Reconciling opposite neighborhood frequency effects in lexical decision: Evidence from a novel probabilistic model of visual word recognition

Thierry Phénix (thierry.phenix@univ-grenoble-alpes.fr)
Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

Sylviane Valdois (sylviane.valdois@univ-grenoble-alpes.fr)
Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

Julien Diard (julien.diard@univ-grenoble-alpes.fr)
Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

Abstract

A new Bayesian model of visual word recognition is used to simulate neighborhood frequency effects in lexical decision. These effects have been reported as being either facilitatory or inhibitory in behavioral experiments. Our model manages to simulate the apparently contradictory findings. Indeed, studying the dynamic time course of information accumulation in the model shows that effects are facilitatory early, and become inhibitory at later stages. The model provides new insights on the mechanisms at play and their dynamics, leading to better understand the experimental conditions that should yield a facilitatory or an inhibitory neighborhood frequency effect.

Keywords: Visual word recognition; Lexical decision; Visual attention; Bayesian algorithmic modeling.

Introduction

Identifying the processes involved in the recognition of printed words and their time course is critical for word recognition models. The lexical decision task is commonly used in behavioral experiments to investigate this issue. In this task, participants have to indicate whether a printed stimulus is a real word (YES response) or not. A standard finding is that high frequency words are processed faster than low frequency words. The frequency effect on response latency is massive for high frequency words but for medium and low frequency words, other effects further modulate performance. Among these effects, a poorly understood and highly controversial effect is the neighborhood frequency effect. Latency on YES responses in lexical decision is modulated by the existence of higher frequency orthographic neighbors, namely the existence of at least one word of higher frequency that differs from the target word by a single letter (e.g., *LIKE* is a higher frequency neighbor of the target word *BIKE*).

Some empirical studies report an inhibitory effect (Carreiras, Perea, & Grainger, 1997; Forster & Shen, 1996; Grainger, 1990; Grainger, O'Regan, Jacobs, & Segui, 1989; Grainger & Segui, 1990; Perea & Pollatsek, 1998) while others report a facilitatory effect (Sears, Hino, & Lupker, 1995; Siakaluk, Sears, & Lupker, 2002), that is to say, respectively lower vs. higher latency for low frequency words with higher frequency neighborhood. Two models of visual word recognition effectively addressed the neighborhood frequency effect, but none of them could account for conflicting experimental results and provide explanations to accommodate these results. The MROM model (Grainger & Jacobs, 1996) is sensitive to the presence of a high frequency neighbor which produces a strong inhibitory effect on visual word recognition. On the other hand, in the Bayesian reader (Norris, 2006), word recognition is facilitated by neighborhood density regardless of their frequencies.

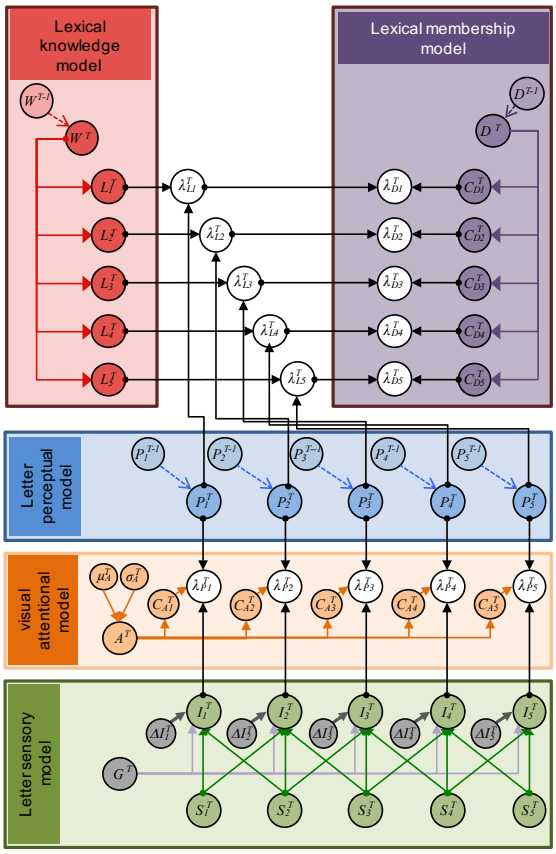

Figure 1: **Graphical representation of the structure of the BRAID model.** Modules are represented as colored blocks, and group together variables of the model (nodes). The dependency structure of the model is represented by arrows.

In this paper, we use a new model of visual word recognition and lexical decision, called BRAID (for "Bayesian word Recognition with Attention, Interference and Dynamics", illustrated Figure 1). It is a structured and hierarchi-

cal Bayesian model that includes, on top of a classical three-layer architecture, an additional attention layer, an interference mechanism between adjacent letters, and dynamic models of the temporal evolution of its variables. The model was previously defined and calibrated, and simulations carried out to account for a large number of classical effects, including the frequency effect, the word superiority effect, and the optimal viewing position effect in word recognition (Phénix, 2018). It was also used to simulate word length and letter transposition effects in lexical decision (Ginestet, Phénix, Diard, & Valdois, submitted). Here, the model is used to account for reference studies of the neighborhood frequency effect in lexical decision.

The rest of this paper is structured as follows. Since a complete mathematical description of BRAID is beyond the scope of the current paper, we first describe the main features of the model, detailing its general structure and providing a short description of its main modules. We then show how Bayesian inference solves the lexical decision task in the model. We then present reference experiments about the neighborhood frequency effects that we aim to replicate, and compare our simulation results with the behavioral findings using the same set of stimuli.

Summary of the BRAID model

The general structure of the BRAID model is illustrated in Figure 1. First, we assume that word recognition relies on three levels of processing as featured in most previous word recognition models. The first level, called the letter sensory sub-model, implements low-level visual mechanisms involved in letter identification and letter position coding (Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger, Dufau, & Ziegler, 2016). Feature extraction is parallel over the input string but the model further implements an acuity gradient, so that lesser information is extracted from letters as distance from fixation increases. As in Whitney (2001), acuity is symmetric around fixation and by default, gaze position is located at word center. Following Gomez, Ratcliff, and Perea (2008) and Davis (2010), location is distributed in the sense that information about the features of one letter extends into adjacent letter positions. This mechanism implements lateral interference between letters. The second level, called the letter perceptual sub-model, implements how information extracted from the sensory input accumulates gradually over time to create a percept, i.e. an internal representation of the input letter string. The third level, called the lexical knowledge sub-model, implements knowledge about the 40.481 English words of the British Lexicon Project (Keuleers, Lacey, Rastle, & Brysbaert, 2012). The probability to recognize a word is modulated by its frequency.

One major originality of BRAID is to assume the existence of a fourth level, called the visual attentional sub-model, that implements an attentional filtering mechanism between the letter sensory sub-model and the letter perceptual sub-model. The transfer of information between these two submodels is modulated by the amount of attention allocated to each letter position. The distribution of attention over the letter string is Gaussian. The peak of the distribution is aligned on gaze position, so that more information on letter identity is transferred for creation of the letter percept at this position. The further the letter from the attention peak, the less attention it receives and hence, the less information is transferred.

BRAID further includes a fifth level, called the lexical membership sub-model, that implements a mechanism to decide whether or not the input letter-string is a known word. This submodel is critical to simulate lexical decision. It can be viewed as an "error model". Assuming the sensory input is a word, all the letter percepts should match an existing word of the lexical knowledge submodel. If the input is not a word, matching should fail on at least one position.

Information propagates dynamically from sensory input, through perceptual representation, to the lexical submodels. Technically, this propagation of information is carried out using "coherence variables" (λ variables in Figure 1, white nodes), which can be interpreted as information switches (Bessière, Mazer, Ahuactzin, & Mekhnacha, 2013). Depending on their state (open/closed or unspecified), the coherence variables allow or do not allow propagation of information between adjacent submodels. In that sense, they allow to connect or disconnect portions of the model. Propagation is bidirectional between the lexical knowledge and the letter perceptual submodel.

Task simulation by Bayesian inference in BRAID

Mathematically, the BRAID model is defined by a joint probability distribution over the high-dimensional state space over all variables appearing in Figure 1. Then, using Bayesian inference, that is to say, applying the rules of probabilistic calculus in the model, mathematical expressions corresponding to letter recognition, word recognition and lexical decision are automatically obtained. Mathematical definitions and derivations cannot be provided here in full, due to lack of space. Instead, we will describe how the model simulates letter recognition, word recognition and lexical decision. The lexical decision task is the task of interest here, but since it involves the two previous ones, we describe them in their nesting order, for clarity purpose.

Letter identification The first task we consider is letter identification, that is, the process of sensory evidence accumulation, from a given stimulus, to perceptual letter identity. This is modeled by the term:

$$Q_{P_n}^{T} = P(P_n^T \mid s_{1:N}^{1:T} \, [\lambda_{P_n}{}^{1:T} = 1] \, \mu_A^{1:T} \, \sigma_A^{1:T} \, g^{1:T}) \, . \qquad (1)$$

We note with uppercase letters probabilistic variables, that is, sets of possible values, and with lowercase letters, specific values (e.g., $s_{1:N}^{1:T}$ is a shorthand for $[S_{1:N}^{1:T} = s_{1:N}^{1:T}]$). $Q_{P_n}^{T}$ is the probability distribution over the letter of interest (P_n^T, at time T and position n, a discrete variable with 26 possible values), given a presented letter stimulus ($s_{1:N}^{1:T}$), given attention distribution (characterized by $\mu_A^{1:T}, \sigma_A^{1:T}$) and gaze posi-

tion ($g^{1:T}$), and given that information propagates through the model from the sensory level to the letter perceptual model only at position n (the only "closed" switch is $\lambda_{P_n}^{1:T}$), and not beyond.

Computing $Q_{P_n}^T$ involves two main components. The first component is the dynamical evolution of knowledge about the perceived letter identity, in which the knowledge about letters at previous time step is combined with an information decay term such that, in the absence of stimulus, the probability distribution over letters decays towards its initial state, that is, a uniform distribution, representing lack of information.

The second component describes sensory evidence accumulation. It relies on the extraction of information from stimulus $s_{1:N}^T$ performed by the letter sensory model. Details are not provided here, but this term features effects of interference from neighbor stimuli, if any, and loss of performance, due to the acuity gradient, when gaze position is not located on the letter under process.

Propagation of sensory information from stimulus to letter percepts is modulated by the visual attention submodel, and, more precisely, by attention allocation. Attention affects the balance between the two components, respectively information decay and sensory evidence accumulation. In other words, at spatial positions that receive sufficient attentional resources to counterbalance information decay, sensory evidence propagates efficiently from the sensory model to the perceptual model, and the probability distribution over letters acquires information. Over time, the probability distribution $Q_{P_n}^T$ converges, so that its maximum probability designates the letter recognized from the stimulus (which, provided enough attention, and except for pathological cases, is the correct letter).

Word identification In a similar fashion as in isolated letter recognition above, we model word recognition by computing:

$$Q_W{}^T = P(W^T \mid e^{1:T}\ [\lambda_{L1:N}^{1:T} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T})\ , \qquad (2)$$

with $e^t = s_{1:N}^t\ [\lambda_{P1:N}^t = 1]\ g^t$. That is, we compute the probability distribution over words W^T, given the same stimulus, gaze and attention characteristics as when computing $Q_{P_n}^T$, but we allow information to propagate further in the model, to the lexical knowledge model, by setting $[\lambda_{L1:N}^{1:T} = 1]$.

Once more, a classical "dynamical system simulation / perceptual evidence simulation" structure is obtained. First, information about words gradually decays, in the absence of sensory information, towards its initial state, which, here, represents word frequency. Second, sensory evidence accumulation is based on the probabilistic comparison between a letter sequence memorized in lexical knowledge and the perceived letter identity, as computed by letter recognition $Q_{P_n}^T$. This comparison, in BRAID, is influenced by the similarity between the letters of the stimulus and all words of the lexicon, so that similar (neighbor) words compete with each other for recognition.

Lexical decision The final task we describe here is our task of interest, lexical decision. It is modeled by computing:

$$Q_D{}^T = P(D^T \mid e^{1:T}\ [\lambda_{D1:N}^{1:T} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T}) \qquad (3)$$

It is a variant of previous questions: here, we allow information to propagate throughout the whole model, by setting $[\lambda_{D1:N}^{1:T} = 1]$, and the target variable is D^T, which is a Boolean variable that represents lexical membership (i.e., it is *true* when the stimulus is perceived to be a known word).

Here, Bayesian inference is more complex, and is best explained by considering, in turn, the two Boolean cases that compete at each time step. First, consider the hypothesis that the stimulus indeed is a word (the $D^T = true$ case): as before, a dynamical decay of stored information (toward an initial state which is a uniform probability distribution) competes with sensory evidence accumulation. Sensory evidence, here, is the whole process of word recognition. In other words, lexical decision proceeds by accumulating evidence from the observation of the lexical knowledge submodel, so that when a word is reliably identified from the stimulus, or when a set of orthographically similar words is activated enough, then the probability that $D^T = true$ is high.

Consider now the hypothesis that the stimulus is not a word (the $D^T = false$ case): this is a variant of the previous case where accumulating evidence from the probability distribution over words relies under the assumption that there would be one error in the stimulus, compared to known word forms. All possible positions for this error are enumerated, and, for a given error position, word recognition proceeds with the input stimulus in all other positions, and alternative letters in this position. In other words, the stimulus is likely not to be a word if changing a letter in the stimulus is required to match it to a known word. Lexical decision results from the competition of the two processes that accumulate information over time in support of a word ($D^T = true$) or against ($D^T = false$).

Simulations

We now describe two experiments that are representative of the inconsistency of observed behavioral outcomes about the neighborhood frequency effect, and how we simulate them with the BRAID model. Experiment A (Perea & Pollatsek, 1998, experiment 1) shows an inhibitory frequency neighborhood effect, whereas Experiment B (Siakaluk et al., 2002, experiment 2) reveals a facilitatory effect.

Behavioral experiments

Design and material The two experiments manipulate neighborhood frequency using quite similar material. In both experiments, the words are English words that have at least one higher frequency neighbor (1HF) or none (0HF). All words are of low-to-medium frequency, except for half of the words that are of very low frequency in Experiment A. Neighborhood density (the number of orthographic neighbors of words) is controlled. All words of Experiment A and half the words of Experiment B have small neighborhood density;

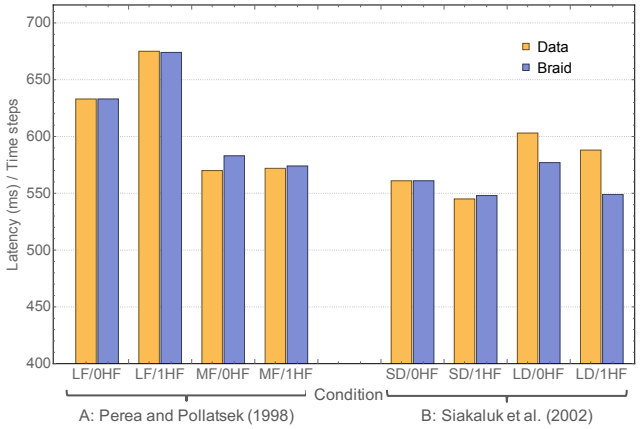

Figure 2: **Behavioral data and simulation results.** On the left, Experiment A (Perea & Pollatsek, 1998, experiment 1) and on the right, Experiment B (Siakaluk et al., 2002, experiment 2). Behavioral data (yellow) and corresponding simulation results (blue) are presented jointly for each experimental condition.

the other half of words in Experiment B have large neighborhood density. Ninety-two lowercase 5-to-6 letter words are used in Experiment A, 60 uppercase 4-to-5 letter words in Experiment B and as many legal and pronounceable pseudowords.

Procedure The two behavioral experiments follow very similar experimental designs. At each trial, a fixation point is presented at the center of the computer screen, followed by a letter-string displayed until the participant responds. Participants have to indicate as quickly and as accurately as possible whether the letter-string is a real word or not by pressing one of two response buttons, for the YES and NO response respectively. Latency (time between the onset of stimulus presentation and the motor response) for the YES responses is the dependent variable.

Experimental results As shown in Figure 2, a significant neighborhood frequency effect is reported in Experiment A but only for the low frequency words. This effect is inhibitory. In contrast, Experiment B reports a facilitatory neighborhood frequency effect for low-to-medium frequency words, which is significant for the two conditions of small and large density.

Simulations with BRAID

Material The stimuli used in the simulations are the same as in the behavioral experiments, with one exception: the word *CASINO* used in Experiment A was removed because it was not part of the British Lexicon Project (BLP; Keuleers et al., 2012) used to identify parameters of the lexical knowledge submodel of BRAID.

Procedure The task simulated by BRAID is the lexical decision task computed by Eq. (3), using default values for all

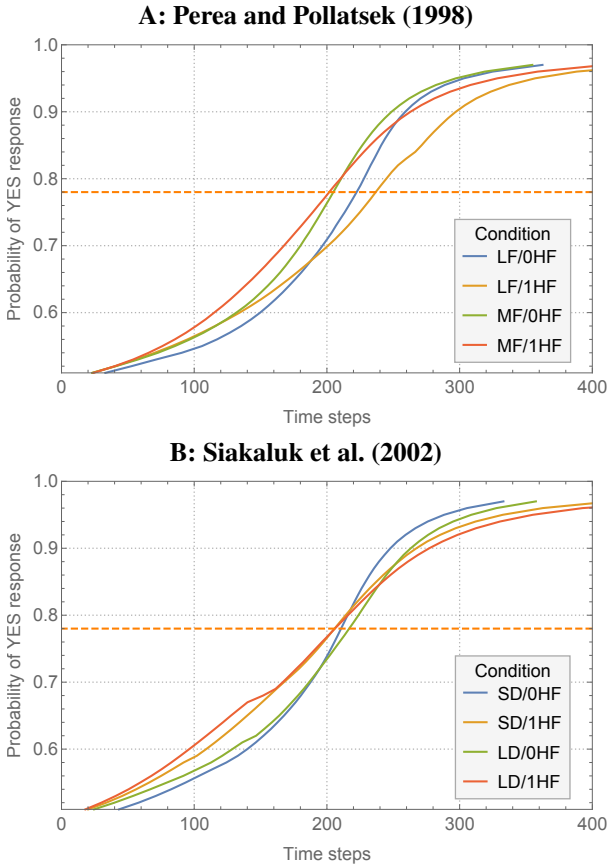

Figure 3: **Simulation of the neighborhood frequency effect in a lexical decision task.** The *x*-axis represents simulation time steps and the *y*-axis the probability of YES response. The dashed line represents the decision threshold used to compare the behavioral and the simulated results.

parameters (Phénix, 2018). For each stimulus, BRAID provides the full time course of evidence accumulation about whether or not the input string is a known word. We first simulate this process for 1,000 iterations. The words that did not reach 0.97 identification probability after 1,000 iterations were removed, resulting in an error rate of 1% and 2.5% for the stimuli used in Simulation A and B, respectively. Then, we reverse each curve to get the number of iterations required for the model to reach a given decision threshold. Finally, for each time step, the average YES response probability is calculated by averaging over all stimuli. Simulation results are shown in Figure 3.

For comparison of the simulation results with the behavioral data, we first align simulation results with the data in the first condition of the behavioral experiment and select the decision threshold that minimizes root-mean-square error (RMSE), so that the remaining three conditions are predicted by the model. The resulting threshold value was found similar in the two Simulations (set at 0.78). Note that the same threshold was adopted for the two conditions of large and small density in Simulation B despite the fact that the

items were presented by blocks in the behavioral experiment which may have induced different decision threshold.

Simulated results Figure 2 provides a comparison of the behavioral and the simulated data, scaling simulated reaction times by a multiplicative factor to align them on the first condition. In Simulation A, the simulated data well fits the human data in showing an inhibitory effect. Words with a higher frequency neighbor are processed more slowly than words with no higher frequency neighbors but only when they are of low frequency. The frequency neighborhood effect is also well captured in Simulation B for the two conditions of neighborhood density. The effect is facilitatory. However, while facilitatory effects of similar size are behaviorally reported, the effect is weaker for words with small neighborhood density in the simulations.

The dynamic curves provide relevant information on the simulated effects (see Figure 3). Exploration of the dynamics of processing time course reveals an inversion of the neighborhood frequency effect with time. The existence of a higher frequency neighbor is facilitatory at the beginning of processing but turns to inhibitory over time. Although it is observed for all the experimental conditions in Simulation A and Simulation B, the variation of the effect from facilitatory to inhibitory is modulated by the target word relative frequency and its neighborhood density. The pattern inversion occurs earlier for the low frequency words (iteration 188; threshold value=0.68) than for the medium frequency words (iteration 213; threshold value=0.81), and earlier for the words that have a small vs. large neighborhood density (iteration 217 vs. 239; threshold value = 0.81 vs. 0.85). As shown on Figure 3 panel B, a slightly lower decision threshold in Simulation B would provide a better fit to the data for the large density condition of Siakaluk et al. (2002).

Discussion

Our purpose in the current study was to provide new insights on the origin of the inconsistent findings reported in the behavioral studies on the neighborhood frequency effect, which was found inhibitory in some experiments (Carreiras et al., 1997; Forster & Shen, 1996; Grainger, 1990; Grainger et al., 1989; Grainger & Segui, 1990; Perea & Pollatsek, 1998) but facilitatory in others (Sears et al., 1995; Siakaluk et al., 2002). For this purpose, two experiments that report opposite inhibitory and facilitatory effects were selected and the BRAID model was used to simulate processing latencies for the same set of words in the lexical decision task. The results nicely mirror the high frequency neighborhood effect in successfully simulating the inhibitory effect for low frequency condition, as reported in Experiment A (Perea & Pollatsek, 1998) and the facilitatory effect reported in Experiment B whatever the density condition (Siakaluk et al., 2002). Inspection of the processing time course further provides an elegant account of the apparently inconsistent findings reported in the behavioral studies. In the two experiments, the neighborhood frequency effect that starts facilitatory turns to inhibitory later on during processing.

What do we learn from these simulation results? Simulation A shows that the degree of low frequency of the target words modulates the apparition of the inhibitory effect. This effect occurs earlier during processing for the words of lower frequency. Simulation B shows that the amplitude of the facilitatory neighborhood frequency effect is higher for words with a large density neighborhood than for words with a small density neighborhood. The facilitatory effect extends over a longer period of time from the beginning of processing when the words have a high density neighborhood, which increases the probability to obtain a facilitatory effect on these words as compared to words with a low density neighborhood. Thus, the degree of low frequency of words and their neighborhood density are factors that modulate the probability of the neighborhood frequency effect to be either facilitatory or inhibitory.

Simulation results further show that the decision threshold is a critical factor. In both Simulation A and B, different neighborhood frequency effects – inhibitory, facilitatory or null – could be obtained, depending on decision threshold. Decreasing the threshold would favor the occurrence of a facilitatory effect while increasing the threshold would increase the probability of an inhibitory effect. The model thus predicts that the effects would be either facilitatory or inhibitory depending on task demand. A facilitatory effect is predicted for the less-demanding task conditions that trigger rapid responses, thus in conditions that emphasize speed over accuracy, or use less word-like pseudowords. In contrast, the effect should turn to inhibitory when stressing accuracy over speed, or when using more word-like pseudowords. These predictions are well in line with previous reports of inhibitory effects in perceptual identification paradigms that stress accuracy and encourage unique word identification (Carreiras et al., 1997) and with evidence for modulations of the inhibitory-facilitatory effects depending on task instruction and on the pseudoword characteristics (Grainger & Jacobs, 1996).

BRAID is the first computational model that provides an account of such opposite effects. It is noteworthy that these opposite effects were here generated while using the model default parameters, thus the same set of parameters in the two experiments.

Inhibitory neighborhood frequency effects have been simulated within the IA framework (Jacobs & Grainger, 1992) and within the Multiple Read-Out model (MROM; Grainger & Jacobs, 1996) but a facilitatory effect could only be simulated at the cost of high error rate (Siakaluk et al., 2002). Opposite effects are simulated in BRAID while keeping the error rate low, depending on the time-course of the two processes that generate the YES and NO responses in lexical decision. The former process accumulates information from the lexical knowledge sub-model, based on similarity with the perceived letters. The later generates the NO response by integrating into the similarity calculation an error-seeking model,

based on detection of a single letter difference between the perceived letters and known words. The presence of at least one more frequent neighbor influences both processes. At the beginning of processing, noisy information from perceived letters matches with both the target word and its higher frequency neighbor, so that the later contributes to increasing lexical contribution in favor of a YES response, which results in a facilitatory effect of neighbourhood frequency. Over time, more and more information on perceived letters accumulates, increasing the probability to detect a mismatch and yielding competition between the target word and its higher frequency neighbor, which results in an inhibitory neighbourhood frequency effect.

The overall findings demonstrate the capacity of BRAID to handle the opposite neighborhood frequency effects that remained unexplained in the last three decades. The simulation results demonstrate that facilitatory and inhibitory neighborhood frequency effects are inherent to the time course of processing in lexical decision, which provides support to the apparently contradictory findings reported in the behavioral studies. The probability to observe a facilitatory or inhibitory effect is further modulated by the target word relative frequency and its number of orthographic neighbors, so that subtle differences in task demand and material characteristics may have yielded contradictory and sometimes confusing behavioral results.

References

Bessière, P., Mazer, E., Ahuactzin, J. M., & Mekhnacha, K. (2013). *Bayesian programming*. Boca Raton, Florida: CRC Press.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: cross-task comparisons. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *23*(4), 857–871.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 696–713.

Ginestet, E., Phénix, T., Diard, J., & Valdois, S. (submitted). Modelling length effect for words in lexical decision: the role of visual attention.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The Overlap model: a model of letter position coding. *Psychological Review*, *115*(3), 577–601.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, *29*(2), 228–244.

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A vision of reading. *Trends in Cognitive Sciences*, *20*(3), 171–179.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*(3), 518–565.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics*, *45*(3), 189–195.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: a comparison of lexical decision and masked identification latencies. *Perception & psychophysics*, *47*(2), 191–198.

Jacobs, A. M., & Grainger, J. (1992). Testing a semistochastic variant of the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1174–1188.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*(1), 287–304.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal bayesian decision process. *Psychological Review*, *113*(2), 327–357.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767–779.

Phénix, T. (2018). *Modélisation bayésienne algorithmique de la reconnaissance visuelle de mots et de l'attention visuelle*. Unpublished doctoral dissertation, Université Grenoble Alpes.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 876–900.

Siakaluk, P. D., Sears, C. R., & Lupker, S. J. (2002). Orthographic neighborhood effects in lexical decision: the effects of nonword orthographic neighborhood size. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(3), 661–681.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–43.