect in non-word naming performance by the DRC model, and by skilled readers.

Besner and Roberts manipulated the stimulus quality of a letter string by varying display contrast. A sharper display contrast like black print on a white background is treated as a clear condition, whereas a black print on a grey background is treated as a degraded condition. In the DRC model, Besner and Roberts simulated the reduction in stimulus quality by slowing the rate of processing in the early part of the model. Specifically, the connections weights between the feature and letter levels in the model were reduced by 40%. This results in a slower rate of activation throughout because the model is cascaded. A detailed rationale for this particular implementation of stimulus quality is provided in Besner and Roberts (2002).

Besner and Roberts conducted two simulations with the model: one in which the lexical route is intact, and one in which the lexical route is lesioned. The lesioning of the lexical route was done by zeroing out the connections between the letter level and the orthographic input lexicon, as well as the connections between the phonological output lexicon and the phoneme system. The stimulus set consists of 64 monosyllabic non-words where letter length varies from short (3 and 4 letters) to long (5 and 6 letters). The results of the intact and lesioned model are shown in the left-hand-panel and the middle panel of Figure 2, respectively.

Figure 2

The results of these two simulations are very similar. Both main effects of stimulus quality and letter length effect are significant. That is, the time to name a non-word increases when the stimulus is degraded, and also increases as the number of letters increase. More importantly, an *under-additive* interaction is observed; slowing the rate of processing significantly affects long stimuli **less** than short stimuli. These simulation results are inconsistent with the human data, as can be seen in the right-hand panel of Figure 2. The two main effects of stimulus quality and letter length effects are significant, but there is no interaction. That is, in skilled readers, slowing the rate of early processing has the **same** impact on short and long stimuli. This qualitative difference in the model and human performance is problematic for the DRC model. Some modification to the model is necessary to remedy the discrepancy.

Before going further, it is necessary to understand why the model produces the under-additive interaction of stimulus quality effect and letter length effect. Besner and Roberts argued the interaction occurs as a result of cascaded parallel processing at the letter level on subsequent serial processing in the model's non-lexical route. Because the GPC module operates serially, the activation of each successive phoneme only starts when the prior phoneme has been receiving activation for a constant number of cycles. Meanwhile, activation at the letter level continues. Because the letter level operates in parallel across all letter positions, all letters continue to receive activation and move closer to asymptote. Consequently, although activation of the first phoneme is affected by slowing the rate of processing, the delay associated with each additional phoneme allows ongoing letter level activation to move closer to asymptote. Thus, the additional phonemes in a long stimulus are less subject to the effect of stimulus quality.

Modification to the DRC

Given the qualitative difference in the model and human performance, some modification to the model is necessary to remedy the discrepancy. Besner and Roberts (2002) suggested a threshold be implemented at the letter level. We implement such a threshold here. In addition, an activation asymptote is set to be the same value as the threshold. That is, the letter level does not start to pass activation forward until a threshold is reached and when that happens, the activation passes forward is bounded by this asymptote. In effect, the modification serves to partition the effects of stimulus quality and letter length into separate stages. Stimulus quality affects the time to reach letter level threshold, whereas letter length affects a process further downstream. Additive effects of processing rate and letter length are therefore expected.

This approach was implemented in DRC. Because the source code of DRC was not available to us, the first step was to implement the non-lexical route of DRC. The feature, letter, GPC, and phoneme modules were implemented. The adequacy of the implementation was verified by comparing the output it produces with the Besner and Roberts' stimulus set with the output from DRC. As seen in the Panel C of Figure 3, the output of the new implementation (labelled nonlexDRC2) matched perfectly with the original model (Panel B).

Figure 3

The next step was to implement the proposed threshold. As described earlier, the threshold is implemented at the letter level such that the activation does not start to pass

forward until a threshold is reached. Further, the activation that is passed forward is fixed at an asymptote equal to the threshold level. This version of the non-lexical route is labelled nonlexDRC/T where T denotes thresholding. It was not clear to us what the threshold value should be. We therefore tested a range of different thresholds (0.6, 0.7, 0.8 and 0.998). As can be seen in Figure 4, all thresholds produced perfect additivity of stimulus quality and letter length. It thus appears that the computational solution presented here is adequate for the identified problem.

Figure 4

Conclusion

By thresholding the letter level module in DRC, the modified non-lexical route is able to simulate what Coltheart et al.'s version of DRC does not. Namely, the additive effects of stimulus quality and letter length on the time to name a non-word. The computational solution implemented here challenges DRC's fundamental assumption of cascaded processing. It remains to be seen whether there is a pay-off to also threshold the lexical route, and further, whether all effects currently simulated by DRC are also simulated by nonlexDRC/T.

References

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 389-401.

Besner, D., & Roberts, M. (in press). Reading nonwords aloud: Results requiring change in the dual route cascaded model. *Psychonomic Bulletin and Review.*

Besner, D., & Smith, M. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 468-482.

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 813-840.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multi read-out model. *Psychological Review, 103*, 518-565.

Jacobs, A. M., & Grainger, J. M. (1994). Models of visual word recognition: Sampling from the state of the art. *Journal of Experimental psychology: Human Perception and*

Performance, 20, 254-266.

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.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P.M.A. Rabbitt & S. Dornic (Eds.) *Attention and Performance* V. (pp 98-118). San Diego, California. Acdemic Press.

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

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behaviour, 14,* 259-264.

Stolz, J. A., & Neely, J. (1995). When target degradation does and does not enhance semantic context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 596-611.

Weekes, B.S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A,* 439-456.
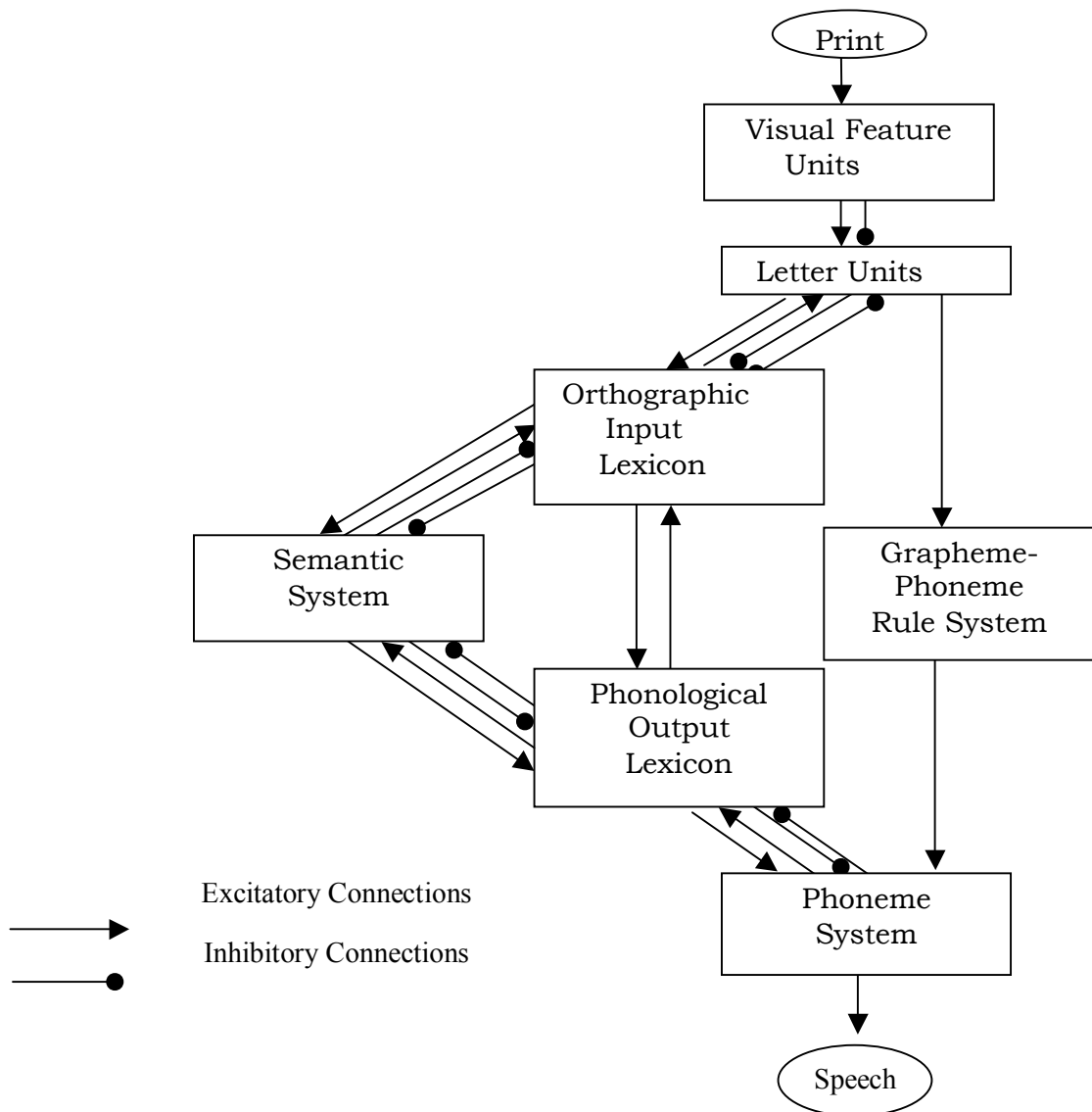
List of Figure Captions

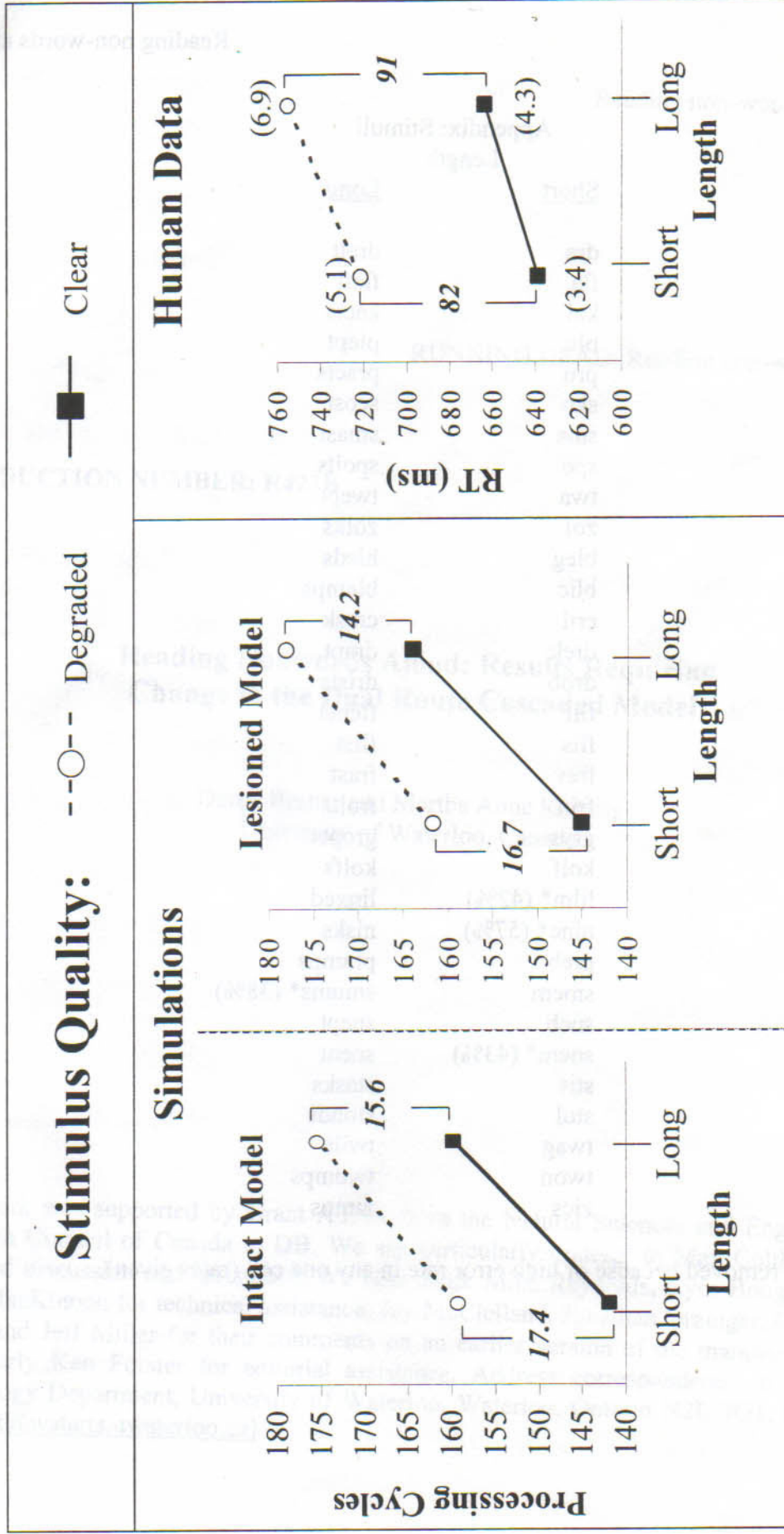Figure 1: The DRC model of visual word recognition

Figure 2: The joint effects of Length and Stimulus Quality for DRC [Processing Cycles] and human readers [RT(ms) and (% Error)].
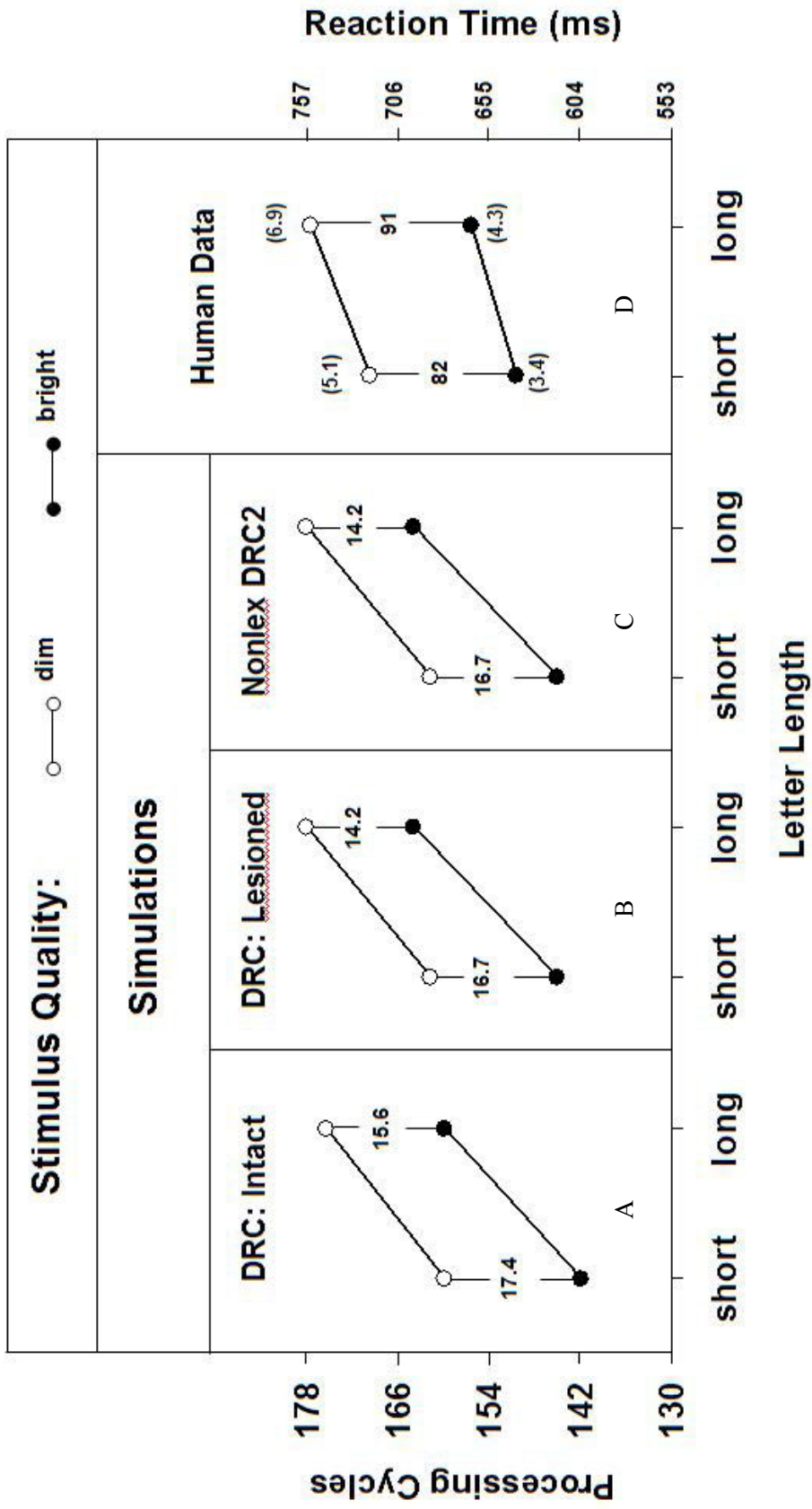
Figure 3: Data from DRC when intact (Panel A), Lesioned (Panel B),
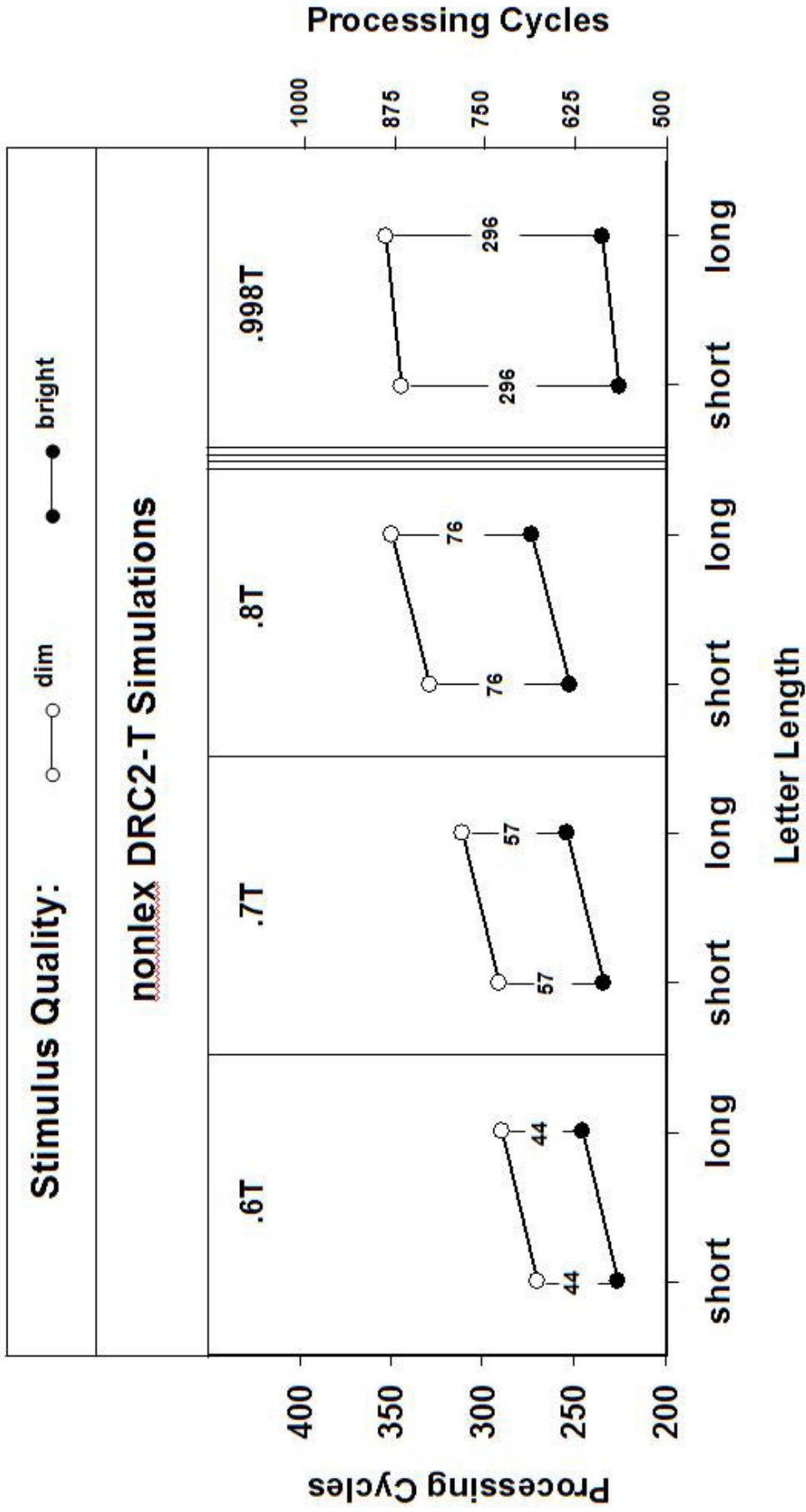and the new version of the nonlexical route (Panel C) along with the Human data (Panel D)

Figure 4: Thresholding the letter level in DRC