Temporal Integration in Visual Word Recognition

Joachim Forget[1], Marco Buiatti[1,2], and Stanislas Dehaene[1,3]

Abstract

■ When two displays are presented in close temporal succession at the same location, how does the brain assign them to one versus two conscious percepts? We investigate this issue using a novel reading paradigm in which the odd and even letters of a string are presented alternatively at a variable rate. The results reveal a window of temporal integration during reading, with a nonlinear boundary around ~80 msec of presentation duration. Below this limit, the oscillating stimulus is easily fused into a single percept, with all characteristics of normal reading. Above this limit, reading times are severely slowed and suffer from a word-length effect. ERPs indicate that, even at the fastest frequency, the oscillating stimulus elicits synchronous oscillations in posterior visual cortices, while late ERP components sensitive to lexical status vanish beyond the fusion threshold. Thus, the fusion/segregation dilemma is not resolved by retinal or subcortical filtering, but at cortical level by at most 300 msec. The results argue against theories of visual word recognition and letter binding that rely on temporal synchrony or other fine temporal codes. ■

INTRODUCTION

Any system that receives multiple pieces of information from different sensors at different times is faced with an integration/segregation dilemma: Should the information be integrated into a single object or event, or should it be taken as evidence for two objects or events? The present article is concerned with how the brain addresses this problem in the case of brief visual displays. Many studies have shown that the perceptual effects of very brief visual stimuli can persist beyond the duration of the stimulus itself. For instance, subjective reports indicate an extended duration of visual perception even when the physical stimulus was brief and is no longer present (visible persistence; Enns, Brehaut, & Shore, 1999; Efron, 1970a, 1970b). Objective tests indicate that participant exhibit an ability to integrate information across two successive visual frames only if they are separated by a short time interval, typically under ~80 msec (Coltheart, 1980a, 1980b; Di Lollo, 1980; Hogben & Di Lollo, 1974; Eriksen & Collins, 1967). Although retinal factors may contribute to this phenomenon, it may also reflect a higher-level decision to integrate or segregate signals, for instance, based on their temporal correlations (Loftus & Irwin, 1998; Di Lollo, Hogben, & Dixon, 1994; Dixon & Di Lollo, 1994; Coltheart, 1980a, 1980b). There may not be a single "perceptual moment" of fixed duration (Stroud, 1955), but a flexible ability to integrate or segregate information across time and space into separate "streams" (Bregman, 1990) or "files" (Kahneman, Treisman, & Gibbs, 1992). The mechanisms by which this integration/

segregation dilemma is resolved remain a major unsolved question.

Here, we probed temporal integration in the domain of visual word recognition, asking how the multiple letters of a word are integrated or segregated. Inspired by Fraisse's (1966) research on letter integration, and by the missing dot paradigm in which the two halves of a dot matrix are presented as two successive time frames (Hogben & Di Lollo, 1974), we divided a letter string into two successive displays, one showing the odd-numbered letters (e.g., B_A_N) and the other showing the even-numbered letters (e.g., _R_I_S; see Figure 1). Of crucial interest was whether viewers would manage to merge the two stimuli and perceive the whole word (in this case, BRAINS), or would only perceive the component strings (e.g., BAN). The experimental variable was the rate of alternation of the two components, as well as the lexical status of the merged and component strings.

One interest of probing temporal integration in the reading domain is that the level at which the integration/segregation problem is resolved can be studied by manipulating the lexical status of the stimuli. If the ability to integrate across time depends solely on a low-level temporal filter, perhaps in the retina or in area V1 (e.g., Duysens, Orban, Cremieux, & Maes, 1985; Levick & Zacks, 1970), or if integration is mostly determined by the temporal correlation between the bottom–up activation patterns evoked by each component (Loftus & Irwin, 1998; Dixon & Di Lollo, 1994), then a high-level variable such as lexicality should have no impact on perception. If, however, temporal integration results from a higher-level process which takes into account the a priori plausibility and lexicality of the integrated percept before stabilizing it, then real words might be expected to resist slower presentation

[1]INSERM-CEA, Gif sur Yvette, France, [2]University of Trento, Mattarello, Italy, [3]Collège de France, Paris

Figure 1. Stimuli and experimental conditions. Top: Stimulus sequence. Two spaced strings of three or four letters, corresponding to the even and odd letters of a six- or eight-letter string, were alternatively presented, three times each, followed by a final mask. The rate of alternation of the even and odd component was varied by manipulating the duration of the blank screen (ISI ranging from 33 to 117 msec). At fast rates, the merged string was perceived. Bottom: Illustration of the main experimental conditions (dashed lines illustrate the actual alignment of the strings on screen). In the whole-word condition, the components were nonwords, but the merged string was a valid French word. In the component-words condition, two French words (e.g., "QUAI" and "ROBE") were presented as components, but the fused string ("RQOUBAEI") was not a word. Finally, in the nonword condition, neither the components nor the fused string was a word.
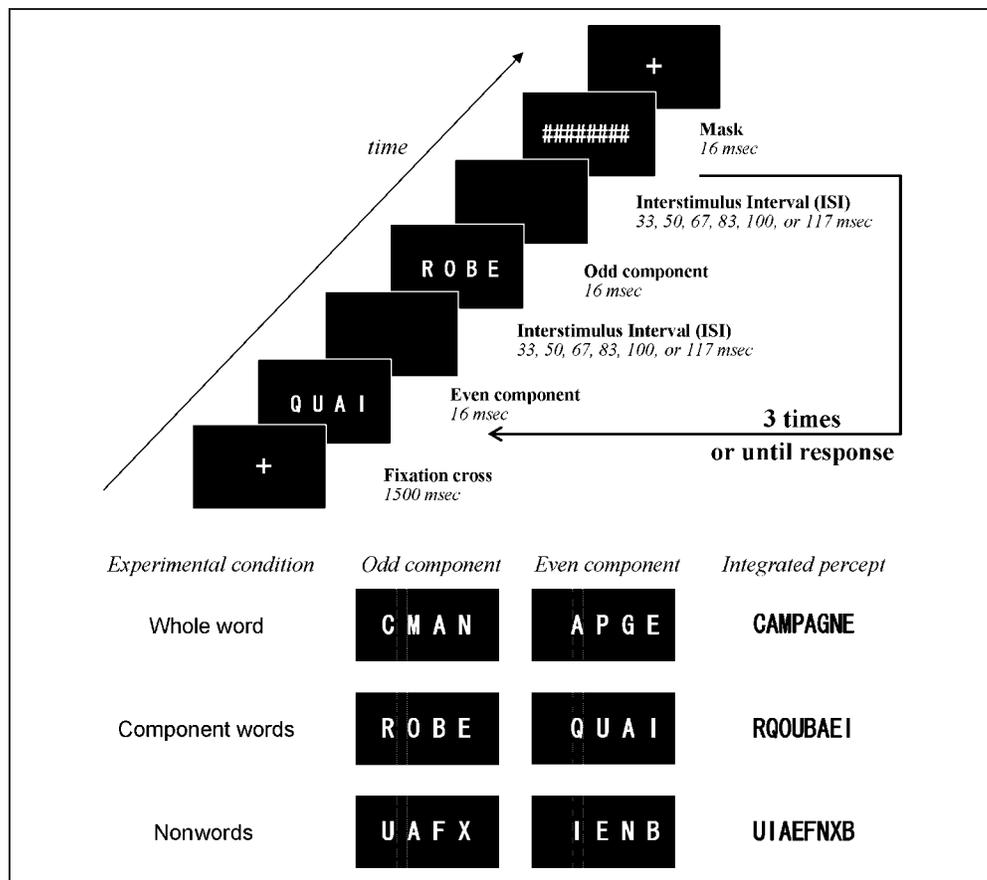
rates better than pseudowords. Although we know of no explicit simulation of this point, interactive activation models (starting with McClelland & Rumelhart, 1981) and more recent Bayesian models of the lexicon (Norris & Kinoshita, 2008), where sensory evidence and top–down plausibility are optimally combined, would seem to naturally make this prediction.

A second interest of our letter stimulus is that, contrary to the perception of dot matrices (Di Lollo, 1980; Hogben & Di Lollo, 1974), there are well-defined models of visual word recognition that can be used to predict the impact of temporally splitting a word into its component letters (assuming that the integration/segregation dilemma is not entirely resolved by low-level pre-orthographic mechanisms). For present purposes, two categories of models can be distinguished: those based on temporal coding and those based on parallel spatial integration. Recently, several temporal coding models of orthographic recognition have been proposed and have claimed to achieve an excellent match to orthographic priming, perceptual confusion, and reading time data (Whitney, 2001; Davis, 1999). They postulate that the position of letters within the input string is determined by the precise discharge time or oscillation phase of input units. Davis's (1999) Self-organizing Lexical Acquisition and Recognition (SOLAR) model encodes letter position by a spatial coding scheme, where

the level of activation of each letter unit varies with its location in the input string (Davis & Bowers, 2006). This code is arrived at through temporal coding—input letters, even when presented simultaneously, are assumed to generate fast serial left-to-right "beats" (every 10 msec) that are combined to create the spatial code. Whitney's (2001) SERIOL model likewise supposes that successive letter positions are encoded by successive delays of 25 msec in the firing of neurons coding for each successive letter. Such models, which are inspired by broader spike-timing models of visual recognition (e.g., VanRullen & Thorpe, 2002), predict a drastic impairment of reading under the temporal splicing used here, even with the shortest temporal interval used in our experiment (50 msec). This is because, assuming a coding scheme by gamma-band frequencies with at most 25-msec firing lag for each letter location (Whitney, 2001), even a 50-msec interval should corrupt the coding of the spatial locations of the letters. The word "BRAINS," for instance, with a 50-msec interval between the substrings B_A_N and _R_I_S, should be erroneously perceived as an unreadable string such as "BARNIS" or "BANRIS."

The situation is quite different with classical connectionist or bigram models which have as their input a bank of parallel letter detectors—at each location, each input letter is encoded by a distinct spatially tuned unit (Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006;

Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger & Whitney, 2004; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 1999; Ans, Carbonnel, & Valdois, 1998; Zorzi, Houghton, & Butterworth, 1998; Grainger & Jacobs, 1996; Mozer, 1987; McClelland & Rumelhart, 1981). These models predict only a moderate impact of temporally separating the odd and even letters. The exact, quantitative prediction depends on the temporal duration of the input letter buffer (a parameter which has not typically received much attention). In principle, if the input buffer has infinite duration, as soon as both odd and even strings have been presented, the visual system is in full possession of the entire letter-based positional code and should be able to use it for lexical access. Thus, these models predict only a linear processing delay as a function of the temporal interval separating the two component strings. In particular, at very short stimulus onset asynchronies [SOAs] (e.g., 50 msec), reading should be essentially normal. Models of the interactive activation type, however, typically postulate that after stimulation, activation relaxes back toward zero with a finite decay constant (e.g., McClelland & Rumelhart, 1981). In this case, spacing of the component strings by a temporal interval greater than this decay constant would create a severe reading deficit. This is because the higher-level units, be they word units as in the early models, or ordered bigrams, as in more recent proposals (Grainger et al., 2006; Dehaene et al., 2005; Grainger & Whitney, 2004; Mozer, 1987), would never simultaneously receive the activation from all letters at once. Even assuming an open bigram code where intermediate letters are allowed (Grainger et al., 2006; Dehaene et al., 2005; Grainger & Whitney, 2004), where bigram B_A would fire for the word "brain," at least half of the bigrams would be disrupted by splitting the string into odd and even letters (e.g., bigram B_R would never be activated). Thus, these models predict a threshold-like disruption of reading by temporal splitting. For SOAs larger than the network's decay constant, reading accuracy and speed should dramatically decrease, with evidence of a switch toward a distinctly slower, serial letter-by-letter strategy—as previously observed with spatial letter s p a c i n g (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008).

To investigate these issues, we performed three experiments, all using this paradigm in which the odd and even letters of a string are presented alternatively at a variable rate. In Experiment 1, using subjective reports similar to Fraisse (1966), we show that there are well-defined windows of integration and of segregation: At fast alternation rates, subjects can read the merged string but not the component strings, whereas at slow rates, the converse occurs. In Experiment 2, we extend this observation to an objective lexical decision task, and demonstrate that outside a window of integration, a sudden disruption of normal reading occurs, suggestive of serial letter-by-letter processing; we also show that the temporal window is larger for words than for pseudowords. Finally, in Experiment 3, we record ERPs evoked by our stimuli to demon-

strate that integration does not arise solely from a low-level visual temporal filter, but presumably results from a late cortical integration mechanism.

GENERAL METHODS

Participants

A total of 39 right-handed students participated (mean age = 20.7 years; 7 women and 3 men in Experiment 1; 9 women and 7 men in Experiment 2; and 7 women and 6 men in Experiment 3 [3 additional subjects were rejected for excessive ERP artifacts]). All had normal or corrected-to-normal vision and were native French speakers unfamiliar with the stimuli and with the aim of the experiment. They were paid for their participation.

Stimuli

Stimuli presentation, timing, and data collection were controlled using E-Prime (Psychology Software Tools Inc., Pittsburgh, PA). Observers were seated 57 cm from the display, and their responses were collected with the keyboard. Stimuli were displayed on a CRT monitor with a vertical refresh rate of 60 Hz, in a size-20 MS Gothic font, white on a black background. All stimuli subtended a maximum of 0.9° (height) × 4° (width) of visual angle. Each trial consisted in alternating presentations of the even (second, fourth, sixth, eighth letters) and odd components (first, third, fifth, seventh letters) of an uppercase letter string. The components were appropriately spaced and centered so that, if merged, they would constitute a single string (Figure 1). On each trial, a fixation cross was presented for 1510 msec. Then the even component was flashed for 16 msec, followed by a blank screen for a variable interstimulus interval (ISI), which determined the SOA separating the onsets of each component string. The same sequence recurred with the odd component. In Experiments 1 and 3, the entire even–odd sequence was repeated three times, and ended with a masking string consisting of 8 "#" signs (16 msec) and the fixation cross. In Experiment 2, where the main measure was response time rather than accuracy, the alternations continued until a response was recorded, thus ensuring an enhanced accuracy. Six SOAs were used: 50, 67, 83, 100, 117, or 133 msec.

EXPERIMENT 1: SUBJECTIVE REPORT

The aim of this experiment was to evaluate the subjective perception of strings whose odd and even components alternated at a variable frequency. We presented stimuli in which either the two component strings, or the merged string, or neither, was a French word. We simply asked subjects how many words they could see.

Methods

In the whole-word condition, the merged string was a word, six or eight letters long (mean $\log_{10}$ frequency per

million = 1.55), whereas the components were nonwords (e.g., components _A_A_E and G_R_G_, merged string GARAGE). In the component-words condition, the components were words, three or four letters long (mean $\log_{10}$ frequency per million = 2.09), whereas the whole string

was not (e.g., components _B_A_R and S_K_I_, merged string SBKARI). Finally, in the nonword condition, neither the components nor the merged string belonged to the French lexicon (e.g., _G_L_A and I_M_R_, merged string IGMLRA). Observers were asked to report, after each trial, whether they had been able to read 0, 1, or 2 words by pressing the corresponding keys on a numeric pad. They were instructed that the words could be either long (6 or 8 grouped letters) or short (3 or 4 spaced letters). Following a training set of 36 stimuli, the target set for Experiment 1 consisted of 360 trials in randomized order, 20 in each factorial combination of six ISIs and three conditions: whole word, component words, and nonword.

Results

Figure 2 shows mean subjective readability for the three stimulus conditions. The results were highly reliable, in the sense that participants hardly ever reported seeing one or two words when neither the components nor the merged string were words (<1% of responses), nor seeing two words in the whole-word condition (<1.5% of responses). ANOVAs on the percentages of 0, 1, or 2-word responses revealed the presence of highly significantly Condition × SOA interactions (all $p < .001$). When the merged string formed a whole word, participants reported being able to read it on a majority of trials with a fast alternation rate (50 or 67 msec SOA, corresponding to oscillation frequencies of 20 Hz or 15 Hz), but ceased to be able to read as soon as SOA exceeded 83 msec (frequencies below 12.1 Hz). When the components were words, the ability to perceive them followed a symmetrical curve: Participants reported not being able to read them at SOAs of 50 or 67 msec, then became able to read one (at SOA = 83 msec) or two of them (at SOA = 100 msec or above).

Discussion

The integration window can be defined as the maximal duration of presentation of fragments of words, beyond which perception of a coherent whole-word percept is impossible. The segregation window can be defined, symmetrically, as the minimal component duration needed to read the components when each of them spells a distinct word. It is interesting and nontrivial that the two definitions converge to a similar value of about 80 msec. Indeed, Experiment 1 suggests that the transition is rather
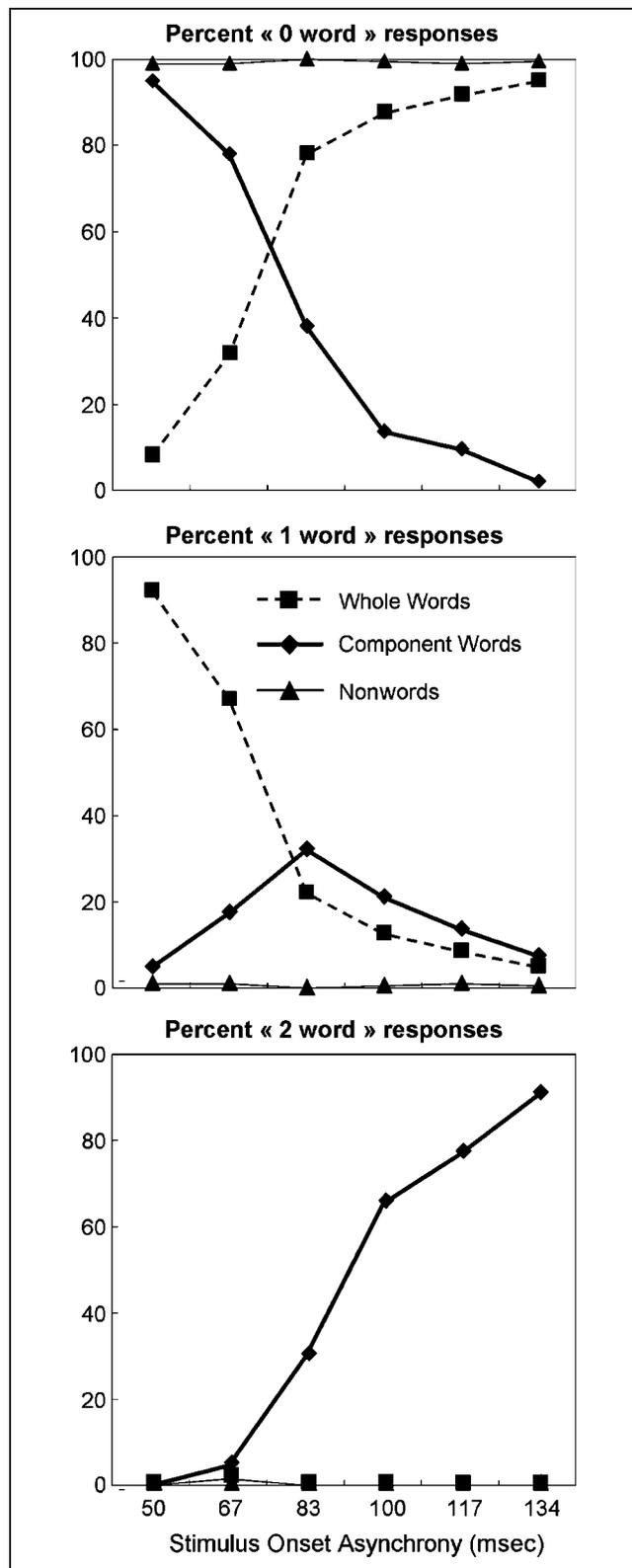

Figure 2. Subjective readability as measured in Experiment 1. Each line type is associated with a stimulus category (whole word, component words, or nonword). The three panels give, respectively, the percentage of responses of each type: 0, 1, or 2 words seen (respectively top, middle, and bottom; the three values sum to 100%). As SOA increases and the rate of presentation slows down, participants increasingly report losing the ability to read the merged string (dashed curve) and gaining the ability to read one or two of the component words (thick curve).

sharp and with essentially no "ambiguity zone" within which no stable percept can be achieved. Integration and segregation appear as mirror images of each other.

The value of ~80 msec that we observe is consistent with reports from the rapid serial visual presentation paradigm, in which sentences or rebuses presented at rates of up to 12 words or pictures per second remain highly readable (Potter, Kroll, Yachzel, Carpenter, & Sherman, 1986). It is only slightly shorter than the value of ~100 msec reported by Holcombe and Judson (2007) as the 75% performance threshold for distinguishing superimposed words such as pump/hell versus pull/hemp. Below this value, and especially at an SOA of 50 msec, participants seem to experience no difficulty in reading a word which has been split into its odd and even letters—a finding that seems incompatible with temporal coding models. A limit of Experiment 1, however, is that it is based solely on subjective readability reports, which may be biased (participants may report as ultimately "readable" stimuli that were, in fact, hard to integrate). We therefore postpone the discussion of its theoretical implications until after Experiment 2, whose goal was to gather objective evidence from reading time and accuracy from the classical lexical decision task.

EXPERIMENT 2: OBJECTIVE ASSESSMENT OF READING

In Experiment 2, response times and error rates were measured in a lexical decision task where participants judged whether the merged string was or was not a French word. Instead of just three alternations, the alternating stimulus was left on until a response was made. For SOAs above the integration threshold, the inability to merge the even and odd components into a single percept should induce a severe slowing down of lexical decision, but given enough time participants should still be able to decipher the word letter by letter. To evaluate this possibility, we examined the impact of word length on lexical decision time. In adult, the reading of normal words is characterized by an independence of reading time on the number of letters. This absence of a word-length effect is found in many tasks including overt reading (e.g., Weekes, 1997), lexical decision (e.g., Lavidor & Ellis, 2002), and semantic decision (e.g., Cohen et al., 2008). However, a linear increase emerges under conditions of stimulus degradation that are thought to prevent parallel processing of the string and call for a deployment of serial attention, including letter s p a c i n g (Cohen et al., 2008). With our alternating stimulus, given Experiment 1's results, we predicted that for SOAs above 80 msec, a word-length effect would appear, indicating a sudden lack of parallel letter integration.

Methods

A total of 360 trials were presented in randomized order, 10 in each factorial combination of SOA (6 levels as in Experiment 1), word length (4, 6, or 8 letters), and lexicality (word or pseudoword). Words of different lengths were equalized in frequency (mean $\log_{10}$ frequency per million $= 1.98$). The pseudowords were generated by cross-splicing the beginning of one word with the end of another, and occasionally altering one or two letters for pronounceability. This method prevented the presence of inadvertent cues to lexical status either in the initial letters or in the even or odd component strings. Participants used a response box and responded bimanually, as accurately and quickly as possible, by depressing the right index button for words and the left index button for pseudowords. Response times were measured from the onset of the second component string, when the full set of letters first became available.

Results

As shown in Figure 3, both RTs and error rates increased steadily with SOA, but with a nonlinear change around SOAs of ~80–100 msec associated with a sudden onset of a word-length effect.

Error rates were much higher for words than for pseudowords [$F(1, 15) = 16.71, p < .001$]. A triple interaction of SOA, length, and lexical status [$F(10, 150) = 9.68, p < .001$] indicated that the task became particularly difficult with words as SOA increased beyond 100 msec and as word length increased. Indeed, there was a main effect of SOA for both words [$F(5, 75) = 34.23, p < .001$] and pseudowords [$F(5, 75) = 4.74, p < .001$], but the SOA × Length interaction was much stronger for words [$F(10, 150) = 8.67, p < .001$] than for pseudowords [$F(10, 150) = 2.25, p < .05$]. Error rate reached ~40% for six- to eight-letter words presented at SOAs 117 and 134 msec, suggesting that, at this point, the inability to integrate the component strings into a coherent word made these stimuli look like pseudowords.

Median correct response times were submitted to a within-subject ANOVA with lexicality, SOA, and length as factors. Responses were overall slower for pseudowords than for words [$F(1, 15) = 31.55, p < .001$], but otherwise showed rather similar profiles for words and pseudowords, as attested by the lack of triple interaction of SOA, length, and lexicality [$F(1, 150) = 1.28, p > .25$]. There was an increase of RT with SOA [words: $F(5, 75) = 35.22, p < .001$; pseudowords: $F(5, 75) = 33.90, p < .001$], with length [words: $F(2, 30) = 15.19, p < .001$; pseudowords: $F(2, 30) = 29.85, p < .001$], and a significant interaction [words: $F(10, 150) = 4.82, p < .001$; pseudowords: $F(10, 150) = 6.08, p < .001$]. Although a significant word-length effect was observed in all subconditions of SOA and lexicality (all $p < .05$), the effect was very small at the shortest SOA $= 50$ msec, and only became large and suggestive of serial letter-by-letter processing for SOAs of 83 msec and above. As shown in Figure 3, this full-blown word-length effect was well established at SOA $= 83$ msec for pseudowords, but only emerged at SOA $= 100$ msec for words. We
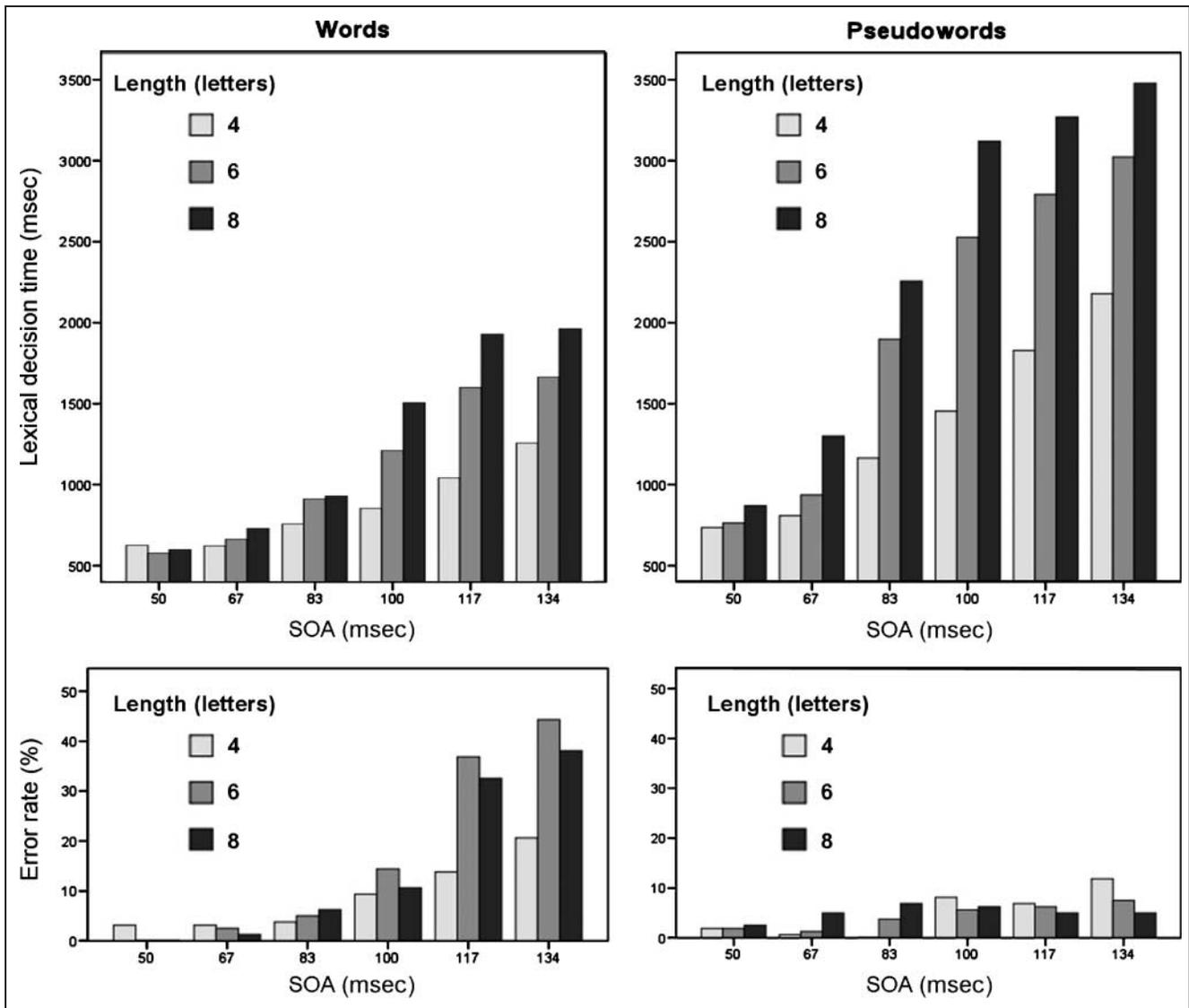
Figure 3. Lexical decision times and error rates in Experiment 2, as a function of the lexical status and length of the merged string. High error rates and a word length effect emerged around SOAs of 80–100 msec. Note the distinct scale for RT in the left and right top panels.

tested this statistically by conducting sliding ANOVAs over consecutive SOAs. Only between 67 and 83 msec was there a significant triple interaction of lexicality, SOA, and length [$F(2, 30) = 5.15, p = .009$], due to the sudden emergence of a large length effect at 83 msec for pseudowords but not yet for words (see Figure 3).

Discussion

The existence of a temporal integration threshold of ~80 msec was confirmed in Experiment 2. Indeed, it is striking that, even with response times over 1 sec and with unlimited stimulus presentation, at the slowest SOAs the participants still made many errors of making a "nonword" response to real words. This observation confirms that successive stimuli separately by 100 msec or more are extremely difficult to integrate as a single percept, even when given time. Lexical decision times concur with this conclu-

sion. Indeed, the curve tracing the envelope of the word length as a function of SOA in Experiment 2 (Figure 3) is very similar to the subjective reports of whole-word readability in Experiment 1 (Figure 2). Both findings converge to suggest that reading suddenly becomes hard and effortful at long SOAs, as the fast integration of the two component strings into a coherent percept is prevented. Subjects appear to default to a nonword response whenever they can no longer see the integrated stimulus, thus explaining the apparent speed–accuracy tradeoff in Figure 3.

Interestingly, the length effect emerges somewhat later for words (SOA ~100 msec) than for pseudowords (SOA ~83 msec). This finding suggests a greater resistance to temporal segregation for words present in the lexicon, which presumably have a more stable cortical representation, and is a plausible first indication that temporal integration is not solely determined by low-level visual processes. However, it might also reflect the well-known

fact that string length affects pseudowords more than words (Lavidor & Ellis, 2002; Weekes, 1997).

Our most important conclusion is that word identification remains fast and accurate as long as the component letters are presented within less than 80 msec of each other—and that it becomes virtually impossible once they are separated by more than 100 msec. As will be further detailed in the General Discussion, the normal pattern of reading times that we observed for short SOAs, is incompatible with temporal coding models, which predicted a severe disruption of reading even at 50 msec SOA. These models might be salvaged, however, if the integration/ segregation dilemma was resolved entirely at an early visual level (e.g., retinal). In this case, the subsequent orthographic stage would be free to operate on the integrated input string with any mechanism, including temporal parsing. Experiment 3 therefore used ERPs to demonstrate that integration is not an early retinal property, but results from a late cortical mechanism.

EXPERIMENT 3: EVENT-RELATED POTENTIALS

In Experiment 3, we used ERPs to study the time course of the process responsible for solving the integration/ segregation dilemma. By contrasting words and pseudowords as the merged strings, we expected to observe massive ERP differences related to lexical status. Their presence would imply that integration of the stimuli is either fully accomplished or well under way, in which case their time of appearance would also provide an upper bound on the duration of temporal integration. We also analyzed the recordings for phase and amplitude modulation at the injected oscillation frequency, with the aim to clarify the role of low-level temporal filtering in the integration/ segregation boundary. If retinal or early visual low-pass filtering is responsible for visual integration, then we should see a stimulus-induced response at the injected frequency only when segregation is reported, not at the SOAs at which integration occurs. If, on the other hand, integration is due to a higher cortical mechanism, then early visual areas should emit phase-synchronous EEG oscillations at all SOAs.

Methods

Stimuli

The SOAs (50, 67, 83, 100, 117, and 133 msec) and stimuli were the same words and pseudowords used in Experiment 2. To maximize differences between conditions, only the 240 eight-letter words and pseudowords were used. Given that Experiments 1 (subjective report) and 2 (objective lexical decision) gave similar information concerning temporal integration thresholds, in Experiment 3 we returned to the subjective report and time-limited presentation used in Experiment 1 (3 alternations of the odd and even components, ending with a masking string of 8 "#" signs). Participants were informed that they should try to read the merged stimulus, and reported whether or not they managed to read it by pressing a left or right key. Response key assignments were varied pseudorandomly across subjects and swapped in the middle of the experiment. Subjects only responded after a dimming of the fixation point, which occurred 800 msec after the stimulus sequence.

Event-related Potentials

ERPs were sampled at 250 Hz with a 129-electrode geodesic sensor net (EGI, Eugene, OR) referenced to the vertex. We rejected trials with voltages exceeding 100 μV, transients exceeding 70 μV, or electrooculogram activity exceeding 60 μV. For ERP analysis, the remaining trials were averaged in synchrony with word onset, digitally transformed to an average reference, band-pass filtered (0.2–35 Hz) and corrected for baseline over a 300-msec window during fixation prior to word onset.

Phase and Amplitude Modulation

For each electrode, the Fourier transform $F_k(f)$ of each epoch k of the nonfiltered EEG signal was calculated using a fast Fourier transform (FFT) algorithm (MATLAB, Natick, MA) on a time window starting from the first stimulus appearance and lasting approximately eight times the SOA, thus broadly covering the stimulus presentation time. For each SOA, the length of the time window was carefully chosen such that the FFT was computed at exactly the stimulation frequency of the component strings $f = 1000/SOA$, which should be the pertinent frequency for area V1 (an ERP should be evoked by each component string). In order to quantify the increase of power (amplitude modulation) and intertrial phase locking (phase modulation) during stimulation, the FFT was also computed on a time window preceding stimulation (hereafter indicated as *baseline*). The baseline time window is exactly of the same length as that used for the stimulation at that SOA, and ends with the first stimulus appearance. Power spectrum was calculated from the Fourier coefficients as the average over epochs of the single-epoch power spectrum: $P_k(f) = F_k(f) \times F_k^*(f)$. The phase-locking factor (PLF) was computed using the following formula (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996): $PLF(f) = \| <F_k(f)/\| F_k(f) \| > \|$, where the average $< >$ is computed over all single epochs k and $\| \|$ indicates the complex norm. PLF values range from 0 (*purely non-phase-locked activity*) to 1 (*strictly phase-locked activity*). Note that Tallon-Baudry et al. used wavelet analysis because they were interested in localizing their effect in both time and frequency dimensions. By contrast, our goal was to identify phase locking at known stimulus frequencies, with the highest frequency resolution (i.e., longest time window), thus justifying the use of a Fourier

transform over a time window covering all the stimulation period.

Cluster Randomization Analysis

The statistical significance of the difference between PLF during stimulation and PLF during baseline was established using a nonparametric randomization test called Cluster Randomization Analysis (CRA) (Maris & Oostenveld, 2007). This test effectively controls the Type I error rate in a situation involving multiple comparisons (such as 129 channels) by clustering neighboring channel pairs that exhibit the same effect. We adapted the implementation of CRA incorporated in the Fieldtrip toolbox (open source software for electromagnetic brain signal analysis; www.ru.nl/fcdonders/fieldtrip) to the spatial arrays of PLF values. The first step of CRA is to identify channels whose t statistics exceeds a critical value when comparing two conditions channel by channel ($p < .05$, two-sided). Channels that exceed the critical value and neighboring in the channel array (separated by less than 5 cm) are then grouped as a cluster. Each cluster is assigned a cluster-level statistic whose value equals the sum of the channel-specific statistics. The cluster p value is estimated as the proportion of the null distribution (obtained by randomizing the order of the two conditions within every participant with 8192 permutations = $2 \wedge$ number of subjects) in which the maximum cluster-level test statistic exceeds the observed statistic.

Results and Discussion

Behavioral readability reports (Figure 4) were analyzed in an ANOVA with factors of lexicality and SOA. Readability was higher for words than for pseudowords [$F(1, 12) = 28.3, p < .001$] and dropped sharply with SOA [$F(5, 60) = 49.1, p < .001$]. Crucially, the two factors interacted [$F(5, 60) = 3.95, p = .004$], indicating that readability varied differently with SOA for words and pseudowords. As seen in Figure 4, the steepest decrease in readability was observed between 67 and 83 msec SOA for pseudowords, and between 83 and 100 msec SOA for words, suggesting that as in Experiment 2, the integration/segregation threshold was lower for pseudowords than for words. To prove statistically that the two curves in Figure 4 were shifted horizontally, rather merely scaled versions of each other, we rescaled each subjects' extreme points (SOAs 134 and 50 msec) to 0 and 1, separately for words and pseudowords, and then reanalyzed the intermediate SOAs. The scaled readability remained significantly higher for words than pseudowords [$F(1, 12) = 19.1, p < .001$; see Figure 4]. Thus, for words the availability of a lexical entry seems to facilitate integration and, therefore, increase the resistance to slower alternation rates.

Figure 5 shows selected ERP time courses for each condition of SOA and lexicality, whereas Figure 6 shows the topography of the observed differences between words
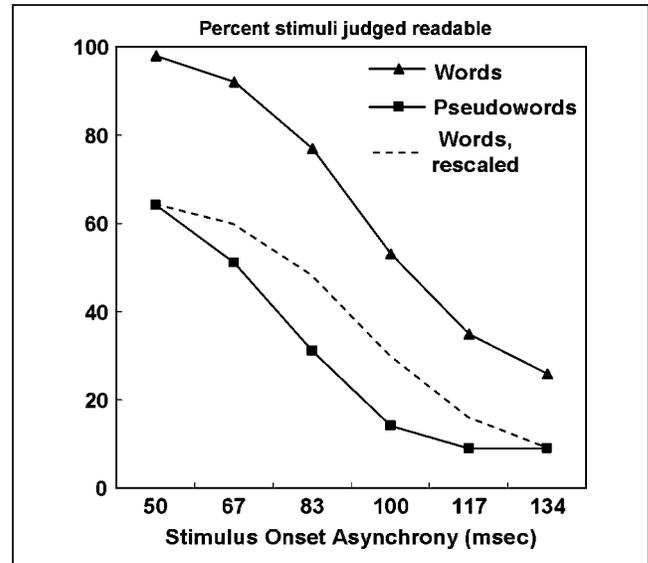

Figure 4. Subjective readability ratings during ERP recordings in Experiment 3. Participants managed to read whole words better than pseudowords. Furthermore, the integration threshold was higher for words (between 83 and 100 msec) than for pseudowords (between 66 and 83 msec), as can be seen once the ratings for words are rescaled to align them with the pseudoword ratings (dashed curved).

and pseudowords. The lexicality of the whole string impacted on two ERP components: a left-lateralized early posterior negativity (EPN) over temporal electrodes, peaking between 350 and 450 msec, and a parietal P3-like positivity, peaking between 500 and 750 msec. Strikingly, both effects were present only at the three shortest SOAs (50, 67, and 83 msec) for which subjects reported a high readability, and vanished for higher SOAs. Indeed, their amplitude as a function of SOA traced a curve very similar to the subject's readability reports (compare Figures 4 and 6).

For statistical validation of these observations, mean ERP amplitude was computed over groups of electrodes representative of the topography and latency of each scalp component: a group of posterior electrodes for EPN (left electrodes: 50 56 57 58 63 64 65 69 70; right electrodes: 91 95 96 97 100 101 102 108) over the latency 350–450 msec, and a group of central electrodes for P3 (left electrodes: 7 31 32 37 38 43 52 53 54 55 60 61 62 67 68 129; right electrodes: 62 68 70 78 79 80 81 86 87 88 92 93 94 106 107) over the latency 500–750 msec. An ANOVA with factors of lexicality, SOA, and hemisphere was then computed separately for the EPN and the P3 component. In both cases, we observed large effects of lexicality [EPN: $F(1, 12) = 68.6, p < .0001$; P3: $F(1, 12) = 10.3, p = .008$], SOA [EPN: $F(5, 60) = 13.3, p < .0001$; P3: $F(5, 60) = 10.5, p < .0001$], and crucially, the Lexicality × SOA interaction [EPN: $F(5, 60) = 3.42, p = .009$; P3: $F(5, 60) = 5.99, p = .0001$], confirming the vanishing of the difference between words and pseudowords once SOA exceeded 100 msec. Only for P3 did we observe a significant Lexicality × Hemisphere interaction [$F(1, 12) = 3.26, p = .015$], suggesting a stronger effect on left-sided electrodes.

Were the EPN and P3 related directly to lexical access? An alternative hypothesis is that they reflected readability, which was higher for words than for pseudowords. For the intermediate SOA = 83 msec, where enough trials were available in each of the readable and nonreadable categories, we redid the above statistical analyses with factors of lexicality, hemisphere, and an additional readability factor (these ANOVAs were performed across the pooled trials of all subjects, as there was not enough data in each cell within each subject). The EPN and P3 differed massively on readable compared to nonreadable trials, both as a main effect [EPN: $F(1, 1456) = 32.67$, $p < .000$; P3:

Figure 5. Time course of the lexical effects observed in ERP recordings. Effects were observed in two time windows (EPN = early posterior negativity; P3 = late central positivity), only for short SOAs where subjects reported perception of the integrated string. Curves show the average voltages computed on electrode clusters corresponding to EPN and P300 peaks (see Experiment 3 Methods), separately in the left and right hemisphere and for each SOA and lexicality condition.
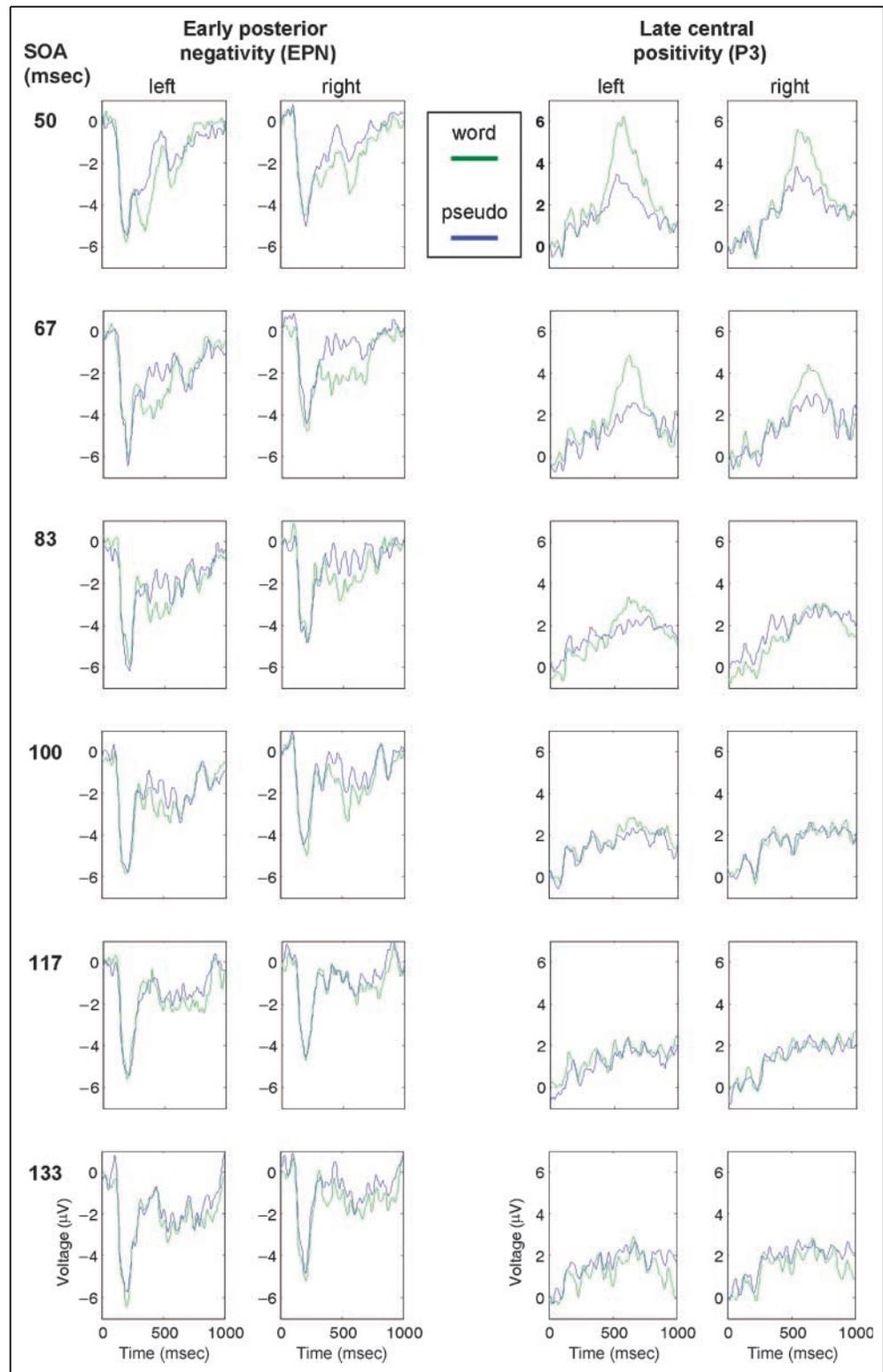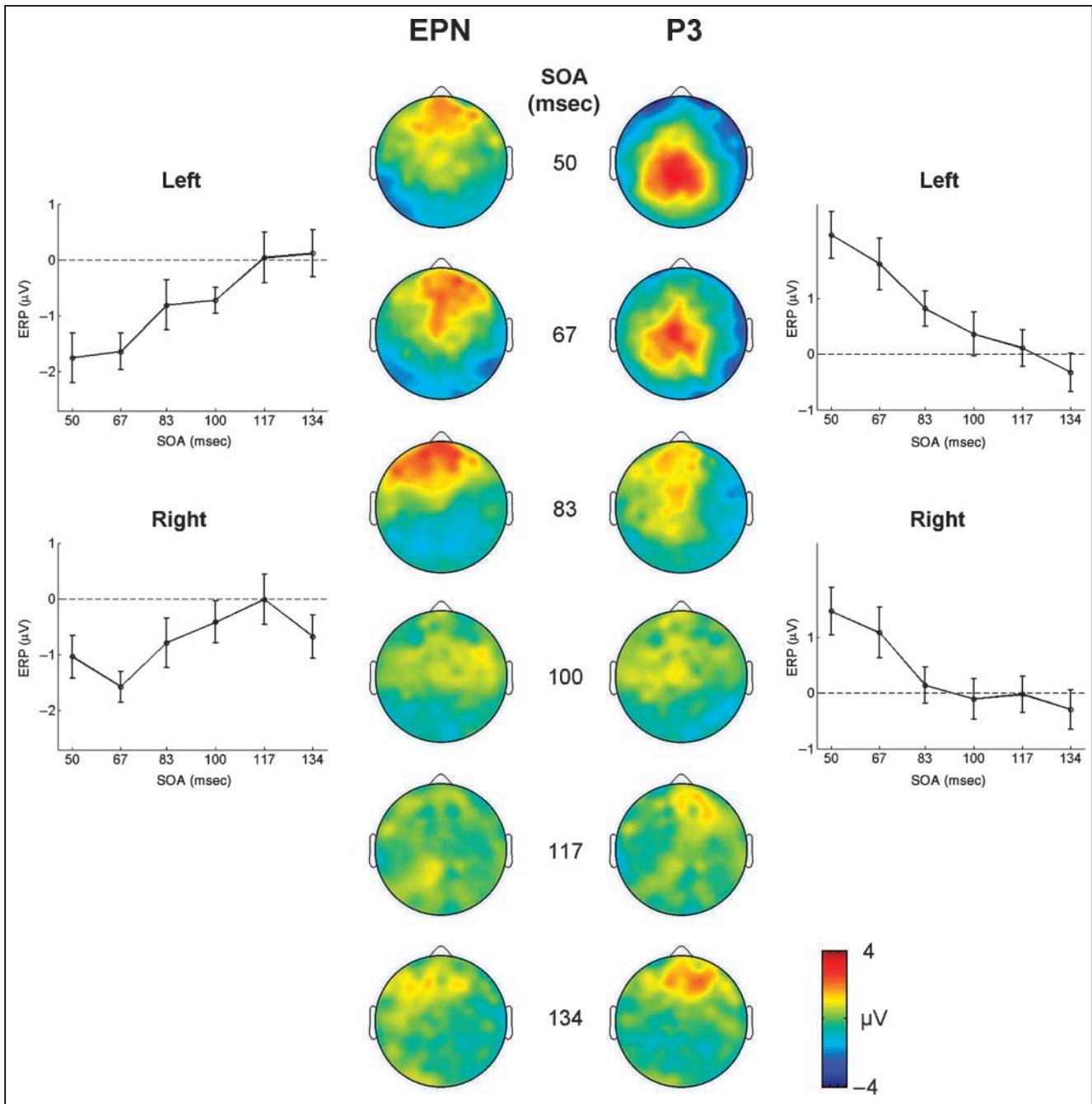
Figure 6. Topography and amplitude of the lexical effects. Each topographical plot shows a scalp map of the difference between word and pseudoword conditions, for each SOA averaged over the latencies corresponding to EPN (350–450 msec) and P3 (500–750 msec) peaks. Average amplitude is shown separately for each hemisphere over groups of electrodes representative of the component topography (see Experiment 3 Methods). Error bars indicate the standard error of the mean.

$F(1, 1456) = 39.65, p < .000$] and within the categories of word [EPN: $F(1, 728) = 18.25, p < .000$; P3: $F(1, 728) = 22.16, p < .000$] and pseudoword [EPN: $F(1, 728) = 14.31, p < .000$; P3: $F(1, 728) = 17.34, p < .000$]. Most importantly, once readability was taken into account, only minor effects of lexicality were seen [EPN: no main effect but a triple interaction, $F(1, 1456) = 5.29, p < .03$; P3: small lexicality effect, $F(1, 1456) = 4.79, p < .03$, but no interactions]. These results indicate that the EPN and P3 mostly

index readability, which in turn is influenced by lexicality, although there may be a small residual lexicality effect as well.

In summary, the ERP results reveal two components whose profile of variation with SOA and correlation with readability indicate that they index the integration of the component string into a readable whole. The latency of the EPN places an upper bound on this integration process: By about 350 msec, the integration/segregation dilemma must

have been resolved. Next, we examined whether a lower bound on integration could be obtained. To this aim, we examined whether early occipital activity evoked by the alternating components could be detected even at frequencies higher than the integration threshold. Such a finding would imply that the components are not fused at retinal, geniculate, or primary visual stages, and therefore, that integration occurs at a cortical level, after early visual processing.

We initially searched for poststimulus increases in power at the predicted frequency ($f_{SOA} = 1000/SOA$ where SOA is in msec). However, it proved difficult to identify this activity both in the raw EEG and in the average ERP because the power spectrum before stimulation was dominated by a large peak in the alpha range (8.5–12 Hz) whose drastic decrease during stimulation masked any stimulus-induced power increase, which was expected in the same frequency range for several SOAs. However, we reasoned that stimulus-induced activity should also be characterized by a phase resetting of oscillations at the stimulation frequency f_{SOA}, time-locked to the stimulation onset (Makeig et al., 2002). The predicted phase locking of this activity at a specific frequency should render it highly discriminable, both from ongoing alpha oscillations (which are not phase-locked to stimulation) and from other ERP waves (which emerge at SOA-independent frequencies).

To evaluate stimulation-dependent phase resetting, we computed the PLF (Tallon-Baudry et al., 1996), a measure that quantifies intertrial phase locking at a specific frequency regardless of the overall variation in amplitude of the oscillations at that frequency. For each SOA and each electrode, phase resetting was evaluated by computing the difference between stimulation and baseline of PLF at the stimulation frequency f_{SOA} (Figure 7, first row). The statistical significance of the difference was evaluated with CRA (Maris & Oostenveld, 2007), a nonparametric randomization test that overcomes the multiple-comparisons problem due to the 129 electrodes being analyzed by clustering together the neighboring electrodes that exhibit similar effects, and assigning a statistical p value to such clusters.

PLF significantly increased for all SOAs ($p < .02$ for SOA = 50 msec, $p < .001$ for all other SOAs), demonstrating that stimulation-dependent phase resetting occurred at all frequencies of stimulation, even above the integration threshold. This effect emerged in an occipital cluster for SOA = 50 msec, and gradually extended to almost all the scalp (SOA = 134 msec) with increasing SOA. The observed pattern is compatible with the idea that the individual component strings are always represented in occipital areas, even at the fastest SOA, and expand into higher cortical areas at the slowest SOAs when their fusion fails to occur.

To show that phase resetting is specific to the stimulation frequency, we computed the PLF differences between stimulation and baseline at all six relevant frequencies and for each of the six SOA conditions by averaging over the occipital channels common to all the clusters observed at each SOA. As shown in Figure 7 (second row), for almost all SOAs, the PLF was highest at the expected frequency f_{SOA} (red bars) compared to other frequencies (blue bars). In only one case (SOA = 100 msec) was the PLF higher at frequencies slightly higher than f_{SOA}, probably due to a distortion induced by the strong decrease in alpha power during stimulation compared to baseline. At SOA = 134 msec, PLF increased at f_{134} (7.5 Hz) but also at f_{67} (15 Hz), which is a harmonic. At SOA = 50 msec, PLF increased at f_{50} (20 Hz) but also at f_{117} (8.6 Hz) and f_{134} (7.5 Hz), most probably due to the massive P300 emerging in that condition. It is also possible that, at the short SOAs where integration was possible, phase synchrony emerged
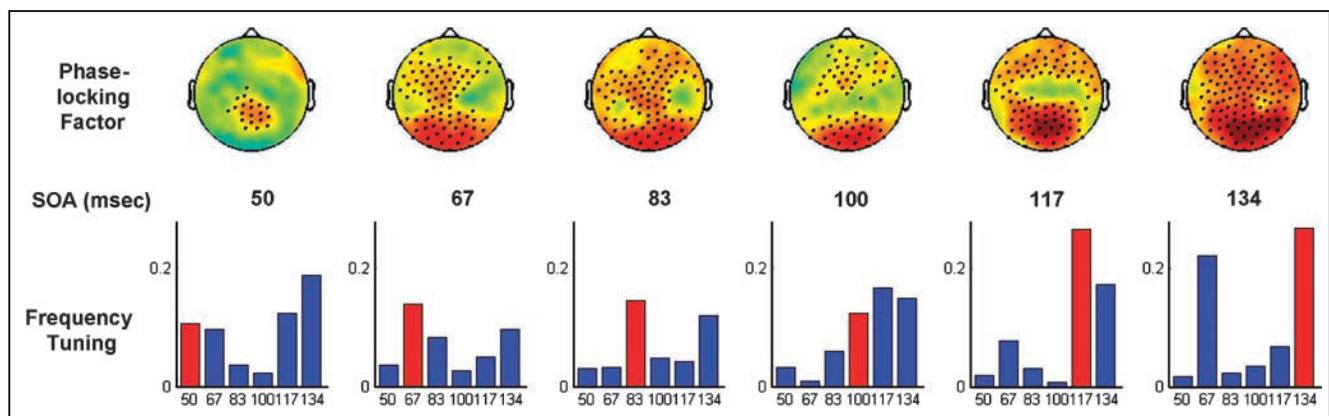

Figure 7. Phase locking of the evoked voltages at the frequency of stimulus alternation. For each SOA, a topographical plot shows the scalp map of the difference in PLF at the corresponding frequency ($f = 1000/SOA$) during stimulation versus during baseline. Channels forming statistically significant clusters are marked with black points ($p < .02$ for SOA = 50 msec; $p < .001$ for all other SOAs). An occipital cluster is present at all SOAs, indicating that the alternating stimuli always enter into occipital visual cortex, even when subjects report perceived the integrated string. Within this occipital cluster, the bar graphs in the second row show the PLF difference at the corresponding SOA frequency (red bars) and at the other SOA frequencies (blue bars).

at a frequency twice smaller than that of the components, that is, the frequency with which the whole word appeared. Indeed, such a frequency halving seems apparent at SOAs 50, 67, and 83 msec in Figure 7, and a similar effect has been observed in the auditory modality (Buiatti, Pena, & Dehaene-Lambertz, 2009).

Most importantly, however, the bar graphs in Figure 7 (second row) clearly show that the amplitude of PLF at f_{SOA} in the common occipital cluster (red bars) increased with SOA. This increase was nonlinear, as PLF became very high at the last two SOAs (117 and 133 msec) where despite the subjects' efforts, all they could see was the alternating component strings. However, the fact that occipital phase locking was still significant even at the shortest SOA (50 msec), where subjects reported close to 100% reading of words, indicates that integration is not entirely resolved by low-pass filtering at an early visual level.

GENERAL DISCUSSION

The three experiments reported here provide converging results. Both subjective and objective assessments of readability indicate the presence of a temporal threshold of about ~80 msec in the capacity to integrate alternating letter strings into a coherent percept. If the alternating components are presented for durations longer than ~80 msec, then the participants still perceive them easily, but they show a striking inability to fuse them into a coherent word, as indicated by a large effect of string length on reading times, characteristic of effortful letter-by-letter reading (Cohen et al., 2008), as well as an absence of early lexical effects in ERP recordings. Below this critical duration, the component strings are easily fused and reading of the integrated string is largely normal, as indicated by early ERP effects of lexicality and by a very small impact of word length on reading times. Yet ERP recordings indicate that the components are still encoded in visual cortex, as they induce a strong phase resetting localized to occipital electrodes. In that respect, our results are strikingly parallel to early neurophysiological findings in motion integration, which indicate that local component motion remains coded in V1 neurons even when a different global percept of motion arises in area MT (Movshon & Newsome, 1996).

The present results concur with the rapid serial visual presentation paradigm (Potter et al., 1986), which demonstrates that streams of words or pictures presented at rates of up to 12 words per second remain understandable. A similar conclusion was reached by Holcombe and Judson (2007) with a slightly different paradigm using English and Chinese words. They presented stimulus sequences in which two distinct words such as "ball" and "deck" were alternatively presented at the same location. Misperceptions (e.g., "back" or "dell") decreased sharply when the alternation slowed down, with a threshold around 5 Hz (only slightly higher than in present experiments, and also seemingly more variable across subjects). They concluded that "by the time visual signals reach awareness, they have been combined over an interval of the order of 100 msec." Alternations of nonlinguistic stimuli also suggest an integration window of ~80 msec highly comparable to the present findings (e.g., Coltheart, 1980a, 1980b; Di Lollo, 1980; Hogben & Di Lollo, 1974).

Two interpretations of temporal integration have been proposed. It might correspond to a late perceptual "decision" of letting serially presented information enter into one or two "object files" or "event files," depending on the probability that it arises from the same item or from two distinct items. Alternatively, and more trivially, it might arise from an early visual low-pass filtering operation, perhaps even retinal, which would prevent the separate encoding of fast alternating stimuli. For instance, V1 neurons are known to respond to a very brief flash of light with a temporally extended response that might, by itself, suffice to explain the perceptual fusion of successive time frames (Duysens et al., 1985; Levick & Zacks, 1970). Yet several arguments favor the higher-level interpretation. First, fast alternating visual stimuli are demonstrated separated in early visual cortex. For instance, in three human patients with an intracranial recording, Krolak-Salmon et al. (2003) demonstrated that computer screen flicker, although occurring at rates of 60 Hz and above, caused corresponding high-frequency oscillations in lateral geniculate, V1 and V2 activity—a result confirmed by single-cell recordings in monkeys (Gur & Snodderly, 1997). Likewise, we observed occipital phase resetting even at the highest frequency tested (20 Hz), indicating that the alternating components that the participants failed to see were still separated in early visual cortex. This observation is compatible with previous evidence that extrastriate and inferotemporal cortices are selectively responsive to extremely brief visual stimuli such as a 16-msec stimulus surrounded by visual masks (Keysers, Xiao, Foldiak, & Perrett, 2001; Kovacs, Vogels, & Orban, 1995), although this response is blocked from accessing higher processing levels (Del Cul, Baillet, & Dehaene, 2007).

A second argument in favor of a late cortical mechanism of temporal integration arises from the finding of lexical influences on visual integration and persistence. In Experiments 2 and 3, we consistently observed that words resisted slower rates of temporal segregation than pseudowords, although the two stimulus sets were carefully matched for orthographic content. Thus, the compatibility of the alternating stimulus with a lexical template seems to bias the integration/segregation dilemma, which would not be possible if it was entirely resolved by early visual low-pass filtering.

One last and more tentative argument stems from the fact that the integration threshold for words was higher in Experiments 2 and 3 than in Experiment 1. This is particularly obvious when comparing the readability reports of the whole-word condition in Experiments 1 and 3 (compare the middle panel of Figure 1 with Figure 4). This condition itself was strictly identical, but a key difference was the context in which it occurred. In Experiment 1,

participants knew that they sometimes had to read the two component words rather than the integrated stimulus, whereas in Experiments 2 and 3 they knew that only the integrated string mattered, and therefore, presumably made a greater effort to integrate rather than segregate the stimuli. Thus, these results open the possibility that attentional effort increases the integration window. Although the present results are only suggestive in this respect, Visser and Enns (2001) demonstrated formally that attention impacts on the duration of visual persistence, increasing the temporal lag at which two successive stimuli could be integrated. Similarly, attention decreases the ability to discriminate two successive flashes (Yeshurun & Levy, 2003) and increases the perceived duration of a brief flash of light or of a temporal gap (Enns et al., 1999). In a related line of research, Noguchi and Kakigi (2008) recently demonstrated that the flash-lag illusion could be shifted when the stimulus traced a Kanji character, whereas no such effect was found in non-Japanese readers or with pseudo-Kanji. Such effects indicate that the perceived temporo-spatial scene does not result from rigid low-level bottom–up processes, but rather involves an active integration process that attempts to make the best sense of incoming inputs, given prior knowledge and sensory evidence (see also Purves, Lotto, Williams, Nundy, & Yang, 2001; Andrews, White, Binder, & Purves, 1996; Dixon & Di Lollo, 1994).

The parallels and differences between temporal integration and masking will be interesting to pursue in further research (see also Groner, Groner, Bischof, & Di Lollo, 1990; Fraisse, 1966). When target–mask SOA is varied, perception of the target typically follows a nonlinear sigmoid curve, with a critical SOA of about 50 msec (Del Cul et al., 2007), slightly shorter than the present critical time of about 80 msec. The masking threshold can also be shifted by higher-level factors such as whether the masked stimulus is an emotional word (Gaillard et al., 2006). These parallels suggest that masking and temporal integration phenomena may be related. Both may arise from the brain's attempt to interpret the observed sequence of visual events as a single image or two (Di Lollo, Enns, & Rensink, 2000; Giesbrecht & Di Lollo, 1998; Dixon & Di Lollo, 1994; Di Lollo et al., 1994), and to converge into fused or differentiated neural assemblies (Dehaene & Changeux, 2005). In masking, a long target–mask delay allows the target to be perceived independently from the mask and to activate highly distributed cortical areas including prefrontal cortex (Del Cul et al., 2007). In the present temporal integration paradigm, likewise, at the slowest SOAs, we found that phase locking extended to the entire scalp (Figure 7), suggesting that the alternating component strings each activated extended areas, thus allowing them to be individually perceived at a conscious level. Conversely, at short SOAs, a form of metacontrast masking may occur, thus explaining that each component string interrupts the processing of the other, eventually preventing both of them from being perceptible. Note,

however, that such masking cannot explain the full pattern of our results. If the alternating letters merely masked each other, this should also decrease the visibility of the whole string—but we find the greatest readability of the whole word in these short SOA conditions. Obviously, masking does not merely act by reducing visibility, but also by increasing the integration of successive stimuli.

In the 1970s, Coltheart and Arthur (1972) and Schultz and Eriksen (1977) argued that the masking that occurs at short target–mask SOAs results, at least in part, from an integration of the mask with the target. Experimentally, they demonstrated that perception of a brief target followed by a second shape could either decrease or increase with SOA (corresponding respectively to integration vs. masking curves), depending on whether the second shape competed with the first or could be integrated with it. They proposed that at short SOAs, integration is seen when the two stimuli are sufficient compatible to receive a single, unified interpretation, whereas masking dominates when the mask is incompatible with the target, thus interrupting (Del Cul et al., 2007; Kovacs et al., 1995) or substituting it (Giesbrecht & Di Lollo, 1998). At long SOAs, the integration/segregation dilemma ceases to arise as the visual system receives enough evidence that two successive events occurred.

Relation to Theories of Visual Word Recognition

As noted in the Introduction, the present methodology also speaks to models of the visual word recognition process. The results argue strongly against theories of temporal coding according to which the precise time-of-arrival or phase of the neuronal responses should play a determinant role in the grouping of visual stimuli. At the highest frequency used in our experiments (20 Hz), subjects easily recognized the integrated stimulus and showed evidence of normal reading, despite receiving half of the letters 50 msec later than the others. Yet, this is still a large delay compared to the gamma-band range (30–100 Hz) typically assumed for cortical fusion or synchronous binding, and it is hard to see how such a temporal disruption in the input should not create havoc in any temporal coding scheme. In Whitney's (2001) model, for instance, successive letter positions are assumed to be encoded by a 25-msec firing delay. By delaying every other letter by 50 msec, our stimulation paradigm should have resulted in a dramatic misperception of letter order, rendering the word unreadable. Likewise, the SOLAR model of Davis and Bowers (2006) and Davis (1999) assumes that serial "beats" spaced every 10 msec, one for each successive letter, are used to encode the spatial layout of incoming strings by a gradient of activation. Again, a 50-msec delay of half of the letters should have prevented this temporal-to-spatial conversion.

More generally, the slowness of the temporal integration window (~80 msec) is hard to reconcile with models

of visual recognition or scene segregation that attribute a key role to the precise temporal ordering of spikes (VanRullen & Thorpe, 2002) or to their precise synchronization in the gamma band (Engel, Konig, & Singer, 1991). These models might be salvaged only if it assumed that the antiphase input oscillation is first filtered out early on in the visual system, and then replaced by an endogenous coding rhythm, largely independent in its phase and frequency from the current input. It remains to be seen whether such a model could be viably developed, given the evidence from Experiment 3 that the input oscillation was not filtered by the retina, but was demonstrably present in occipital cortices.

Other connectionist models of reading postulate that the input strings are encoded by a bank of letter detectors that then feed into higher-level word detectors (Coltheart et al., 2001; Harm & Seidenberg, 1999; Ans et al., 1998; Zorzi et al., 1998; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981) or bigram detectors (Grainger et al., 2006; Dehaene et al., 2005; Grainger & Whitney, 2004; Mozer, 1987). Our data are compatible with these models, but place constraints on the input letter buffer. If the buffer had a very slow decay constant (e.g., a few hundreds of milliseconds, as assumed in iconic memory models), it is not clear why such an input code should be severely disrupted by alternation delays of 80 msec or more. The input buffer would merely have to wait until the two successive letter subcomponents have been presented, at which point it would be in possession of the full letter information. The data, therefore, require the postulation of a short-duration letter buffer. With a short decay constant, there would be very little activation remaining of the first substring when the second string appears after 80 msec or more. As a result, higher-level word or bigram units would never "see" a joint letter array, but only the component strings. This hypothesis could be tested physiologically by probing the state of activity of regions thought to correspond to the bank of letter detectors (Dehaene et al., 2004, 2005).

The results fit most easily within recent models that postulate a role for intermediate graphemic units such as bigrams (pairs of letters) in visual word recognition (Grainger et al., 2006; Dehaene et al., 2005; Grainger & Whitney, 2004). At slow alternation rates, a bigram code would be drastically disrupted by our stimulus, which segregates the odd and even letters into distinct substrings. More experimentation, however, will be needed to probe the specific predictions of this bigram model. One such prediction is that the temporal disruption of frequent letter bigrams should have a greater impact than an equivalent disruption of rare bigrams. The present research merely lays the methodological ground for testing such predictions.

REFERENCES

Andrews, T. J., White, L. E., Binder, D., & Purves, D. (1996). Temporal events in cyclopean vision. *Proceedings of the National Academy of Sciences, U.S.A., 93,* 3689–3692.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105,* 678–723.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

Buiatti, M., Pena, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage, 44,* 509–519.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage, 40,* 353–366.

Coltheart, M. (1980a). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183–228.

Coltheart, M. (1980b). The persistences of vision. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 290,* 57–69.

Coltheart, M., & Arthur, B. (1972). Evidence for an integration theory of visual masking. *Quarterly Journal of Experimental Psychology, 24,* 262–269.

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.

Davis, C. J. (1999). *The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition.* Australia: University of New South Wales.

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 535–557.

Dehaene, S., & Changeux, J. P. (2005). Ongoing spontaneous activity controls access to consciousness: A neuronal model for inattentional blindness. *PLoS Biology, 3,* e141.

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

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychological Science, 15,* 307–313.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5,* e260.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General, 109,* 75–97.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129,* 481–507.

Di Lollo, V., Hogben, J. H., & Dixon, P. (1994). Temporal integration and segregation of brief visual stimuli: Patterns of correlation in time. *Perception & Psychophysics, 55,* 373–386.

Dixon, P., & Di Lollo, V. (1994). Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing. *Cognitive Psychology, 26,* 33–63.

Duysens, J., Orban, G. A., Cremieux, J., & Maes, H. (1985). Visual cortical correlates of visible persistence. *Vision Research, 25,* 171–178.

Efron, R. (1970a). Effect of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics, 8,* 231–234.

Efron, R. (1970b). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia, 8,* 37–55.

Engel, A. K., Konig, P., & Singer, W. (1991). Direct physiological evidence for scene segmentation by temporal coding. *Proceedings of the National Academy of Sciences, U.S.A., 88,* 9136–9140.

Enns, J. T., Brehaut, J. C., & Shore, D. I. (1999). The duration of a brief event in the mind's eye. *Journal of General Psychology, 126,* 355–372.

Eriksen, C. W., & Collins, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology, 74,* 476–484.

Fraisse, P. (1966). Visual perceptive simultaneity and masking of letters successively presented. *Perception & Psychophysics, 1,* 285–287.

Gaillard, R., Del Cul, A., Naccache, L., Vinckier, F., Cohen, L., & Dehaene, S. (2006). Nonconscious semantic processing of emotional words modulates conscious access. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 7524–7529.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1454–1466.

Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & van Heuven, W. J. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 865–884.

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

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences, 8,* 58–59.

Groner, R., Groner, M. T., Bischof, W. F., & Di Lollo, V. (1990). On the relation between metacontrast masking and suppression of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 381–390.

Gur, M., & Snodderly, D. M. (1997). A dissociation between brain activity and perception: Chromatically opponent cortical neurons signal chromatic flicker that is not perceived. *Vision Research, 37,* 377–382.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106,* 491–528.

Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research, 14,* 1059–1069.

Holcombe, A. O., & Judson, J. (2007). Visual binding of English and Chinese word parts is limited to low temporal frequencies. *Perception, 36,* 49–74.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object specific integration of information. *Cognitive Psychology, 24,* 175–219.

Keysers, C., Xiao, D. K., Foldiak, P., & Perrett, D. I. (2001). The speed of sight. *Journal of Cognitive Neuroscience, 13,* 90–101.

Kovacs, G., Vogels, R., & Orban, G. A. (1995). Cortical correlate of pattern backward masking. *Proceedings of the National Academy of Sciences, U.S.A., 92,* 5587–5591.

Krolak-Salmon, P., Henaff, M. A., Tallon-Baudry, C., Yvert, B., Guenot, M., Vighetto, A., et al. (2003). Human lateral geniculate nucleus and visual cortex respond to screen flicker. *Annals of Neurology, 53,* 73–80.

Lavidor, M., & Ellis, A. W. (2002). Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. *Brain and Language, 80,* 45–62.

Levick, W. R., & Zacks, J. L. (1970). Responses of cat retinal ganglion cells to brief flashes of light. *Journal of Physiology, 206,* 677–700.

Loftus, G. R., & Irwin, D. E. (1998). On the relations among different measures of visible and informational persistence. *Cognitive Psychology, 35,* 135–199.

Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., et al. (2002). Dynamic brain sources of visual evoked responses. *Science, 295,* 690–694.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190.

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

Movshon, J. A., & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience, 16,* 7733–7741.

Mozer, M. C. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 83–104). Hillsdale, NJ: Erlbaum.

Noguchi, Y., & Kakigi, R. (2008). Knowledge-based correction of flash-lag illusion. *Journal of Cognitive Neuroscience, 20,* 513–525.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General, 137,* 434–455.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 115,* 281–294.

Purves, D., Lotto, R. B., Williams, S. M., Nundy, S., & Yang, Z. (2001). Why we see things the way we do: Evidence for a wholly empirical strategy of vision. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 356,* 285–297.

Schultz, D. W., & Eriksen, C. W. (1977). Do noise masks terminate target processing? *Memory & Cognition, 5,* 90–96.

Stroud, J. M. (1955). The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology* (pp. 174–207). Glencoe, IL: Free Press.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience, 16,* 4240–4249.

VanRullen, R., & Thorpe, S. J. (2002). Surfing a spike wave down the ventral stream. *Vision Research, 42,* 2593–2615.

Visser, T. A., & Enns, J. T. (2001). The role of attention in temporal integration. *Perception, 30,* 135–145.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology, 50A,* 439–456.

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,* 221–243.

Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science, 14,* 225–231.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1131–1161.