A NEW PERSPECTIVE ON VISUAL WORD PROCESSING EFFICIENCY

James T. Townsend
Department of Psychological and Brain Sciences, Indiana University, Bloomington, Indiana, USA
Joseph W. Houpt
Department of Psychology, Wright State University, Dayton, Ohio, USA

Abstract

As a fundamental part of our daily lives, visual word processing has received much attention in the psychological literature. Despite the well established perceptual advantages of word and pseudoword context using accuracy, a comparable effect using response times has been elusive. Some researchers continue to question whether the advantage due to word context is perceptual. We use the capacity coefficient, a well established, response time based measure of efficiency to provide evidence of word processing as a particularly efficient perceptual process to complement those results from the accuracy domain.

Due to the relative ease with which most adults read, it is reasonable to assume that word perception is an efficient process. This is further supported by the intuition that with more experience with a process we become more efficient and we are quite experienced with the written word. Often, the efficiency is measured using single letter perception as a base line. When word context offers an advantage in the accuracy or processing time of perceiving a letter, this supports the claim that word perception is efficient.

From the early days of experimental psychology, researchers have been interested in the value of a word context for perceiving letters. In one study, letters were displayed sequentially to participants at faster and faster rates until they could no longer correctly identify the letters. They found that participants maintained accuracy with shorter durations when the letters were presented as part of a word compared with random letter sequences (Cattell, 1886). One problem with studies of this nature is that they do not control for the constraint on possible letters that a word context puts on the possible letters. Hence it is not clear from those early results whether the advantage is a perceptual advantage or a decisional advantage. In the late 1960's an alternative task was designed to eliminate the decisional advantage of word context so as to examine the perceptual effects. In this task a letter or word was tachistoscopically displayed to a participant. They then chose from two possible choices, one of which was correct. In the letter condition, the choices were letters. In the word condition, both choices were words that differed in only a single letter. Since both alternatives were words, the word context was no longer informative as to the identity of the letter. Participants were still more accurate at perceiving letters in the word condition than the letter condition (Reicher, 1969). Furthermore, they found that participants are also more accurate with word contexts than random letter sequence contexts (the word superiority effect). An efficiency gain of context over letters alone is not unique to words though. If a sequence of letters conformed to the pronunciation rules of English, strings referred to as pseudowords, then participants were again more accurate than letters alone (the pseudoword superiority effect, e.g., McClelland & Johnston, 1977).

Despite the robustness of the word and pseudoword superiority effects, a comparable effect using response times (and controlling for decisional information due to context) has been elusive. This may be in part explained by the possibility that people will read an entire word even if the task does not require it. Indeed, this has been put forth as further evidence that word perception is special (LaBerge & Samuels, 1974). One of the goals of this paper is to demonstrate a response time based word superiority effect, and possibly a pseudoword superiority effect as well. In the next section we describe the capacity coefficient, a response time based measure of efficiency. We propose that this measure, along with a task that controls for both the available information and possibly mandatory word reading, provides evidence of word processing as a particularly efficient process to complement the accuracy results.

The Capacity Coefficient

The capacity coefficient, $C(t)$ is an established response time based measure of the effect of increased load on processing efficiency (Townsend & Nozawa, 1995; Townsend & Wenger, 2004). Specifically, $C(t)$ is a measure of the change in processing rates as the task requires attention to more targets, or possibly more dimensions of a single target. The basic idea of the measure is to compare response times when reading the full string to the times that would be predicted if each character took the same amount of time, whether or not it was in a string.

The capacity function for an exhaustive task is defined using the natural log of the cumulative distribution function, $K(t) = \ln F(t); F(t) = \Pr\{\mathrm{RT} \leq \mathrm{t}\}$, and is similar to the cumulative hazard function used in survival analysis. If K_{c1} is the cumulative hazard for the first character response times, K_{c2} is the cumulative hazard for the second character, etc., and K_S is the cumulative hazard for the string condition, the capacity coefficient is given by $C(t) = \left[\sum_{i=1}^{4} K_{c_i}\right] / K_S$.

This formulation is based on the predictions of the unlimited capacity, independent, parallel (UCIP) model. The assumptions of the UCIP model are sufficient conditions for there to be no change in the rates of processing with increased load. If these assumptions hold then the relationship between the processing times of the string to the processing times of the individual characters is as follows:

$$\Pr\{RT_S \leq t\} = \Pr\{RT_{c_1} \leq t\}\Pr\{RT_{c_2} \leq t\}\Pr\{RT_{c_3} \leq t\}\Pr\{RT_{c_4} \leq t\}$$

By taking the natural log of both sides of this equation, then dividing by the left hand side, we see that the UCIP model predicts $C(t) = 1$ for all $t \geq 0$. This gives us a baseline for comparison. If a person performs better than the baseline model, $C(t) > 1$, their performance is referred to as super-capacity. There are multiple ways performance could be super-capacity. For example, if there is facilitation between the characters, or in more extreme cases if the information from the characters is accumulated together toward a single decision (Townsend & Wenger, 2004). Performance worse than the baseline model, $C(t) < 1$, is limited-capacity. In contrast to the case of super-capacity, inhibition between characters could result in limited-capacity. When performance is about the same as the baseline model, $C(t) \approx 1$, then we refer to it as unlimited capacity.

The capacity coefficient measures processing efficiency in isolation by comparing the capacity coefficient to predicted values of unlimited capacity, independent, parallel models. Thus, this measure also allows us to compare the efficiency of a variety of processes despite any possible differences in difficulty due to component processes. In particular, we are able to draw conclusions about the efficiency of word processing relative to pseudoword, non-word, upside-down non-word, and unfamiliar character string processing.

Method

To compare perceptual efficiency across words, pseudowords, non-words, upside-down words and unfamiliar characters, our task must eliminate the extra information available given a word

	Target	Distractors				Single Character							
Word	care	bare	cure	cave	card	c	b	a	u	r	v	e	d
Pseudoword	lerb	nerb	larb	lemb	lerf	l	n	e	a	r	m	b	f
Non-Word	rlkf	vlkf	rtkf	rlhf	rljk	r	v	l	t	k	h	f	k
Upside-down	ɟʞlɹ	ɟʞlʌ	ɟʞʇɹ	ɟɥlɹ	ʞɾlɹ	ɹ	ʌ	l	ʇ	ʞ	ɥ	ɟ	ʞ
Katakana	サイクオ	ヘイクオ	サナクオ	サイフオ	サイクノ	サ	ヘ	イ	ナ	ク	フ	オ	ノ

Table 1: Stimuli used for capacity analysis.

context. The possibility that words are exhaustively processed automatically may lead to a disadvantage for words on response time measures. To address these issues, we used a task which forces exhaustive processing of the characters in a string. This experiment consists of two components. We measured the participants' response times to correctly identifying the target string. To ensure that participants identified targets using the entire string and not any subset, we included a distractor of a string with a single character in each position different from the string. For example, if the target is "care" then "bare," "cure," "cave" and "card" were used as distractors (see Table 1). The participants also distinguished between each letters in isolation. Whereas in the exhaustive case the participant needed to distinguish between "bare" and "care," participants distinguished between "b" and "c" in this condition.

Participants were recruited from the Indiana University population. Eight females and two males participated in this study, all of whom were native English speakers and reported that they did not read or speak Japanese. Their ages ranged from 19-34. All participants reported having normal or corrected to normal vision, no difficulty reading English, and no prior diagnoses of a reading disorder.

Table 1 summarizes the stimuli used for both the single character and exhaustive trials for each type. There were five types of stimuli used: words, pronounceable non-words (pseudowords), unpronounceable non-words, upside-down unpronounceable non-words, and strings of Katakana characters. All strings used were four characters long. Words were chosen so that the frequency of the target was roughly equal to the average frequency of the distractors. Pseudowords were taken from the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002). Strings and characters were presented in black Courier font on a gray background.

Participants were paid \$8 per session, and received a \$20 bonus upon completion of all 10 sessions. Each session lasted between 45 and 60 minutes and was dedicated to one of the five types of stimuli (e.g., word, pseudoword, . . .), so there were two sessions of each type. At the beginning of each session, we read the participant the general instructions for the task while those instructions were presented on the screen. The instructions encouraged participants to respond as quickly as possible while maintaining a high level of accuracy. Each session was divided into five blocks, one block of string stimuli and a block for each of the corresponding single character stimuli.

Each block began with a screen depicting the button corresponding to each of the categories. Participants had 40 practice trials, 20 of each category. Next participants were given 240 trials divided evenly between the two categories, the first 40 of which were not used in the analysis. Each trial began with a 30 ms presentation of a fixation cross. After a random delay (300-600 ms), the stimulus was presented for 80 ms. Participants had a maximum of 2500 ms to respond. If the participant responded correctly, the next trial started after a 400 ms delay. If the participant responded incorrectly, a tone was played during the 400 ms delay. The session order was counterbalanced among the participants so that participants completed the different

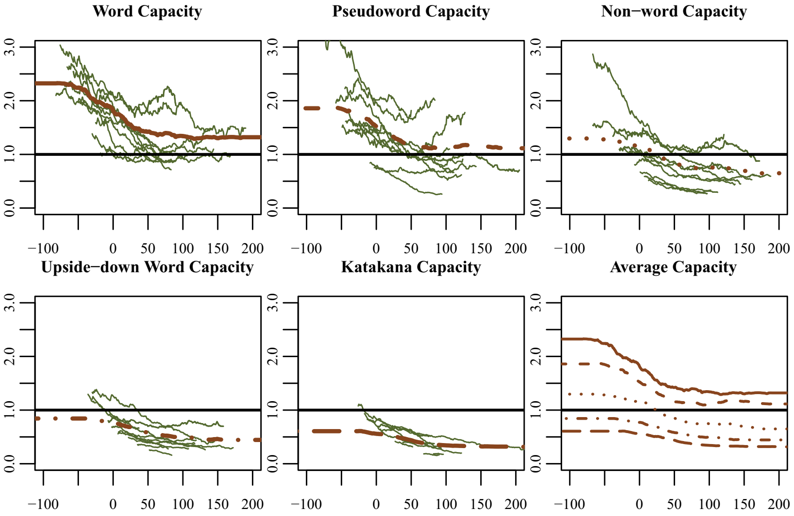

Figure 1: Capacity coefficient functions for each version of the task. The thin lines represent individual participants' data and the thick line represents th average across functions.

types on different days and in different orders.

Results

Individual capacity coefficients are shown in Figure 1. Z-scores for individual and group data, using the statistic in Houpt and Townsend (2012) are shown in Table 2. Each Z-score indicates a test of the null-hypothesis that a participant performs equally to a UCIP model. Significance values are based on a two-sided test. Nearly all participants are significantly different from UCIP, usually better in the word and pseudoword conditions and worse in the other conditions.

Using repeated measures ANOVA, we found a significant effect of condition on capacity ($F(4,36) = 22.64, p < 0.05, \eta^2_G = 0.58$). For post-hoc analyses, we used the z-scores resulting from the mean difference between subjects' capacity z-scores in each pair of conditions. Word capacity was significantly higher than pseudoword capacity ($z = 7.27, p < 0.0025$), random letter capacity ($z = 22.9, p < 0.0025$), upside-down capacity ($z = 36.7, p < 0.0025$), and Katakana capacity ($z = 45.9, p < 0.0025$). Pseudoword capacity was significantly higher than random letter capacity ($z = 15.6, p < 0.0025$), upside-down capacity ($z = 29.4, p < 0.0025$), and Katakana capacity ($z = 38.6, p < 0.0025$). Random letter capacity was higher than upside-down capacity ($z = 13.8, p < 0.0025$), and Katakana capacity ($z = 22.9, p < 0.0025$). Upside-down capacity was significantly higher than Katakana capacity ($z = 9.19, p < 0.0025$).

Discussion

Due to space limitations, we limit the majority of our discussion to the word and, to a lesser extent, the pseudoword results. We have demonstrated clear evidence of super-capacity processing of the word stimuli for nine of the ten participants. These participants are efficiently perceiving the whole word in comparison to individual letter perception. As mentioned earlier, evidence for the word superiority effect has been difficult to demonstrate with response times. These findings provide that evidence and thus agree with the majority of the word perception literature based on accuracy results. Based on comparisons across conditions, it is also clear that the

	Word	Pseudoword	Random	Upside-Down	Katakana
1	9.97^{***}	3.92^{***}	7.19^{***}	-2.62^{**}	-4.43^{***}
2	11.92^{***}	4.44^{***}	-0.73	-5.95^{***}	-10.02^{***}
3	8.19^{***}	-6.29^{***}	-6.88^{***}	-10.88^{***}	-12.34^{***}
4	0.13	-3.38^{***}	-7.34^{***}	-6.60^{***}	-10.58^{***}
5	0.79	10.70^{***}	-2.36^{*}	-6.27^{***}	-6.86^{***}
6	7.34^{***}	5.19^{***}	10.61^{***}	-2.58^{**}	-11.99^{***}
7	9.34^{***}	3.25^{**}	-2.27^{*}	-2.49^{*}	-5.78^{***}
8	7.17^{***}	7.84^{***}	4.68^{***}	2.86^{**}	-1.79
9	5.71^{***}	13.34^{***}	-8.43^{***}	-9.52^{***}	-7.37^{***}
10	3.88^{***}	2.45^{*}	-2.46^{*}	-7.44^{***}	-9.40^{***}
Group	20.38^{***}	13.11^{***}	-2.52^{*}	-16.28^{***}	-25.47^{***}

Table 2: Workload capacity statistical results for each participant. ($\ast\ast\ast : p < .001$; $\ast\ast : p < .01$; $\ast : p < .05$)

word perception was more efficient than non-word, upside-down word, and strings of Katakana perception, findings that again match with the results reported for accuracy (e.g., McClelland & Rumelhart, 1981). There is also evidence for a pseudoword superiority effect, another well established effect in the accuracy domain (McClelland & Johnston, 1977). Although the evidence was not as consistent as the word results, eight of the ten participants were super capacity for some time, with only two participants showing significantly limited capacity processing for most times.

There are multiple plausible explanations for the capacity coefficient results demonstrating particularly efficient processing of words. At least one of the assumptions of the UCIP model must have been violated, so we examine each of those assumptions in turn. Each of these violations have been considered previously for modeling the accuracy based superiority effects.

One assumption that may have been violated is that of independence. If there is any type of facilitation between the letter processes, each letter would be processed faster within a word which would explain the capacity coefficient values above one. There could be many explanations of this facilitation. For example, word processing mechanisms may in fact take advantage of the considerable amount of co-occurrence between letters in English. As is often observed, there are only a fraction of possible four letter combinations used for words and it would be surprising if we did not take some advantage of this reduction in uncertainty. This correlation between letters is an important part of how connectionist models explain the word superiority effect (McClelland & Rumelhart, 1981; Plaut, McClelland, Seidenberg, & Patterson, 1996; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001).

Many visual word processing models include a separate, phonological pathway. If a phoneme is activated as a possible interpretation of some letter combination, then it may in turn send positive feedback to those letters, speeding up their processing. Hence a phonological component of visual word processing could also lead to capacity coefficient values above one. Both the correlation between letters and the lack of a regular pronunciation of the non-words imply that these predictions are consistent with lack of evidence against the UCIP model of non-word processing. The phonological explanation is also supported by the evidence of a pseudoword superiority effect.

Another assumption of the UCIP model is that the letters are processed in parallel, with a separate detection of each letter. An alternative architecture that does predict capacity coefficient values above one is the coactive architecture. By pooling activation from each of the letters when processing a word, the word is processed much faster than if each letter is processed separately. A coactive architecture in this sense can be thought of as an extreme version of a facilitatory parallel model, in which all activation in each of the letters is shared. Many connectionist models of visual word perception assume a type of coactive architecture. In these models the activation accumulated in favor of a letter is immediately passed on to the word level. In this framework the type of parallel model assumed in the UCIP would not pass on any activation until the letter process is complete. A coactive architecture could also lead to violations of the assumption of unlimited capacity, so that seemingly more resources are available to each component when more components are present.

There were clearly individual differences present in these data, particularly in word and pseudoword processing capacity. This finding mirrors results reported in accuracy based studies (e.g., Reicher, 1969) and it will be an interesting extension of this work to compare the capacity measure to established measures of individual differences in reading.

Finally, we reiterate the importance of going beyond the simple ANOVA analysis of these data. Merely finding an ordering of the means in the string conditions says nothing about the relative processing efficiencies. For example, faster word processing than non-word processing could be due to the letters in "care" being relatively faster to process than the letters "rlkf". Workload capacity analysis, however, takes the processing of the components into account in estimating efficiency.

References

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, *11*, 63–65.

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.

Houpt, J. W., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, *56*, 341–355.

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception and Psychophysics*, *22*, 249–261.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter percpetion: Part 1. an account of basic findings. *Psychological Review*, *88*, 375–407.

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

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology*, *55A*, 1339–1362.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 274–280.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology*, *39*, 321–360.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, *111*, 1003–1035.