An Explanation of the Length Effect for Rotated Words

Carol Whitney (cwhitney@cs.umd.edu)
Neural and Cognitive Sciences Program
Philosophy Department
University of Maryland
College Park, MD 20742

Abstract

We review a model of letter-position encoding, wherein position is tagged by timing of firing relative to an underlying oscillatory cycle. We show how this model can account for data concerning reaction times for lexical decision on rotated letter strings.

Introduction

Little is known of the representations used by the brain for cognitive processing. We suggest that the problem of how letter position within a string is neurally encoded provides a tractable area of investigation into this realm. This problem is circumscribed, yet involves important higher level processes such as the composition of a representation from constituent entities, and the formation of a representation that is independent of absolute location in visual space.

It is commonly assumed that a string's letters are processed in parallel. In contrast, we propose that letters are activated serially. This interpretation is based on the SERIOL framework for letter-position representation (Sequential Encoding Regulated by Inputs to Oscillating Letter units). The serial activation provides a natural way to encode letter order, and accounts for a variety of experimental results (Whitney & Berndt, 1999; Whitney, 2001a; Whitney 2001b).

Here we show how the model can explain the interaction of word length and angle of rotation on reaction time for a lexical decision experiment involving rotated letter strings (Koriat & Norman, 1985). First we recount those experimental results. Next, we review the SERIOL model. We then show, via computational modeling, how the SERIOL framework can account for the experimental data.

Letter string rotation experiment

Subjects were to determine whether or not a letter string was a word. The strings varied in length from two to five letters, and were presented rotated at various angles (in multiples of 20°) from normal horizontal presentation. Each string was rotated as a whole. The researchers found an interesting interaction between string length and rotation angle on reaction times (Koriat & Norman, 1985). The following results are for the word condition. (Non-words evoked a similar pattern.) See Figure 1. For
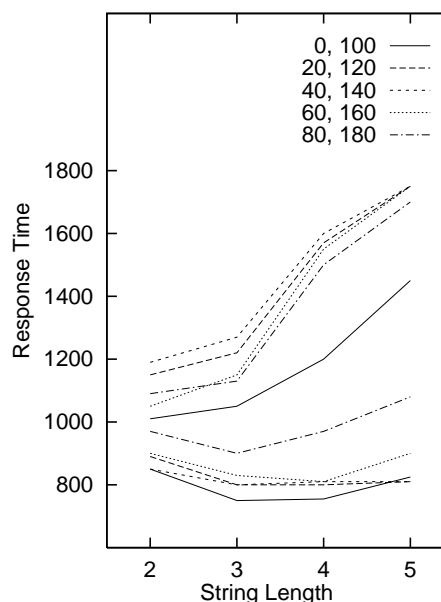

Figure 1: Experimental response times (in milliseconds) for the rotated string lexical decision task. Each line represents one angle of rotation, where the lower lines correspond to 0° through 80°, and the upper lines correspond to 100° to 180°.

angles of 60° or less, reaction time did not increase with string length. For 80°, reaction times were similar for words of two to four letters, while five-letter words incurred an increased reaction time. For 100°, two- and three-letter words had similar reaction times, with reaction times increasing for four- and five-letter words. For angles of 120° to 180°, reaction time varied approximately linearly with word length, with each additional letter adding 200ms.

Because of this non-uniform pattern of interaction between string length and rotation angle, the authors conclude that the results cannot be explained solely in terms of processing time related to mentally rotating the string to the horizontal position. Nor can the data be explained by supposing that processing switches from parallel to serial at some rotation angle, due to the intermediate region (80° and 100°) where reaction times are neither

constant nor linear. In fact the authors state, "it is difficult to propose an interpretation of the results in terms of one unitary principle" (Koriat & Norman, 1985, p. 504). However, the SERIOL framework does offer such an interpretation based on how letter-position is encoded, as discussed below,

The SERIOL model

The SERIOL model is theoretical framework for letter-position coding. It unifies and accounts for a wide range of experimental data on reading, including positional perceptability of letters in strings, visual field differences in letter perceptability, the location of the optimal viewing position and its relationship to reading direction, hemispheric modes of processing, and positional patterns of letter priming within a string (Whitney & Berndt, 1999; Whitney, 2001a; Whitney, 2001b).

The theoretical framework consists of five layers, ranging from the retinal level to the word level. For this discussion, we concentrate on the letter level. The letter level is comprised of computational units (nodes) that represent individual letters. We assume that a feature level (nodes representing letter features) provides input to the letter level. This input is such that the features of the first letter are more highly activated than the features of the second letter; the features of the second letter are more highly activated than those of the third, and so on. That is, feature level activation decreases from left to right across the string. We denote this pattern of activation the *locational gradient*. The proposed mechanisms underlying formation of the locational gradient and experimental evidence for its existence are given in Whitney (2001a) and are not important for this discussion.

We propose that the locational gradient induces a temporal firing pattern across letter nodes wherein position is represented by the precise timing of firing relative to other letter nodes. That is, the first letter fires, then the second letter, then the third, etc. Thus, the temporal sequence of firing encodes the spatial order of the letters. See Figure 2.

This idea is consistent with current neurobiological models of information encoding. Hopfield has proposed that quantities are represented by the explicit timing of action potentials, rather than by their firing rate (Hopfield, 1995). In that model, encoding neurons undergo internal, sub-threshold[1] oscillations of excitability. The magnitude of an input to such a neuron determines when threshold is exceeded. For a small input, threshold is not exceeded until late in the cycle when the cell's oscillation brings its potential near threshold. For a larger input, threshold is exceeded earlier in the cycle. Thus, the size of an input is represented by spike timing relative to the oscillatory cycle. See Figure 3.

It has been suggested that oscillatory activity in the 40 Hz range is related to cognitive processing (Tiitinen, Sinkkonen, Rainikainen, Alho, Lavi-Kainen, & Naatenen, 1993) and that short-term memories are encoded on 40 Hz sub-cycles of a low-frequency (5 Hz) oscillation (Lisman & Idiart, 1995). We propose that a simi-
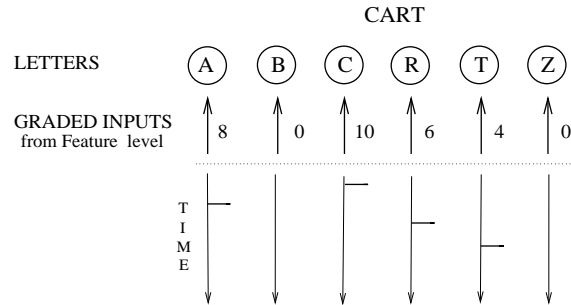

Figure 2: Proposed encoding of letter position; example for the word "cart". A subset of the letter nodes is shown with input levels (in abstract units). The simultaneous graded inputs, via interaction with the dynamics of the letter nodes, create the temporal firing pattern shown in the lower portion of the figure.
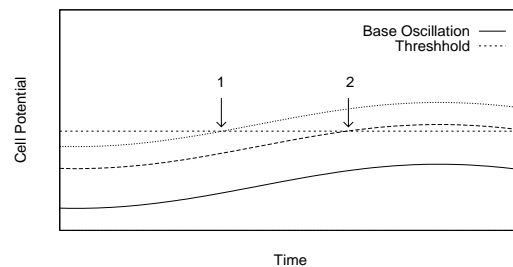

Figure 3: When a relatively large input is added to the base sub-threshold oscillation (top curving line), the cell potential crosses threshold at time 1 (action potential not illustrated). If instead, a smaller input is added, the cell potential crosses threshold later in the cycle, at time 2.

lar scheme underlies letter-position coding, wherein each letter position corresponds to a successive 40 Hz sub-cycle (i.e., lasting 25ms) within a 5 Hz oscillation (i.e., lasting 200ms).

In our model, all letter nodes are assumed to undergo synchronous, periodic oscillations of excitability. Due to the locational gradient, the letter node representing the letter in the first position receives the highest level of excitatory input and fires first, the second receives the next highest amount and fires next, and so on. Suitable input levels and lateral inhibition assure that only one letter node fires at a time. Once a letter node has received lateral inhibition following firing, its input from the feature level is inhibited, so it does not start firing again in the same oscillatory cycle. This coding scheme does not employ position-specific letter detectors; all nodes at the feature level provide input to all letter nodes. Any letter node can represent any position, depending on the level of input that it receives, and the resulting timing of firing. A precise description and simulation of this temporal encoding process is given in (Whitney & Berndt, 1999).

At the next level of the framework, this serial encod-

ing is converted to a non-temporal, contextual encoding (bigrams). A bigram node is tuned to the ordered firing of letter pairs within an oscillatory cycle. For example, the sequence of letters C, A, R, T would activate bigram nodes CA, AR, RT, CR, AT, CT. At the final level of processing, bigram nodes activate word nodes.

Because reaction times for lexical decision for three- to six-letter words do not vary with word length (Fredriksen & Kroll, 1976), it has been assumed that letters are processed in parallel. How then can the SERIOL model be reconciled with this result? We have suggested that a minimal reaction time occurs, based on completion of the oscillatory cycle. That is, all words that can represented within a single cycle have similar reaction times (Whitney, 2001a). A similar suggestion has been made regarding reaction times in the Sternberg task. In that task, a subject is given a list of items, followed by a probe. The reaction times for 'yes' responses to the probe are the same as for 'no' responses, suggesting that subjects always perform an exhaustive search of the mental representation of the list. However, this explanation seems implausible, as it would be more efficient to initiate a 'yes' response as soon as the probed item is encountered in memory. Jensen and Lisman (1998) have proposed that the list of items is encoded serially within an oscillatory cycle, and that the search does indeed stop if the probe item is encountered, but that a motor response can only be initiated at the trough of the oscillatory cycle. This explains why there is no positional effect for the Sternberg task, which seems to require serial checking of the list members. Similarly, we propose that under normal presentation conditions, there is no length effect on lexical decision for words that can be represented in a single oscillatory cycle (i.e. words of seven or fewer letters).

However, we suggest that length effects can arise under conditions of degraded presentation, when input levels to letter nodes are reduced such that it takes multiple oscillatory cycles to represent a sequence of letters that is normally represented in a single cycle. We suggest that such a phenomenon underlies the reaction time results from the rotated word experiment. This analysis implies that such length effects should on the time scale of the oscillatory cycle. Recall that for the largest rotation angles, each additional letter increased reaction times by 200ms, exactly the length of the proposed oscillatory cycle. [1] Thus we propose that the unitary principle which can explain this data is that letter position is encoded temporally. When the input is degraded (by rotating the letter strings) the underlying temporal nature of the encoding is exposed. In the following section, we give the details of this analysis in terms of the SERIOL model.

---

[1]We do not address how words that are too long to be represented in a single cycle are normally represented. For horizontal presentation, there is no evidence for a sharp jump in reaction times as word length increases past seven or eight letters, as might be expected for going from one cycle to two cycles. Presumably a mechanism exists for smoothly integrating information across cycles under normal conditions. That mechanism is beyond the scope of this paper.

A SERIOL account of the rotated word data

In performing the experimental task, we assume that subjects mentally rotated the string to the canonical horizontal orientation, and processed the string as usual. This assumption is consistent with the fact that reaction times for two-letter words were smoothly increasing with rotation angle. We also assume that the act of mental rotation decreases the amount of input reaching the letter nodes, and that this degradation increases with the amount of rotation. These assumptions, in conjuction with the SERIOL model, provide a natural explanation for the experimental data. As the level of input to letter nodes declines, the number of oscillatory cycles required to represent the word increases. Up to a certain amount of rotation, there is still sufficient input to activate all the letters within a single oscillatory cycle (i.e., up to 60°). After that point, there is sufficient input to activate all of the letter in shorter words, while longer strings require an additional cycle (i.e., for 80° and 100°). This accounts for the intermediate region where reaction times are neither constant nor linear. With further degradation, only two-letter words can be represented in a single cycle; each additional letter incurs an additional cycle (i.e., 120° to 180°).

In the SERIOL framework, the serial encoding of letter order is converted to a non-temporal, bigram encoding at the next level of processing. Bigram activation depends on the ordered firing of letter nodes within a single oscillatory cycle. If, as we propose, severely degraded input causes each letter node to fire on a separate cycle, how then could the bigram nodes become activated? We propose that letters which have previously fired can refire again in succeeding cycles. However, this refiring can't be triggered by the feature-level input, since we assume that external input is inhibited once it activates a letter node (so that the letter node will not continually refire.) How then could a previously activated letter node refire?

It has been proposed that an afterdepolarization (ADP) can maintain short-term memory across oscillatory cycles (Lisman & Idiart, 1995). We suggest that this mechanism could maintain the order of letter nodes that have been previously activated when there is insufficient input to activate all the letter nodes within a single cycle. The ADP has been observed in cortical pyramidal cells, and is a slowly increasing excitability that peaks at approximately 200ms after spiking. As such, it could cause refiring on successive oscillatory cycles without external input. The slowly increasing ramp of the ADP can maintain the firing order of elements across oscillatory cycles, as demonstrated by a mathematical model (Lisman and Idiart, 1995).

We have implemented a mathematical model of the reaction time for the rotated word experiment based on these ideas. We modeled the interaction between the underlying oscillatory cycle, the input levels, the lateral inhibition, and the ADP to arrive at an initial firing time for the final letter of the string. This firing time, combined with other quantities, gives the modeled reaction time. The results of the simulation fit to the experimental data
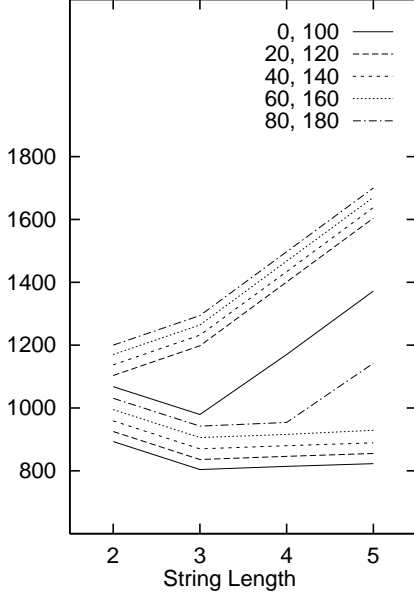
Figure 4: Simulated response times (in milliseconds) for the rotated string lexical decision task. Notation is the same as in Figure 1.

are shown in Figure 4.

Next we present the details of the computational model. Instantiating the theoretical framework in a simulation entails the specification of quite a few parameters. Most of these parameters are related to the neuronal dynamics (ADP, oscillations, and inhibition) and are set to physiologically plausible values, similar to those used by Lisman and Idiart (1995). In fitting the computational model to the experimental data, the primary focus of optimization was the function specifying the inputs from the feature level. This function was hand tuned.

The modeled reaction time, R, is given by:

$$R(\theta, l) = BR + H(\theta) + L(l) + W(\theta, l)$$

where θ denotes the angle of rotation and l denotes the string length.

BR denotes a base reaction time set to 730ms. H denotes the time required to mentally rotate the string; it is a linearly increasing function of θ. Fitting to the reaction times for two-letter words gives:

$$H(\theta) = \theta/20° * 30\text{ms}$$

.

L reflects a length effect that is not explicitly modeled. The experimental data indicate that two-letter words suffer a processing disadvantage. The SERIOL model does not account for this particular length effect. However, assuming constant, extra processing time for two-letter words on top of the explicitly modeled effects gives a good fit to the data. We have chosen to include this increment in the modeled results, in order to facilitate compar-

ison between the simulation and the experimental data. $L(l)$ is 100ms for $l = 2$ and 0 otherwise.

W denotes the time required to activate all the letter nodes corresponding to the string; that is, W is the first time at which the final letter node fires. The functions which determine W are the instantiation of the SERIOL framework. These functions specify the activation of the letter nodes.

Following Lisman and Idiart (1995), letter nodes are modeled as units that undergo a sub-threshold oscillatory drive, exhibit an increase in excitability after firing (ADP), and send lateral inhibitory inputs to each other. We use i to denote the letter node representing the ith letter of the word. The membrane potential, V, of a letter node is given by:

$$V(\theta, i, t) = O(t) + A(i, t) - I(t) + E(\theta, i, t)$$

where O denotes the oscillatory drive, A denotes ADP, I denotes inhibitory input, and E denotes excitatory external input (originating from the feature level). A node fires when V exceeds a threshold, TH. TH is specified relative to resting potential, and set to 10mV. Firing causes the node's ADP component to be reset, and inhibition to be sent to the other nodes.

The oscillatory function O has a cycle length of 200ms, and linearly increases from -5mV to 5mV during the first half of the cycle, and decreases back to -5mV during the second half.

The ADP and inhibition are modeled by functions of the form:

$$F(t; M, T) = M * (t/T)^{1.5} * \exp(1 - t/T)$$

which increases to a maximal value (controlled by M) and then decreases (on a time scale controlled by T). The ADP is given by:

$$A(i, t) = F(t - t_i; M_A, T_A)$$

where t_i denotes the time at which the ith node last fired. ($A(i, t)$ is 0 if the node has not yet fired.) The inhibition is given by:

$$I(i, t) = \sum_{j=1}^{l} F(t - t_j; M_I, T_I).$$

The following values were used: $T_A = 200\text{ms}$, $M_A = 11\text{mV}$, $T_I = 3\text{ms}$, $M_I = 3\text{mV}$. These quantities were hand tuned within a narrow range of values, in conjunction with specification of the external input functions, to give the desired firing pattern.

The external input E is a decreasing function of position i; this corresponds to the locational gradient in the SERIOL framework. The following function was used:

$$E(0°, i, 0\text{ms}) = 10.6\text{mV} - i * 0.5\text{mV}$$

We assume that mental rotation degrades the external input, so E decreases as θ increases. The following function was used:

$$E(\theta, i, 0\text{ms}) = E(\theta - 20°, i, 0\text{ms}) - 0.65\text{mV} * \sin(\theta)$$

Firing inhibits the external input, so $E(\theta, i, t) = 0$ if node i has fired prior to time t. We also assume that the external input builds up over time; if node i has not yet fired, $E(\theta, i)$ increases by 0.2Mv after each oscillatory cycle.

A simulation for each combination of l and θ was run, starting at time $t = 0$ and using a time step of 1ms. At each time step, the potential of each letter node was calculated using the equation for V. The value of $W(\theta, l)$ was taken to be the first t at which $V(\theta, l, t)$ exceeded threshold. The total reaction time was then calculated using the equation for R. For all rotation angles and string lengths, all active letters of the string fired in the correct sequence on each cycle.

For example, for $\theta = 0°$ and $l = 4$, nodes 1, 2, 3, and 4 fired at 49, 63, 74, and 84 milliseconds, respectively, giving $W(0°, 4) = 84$ms. For $\theta = 180°$ and $l = 4$, nodes 1 and 2 fired at 86 and 100 milliseconds. In the second cycle, nodes 1, 2, and 3 fired at 52, 65, and 94 milliseconds (relative to the start of that cycle.) In the third cycle, nodes 1, 2, 3, and 4 fired at 43, 55, 66, and 97 milliseconds, giving $W(180°, 4) = 2 * 200 + 97 = 497$ms. Each node refired earlier in successive cycles due to the ADP. This earlier firing, in conjunction with increasing external input, allowed more letters to fire on each cycle. The slowly increasing ramp of the ADP, along with lateral inhibition, maintained the proper firing sequence across cycles.

Conclusion

Koriat and Norman (1985) were unable to explain their data purely in terms mental rotation. By taking into consideration a possible scheme for the coding of letter position, we have accounted for this data. We propose that the rotation of the stimuli reveals the underlying temporal nature of the encoding. Although the computational model of this experimental data may seem complex, it is keyed on two basic assumptions: letter order is encoded temporally within an oscillatory cycle, and input levels to letter units are reduced for rotated input. The theoretical framework specifying the mechanism that induces this temporal encoding was developed previously; it accounts for a wide range experimental data on the interactions of string position, reading direction, and presentation duration on letter perceptability (Whitney & Berndt, 1999; Whitney, 2000a; Whitney, 2000b). Applying the SERIOL model to the experimental data of Koriat and Norman (1985) yields a natural explanation of those data. It accounts for the finding that there is an intermediate region of rotation angles where processing seems neither fully parallel nor fully serial, and predicts the finding that the increase in reaction time per letter for large rotation angles is 200ms (the proposed length of the oscillatory cycle).

Acknowledgements

References

Frederiksen, J.R., & Kroll, J.F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*: 361-379.

Hopfield, J.J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376: 33-36.

Koriat, A., & Norman, J. (1985). Reading Rotated Words. *Journal of Experimental Psychology: Human Perception and Performance, 11*: 490-508.

Jensen, O., & Lisman, J.E. (1998). An oscillatory short-term memory buffer can account for data on the Sternberg task. *Journal of Neuroscience, 18*: 10686-10699.

Lisman, J.E., & Idiart, M.A.P. (1995). Storage of 7 +- 2 short-term memories in oscillatory subcycles. *Science, 267*: 1512-1515.

Tiitinen, H., Sinkkonen, J., Rainikainen, K., Alho, K., Lavi-kainen, J., & Naatanen, R. (1993). Selective attention enhances the 40-hz response in humans. *Nature, 364*: 59-60.

Whitney, C., & Berndt, R.S. (1999). A new model of letter string encoding: Simulating right neglect dyslexia. *Progress in Brain Research, 121*: 143-163.

Whitney, C. (2001a). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. To appear in *Psychonomic Bulletin & Review.*

Whitney, C. (2001b) Position-specific effects within the SERIOL framework of letter-position encoding. Submitted.