PSPs and ERPs: Applying the dynamics of post-synaptic potentials to individual units in simulation of temporally extended Event-Related Potential reading data

Sarah Laszlo[a] and Blair C. Armstrong[b]

(a) Department of Psychology, State University of New York, Binghamton
4400 Vestal Parkway East, Binghamton, NY, 13902

(b) Basque Center on Cognition, Brain, and Language
Paseo Mikeletegi 69, Piso 2, San Sebastian, Spain, 20009

Address Correspondence To:
Sarah Laszlo
4400 Vestal Parkway East, Binghamton, NY, 13902
Phone: 607.777.3380
Email: cogneuro@alum.mit.edu

Abstract

The parallel distributed processing (PDP) framework is built on neural-style computation, and is thus well-suited for simulating the neural implementation of cognition. However, relatively little cognitive modeling work has concerned neural measures, instead focusing on behavior. Here, we extend a PDP model of reading-related components in the event-related potential (ERP) to simulation of the N400 repetition effect. We accomplish this by incorporating the dynamics of cortical post-synaptic potentials—the source of the ERP signal—into the model. Simulations demonstrate that application of these dynamics is critical for model elicitation of repetition effects in the time and frequency domains. We conclude that by advancing a neurocomputational understanding of repetition effects, we are able to posit an interpretation of their source that is both explicitly specified and mechanistically different from the well-accepted cognitive one.

Keywords: Parallel Distributed Processing; ERPs; N400; Visual Word Recognition; Repetition Effects; Post-Synaptic Potentials; Neural Computation

1.1 Introduction

Visual word recognition is a temporally extended process requiring integration of information at multiple levels of representation (e.g., orthography, phonology, semantics). Because visual word recognition requires the interaction of information at many levels, attempts to specify its mechanisms have benefitted from the instantiation of computational models, which allow theories about complex interactions between representational levels to be made explicit. In computational models, theories about how representations interact can be implemented formally, and simulations can be conducted to determine whether a particular theory can produce the data it has been formulated to explain. Models from numerous theoretical frameworks have been applied to the problem of visual word recognition, most notably the Parallel Distributed Processing (PDP; Seidenberg & McClelland, 1989, Plaut, McClelland, Seidenberg, & Patterson, 1996; Harm & Seidenberg, 2004), Dual-Route (Coltheart, Rastle, Perry, Langdon, & Zeigler, 2001; Perry, Ziegler, & Zorzi, 2007), and Bayesian (Norris, 2006) frameworks.

PDP, Dual-Route, and Bayesian models differ substantially in their implementation and theoretical stance, varying in details such as how information is represented, the importance of learning, the type and richness of information available to the models, the nature of models' internal computations, and how important it is for those computations to be analogous to neural computations. However, one point on which proponents of most frameworks agree is that word recognition models could benefit from more contact with cognitive neuroscience (see Harm & Seidenberg, 2004; Perry, et al., 2007, Griffiths, Chater, Kemp, Perfors, & Tannenbaum, 2010).

There is good reason for this agreement: typically, computational models of visual word recognition have the goal of simulating behavioral data such as reaction time or accuracy. These are fundamentally *end state* measures, which means they provide information about the outcome of visual word recognition, not its temporally extended internal working. The agreement that models could benefit from more contact with cognitive neuroscience stems from the sense that measures of brain activity collected while visual word recognition is taking place might provide important constraint to models' internal dynamics—and, thereby, a better understanding of visual word recognition's internal workings.

Precisely because of the temporally extended nature of visual word recognition, one method that has been informative in its study is the Event-Related Potential (ERP) technique. Because the neural source of ERPs is excitatory and inhibitory[1] Post-Synaptic Potentials (PSPs) that can be measured essentially instantaneously at the scalp, the temporal resolution of ERPs is on the order of milliseconds; this scale is appropriate for the study of visual word recognition, which is known to unfold within, at most, 500 ms after stimulus onset (Grainger & Holcomb, 2009).

---

[1] Note that IPSPs and EPSPs do not cancel each other out, despite having opposite sign, because when measured distally, at the scalp, EPSPs typically outweigh IPSPs, or, at least, an equal number of IPSPs and EPSPs do not typically occur in synchrony; for review, see Fabiani, Gratton, & Federmeier, 2007.

That ERPs are recognized as an excellent tool for the study of visual word recognition has resulted in a voluminous literature on the topic. Recently, it has been expressed by many authors that this literature would benefit from the theoretical unification that can be achieved by the application of a computational model (Barber & Kutas, 2007; Van Berkum, 2008; Laszlo & Federmeier, 2011; Laszlo & Plaut, 2012). Especially, it has been noted that whereas substantial understanding exists regarding the *functional* causes of ERP effects, there is much less known about the *mechanistic* sources of these effects at the level of neural implementation (Laszlo & Armstrong, 2013). Thus, just as agreement is emerging in the modeling literature that word recognition models could benefit from more contact with cognitive neuroscience, complementary agreement is emerging amongst electrophysiologists that the ERP word recognition literature could benefit from theory building through the use of computational models.

1.2 The ERP Model

In prior work, we began bridging the gap between computation and cognitive electrophysiology through development of the ERP model (Laszlo & Plaut, 2012). The ERP model is heavily based on PDP models that preceded it (e.g., Seidenberg & McClelland, 1989; Plaut et al., 1996; Harm & Seidenberg, 2004); Figure 1 displays the architecture of the model, which, like its predecessors, takes a distributed pattern of orthographic input, and after multiple non-linear transformations in hidden layers, produces a distributed semantic output.

An important difference between the ERP model and its predecessors is that it is required to simulate not only behavior, but also ERP component effects—specifically, effects pertaining to the N400 (a component thought to represent attempted lexical-semantic access; see Kutas & Federmeier, 2011). In order to enable this, the ERP model was given neurally realistic properties not typical in PDP reading models. Primarily, the ERP model's departure from its predecessors comes in its separation of excitation and inhibition. This separation confers several neurally realistic properties, such as more excitatory than inhibitory units (important for simulation of ERPs, where EPSPs are thought to dominate), separate time courses of excitation and inhibition, and fast and slow populations of inhibition. However, the ERP model lacks numerous characteristics of a true cortical system. We theorized that bringing additional neural realism to the model would enable it to simulate more N400 effects, and, further, that providing this realism could provide insight into the neural mechanisms of the simulated effects—an area essentially unexplored in the N400 literature.

The ERP model simulated N400 effects observed in response to unconnected text. Of course, unconnected text is dissimilar to natural reading in that it does not involve context. To extend the ERP model's relevance to realistic reading, therefore, it is important to extend its sensitivity to context. The simplest form of context, and a form that exerts a robust effect on the N400, is immediate repetition of a wordform (e.g., Rugg & Nagy 1987; Rugg, 1985; Rugg, 1990; Nagy & Rugg, 1989). This minimal context requires that processing a wordform, in the simplest possible manner, be dependent on what has come before it. Consequently, providing an explicit mechanistic account of this phenomenon is an important first step in making the bridge between understanding the ERP response to isolated items and understanding the response to items in context.

The N400 repetition effect is characterized as a positivity in response to second presentation of a wordform. The accepted cognitive theory of this effect is that after an initial presentation of a wordform, its semantics do not immediately deactivate; rather, they decay over time (Rugg, 1985). Thus, when an item is repeated, its associated semantics are still active, and, therefore, less elaborate semantic processing is required, resulting in a smaller N400 to repetitions. This theory is widely accepted, and essentially has not been challenged since its formulation (e.g., Rugg, 1990; Besson, Kutas, & Van Petten, 1992; Laszlo & Federmeier, 2007). In extending the neural realism of the model to allow it to simulate N400 repetition effects, we sought to explore whether a neuro-mechanistic explanation would provide novel insight into their generally-assumed cognitive basis.

Additionally, although much of the word recognition literature has focused on time-domain ERPs, there is a growing body of work highlighting insights gained from ERPs in the frequency domain. For instance, recent work has demonstrated that time-domain N400 effects may be due to changes in particular frequency bands and not to changes in the full power spectrum, as would be observed if all neurons in the population generating the N400 modulated their activity to the same degree following a repetition (Roehm, Bornkessel-

Schlesewsky, & Schlesewsky, 2007). Frequency-dependent changes may also have important ramifications for the brain's capacity to synchronize local and distal neural populations that code for a particular representation in a distributed fashion (Weiss & Mueller, 2003; Mellem, Friedman, & Medvedev, 2013). Frequency-domain effects are therefore of theoretical interest, but have been completely unstudied in the computational reading literature. For these reasons, we sought to incorporate frequency-domain analysis into our simulations—to our knowledge, for the first time in this literature.

2.1 The Alpha Model

In the ERP model, mean semantic activation is linked to mean N400 amplitude.[2] Thus, for the model to simulate reduced N400s with repetition, it must display reduced mean semantic activation when repetitions occur. That is, units must have the capacity to become *fatigued.* It is important that this fatigue occur selectively, acting on single units as opposed to the entire semantic layer, because units that have not recently been active must be able to activate to maximum, as when a novel item is presented instead of a repetition.

The desired dynamic of activation for individual semantic units is thus one where a peak of activation (response to a first presentation) is followed by gradual decay, as posited by the cognitive theory of N400 repetition effects and also as necessary to reduce N400 amplitude with repetition. Crucially, this

[2] A voltage determined by distal summation of numerous cortical IPSPs and EPSPs.

dynamic can be formally expressed by the *alpha function;* used in neural computation to simulate PSPs:

$$V = \alpha t e^{\frac{-t}{T}} \qquad (1)$$

Equation 1: The alpha function

In the alpha function as used classically (e.g., Bugmann, 1997), *V* is a measure of membrane potential (voltage), α a scaling constant, *t* the number of time steps since a unit became active, and *T* a free parameter that determines when *V* peaks (e.g., David, et al., 2006). The shape of the alpha function, as defined in Equation 1 and as used in our simulations, is displayed in Figure 1.

That the alpha function is used in simulation of PSPs makes it especially appropriate for use in our model, not only because it produces the desired dynamic, but also because cortical PSPs are the source of the ERP signal (Fabiani, Gratton, and Federmeier, 2007). The result, *V,* of the alpha function represents a voltage, as does the N400, to which semantic unit activations are linked. Indeed, the appropriateness of this function is supported by use of an analogous function in dynamic causal modeling of evoked responses (see Daunizeau, David, & Stephan, 2011), where this type of function has been shown to approximate activation dynamics in actual neurons (David et al., 2006).

Thus, independent observations about the dynamics of the function needed to implement N400 repetition effects, the neural source of those effects, and the computational properties of the alpha function converge to suggest a mechanism for simulation. Therefore, in our attempt to extend the ERP model to simulation of N400 repetition effects, we constrained excitatory unit activations to the envelope of the alpha function (as specified in Equation 1).

In what follows, we will continue to refer to the Laszlo & Plaut (2012) model as the ERP model, but we will refer to the model constrained with the alpha function as the *alpha model*. Application of the alpha function (Equation 1) to excitatory unit activation in the alpha model is the only distinction between the two models[3]. The goal of the simulations presented below was to determine whether a selective fatigue factor, as implemented with the alpha function, constitutes a formally sufficient mechanistic explanation for N400 repetition effects.

2.2 Methods

ERPs. Target data were drawn from the single-item ERP corpus, which has been described in detail elsewhere (Laszlo & Federmeier, 2011). Briefly, EEG was collected from 120 participants who read an unconnected sequence of text including 75 each words (DOG), acronyms (DVD), pseudowords (DOD) and illegal strings (XFQ). Each of these items repeated once.

Model Architecture. The architecture of the alpha model is identical to that used in the ERP model (see Figure 1). Exhaustive details of model implementation, including connectivity, learning algorithms, parameter values,

[3] Note that, because the architecture of the models is in all other ways identical, the alpha model formally retains the ability to simulate any phenomenon the ERP model can simulate.

activation equations, training, and testing, are available in the Supplementary Materials; the simulation code is also available from the first author.

In the ERP model, excitatory units activated according to the sigmoid (see Figure 1). In the alpha model, excitatory units activate as the product of the sigmoid and alpha functions—this is the only difference between models. Importantly, scaling by the alpha function is performed on a *per unit*, rather than *global* basis. That is, *V* is calculated separately for each unit, and in these calculations, *t* is incremented *not* with every time step, *but only* when a unit's activation on the prior time step exceeds a threshold. The result of this is that only units that respond to an input strongly become fatigued. Units that do not respond strongly to an input do not become activated above threshold, and therefore do not become fatigued (see Figure 1). Thus, if repetition effects are observed in the model, they are due *not* do a global fatigue mechanism that ensures that activations decrease over time, but rather to selective fatigue of *only* units that are activated in response to a given input, despite the model never being trained that this is the appropriate dynamic (see below).

Training. On each trial, one orthographic input was clamped on, and activation was allowed to propagate through the network. The activation in semantics was then compared to the target semantic activation associated with that orthographic input to determine how the weights should be adjusted to improve future performance using backpropagation (Rumelhart, Hinton, & Williams, 1986), with the constraint that the excitatory connection weights had a lower bound of 0. Training items consisted of 62 words and 15 acronyms (details of all representations and of the modified backpropagation procedure are available in the Supplementary Materials). Importantly, the network's activation reset after each item during training; *the model received no training on repetitions*. The model's output dynamics in response to repeated items must therefore be an *emergent* characteristic of its architecture, *not* the result of training on the desired response to repetitions. After training, the model's weights were fixed and not modified during testing (for a complementary approach, see Oppenheim, Dell, & Schwartz, 2010).

Testing. After training, the network was presented with input pairs of the form AA (repetitions) or AB (non-repetitions). A single time step of blank input intervened between items in each pair. In testing, the network was *not* re-initialized between items in a pair.

In addition to being tested on the trained items, the network was tested on repetitions and non-repetitions of pseudowords and illegal strings. The nonwords provide a difficult test for the model, because they were not trained. Thus, when presented with nonwords at test, the model must produce repetition dynamics it has never been trained on, in response to items it has never been exposed to—and, in the case of frequency-domain analysis, in a domain it was never originally designed to simulate.

A control simulation, performed in order to assess whether application of the alpha function is *necessary* for simulation of repetition effects, was also conducted using the original ERP model—in this simulation, all methods were identical to those described above, but the alpha function was not applied.

2.3 Results

2.3.1 ERPs

Grand-averaged ERPs (Figure 2) were computed over the middle parietal electrode for each item type on first and second presentation. Data were trimmed to the time-window corresponding to N400 effects in this data set: 250-450 ms (Laszlo & Federmeier, 2011). To maximize the consistency of ERP and simulation analyses, these data were again trimmed to a statistically-defined window of interest, the full width at half-maximum (FWHM) of the N400 window.

Time Domain. Repetition effects were assessed in the time domain by analyzing mean amplitudes for each item type at each level of repetition using Linear Mixed Effects Regression (LMER; Baayen, 2008), with item as a random factor and item type as a fixed factor. P-values were generated via Markov-Chain Monte Carlo (MCMC) sampling. Analyses replicated the standard finding: N400 mean amplitudes were reduced for all item types on second presentation (all $p_{MCMC} < .0002$).

Frequency Domain. Data were also analyzed in the frequency domain—this provides additional constraint on the model—especially since the model was not trained on any aspect of the appropriate frequency response—as it requires that simulated ERPs (sERPs) have not only the same qualitative properties as ERPs, but also similar waveshapes. For these analyses, the fast Fourier transform (FFT) was applied to averaged, unfiltered ERPs corresponding to the first and second presentation of each item type (Figure 2). As is evident in Figure 2, repetitions are characterized in the frequency domain as an enhancement of low frequency (2-4 Hz; alpha band) power for all item types (all pMCMC < .0003).

2.3.2 The Alpha Model

Model sERPs were generated by averaging semantic activation on each time step for the second item in each AA or AB pair (Figure 2). As is evident from the Figure, sN400 amplitudes were reduced for repetitions of each item type. For quantification of these effects, the sN400 was defined as the temporal region that corresponded to the FWHM of the simulated waveform.

Time Domain. As in the ERPs, each lexical type showed a reduction in sN400 amplitude with repetition (Words: pMCMC < 0.0001; Pseudowords: pMCMC = 0.0001; Illegal strings: p < 0.0001; Acronyms: pMCMC = 0.0364). Analyses were conducted on identical portions of the data for all item types (i.e., word and nonword repetition effects were not quantified differently)[4]. No reliable repetition effects were observed, for any item type, in the control simulation.

Frequency Domain. Model power-frequency curves for each item type are presented in Figure 2. Note that although the analogous frequency data were

[4] For details on how the model performs lexical decisions (i.e., tells apart words from nonwords), see Laszlo & Plaut, 2012.

extracted for the model, this was done using an alternative but formally identical mathematical technique made possible by having direct access to the individual unit activations (see the technical appendix for details). Note the similarity between simulation and empirical data, despite the fact that the model was never trained in the frequency domain. Moreover, note that there is an increase in low-frequency power in model spectra, as in the ERPs. In the lowest model frequency band, all lexical types show a numerical increase in power on second presentation although the effect was not significant for acronyms (which consist of only 15 items and therefore suffer from the lowest statistical power; less conservative tests reveal significant effects for acronyms as well). Because model units are arbitrary, we are not able to specifically link this result quantitatively with the 2-4 Hz (alpha) band in the ERPs; this is why we phrase analysis of the model data in terms of its “lowest” frequency band; emphasizing the qualitative similarity of model and ERPs even though no specific quantitative link is possible at this stage of the research. Again, no reliable repetition effects were observed for any item type in the control simulation.

4.1 Discussion

The goal of the present simulations was to advance a neuro-mechanistic account of a cognitively well-understood phenomenon: the N400 repetition effect. We worked under the assumption that one way to improve the cognitive power of the ERP model would be to improve its neural realism. To this end, we considered several sources of information: the cognitive theory of N400 repetition effects (priming due to slow activation decay), the desired empirical dynamic of N400 repetition effects (reduced N400s with repetition), and the computational properties of the neural system underlying the ERP signal. Together, these considerations suggested that by applying a an approximation of the shape of PSPs to the ERP model, we could produce the desired empirical effects.

Thus, we created the alpha model. The alpha model was able to simulate N400 repetition effects in both the time and frequency domains. In the frequency domain, it was able to do so despite never receiving training on 1) repetitions, 2) nonwords, or 3) the desired frequency response, and not having been originally designed to study the frequency domain. This success in simulating phenomena far afield from those it was trained on (single presentations of words and acronyms in the time domain only) suggests that core characteristics of the model's architecture are deeply similar to characteristics of the analogous neural system. A control simulation, without the alpha dynamic, did not simulate any repetition effects. This suggests that it was, specifically, application of the alpha function that enabled the alpha model's success.

Analysis of the model in the frequency domain revealed not only that it provided a good explanation of the empirical data, but also emphasized the importance of low frequency modulation (< 5 Hz) in driving repetition effects. Interestingly, these findings are roughly consistent with other studies employing simple semantic priming manipulations, which have found changes at approximately 8 Hz (Kujala et al., 2011), but differ from studies in semantic paradigms employing richer stimuli—such as sentences—which have typically reported higher-frequency changes in coherence (> 30 Hz), as opposed to spectral power (see Mellem et al., 2013). One clear explanatory hypothesis for this distinction is that the frequency response of neurons involved in semantic access is modulated by the degree of context available—additional work, such as that proposed for investigation of semantic priming with the alpha model discussed below, will be needed to explore this possibility.

The canonical view of N400 repetition effects suggests that, when an item is repeated, it benefits from priming of its semantic features, meaning that less semantic processing is required on second presentation, leading to reduced N400s (e.g., Rugg, 1985; for review, see Kutas & Federmeier, 2011). In the model, in contrast, activation of semantic features by a first presentation is responsible for reduced sN400s in repetition, but the mechanism of this reduction is neural fatigue, not a mechanistically unspecified decreased need for semantic processing. That is, in the model, it is not that semantic features do not need to be activated as much on second presentation because they are still partially active, it is that they *cannot* activate as much on second presentation due to fatigue. In producing this insight, the model extends its success beyond simply reproducing numerical patterns, to making a novel contribution to what is understood about the mechanism of a phenomenon that has been cognitively understood for nearly 30 years.

As a mechanism, the alpha dynamic is domain general: that is, it does not apply only to single repetition priming. For example, *semantic satiation* is the extreme case where multiple repetitions of an item cause that item to seem to lose all meaning (for review of this phenomenon and its implications, see Amster, 1964). In the model, this phenomenon is, in principle, transparently explained as being caused by asymptotic fatigue of semantic features (see Figure 1).

As another example, *semantic priming* effects on the N400 are well-characterized as reductions in amplitude to targets that are semantically related to a prime (e.g., a smaller N400 to CAT in DOG-CAT than to FORK in DOG-FORK; see Kutas & Hillyard, 1989); these effects are typically similar to, but smaller than, N400 repetition effects. In the model, semantic primes would essentially be treated as “partial” repetition primes: *some* but not *all* of the target semantic features would already have been activated (and thus fatigued) by the prime; thus, semantic priming in the model would be similar to, but smaller in amplitude than sN400 repetition effects, as in the empirical data. Similarly, the increase in low frequency power would be expected to be smaller (and possibly occur at slightly higher frequencies, consistent with Kujala et al., 2011) because only a partially overlapping semantic code would be engaged by the prime and consequently be fatigued.

5.1 Conclusion

The success of the present model in capturing a range of novel effects via a simple neural fatigue dynamic suggests that fatigue is a powerful mechanism for understanding multi-word integration, as exemplified here by single-word repetition. The use of a domain-general, biologically-plausible variation of the PDP framework provides a basis for extending the model to more nuanced simulation of language electrophysiology—such as that observed in semantic satiation and semantic priming. The model presented here certainly is not a complete characterization of all the neural computations that occur during a complex cognitive event, like word recognition in context (e.g., it does not consider spike trains)—more work will be needed to even further improve the neural realism of the model. However, it 1) is more neurally realistic than any reading model that has preceded it and 2) demonstrates the general principle that improving the neural realism of reading models empowers them to simulate a broader range of effects.

Acknowledgments

The authors acknowledge M. Monk, K.D Federmeier, and members of Binghamton University Modeling Meeting—especially K. Kurtz—for discussion. This research was supported by NSF-CAREER BCS-1252975 (S.L.), and by Marie Curie IIF 627784 (B.C.A.).

Figure Captions

Figure 1. (A) Architecture of the ERP model. INH stands for “inhibitory”. (B) The shape of the sigmoid function (inset), and of the alpha function above and below threshold. Note that for alpha units, as $t \rightarrow \infty$, $V \rightarrow \Theta$.

Figure 2. ERP and model (sERP) data in the time and frequency domains. Time-domain ERP data consists of grand-averaged responses to first and second presentations of words, acronyms, pseudowords, and illegal strings, over the middle parietal electrode site; the same data is presented in the frequency domain. Time-domain sERP data consists of responses, averaged over all semantic units, to first and second presentations of the same item types. The same data is presented in the frequency domain.

References

Amster, H. (1964). Semantic satiation and generation: Learning? Adaptation?.*Psychological Bulletin*, *62*(4), 273.

Baayen, R.H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press.

Barber, H.A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. Brain Research Reviews, 53, 98-123.

Besson, M., Kutas, M., & Petten, C.V. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences.*Journal of Cognitive Neuroscience*, *4*, 132-149.

Bugmann, G. (1997). Biologically plausible neural computation. *Biosystems,* 40, 11-19.

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.

Dauizeau, J., David, O., & Stephan, K.E. (2011). Dynamic causal modeling: A critical review of the biophysical and statistical foundations. *Neuroimage, 58*, 312-322.

David, O., Kiebel, S.J., Harrison, L.M., Mattout, J., Kilner, J.M., & Friston, K.J. (2006). Dynamic Causal Modeling of Evoked Responses in EEG and MEG. Neuroimage, 30, 1255-72.

Fabiani, M. Gratton, G, and Federmeier, K.D. (2007). Event-related brain potentials: Methods, theory, and application. In J.T. Cacioppo, L. Tassinary, and G. Berntson (Eds.), Handbook of Psychophysiology, 3rd edition (pp.85-119). Cambridge:Cambridge University Press.

Grainger, J., & Holcomb, P.J. (2009). Watching the Word Go by: On the Time-course of Component Processing in Visual Word Recognition. *Language and Linguistics Compass,* 3, 128-156.

Griffiths, T.L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J.B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences,* 14, 357-364.

Harm. M.W., & Seidenberg, M.S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review,* 111, 662-720.

Kujala, J., Vartiainen, J., Laaksonen, H., & Salmelin, R. (2011). Neural interactions at the core of phonological and semantic priming of written words. *Cerebral Cortex, 22*(10), 2305-2312.

Kutas, M., & Hillyard, S.A. (1989). An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience*, *1*(1), 38-49.

Kutas, M., & Federmeier, K.D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology.

Laszlo, S., & Federmeier, K.D. (2007). Better the DVL you know: Acronyms reveal the contribution of familiarity to single word reading, *Psychological Science,* 18, 122-126.

Laszlo, S., & Federmeier, K.D. (2011). The N400 as a snapshot of interactive processing: evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology,* 48, 176-186.

Laszlo, S., & Plaut, D.C. (2012). A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain and Language*, 120, 271-281.

Laszlo, S.,& Armstrong, B.C. (2013). Applying the dynamics of post-synaptic potentials to individual units in simulation of temporally extended ERP reading data. *Proceedings of the 35th Annual Conference of the Cognitive Science Society.*

Nagy, M.E., & Rugg, M.D. (1989). Modulation of Event-Related Potentials by Word Repetition: The Effects of Inter-Item Lag. *Psychophysiology*, *26*(4), 431-436.

Norris, D. (2006). The Bayesian Reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*(2), 327.

Oppenheim, G.M., Dell, G.S., & Schwartz, M.F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227-252.

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

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.

Roehm, D., Bornkessel-Schlesewsky, I., & Schlesewsky, I. (2007). The internal structure of the N400: Frequency characteristics of a language related ERP component. *Chaos and Complexity Letters, 2*, 365-395.

Rugg, M.D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology,* 22, 642-647.

Rugg, M.D., & Nagy, M.E. (1987). Lexical contribution ot nonword repetition effects: Evidence from event-related potentials. *Memory & Cognition,* 15, 473-481.

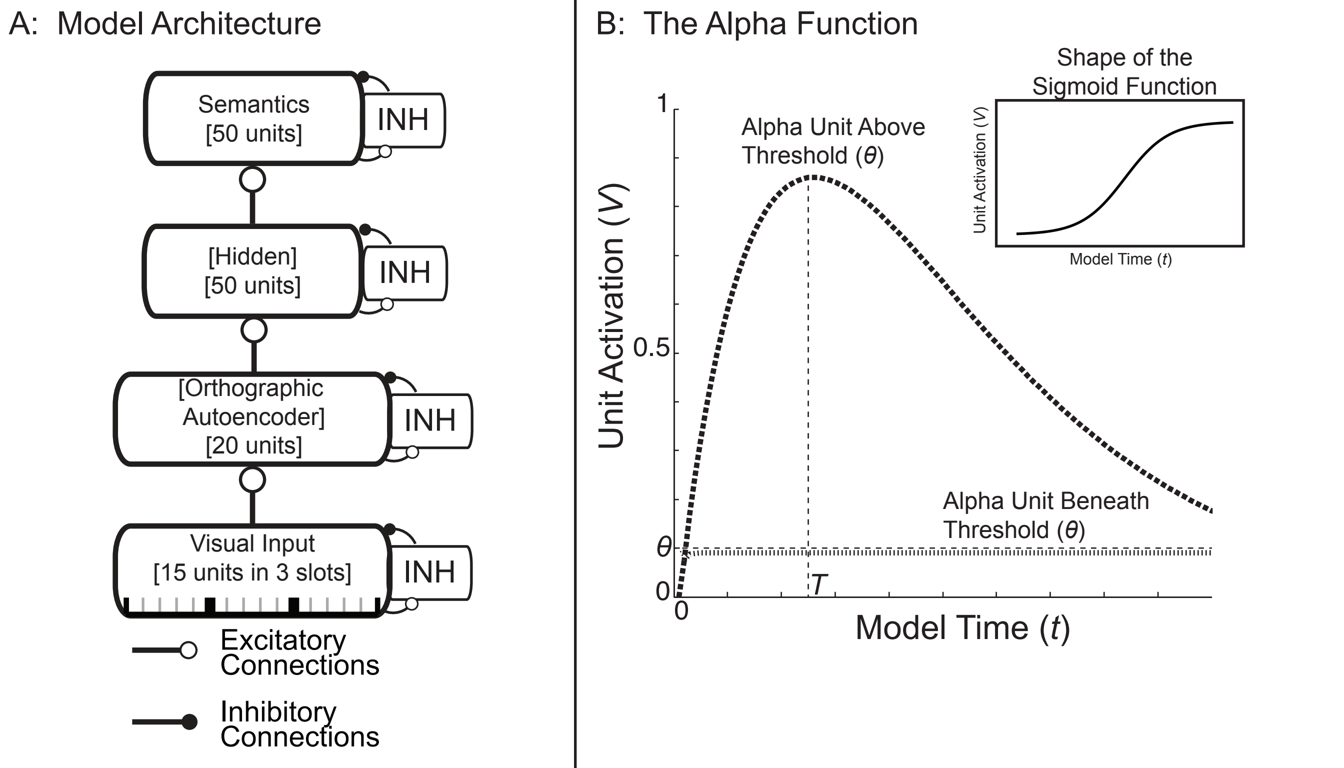
Rugg, M.D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory and Cognition,* 18, 367-379.

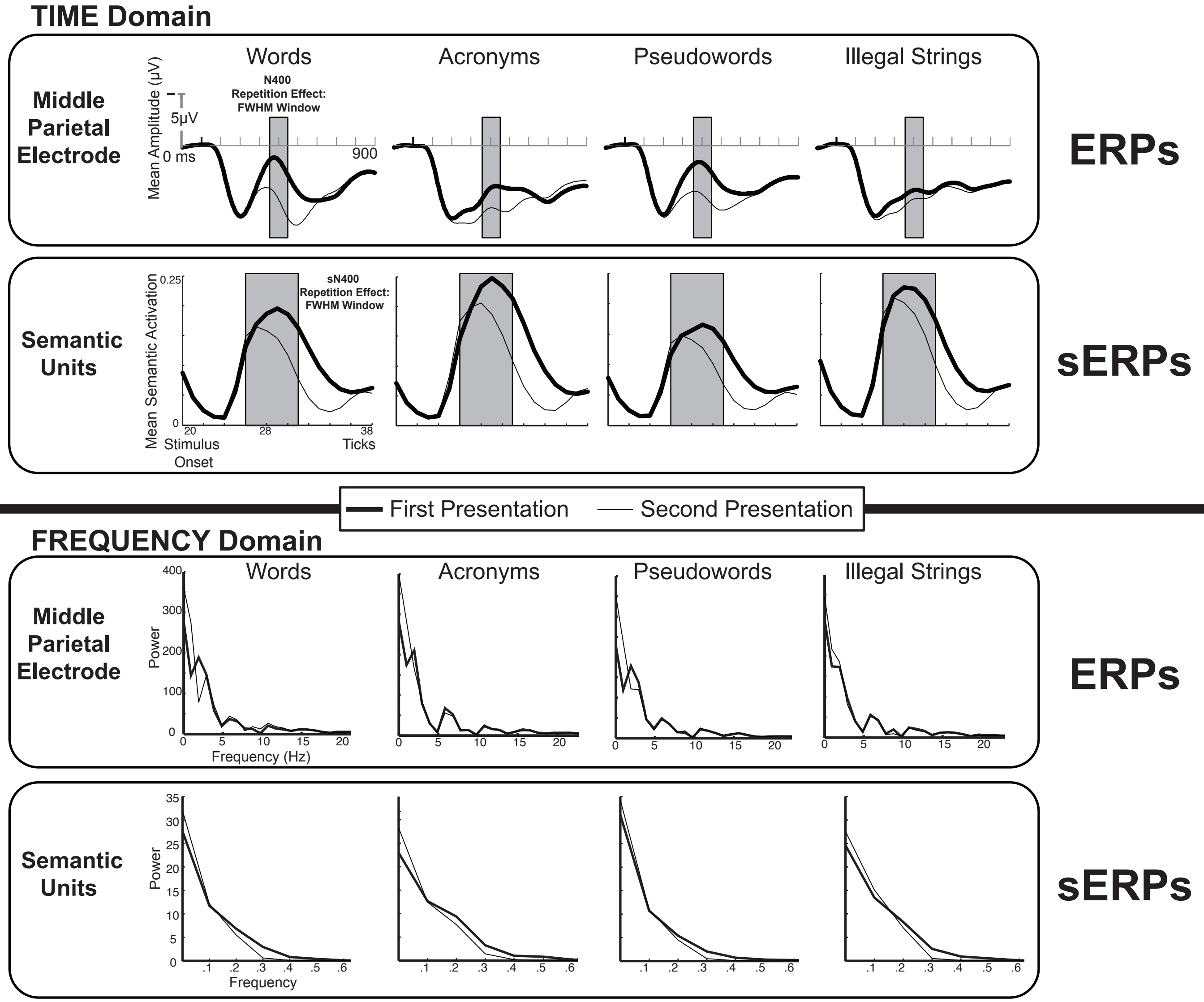
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533-536.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review,* 96, 523-568.

VanBerkum, J.J.A., (2008). Understanding Sentences in Context: What Brain Waves Can Tell Us. *Current Directions in Psychological Science,*17, 376-380.

Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. *Brain and Language, 85*, 325-343.


A: Model Architecture
Semantics
[50 units]
INH
[Hidden]
[50 units]
INH
[Orthographic
Autoencoder]
[20 units]
INH
Visual Input
[15 units in 3 slots]
INH
Excitatory
Connections
Inhibitory
Connections
B: The Alpha Function
Shape of the
Sigmoid Function
Unit Activation (V)
Model Time (t)
1
0.5
θ
0
0
T
Alpha Unit Above
Threshold (θ)
Alpha Unit Beneath
Threshold (θ)
Unit Activation (V)
Model Time (t)



TIME Domain
Words
Acronyms
Pseudowords
Illegal Strings
Middle Parietal Electrode
Mean Amplitude (µV)
5µV
0 ms
900
N400 Repetition Effect: FWHM Window
ERPs
Semantic Units
Mean Semantic Activation
0.25
0
20
28
38
Stimulus Onset
Ticks
sN400 Repetition Effect: FWHM Window
sERPs
First Presentation
Second Presentation
FREQUENCY Domain
Words
Acronyms
Pseudowords
Illegal Strings
Middle Parietal Electrode
Power
400
300
200
100
0
0
5
10
15
20
Frequency (Hz)
ERPs
Semantic Units
Power
35
30
25
20
15
10
5
0
.1
.2
.3
.4
.5
.6
Frequency
sERPs