Letter perception: from item-level ERPs to computational models

Arnaud Rey, Stéphane Dufau, Stéphanie Massol, Jonathan Grainger

LETTER PERCEPTION: FROM ITEM-LEVEL ERPS TO COMPUTATIONAL MODELS

Arnaud Rey, Stéphane Dufau, Stéphanie Massol, and Jonathan Grainger

Laboratoire de Psychologie Cognitive – CNRS - Université de Provence
3, place Victor Hugo - case D
13331 - Marseille Cedex 3 - France
arnaud.rey@univ-provence.fr
stephane.dufau@univ-provence.fr
stephanie.massol@univ-provence.fr
jonathan.grainger@univ-provence.fr

ABSTRACT

In the present study, online measures of letter identification were used to test computational models of letter perception. Event-related potentials (ERPs) were recorded to letters and pseudo-letters revealing a transition from feature analysis to letter identification in the 100-200 ms time window. Measures indexing this transition were then computed at the level of individual letters. Simulations with several versions of an interactive-activation model of letter perception were fitted with these item-level ERP measures. The results are in favor of a model of letter perception with feedforward excitatory connections from the feature to the letter levels, lateral inhibition at the letter level, and excitatory feedback from the letter to the feature levels.

KEY WORDS

letter perception, ERP, computational modeling, feedback, lateral inhibition

1. Introduction

Recent MEG and ERP studies on letter perception revealed an occipital activation at 100 ms after stimulus onset that was not sensitive to the specific content of the stimulus and that was interpreted as reflecting low-level visual feature processing [1-3]. Subsequent inferior occipitotemporal activation was found at around 150 ms post-stimulus onset and was interpreted as reflecting the earliest stage of stimulus-specific processing. This result was consistent with several other ERP studies showing that a similar amount of time is needed to begin the identification of a visual target in a natural scene [4-6]. These data suggest that letter identification progressively takes place within a transitory 100-200 ms time window and one can assume that the dynamics of this transition varies across individual letters. In the present study, the ERP properties for individual letters within this time window were used as an online index of letter identification processes.

The first – empirical – goal was therefore to record ERPs for a restricted set of letters that are sufficiently repeated to extract from the ERP signal a stable measure differentiating each letter throughout the 100-200 ms time window. The second – theoretical – goal is to use these variations in ERPs to individual letters in order to test the predictions of different computational models of letter perception. The core representational assumptions of these models are derived from the interactive-activation model of McClelland and Rumelhart (1981) [7]. In this model, a hierarchical organization is assumed with two levels of processing: a feature and a letter level. Localist representations of features and letters are used at each processing level as simplified instantiations of the pattern of activity related to elementary, feature-based visual processes on the one hand, and higher level, item-specific identification processes on the other. These representations are inter-connected by feedforward, feedback and lateral connections, each being characterized by a fixed parameter that determines its weight. By systematically varying these parameter values, we test the predictions of the different computational instantiations of this general architecture against the ERP data.

2. ERP data

As for the ERPs, we first report the comparison between the averaged ERPs obtained for the set of letters and a set of pseudo-letters that were matched on low-level visual dimensions. As shown in Figure 1, the onset of the divergence between the letter and pseudo-letter ERPs was reached by 145 ms replicating previous findings and again suggesting that letter identification processes take place within the 100-200 ms time-window.

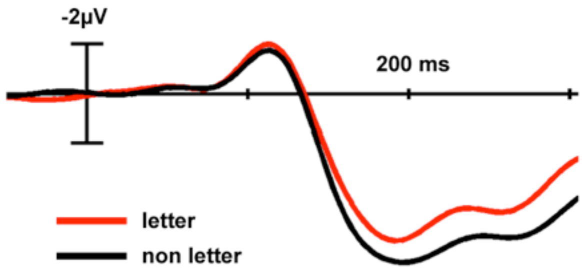

Figure 1. Averaged ERPs for letters and pseudo-letters.

The mean ERP for each letter during the 300 ms that follow stimulus onset was then computed and are displayed in Figure 2. Although the average ERP signals for each letter do display some variation, visual inspection reveals a consistent pattern of activity in the 100-200 ms time-window. This pattern is characterized for all letters by a peak of negativity around 100 ms (i.e., N_{100}) followed by a systematic transition and a positive peak around 200 ms (i.e., P_{200}). Given the result of the grand average letter/pseudo-letter comparison, one can assume that individual letter identification takes place, on average, between these N_{100} and P_{200} peaks of activity. We therefore decided to extract a representative index of this N_{100}/P_{200} transition by measuring the latency and amplitude of the N_{100} and P_{200} peaks for each letter. These measures were then *z*-transformed and used to test several instantiations of the interactive activation model.

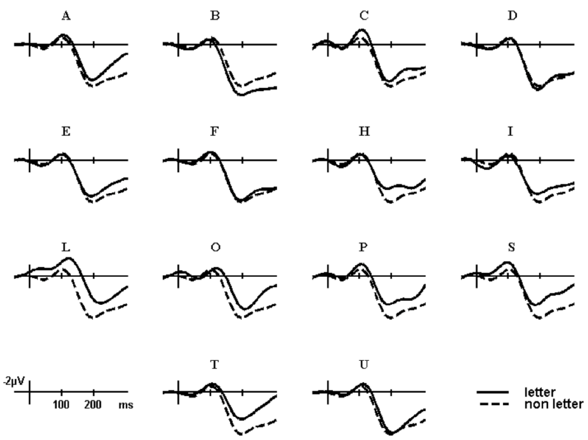

Figure 2. ERPs for each of the 14 target letters used in the experiment. Note that the dashed line corresponds to the average ERP for all non-letters and simply provides a visual baseline.

3. Modeling

By systematically varying the connection parameters of the interactive activation model presented in Figure 3, 9200 version of the model were tested. Each model generated activation curves for each of the target letters as shown in Figure 4. From these curves, activation and latency values were computed and characterized the dynamic of letter identification. More specifically, the activation value for a given letter and a given parameter set corresponded to 80% of the activation of this letter at cycle 50. The latency value was then directly derived from this activation value. These values were *z*-transformed and respectively confronted to the ERP amplitude and latency measure. Root mean square deviations (RMSD) between theoretical and empirical measures were finally calculated.

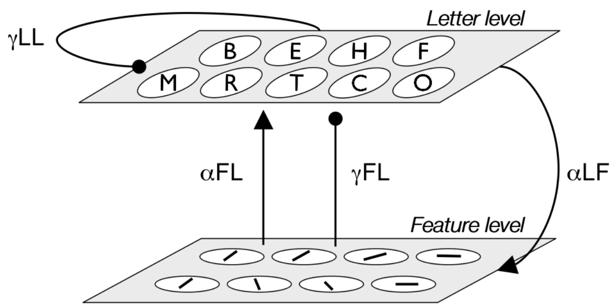

Figure 3. The interactive activation model of letter perception. Feature and letter levels are related by excitatory and inhibitory connections characterized by α and γ parameters, respectively.

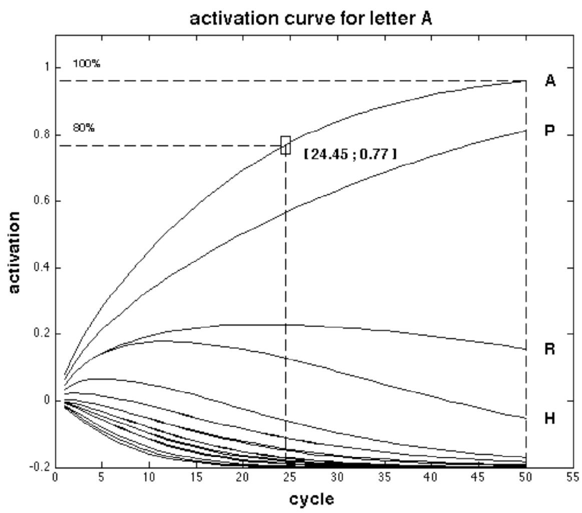

Figure 4. Activation curves obtained with the model for a given set of parameters when letter A was presented.

4. Results

Preliminary results indicate that the lowest RMSDs are obtained both on the latency and the amplitude measures for models with no feature-to-letter inhibition. Significant correlations are obtained both between the predicted letter activations and the ERP letter amplitudes ($r(13)=.55$, $p=.039$) and between the predicted letter cycles and the ERP letter latencies ($r(13)=.52$, $p=.058$). Models having no feedback lead to the worst RMSDs suggesting that feedback from the letter to the feature level is an important factor for the dynamics of the present interactive activation model. Similarly, models with no lateral inhibition outperformed models with no feedback, and this result again suggests that lateral inhibition is probably an important factor but to a lesser extent than the letter-to-feature excitatory feedback.

Overall these results showed that, for some versions of the IA model of letter perception, there were significant correlations between the ERP measures and individual letter identification measures derived from the simulations. This result was, a priori, not obvious because of the integrated nature of the ERP signal. Finding such correlations suggests that computational modeling might be a useful tool for interpreting ERP waveforms and for linking this electrophysiological measure to cognitive mechanisms (for a similar approach, see [8]).

A systematic comparison of the descriptive adequacy of different categories of models showed an advantage of a model composed of feature-to-letter excitatory connections, lateral inhibition, and letter-to-feature feedback, but no feature-to-letter inhibition. This result is globally consistent with recent propositions in the computational neuroscience of pattern recognition, suggesting that the visual system is organized hierarchically with excitatory feedforward connections, lateral inhibition and feedback loops (e.g., [10-12]). Conversely, it is totally inconsistent with the current description of computational models of word reading derived from the original interactive activation model that assume strong feedforward inhibitory connections, no lateral inhibition at the letter level and no feedback from the letter to the feature level (e.g., [13-15]).

4. Conclusion

In conclusion, the present study used item-level measures from ERPs in order to obtain an online index of individual letter perception processes. Because letters are highly frequent visual patterns, they can be intensively repeated during an experiment, yielding a stable item-level ERP signal that closely characterizes individual letter processing. Increasing the set of letters and comparing upper and lower case letters and different letter fonts having different visual properties will certainly provide empirical constrains for further tests of computational models of letter perception. The present study suggests that future directions in computational modeling can certainly benefit from the neurophysiological information available from ERP studies together with studies of the structure and dynamics of the primate visual cortex.

Authors' note

An extended version of this study can be found in Rey, A., Dufau, S., Massol, S., & Grainger, J. (in press). Testing computational models of letter perception with item-level ERPs. *Cognitive Neuropsychology*.

References

[1] J.P. Petit, K.J. Midgley, P.J. Holcomb, & J. Grainger, J., On the time-course of letter perception: A masked priming ERP investigation. *Psychonomic Bulletin & Review*, *13*, 2006, 674-681.

[2] A. Tarkiainen, P. Helenius, P.C. Hansen, P.L. Cornelissen, & R. Salmelin, Dynamics of letter string perception in the human occipitotemporal cortex. *Brain, 122*, 1999, 2119-2131.

[3] A. Tarkiainen, P.L. Cornelissen, & R. Salmelin, R, Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain, 125*, 2002, 1125-1136.

[4] S.J. Thorpe, D. Fize, & C. Marlot, Speed of processing in the human visual system. *Nature, 381*, 1996, 520-522.

[5] H.E. Schendan, G. Ganis, & M. Kutas, Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology, 35*, 1998, 240-251.

[6] R. VanRullen, The power of the feed-forward sweep. *Advances in Cognitive Psychology, 3*, 2007, 167-176.

[7] J.L. McClelland, & D.E. Rumelhart, An interactive activation model of context effects in letter perception: Part 1. An account of Basic Findings. *Psychological Review, 88*, 1981, 375-407.

[8] M. Braun, A.M. Jacobs, A. Hahne, B. Ricker, M. Hofmann, & F. Hutzler, Model-generated lexical activity predicts graded ERP amplitudes in lexical decision. *Brain Research, 1073-1074*, 2006, 431-439.

[9] J. Bullier, Integrated model of visual processing. *Brain Research Reviews, 36*, 2001, 96-107.

[10] S. Dehaene, L. Cohen, M. Sigman, & F. Vinckier, The neural code for written words : a proposal. *Trends in Cognitive Sciences, 9*, 2005, 335-341.

[11] E.T. Rolls, A computational neuroscience approach to consciousness. *Neural Networks, 20*, 2007, 962-982.

[12] D.C. Van Essen, Corticocortical and thalamocortical information flow in the primate visual system. *Progress in Brain Research, 149*, 2005, 173-185.

[13] M. Coltheart, K. Rastle, C. Perry, R. Langdon, & J.C. Ziegler, DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 2001, 204-256.

[14] J. Grainger & A.M. Jacobs, Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 1996, 518-565.

[15] C. Perry, J.C. Ziegler, & M. Zorzi, Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review, 114*, 2007, 273-315.