Consciously detecting and recognizing a stimulus without knowing what it looks like

Daphné Rimsky-Robert[1,2], Matteo Lisi[3], Camille Noûs[4], Claire Sergent[1]

[1] Integrative Neuroscience and Cognition Center - UMR 8002 CNRS/Université de Paris

[2] Brain institute—ICM, Inserm U1127, CNRS UMR 7225, Sorbonne Université, Paris, France

[3] Department of Psychology, University of Essex, Colchester, UK

[4] Laboratoire Cogitamus

Corresponding Author: Daphné Rimsky-Robert, address: INCC bureau H418b, 45 rue des Saints-Pères 75006 PARIS, +3360850218, e-mail: drr@tuta.io

Keywords: Consciousness, perception, sensory processing

Category: Biological Science, Cognitive and Psychological Science

Abstract

Does conscious perception occur during initial sensory processing, or does it arise later in a supra modal fashion? If conscious access truly depends on supra modal processes, we may be able to induce "asensory perception", where only the semantic features of a meaningful stimulus are accessed, untied to its sensory attributes. Here we tested this prediction by degrading the low-level sensory representations of visual words in the brain using pattern masking, and subsequently presenting audio words that were either semantically related to the masked word or not. We hypothesized these retrospective semantic cues would reactivate the remaining traces of the masked word in the brain, and induce awareness of any information that was not disrupted by masking. In three separate experiments we show that, when presented with retrospective cues that are semantically related to the masked word, participants are better at detecting the presence of the preceding masked word and naming it, while at the same time being unable to report its visual features. In other words, participants could consciously detect and recognize the preceding masked word, without knowing what it looked like. These findings suggest that non-sensory information can be consciously accessed in relative independence from the build-up of sensory representations.

Introduction

Becoming conscious of a stimulus usually means becoming aware of all its different attributes, not only its physical features but also its more abstract attributes such as its meaning(Baars, 2005; Dehaene et al., 2006; Dehaene & Naccache, 2001; Kouider et al., 2010; V. A. Lamme et al., 2000; V. A. F. Lamme & Roelfsema, 2000). A much-debated issue is whether conscious perception occurs during the initial phase of stimulus processing where these different representations are constructed within distinct brain areas(Fisch et al., 2009; V. A. F. Lamme & Roelfsema, 2000; Pitts et al., 2014), or whether it corresponds to a secondary supra-modal process that amplifies and broadcasts existing representations to a global workspace(Baars, 2005; Dehaene & Naccache, 2001). If conscious access is truly independent from early sensory processing, then an extreme prediction is that we could consciously perceive abstract representations whilst failing to register low-level features. In this study we tested this prediction taking advantage of a phenomenon called retro-perception. Retro-perception is said to occur when conscious access to an initially unconscious stimulus is triggered retrospectively after its disappearance(Rimsky-Robert et al., 2019; Sergent, 2018; Sergent et al., 2013; Thibault et al., 2016; Xia et al., 2016), by means of a retro-cue that reactivates the latent representations left by the stimulus. We reasoned that if, before the retro-cue, we erased the low-level sensory representations of a stimulus using visual masking, then the retro-cue could induce "asensory perception", comprising only an abstract representation of the stimulus. In our protocol, visual words were rendered unconscious using visual masking(Enns & Di Lollo, 2000; V. A. F. Lamme & Roelfsema, 2000; Weiskrantz et al., 1995). They were then followed by an auditorily presented word, which served as a semantically congruent or incongruent retro-cue(Bernstein et al., 1989; Briand et al., 1988; Dark, 1988; Neely, 1991; Thomas et al., 2012; VanVoorhis & Dark, 1995) through associative priming(e.g. "porcupine" – "hedgehog"). If the auditory cue primed the relevant semantic links(Collins & Loftus, 1975) after the occurrence of the target, it should specifically boost semantic processing and make this feature of the stimulus salient, whilst the visual masks should strongly degrade the target's low-level visual features and make them inaccessible to consciousness(Enns & Di Lollo, 2000; V. A. F. Lamme & Roelfsema, 2000; Weiskrantz et al., 1995).

Figure 1 illustrates our hypotheses and predictions. Once a visual word is presented, it is quickly processed through a hierarchy of processing levels, from the extraction of local edges (V1), to the recognition of letter forms (V2-V4-V8), groups of letters (Occipito-Temporal Sulcus, OTS), and finally up to the activation of phonological and semantic representations(Cohen & Dehaene, 2004; Dehaene et al., 2005; Vinckier et al., 2007) **(Figure 1a)**. The presentation of a pattern mask shortly after the stimulus degrades or replaces the low-level visual representation of the stimulus, presumably up to V8(Callaway, 1998; Enns & Di Lollo, 2000; Felleman & Van, 1991; V. A. F. Lamme et al., 2002; V. A. F. Lamme & Roelfsema, 2000), while sparing the higher level semantic and phonological representations, as suggested notably by experiments on masked priming (Abrams & Grinspan, 2007; Fowler et al., 1981; Humphreys et al., 1982; Kouider & Dehaene, 2007; Naccache & Dehaene, 2001) **(Figure 1b)**. Thus, after the word has been masked, the brain presumably retains a trace of the semantic and phonological attributes of the stimulus, but no trace or a degraded trace of its physical attributes. The subsequent presentation of a semantic context, such as that provided by a semantically congruent word, should reactivate the semantic trace left by the masked word and thus improve its identification, as suggested by previous studies using retrospective semantic priming(Bernstein et al., 1989; Briand et al., 1988; Dark, 1988; Neely, 1991; Thomas et al., 2012; VanVoorhis & Dark, 1995) **(Figure 1c)**. Our prediction is that such reactivation of a latent semantic representation could trigger conscious access to this representation. Semantic cues presented before a masked object have been shown to boost perception and increase the probability of the visual object reaching awareness(Boutonnet & Lupyan, 2015; Lupyan & Ward, 2013). In the case of retro-cueing, however, the cue occurs after the target has been presented and masked: at this stage, there is a mismatch in the information present at high and low levels of processing, the first containing information about the word, and the second information about the mask. This would leave the semantic representation stranded, and broadcasted on its own to a global workspace of areas including fronto-parieto-cingulate hubs of consciousness **(Figure 1d)**.

We tested our hypotheses in three separate experiments. All experiments included the presentation of visual words with varying physical features, embedded within a sequence of masks, and followed by an auditory word-cue that could be semantically congruent or incongruent with regards to the visual word. In experiments 1 and 3 we used letter-casing as our target visual feature, while in Experiment 2 we used location-on-

screen. All Experiments showed that congruency of the auditory cue affected the probability of correct meaning-identification and correct detection, but not that of correct visual-feature-discrimination. Overall, our results indicate that congruent semantic retro-cues, provided after a masked visual target word, lead to decoupling the awareness of the presence and meaning of the word and the awareness of its visual attributes.

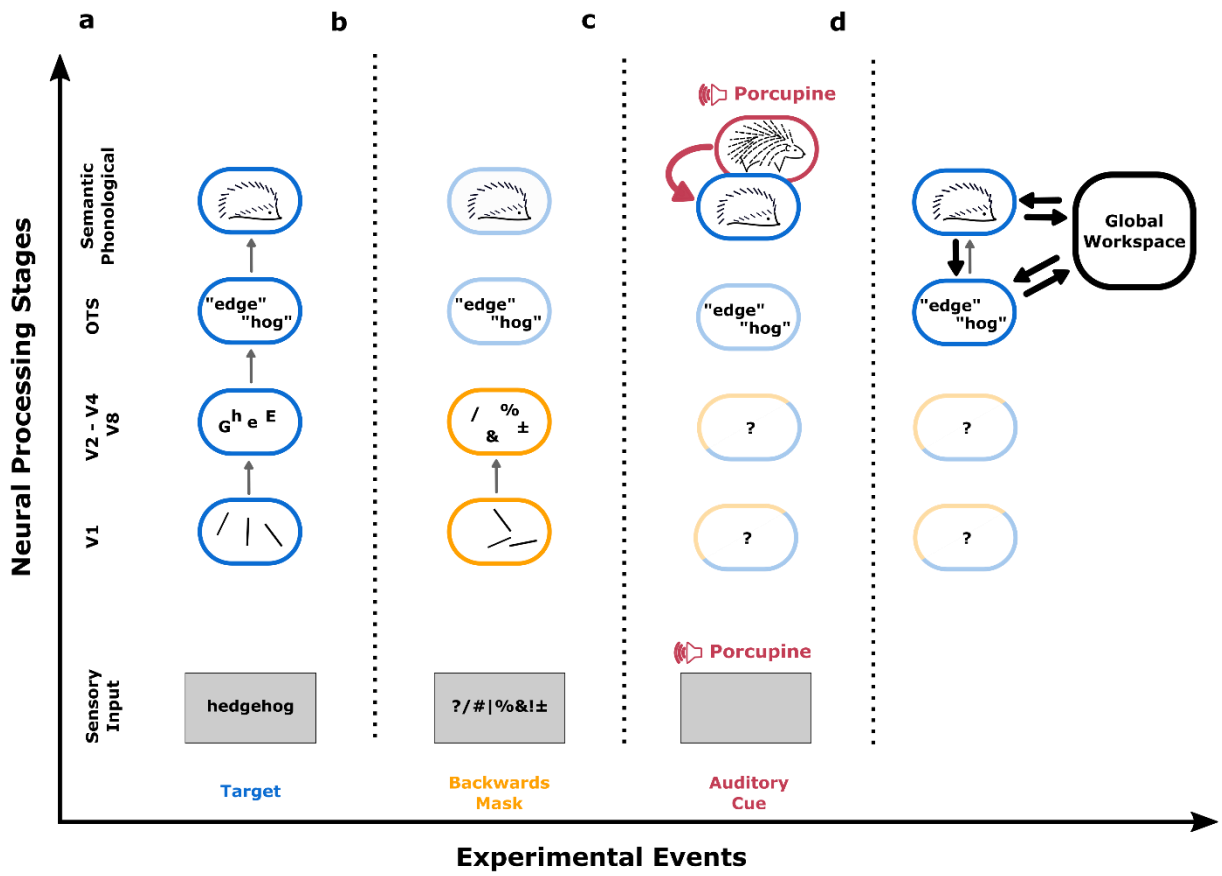

Figure 1: Hypothesis. *Visual word processing is organized hierarchically along the ventral pathway up to semantic processing. (**a**). The presentation of a backwards mask which shares most visual properties of the target word would overwrite all buffered information up to the OTS. However it should not affect information retained at higher levels (semantic, phonological) since it contains no such information (**b**). The presentation of a contextually congruent auditory cue may reactivate the semantic trace left by the masked word (**c**) and bring it to consciousness, for example via broadcasting through the fronto-parietal Global Workspace*(Baars, 2005; Dehaene & Naccache, 2001)*, leading to strong recurrent processing and consciousness of the remaining information. Given the interference of the backwards mask, only the high-level representations would still be available for broadcasting (**d**).*

Results

Experiment 1: word identification, case discrimination and visibility rating

In order to test our hypothesis, in Experiment 1 we presented 15 participants with masked words that could last 12, 24, 36 or 48ms, thus randomly varying the strength of masking across trials (**Figure 2a**). Target-absent catch trials were included in the experiment (in the same proportion as other sub-conditions, namely 11% of the trials) in order to test participants' ability to simply detect the presence of a word. Each stimulus sequence was followed by an auditory cue, which could be semantically associated to the target word (congruent trials) or not (incongruent trials) with equal probability (**Figure 2a**). After each trial, participants had to report, in this order: the identity of the target word (by speaking into a microphone), the word's case (upper or lower) using the keyboard, and rate the overall visibility of the stimulus on a 9-points scale with 0 being "I have not seen anything but the masks", and 8 "I have perfectly seen the target word" (see **Supplementary Note 1** for the full instructions). In total each participant was presented with 432 different visual words with no repetition (see **Supplementary Table 1** and **Supplementary Table 2**). Congruency and word-pairs were balanced within and across participants as described in the Methods section.

In all 3 experiments, word identification performance and case or position discrimination were jointly analyzed using a multivariate hierarchical Bayesian modeling approach that included a correction for guessing (see below, and **Supplementary Note 2**). Using a multivariate approach(Cox, 1972) allow to model both responses simultaneously and examine their trial-by-trial correlation. Visibility ratings and detection sensitivity were analyzed using standard univariate analyses (see Methods). For the latter approach, Bonferroni corrections for multiple comparisons were used throughout.

An important step in analyzing word identification performance was to correct for guesses. Indeed, given that our task involves semantically related retro-cues, a non-negligible fraction of correct identifications of the target word in congruent trials might actually correspond to 'guesses': participants will tend to report words from the same

semantic field as the cue and thus could name the target correctly just by chance. These guessing probabilities can be directly estimated from the errors made in incongruent trials. Indeed, in incongruent trials when participants report the semantically associated word in the list instead of the target word, this corresponds to an error, and it is thus not confounded with correct responses. The rate of "semantically associated errors" in incongruent trials provides an estimate of the participant's tendency to name the semantically associated word even when this word was not presented. This rate was estimated for each participant at each target duration (with longer-lasting targets, they are likely to rely less on the cue and more on visual information, see **Supplementary Figure 1**), then used to constrain estimates of guessing rates in the hierarchical Bayesian model. The correction performed by the model can be directly observed in **Figure 2b, 3b and 4b**, by comparing the data indicated with circles with the model predictions shown by a shaded line: in congruent trials, model predictions are systematically lower than the observed proportion of correct responses as they are corrected for guesses. Unless indicated otherwise, in the following estimates from the model or their Bayesian credible intervals (BCI) are reported as odds-ratios: for example, an interval of [5 20] for the effect of congruency would indicated that the probability of a correct response was between 5 and 20 times more likely in the congruent condition than the incongruent one.

1.A: Word identification

Modeling performance for word identification **(Figure 2b)** revealed a positive effect of congruency (95%BCI [4.31, 20.91]), and target duration (95%BCI [3.13, 4.76]; expected change in odds-ratio of correct identification for each additional 12ms of presentation duration), as well as a significant interaction (95%BCI [0.69, 0.96]), suggesting that congruency had less influence when the words were presented for longer durations and were thus easier to see. Considering each duration separately, the hierarchical Bayesian model predicted a robust, positive effect of word congruency on the accuracy of word identification model for all durations in favor of higher performance in the congruent condition (difference in expected accuracy, congruent minus incongruent: 12ms, 95%CI [0.018, 0.07]; 24ms, 95%BCI [0.06, 0.17]; 36ms, 95%BCI [0.12, 0.28]; 48ms, 95%BCI [0.04, 0.19]).

1.B: Visibility Ratings

We next examined whether the increased performance was associated with greater awareness of the visual stimulus. We first conducted a repeated measures two-way ANOVA on mean subjective visibility ratings obtained on the visibility scale (**Supplementary Figure 2**), which revealed a significant increase in visibility with target duration ($F(1.23,17.29)=81.34$, $p<0.001$) and congruency ($F(1.00,14.00)=38.43$, $p<0.001$), with an interaction between these two factors ($F(3.00,42.00)=15.812$, $p<0.001$). These results suggest that participants subjectively felt that they were seeing the target word better when it was followed by a congruent cue. We next used the distributions of visibility ratings to study participants' objective ability to report whether a word has been presented or not: detection sensitivity was derived by computing the area under the Receiver Operating Characteristics curve (AROC) i.e. when comparing the distribution of ratings for stimulus-present versus stimulus-absent trials in each condition **(Figure 2c)**. A repeated measures two-way ANOVA revealed a significant increase in detection sensitivity with target duration ($F(3.00,42.00)=92.00$, $p<0.001$), and congruency ($F(1.00,14.00)=15.79$, $p<0.005$) with an interaction between these two factors ($F(3.00,42.00)=3.62$, $p<0.05$). Bootstrapped 95% confidence intervals of the mean difference between congruent and incongruent conditions revealed a significant difference in performance for target durations of 36ms and 48ms, and a trend in the same direction for target duration of 24ms (target dur. 12ms: CI = [-0.03, 0.04], 24ms: CI = [-0.004, 0.08], 36ms: CI = [0.04, 0.1], 48ms: CI = [0.01, 0.05]). This suggests that beyond improving performance at the word identification task, congruent retrospective cueing increased the probability of consciously detecting the masked word.

1.C: Case discrimination

Modeling performance for case discrimination **(Figure 2d)** revealed a positive effect of target duration on performance (95%BCI [1.57, 2.13]), but no effect of congruency (95%BCI [0.81, 1.55]), nor any interaction between these factors (95%BCI [0.88, 1.18]). This result, in agreement with our prediction, contrasts with the positive effect of congruency on word identification, subjective visibility, and detection sensitivity.

1.D: Task interactions

Our multivariate modeling approach enabled us to jointly model the word identification and case discrimination responses, as well as examine their trial-by-trial correlation. Importantly this allows us to assess how this correlation changes with congruency, which can be read as an interaction between congruency and accuracy at the second task. The interaction factor between congruency and accuracy at the second task was significant, in favor of a smaller effect of congruency on word identification in trials with correct case discrimination response (95%BCI [-1.12, -0.41], change in log-odds of correct identification). The trial-by-trial association between the accuracies of the two responses is quantified by the parameter ρ (**Figure 2e,** see **Supplementary Note 2**), which describes how likely participants are to either be correct/incorrect at both tasks, rather than be correct at one and wrong at the other (more precisely, ρ is the logarithm of the odds that the two responses are either both correct or both wrong in the same trial). For easier interpretation, we report the results of this analysis as the probability of having both objective responses correct or incorrect in each trial (response consistency probability, RCP). In congruent trials, participants were less likely than chance to provide answers that were either both correct or both wrong (95%BCI [0.29, 0.48]). We found a trend in the opposite direction for incongruent trials (95% BCI [0.44, 0.69]). The responses tended to be both correct or wrong more frequently in incongruent trials. Importantly, we found a significant difference between the two conditions, in favor of higher values of response consistency probability in the incongruent condition (95%BCI of the difference [-0.27, -0.10]). These results point to a decorrelation between participants' ability to report visual feature and word meaning specifically when the cue was congruent.

Our last analysis further investigated the origin of this decorrelation. According to our hypothesis, congruent retro-cues trigger conscious access to the latent semantic representation of a previously unconscious word, while conscious access to its visual features is no longer possible due to masking. In this hypothesis, the decorrelation between the two tasks in congruent trials is due to these "retro-perceived" trials on which the word is correctly identified, but its case not. In order to test for this proposition, we computed the conditional probability of correctly reporting the word's case when its

meaning was correctly identified for congruent and incongruent cues respectively. For this analysis, since these conditional probabilities require to examine the success or failure at both task trial by trial, we corrected for guesses in congruent trials on an individual trial basis, by using participant's trial by trial visibility ratings: for each participant we determined a visibility threshold above which we could consider that most guesses had been excluded (see details in the Methods section) **(Figure 2f)**. Performance for case discrimination on these "word-correct" trials was high for incongruent cues (above 90%). It significantly decreased for congruent cues (95%CI of the mean difference in %correct = [-3.33, -10.00]). Since the estimated performance in congruent is a mixture of trials that would have been perceived with no help of the cue (as in incongruent trials) and retro-perceived trials, we also computed an estimate of performance in retro-perceived trials only (see details in the Methods section and **Supplementary Note 3 and Supplementary Figure 3**). This estimate goes down to 69.2% average correct case discrimination, with a confidence interval including 50% (95%CI = [48.85, 78.42]), compared to 86.3% in all congruent and 92.9% in incongruent trials. This suggests that, on those retro-perceived trials where the visual word was identified thanks to the retrospective influence of the congruent cue, participants not significantly better than chance at reporting the word's case.

In conclusion, the results of this first experiment are consistent with the prediction that a retrospective semantic context can yield awareness of an "asensory" word, devoid of visual attributes.

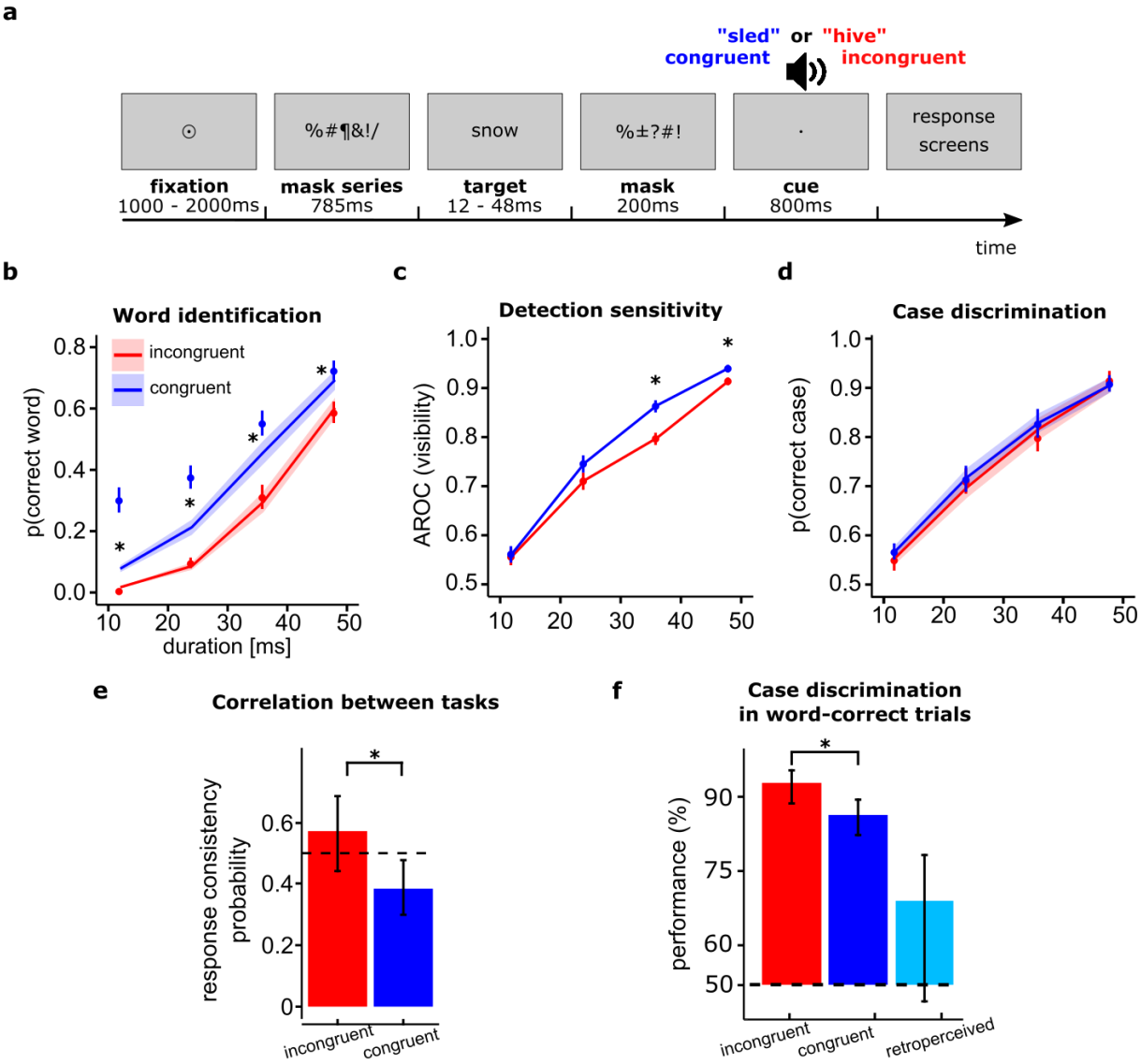

Figure 2: Experiment 1. *Procedure. The response screens were: (1) word identification, (2) case discrimination, (3) overall visibility rating* ***(a)****. Word identification performance: points are the uncorrected data, lines represent performance estimated by the model after correcting for guesses. Error bars are SE of the mean across observers, and error bands are the SE of model predictions across observers.* ***(b)*** *Detection sensitivity derived from the AROC of visibility ratings: error bars are standard errors of the mean across observers.* ***(c)*** *Case discrimination performance: lines represent performance expected by the model, points represent the uncorrected data. Error bars are SE of the mean across observers, and error bands are the SE of model predictions across observers.* ***(d)****. Correlation between tasks expressed as probability of having both responses correct or incorrect. Error bars are 95% Bayesian credible intervals.* ***(e)****. Case discrimination performance when word identification is correct, with a specific estimate for retro-perceived trials. Error bars represent 95% bootstrapped confidence intervals.* ***(f)*** *Stars indicate significance of the statistical tests as described in the text.*

Experiment 2: word identification, position discrimination and visibility ratings

In order to test the generality of the effect evidenced in Experiment 1, we conducted a second experiment testing another fundamental visual property of the stimulus: its location in space. In experiment 2, case discrimination was replaced by position on-screen, with two streams of masks appearing above and below the central fixation point. Across trials the target appeared randomly in one or the other stream **(Figure 3a)**. In this experiment, we also reversed the order of the questions, asking participants to report the visual feature first, then word identity, so as to make sure that the observed dissociation between the two tasks is not dependent on task-order. Finally, we also wanted to make sure that visibility ratings truly assessed the sensation of having seen the target word, and not some form of confidence in the word identification response. To this aim, we asked participants to provide separate visibility ratings for each task: they were instructed to report the position of the target word, then read the target out loud into a microphone, and provide a visibility rating after each response. The instructions for the two visibility scales put different emphasis on the position versus the identity of the word, but both requested that the participant report the sensation of having seen a word versus not at all (distinction between zero visibility and non-zero visibility, see **Supplementary Note 1** for the full instructions). Eighteen participants took part in this experiment. To check that participants did not direct their attention to one stream of letters as a strategy to improve their detection performance (report the stimulus as being in the other position if its wasn't detected in the attended one), we analyzed stimulus absent trials first: if participants only paid attention to a single stream, they would report all stimulus-absent trials as originating from the unattended position. Participants responded that the target was "up" on average on 43.8± 16.4% of catch trials. Individual values ranged from 29.2% to 77% which did not, in our opinion, reflect more than individual bias toward one key or location in choosing a response. **Figure 3b-f** shows the results of Experiment 2.

2.A: Word identification & visibility

Modeling performance for the word identification task **(Figure 3b)** revealed a positive effect of congruency (95%BCI [7.14, 56.71]), and target duration (95%BCI [2.76, 4.63]). Their interaction was significant, indicating a decrease in the effect of congruency with target duration (95%BCI [0.53, 0.92]). The expected accuracy of word identification was systematically different across congruent and incongruent trials for all durations in favor of higher performance in the congruent condition as shown in **Figure 3b** (12ms: 95%BCI [0.005, 0.03]; 24ms: 95%BCI [0.02, 0.07]; 36ms: 95%BCI [0.04, 0.15]; 48ms: 95%BCI [0.02, 0.19]). These results replicate those of Experiment 1.

A repeated measures two-way ANOVA on the visibility ratings for word identification (**Supplementary Figure 4**) revealed a significant increase with target duration ($F(1.19,20.28)=41.70$, $p<0.001$) and congruency ($F(1.00,17.00)=9.57$, $p<0.001$), with an interaction between these two factors ($F(3.00,51.00)=24.05$, $p<0.001$). Analysis of the AROC for detection sensitivity **(Figure 3c)** with a repeated measures two-way ANOVA revealed a significant increase in detection sensitivity with target duration ($F(3.00,51.00)=38.84$, $p<0.001$), and congruency ($F(1.00,17.00)=9.24$, $p<0.01$) with an interaction between these two factors ($F(3.00,51.00)=8.51$, $p<0.001$). Bootstrapped 95% confidence intervals of the mean difference between congruent and incongruent conditions revealed a significant difference in performance for target durations of 36 and 48ms, after correction for multiple comparisons (target dur. 12ms: CI = [-0.05, 0.01], 24ms: CI = [-0.02, 0.04], 36ms: CI = [0.02, 0.09], 48ms: CI = [0.03 0.09]). As in Experiment 1, congruent retro-cues improved the ability of participants to report the identity of the masked word as well as detect whether it had been presented.

2.B: Case discrimination & visibility

Modeling performance for position discrimination **(Figure 3d)** revealed a positive effect of target duration on performance (95%BCI [1.53, 1.91]), but no effect of congruency (95%BCI [0.86, 1.51]), nor any interaction between these factors (95%BCI [0.87, 1.10]) as in Experiment 1, but with another physical attribute of the target.

We performed the same analysis on the visibility ratings for position discrimination, to make sure that this effect truly reflected visibility and not some form of confidence in word identification. A repeated measures two-way ANOVA on these

visibility ratings (**Supplementary Figure 4**) revealed a significant increase with target duration ($F(1.35,22.94)=64.10$, $p<0.001$) and congruency ($F(1.00,17.00)=30.30$, $p<0.001$), with an interaction between these two factors ($F(3.00,51.00)=6.72$, $p<0.001$). Analysis of the AROC for detection sensitivity **(Supplementary Figure 5)** with a repeated measures two-way ANOVA revealed a main effect target duration: $F(3.00,51.00)=48.61$, $p<0.001$, and congruency ($F(1.00,17.00)=7.54$, $p=0.01$) but no interaction between these two factors ($F(3.00,51.00)=2.35$, $p=0.08$). Bootstrapped 95% confidence intervals of the mean difference between congruent and incongruent conditions revealed a significant difference in visibility for the two easiest conditions (target dur. 12ms: CI = [-0.03, 0.02], 24ms: CI = [-0.04, 0.04], 36ms: CI = [0.006, 0.07], 48ms: CI = [0.008, 0.05]). In conclusion, the results were very similar to those observed with visibility ratings for word identification, comforting the interpretation that congruent retro-cues increased the sensation of having seen the target word, a sensation that was veridical since it was accompanied by increased objective detection performance.

2.C: Task interactions

The interaction factor between congruency and accuracy at the second task was significant, in favor of a smaller effect of congruency on word identification in trials with correct case discrimination response (95%BCI [-1.85 -0.92], change in log-odds of correct identification). In congruent trials, response consistency probability **(Figure 3e)** was at chance (95%BCI [0.33 0.457]), whereas it was above chance in incongruent trials (95% BCI [0.64 0.87]), with a significant difference between these conditions (95% BCI [-0.39 -0.21]), indicating that participants are more likely to have consistent responses at both tasks in incongruent trials.

The conditional probability of correctly reporting the word's position when it was identified and participant were confident in their identification (see details in Methods) **(Figure 3f)** was also in favor of more errors in the congruent condition than in the incongruent condition (95%CI = [7.85, 14.11]), with an estimated performance at position discrimination in retro-perceived trials of 75.4% (95%CI = [67.21, 81.98]), compared to 86.8% in all congruent and 97.8% in incongruent trials.

Overall, Experiment 2 showed the same pattern of results, even when position was used instead of case. When provided with separate visibility ratings for both dimensions of the stimulus, participants gave much higher ratings for congruent stimuli than

incongruent ones, even when asked to rate specifically their perception of the word's position, which they did not discriminate any better. Interestingly this was also true for objective detection sensitivity, meaning that participants were overall better at detecting the presence of either feature in congruent trials. This confirms that congruent retro-cues increased the general sensation of having seen the stimulus, as well as the objective sensitivity to its presence, independently of the specific representations associated with this sensation (position or word identity). Still, one question remains: are these increased sensations of having seen and detection sensitivity accompanied by an illusion of having identified all features of the stimulus?

In order to address this possibility, we conducted another experiment, this time asking participants to rate their *confidence* in their response to the word identification task on the one hand, and to the visual discrimination task on the other. If participants were aware that, although they detected the stimulus presence, they did not specifically perceive the word's visual feature, their confidence in their response to the visual feature should match their actual performance to discriminate that visual feature. In contrast, if they hold the illusory belief that they, in fact, consciously and specifically perceived the visual feature, they should show overconfidence in this response.

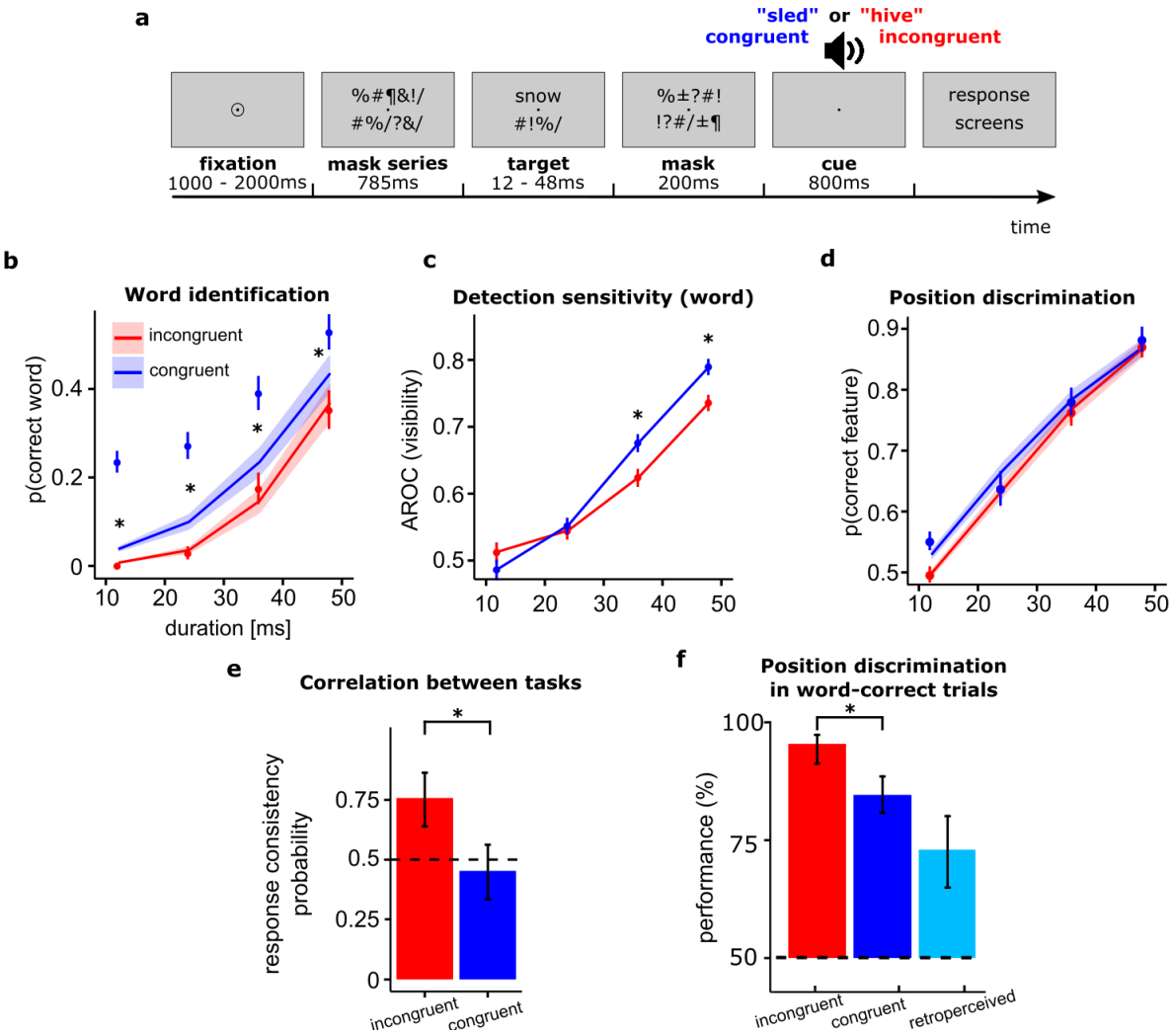

Figure 3: Experiment 2. *Participants were presented with 2 streams of stimuli above and below the fixation center. Response screens asked for: position discrimination -> visibility/confidence on position -> word identification -> visibility/confidence on position* ***(a)****. Word identification performance* ***(b)*** *Detection sensitivity for word identification* ***(c)*** *Case discrimination performance* ***(d)****. Detection sensitivity for position discrimination* ***(e)*** *Response consistency probability* ***(f)****. Case discrimination performance when word answers are correct* ***(g)*** *See Figure 1 for details.*

Experiment 3: word identification, case discrimination and confidence ratings

In order to determine whether or not there is an illusion of having identified the visual properties of the stimulus in retro-perceived trials, we conducted a third experiment testing word identification and case discrimination, as in Experiment 1, this time probing confidence ratings at each response, instead of visibility **(Figure 4a)**. On these confidence scales "0" corresponded an absence of confidence in the participant's response, regardless of whether they perceived the stimulus (see **Supplementary Note 1**). Participants were asked to respond first about letter casing, then provide their confidence rating on this response, followed by the word naming task and its own confidence rating. All analyses were identical to Experiment 2.

3.A: Word identification

The analysis of performance at the word identification task **(Figure 4b)** revealed a positive effect of congruency (95%BCI [1.37, 5.80]), and target duration (95%BCI [3.60, 5.60]). Their interaction was not significant (95%BCI [0.78, 1.20]). We found that accuracy was systematically higher in congruent trials than in incongruent ones at all durations as indicated in **Figure 4b** (12ms: 95%BCI [0.004 0.03]; 24ms: 95%BCI [0.03 0.09]; 36ms: 95%BCI [0.11 0.22]; 48ms: 95%BCI [0.06 0.20]).

3.B: Case discrimination

The analysis of performance at the case discrimination task using the model (**Figure 4c**) revealed a positive effect of target duration (95%BCI [1.25 1.55]), no effect of congruency (95%BCI [0.69, 1.20]), and no interaction between these factors (95%BCI [0.92, 1.16]). This result replicates that of Experiment 1 & 2.

3.C: Confidence ratings

A repeated measures two-way ANOVA on confidence ratings for word identification (**Supplementary Figure 6**) revealed a significant increase in confidence with target duration ($F(1.24,21.16)=132.04$, $p<0.001$) and congruency ($F(1.00,17.00)=65.54$, $p<0.001$), with an interaction between these two factors ($F(1.24,21.16)=28.22$, $p<0.001$). A repeated measures two-way ANOVA on confidence

ratings for case discrimination (**Supplementary Figure 6**) revealed a significant increase in confidence with target duration ($F(1.31,22.31)=69.95$, $p<0.001$) and congruency ($F(1.00,17.00)=9.07$, $p<0.01$), with an interaction between these two factors ($F(1.31,22.31)=7.05$, $p<0.001$). These results indicate that there might be a slight illusion of having a better knowledge of the word case in congruent as opposed to incongruent trials; still participants showed much stronger improvement in their confidence with congruent retro-cues for word identification than for letter casing, matching their pattern of performance at both tasks.

The analysis of confidence sensitivity between target present and target absent trials revealed a similar pattern (AROC, **Figure 4d and 4e)**. A repeated measures two-way ANOVA on confidence sensitivity for word identification revealed a significant increase with target duration ($F(1.93,39.89)=121.19$, $p<0.001$), and congruency ($F(1.00,17.00)=14.90$, $p=0.001$) with an interaction between these two factors ($F(3.00,51.00)=6.26$, $p=0.001$). Bootstrapped 95% confidence intervals of the mean difference between congruent and incongruent conditions revealed a significant difference for the all conditions apart from the most difficult, after correction for multiple comparisons (target dur. 12ms: CI = [-0.03, 0.07], 24ms: CI = [0.01, 0.6], 36ms: CI = [0.01, 0.1], 48ms: CI = [0.03, 0.11]). The same analysis on confidence sensitivity for case discrimination **(Figure 4e)** revealed a significant increase with target duration ($F(3.00,51.00)=96.94$, $p<0.001$), but no effect of congruency ($F(1.00,17.00)=2.84$, $p=0.11$) with, nonetheless, an interaction between these two factors ($F(2.45,41.69)=4.45$, $p=0.02$). Bootstrapped 95% confidence intervals of the mean difference between congruent and incongruent conditions revealed a significant difference in performance for the easiest condition only, after correction for multiple comparisons (target dur. 12ms: CI = [-0.01, 0.03], 24ms: CI = [-0.05 0.1], 36ms: CI = [-0.03, 0.5], 48ms: CI = [0.02 0.08]). In conclusion, congruency appeared to have a much stronger and more reliable effect on confidence sensitivity at the word task than the visual feature one, where, in spite of a slight tendency for overconfidence for long target durations, participants appear to have some knowledge of their inability to report the word's visual feature when it was difficult to perceive.

3.D: Task interactions

As a last analysis of sensitivity, we performed a 2-way repeated measures ANOVA between congruency and type of rating (word or visual feature), specifically aimed at

assessing whether there would be a 2-way interaction between these factors. We tested this interaction in Experiment 2 (visibility ratings) and Experiment 3 (confidence ratings) separately. We found this interaction to be significant in Experiment 3 ($F(1.00,17.00)=9.687$, $p<0.01$), but not Experiment 2 ($F(1.00,17.00)=0.560$, $p=0.46$), which is further indication of a strong dissociation between visibility and confidence in the context of these experiments. We interpret this dissociation as the result of participants detecting the stimulus more often thanks to the retro-cue, but having no trace of its visual feature to report with confidence.

The interaction factor between congruency and accuracy at the second task showed a trend in favor of a smaller effect of congruency on word identification in trials with correct case discrimination response (95%BCI [-0.48 0.05], change in log-odds of correct identification). The analysis of response consistency probability across tasks **(Figure 4f)** showed that participants were less likely to have both answers correct or incorrect on congruent trials (95%BCI [0.31 0.48]), and this measure was at chance for incongruent trials (95%BCI [0.33 0.55]). The difference between congruent and incongruent conditions was weaker than what was observed in Experiment 1 and Experiment 2 (95%BCI [-0.12 0.01]). We sought to characterize more precisely the evidence provided by Exp. 3 using Bayes factors (see **Supplementary Note 4** for details and prior selection). The Bayes factor indicated that the data was approximately 1.78 times more likely under the hypothesis of a decreased correlation in congruent trials than the null hypothesis. Hence, although the evidence provided by Exp. 3 alone is not decisive, it does present a trend consistent with what observed in Exp. 1 and 2.

The analysis of conditional probabilities confirmed a significant effect of congruency on the association between performances at the two tasks. The conditional probability of correctly reporting the word's case when it was identified and participant were confident on their identification **(Figure 4g)** decreased significantly for congruent trials compared to incongruent trials (95%CI = [-2.44, -7.52]), with an estimated case discrimination performance in retro-perceived trials of 60.1% that was not statistically different from 50% (95%CI = [39.7, 71.6]), compared to 80.1% in all congruent and 85.2% in incongruent trials.

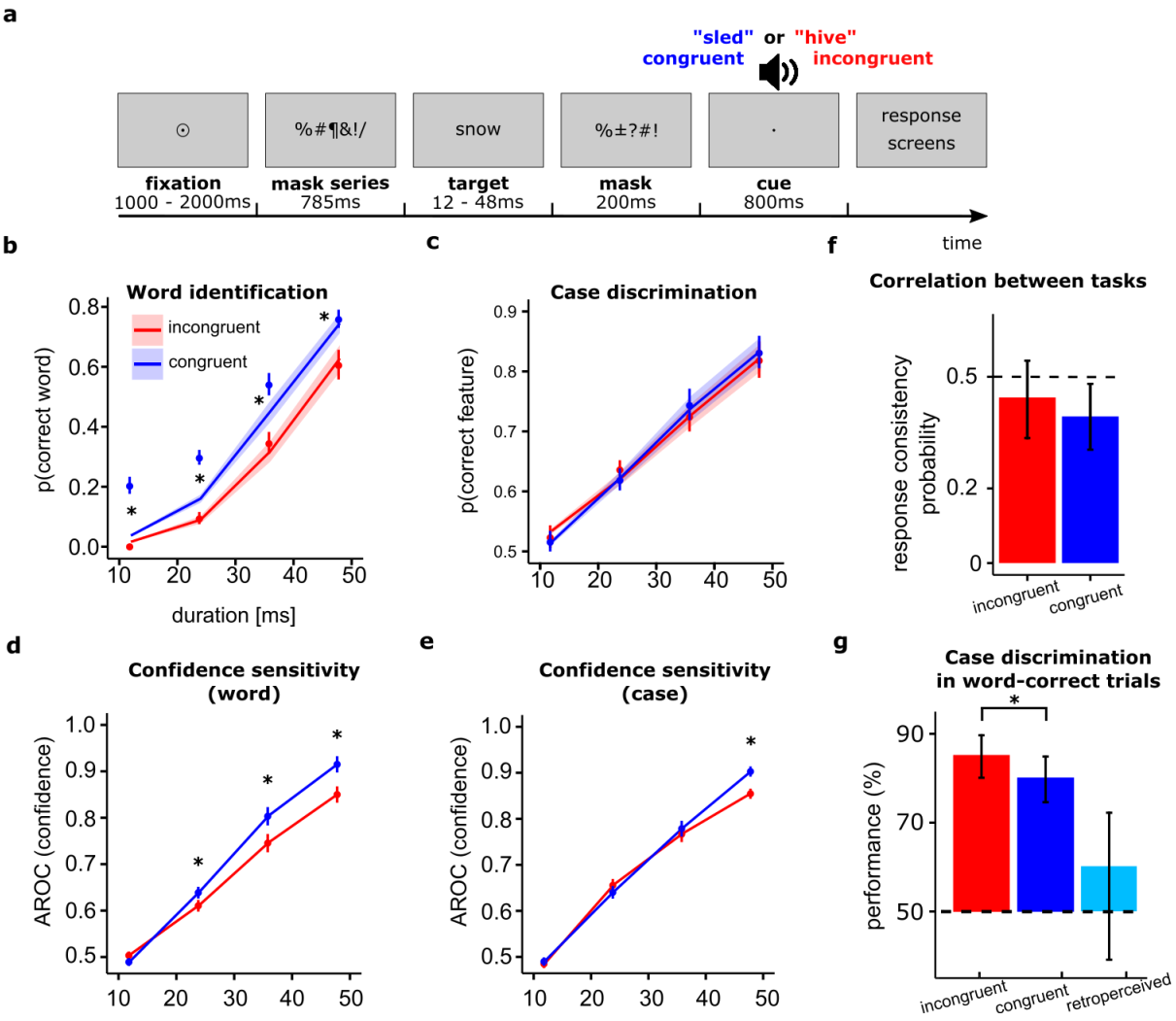

Figure 4: Experiment 3. *Participants were presented with one stream of centered stimuli. Response screens asked for: case discrimination -> confidence on position -> word identification -> confidence on position* ***(a)****. Word identification performance* ***(b)*** *Detection sensitivity for word identification* ***(c)*** *Case discrimination performance* ***(d)****. Detection sensitivity for position discrimination* ***(e)*** *Response consistency probability* ***(f)****. Case discrimination performance when word answers are correct* ***(g)*** *See Figure 1 for details.*

Controlling for the effects of partial awareness

One possibility that remains to be addressed concerns whether participants perceived one or a few letters of the target stimuli and were able to guess the correct answer in congruent trials based on this information. Indeed, since the stimuli here are presented around the perceptual threshold, this requires controlling for "partial awareness" of sensory information(Kouider et al., 2010): participants might have perceived, at least in some instances, fragments of stimuli such as a couple of letters, from which they inferred the correct word response. Combining a few letters from the word with a semantic context could lead to a correct guess in congruent trials, while the same process would lead to an error in incongruent trials where the semantic context is not relevant. For example if the participant caught the following partial stimulus: "FL???R", on incongruent trials (e.g. "sled") it would equally lead to naming "flower" and "flavor", whereas a congruent cue like "tulip" would bias the naming response towards "flower". In this case, the effect of congruency on word identification would not be a retrospective effect of the semantic cue, but just the effect of the combination of an initial conscious information (a few letters) with a subsequent one (the semantic context).

The most probable interpretation of this kind is that partial awareness of a few letters comes with the awareness of the visual features of these letters. In this case, if the improvement of word identification with congruent cues relies on trials with partial awareness, then this implies that on those trials participants should also be aware of the target's visual feature. In other words, we should observe an increased correlation between success at the visual feature task and success at the word identification task for congruent trials as opposed to incongruent trials. However the data showed the opposite pattern: participants generally displayed a higher probability of having both answers correct or incorrect in incongruent trials than they did in congruent ones, where participants often had the word correct but not the visual feature, and the correlation between task performances dropped **(Figures 2f, 3f, 4g in the main text)**.

We also considered another, subtler version of the partial awareness interpretation in which the awareness of a few letters would not come with the awareness of visual features, but would only correspond to the awareness of more abstract orthographic representations (a few letters without their visual attributes). In this case we should observe a correlation not between word identification and visual feature

discrimination, but between word identification and orthographic distance. More specifically, in congruent trials, incorrect naming should systematically correspond to trials without any partial awareness, thus leading to a very poor orthographic correspondence between the word reported and the target word, while in incongruent trials, errors in word identification should include more partial awareness trials. Therefore orthographic distance between the response and the target when word identification failed should be lower for incongruent that congruent trials. To test for this possibility we calculated the relative orthographic distance between response and target for each trial by using the Levenshtein distance (see **Supplementary Note 5**) and used this metric as a proxy to assess word fragment or letter perception. **Supplementary Figure 7** shows that, contrary to what would be expected from this partial awareness interpretation, orthographic distance did not significantly depend on congruency (repeated measures ANOVAs : **Exp.1 –** congruency: $F(1.00,14.00)=0.051$, $p=0.82$, target duration: $F(2.45,34.23)=40.157$, $p<0.001$, interaction: $F(2.26,31.62)=1.055$, $p=0.37$**,** average number of trials per condition (ATC): C = 18±11, IC = 34±12. **Exp.2 –** congruency: $F(1.00,17.00)=1.891$, $p=0.19$, target duration: $F(2.17,36.93)=22.790$, $p<0.001$, interaction: $F(1.68,28.57)=0.960$, $p=0.38$, ATC: C = 17±12, IC = 28±18, **Exp.3 –** congruency: $F(1.00,17.00)=0.209$, $p=0.65$, target duration : $F(2.31,39.32)=3.584$, $p=0.03$, interaction: $F(2.23,37.93)=1.150$, $p=0.33$, ATC: C = 27±8, IC = 39±10). Note that the significant effect of target duration on orthographic distance indicates that partial awareness does occur in our experiment, but it does not account for the effect of cueing on identification.

In conclusion, although there are indications that participants do experience partial awareness in this protocol, our analyses indicate that this cannot account for the effect of retrospective semantic cueing on word identification.

Finally, it is important to note that this partial awareness hypothesis cannot explain the improved detection sensitivity that we observed for congruent retro-cues because it should entail identical visibilities and detection sensitivities for semantically congruent and incongruent cues.

Discussion

In this study, we investigated the effect of a congruent retro-cue on perceptual awareness on two distinct attributes of a single stimulus, each thought to arise at very different stages along processing hierarchy. We have shown that retrospectively cueing a masked word with a semantically related auditory cue increases participants' probability of reporting the target's meaning around the perceptual threshold. However, this was not true for the visual features of our target words: the cue had no effect on participants' ability to report the target's case or location. We tested various types of subjective ratings: Experiment 1 showed an increase in participants' ability to detect the presence of the stimulus with congruency, Experiment 2 showed this increase in detection sensitivity was present for both meaning and visual feature although task performance was not always matching, and Experiment 3 revealed that confidence sensitivity overall faithfully reflected task performance, suggesting that the participants do not fall under the illusion of holding a true representation of the visual feature when they don't: confidence ratings show radically different patterns of results apart from the highest target duration, at which the stimulus may be supra-liminal for many participants as performance at the case discrimination task indicates.

We interpret these results as evidence that retrospective cueing led to conscious access to high-level semantic properties of the stimulus with no access to its basic sensory features. This suggests that a retrospective semantic context can bring into awareness the semantic trace left by a previous, masked stimulus even if the masks have degraded the low-level representations of this stimulus (**Figure 1**). Overall, these findings suggest that conscious access mechanisms can operate independently from the build-up of sensory representations of a stimulus, yielding "asensory" percepts.

Research on visual masking has shown that, when visual words are masked, participants can sometimes retain partial awareness of these words, at least for lower-level features. In other words, participants may consciously see only a few letters, and combine this information with a semantic cue to guess the identity of the masked word(Kouider et al., 2007; Kouider & Dupoux, 2004). Although this effect is likely present to some extend in our data, we believe it cannot account for the pattern of results that we observed, because the orthographic distance between the target words and participants' incorrect responses did not show any effect of congruency. Furthermore, contrary to what

605 we observed, partial awareness would predict similar visibilities and detection
606 sensitivities for congruent and incongruent trials.

607 Several previous studies have shown that a retrospective context can influence the
608 lexical processing or identification of a preceding masked prime word(Bernstein et al.,
609 1989; Briand et al., 1988; Dark, 1988; Neely, 1991; Thomas et al., 2012; VanVoorhis &
610 Dark, 1995) but none of them has investigated whether conscious perception *per se* was
611 modified by the retrospective context word, or tested more than one attribute of the
612 prime word. Here, we directly addressed conscious perception by collecting subjective
613 visibility on each trial. Both mean subjective visibility, and the objective detection
614 sensitivity measure that derives from it, demonstrate that congruent retro-cue directly
615 improve conscious perception of the masked word: it increases the number of trials
616 where participants report that they have seen the word. Since this effect is observable in
617 better detection sensitivity (as revealed by greater AROC), this cannot be due to a bias in
618 responding that one has seen.

619 Such asensory conscious perception raises important questions for future
620 research, and notably how do participants know that this asensory perception actually
621 comes from the external environment? How come they can report, and rightfully so, that
622 they have seen the stimulus, that the stimulus was indeed presented although they have
623 no associated *visual* representation of this stimulus? According to our schematic model,
624 the latent semantic trace left by the masked word is reactivated by contextual influence
625 and this reactivation allows broadcasting this representation to a global workspace of
626 areas associated with conscious access. But in principle spontaneous fluctuations in the
627 activity of semantic areas independently of any external input could also take the form of
628 a "mock trace" that could be mistaken for a trace left by an actual external stimulus. One
629 possibility is that a genuine trace left by a stimulus might be more precise in its encoding
630 than spontaneous fluctuations and that this difference in encoding is the index used to
631 produce a visibility rating and an accurate judgment about the presence of an external
632 stimulus, even in the absence of any associated visual information.

633 The present results also interrogate the relationship between awareness and
634 sensory integration. It shows that an external stimulus can be consciously perceived even
635 when its different features are not bound together, and even if some of its basic features
636 are actually missing. However it does not allow concluding definitively on the role of top-

down attention and conscious access mechanisms on feature integration(Treisman & Gelade, 1980). Whether feature binding occurs pre-experientially or not is still very much debated(Bouvier & Treisman, 2010; Chen et al., 2018; Keizer et al., 2015; Koivisto & Silvanto, 2011). One could argue that current results provide some evidence in favor of the idea that representations are not bound pre-experientially(Treisman & Gelade, 1980), otherwise all features should reach consciousness together at the time of retrospective cueing. However, it can also be argued that because of the short delay between target and pattern mask, there was not enough time for recurrent processing to occur before low level information was erased(Di Lollo, 2018; V. A. F. Lamme & Roelfsema, 2000), and that if recurrent processing had occurred normally, it would have bound the features pre-experientially.

This experiment, in line with others on retro-perception(Sergent et al., 2013; Thibault et al., 2016; Xia et al., 2016), shows that it is possible to influence whether a stimulus reaches awareness or not, even several hundred milliseconds after its disappearance. This highlights high flexibility in the way external events are perceived, not only in terms of whether an event is accessed or not, but goes to show that what is accessed is also subject to change according to the current context. Our experimental findings can be viewed as evidence that consciousness is an intrinsically dynamic process. Perceived events are not sequentially processed, bound and propelled through experience after waiting in line: rather it seems that different features can become conscious if and when they become relevant according to the ongoing context.

In conclusion, this study demonstrates the extreme flexibility of the form that can take conscious perception of an external stimulus, both in terms of timing and in terms of content. Conscious perception of a word can be reduced to its most essential property: its meaning. This can be seen as a form of quintessential conscious access: we know that a stimulus has been presented; we know what it is; we don't know what it looks like. These results therefore challenge our understanding of the relationship that conscious experience entertains with external events, and interrogate the very nature of what we call a conscious percept.

Methods

Participants. For Experiment 1, 23 healthy participants underwent a pre-selection session (see details below). This allowed excluding subjects whose overall performance was at chance (or at ceiling) with our masked stimuli. Fifteen participants (10 females and 5 males) matched our selection criterion and were included in the main experiment. They were aged 23 ± 3.3. For Experiment 2, 23 participants were recruited, out of which 5 were excluded prior to analysis because their performance was at chance or they did not follow the instructions. The remaining 18 participants (9 females and 9 males) had an average age of 27.5 ± 7.1. In Experiment 3, 26 participants were recruited, out of which 8 were excluded prior to analysis because their overall case discrimination performance was at chance or they did not follow the instructions. Remaining participants (10 females and 8 males) had an average age of 24.5 ± 3.7. All participants gave informed consent in writing prior to participation, and the Université Paris Descartes Review Board CERES approved the protocols for the study in accordance with French regulations and the Declaration of Helsinki. Participants received a compensation of 10€ per hour of their time.

Stimulus-list. Our stimulus list was composed of 432 unique words (targets) from the French language. These words ranged from 3 to 8 letters long (+ one 2-letter word, "os" which translates to "bone"), and did not carry accents of cedillas. They were selected for being strongly associated to another word from the French language less than three-syllables-long, which could then be used as a congruent cue for that word. These associations were determined through 5 different free-association experiments (two experiments(Ferrand, 2001; Ferrand & Alario, 1998), the online database *dictionnaire des associations verbales du français*(Debrenne, 2011) (www.dictaverf.nsu.ru), and our own online experiments which used the same methodology as this database on additional words described in **the Supplementary Note 6**). Cue-target pairs were considered "strongly associated" when, in one free-association experiment, more than 25% of participants first responded the target when presented with the cue. Each word in the target stimulus list was associated with only one possible congruent cue, and only one possible incongruent cue. Incongruent cues were produced by shuffling the list of congruent cues, and checking the absence of semantic relation with the corresponding target word.

Experimental design. This experiment followed a 2x4 factorial design: congruency (2 levels) and target duration (4 levels), plus an extra level of stimulus absent trials. Each of these 9 conditions included 48 individual trials, split in 15 or 12 blocks. Participants knew that the target was absent on some trials. Word-lists for all participants were established prior to the experiments so as to avoid or balance cross-trial priming and prior exposure effects (**See Supplementary Note 7** for Participant Selection and Training & Stimulus Presentation Balance).

Procedure. In Experiment 1, each trial began with the onset of a circle around the central point, when fixation was successfully maintained for 200ms using an EyeTracker (see **Supplementary Note 8** for Materials). Then, following a 1-2s jittered delay, 6 random strings of non-alphanumerical characters (picked randomly without replacement among the following: *%#?&/!|±¶*) with the same font and size as target words would rapidly flash on screen with no inter-stimulus interval. These strings would vary both in length (3 to 8 items, all six possible length being displayed once in each trial) and duration (two masks of each possible duration: 12, 96 or 192ms). A 7th non-alphanumerical mask would follow suit, be 9-characters long and last for 200ms immediately before target onset. Targets would be displayed for 12, 24, 36 or 48ms. There were 11% of "target absent" trials where a mask was presented instead of a target, with the same length and duration probability. The target was then replaced with an 8th and last mask that was again 9-characters long and lasted for 200ms. All the sequence unfolded with no blank interval between the stimuli. Fifteen milliseconds after the offset of the last mask, i.e. 215ms after the offset of the visual word target, an auditory cue was presented to the participant. The auditory cue lasted less than 800ms and was either semantically related to the target (congruent trials), or unrelated (incongruent trials). All these audio words were pre-recorded with the same female voice and a neutral tone. Participants waited until the offset of the cue before they could respond. They first named the target visual word aloud and their response was recorded via a microphone. If they had not seen any word, they were instructed to provide a random word to the best of their ability. Following this first response, participants had to press the space bar or wait for the end of the recording delay and were then prompted to report whether the word was written in upper or lower case (and guess if they did not know) by selecting an option on the screen using the keyboard. Finally they had to rate the word's visibility on a 9-point horizontal scale (a small filled rectangle could be moved within a larger box with ends flanked with the labels "min" and

731 “max”, respectively to the left and the right. Details about the instructions to use the scale
732 are available in **Supplementary Note 1**). Participants were instructed to use the lowest
733 rating when they had not seen the stimulus at all, and to rate overall visibility using the
734 rest of the scale if they had.

735 In Experiment 2, the procedure was identical to Experiment 1 except for the
736 following: two synchronous streams of masks were presented just above and below a
737 small fixation circle. Simultaneous masks were different strings of symbols of different
738 lengths that lasted the same duration. On each trial a single target word appeared, in one
739 of the streams, while on the other stream another mask of the same length would be
740 displayed. Note that while there is a known preference for reading in the right visual
741 hemifield(Ellis, 2004) compared to the left, no such strong preference seems to exist
742 between the upper and lower hemifields(Hagenbeek & Van Strien, 2002). In 11% of trials,
743 no target word was presented and the trial only consisted in two streams of masks. While
744 in Experiment 1 the interval between the target offset and the cue onset (ISI) was kept
745 constant at 215ms, in Experiment 2 and 3 it was the interval between target onset and
746 cue onset (SOA) that was kept constant at 248ms, with the introduction of a blank period
747 between last mask offset and cue onset with a duration adjusted according to target
748 duration (between 0 and 36ms). Immediately after the offset of the auditory cue,
749 participants were prompted by a response screen to report where the target had
750 appeared (top or bottom). They then had to provide a visibility rating on this feature using
751 a cursor on a 9-point visual scale similar to Exp. 1 taking into account both detection and
752 position discrimination (see detailed instructions in **Supplementary Note 1**). Then they
753 were required to report the target word verbally (recorded via a microphone), and then
754 provide a visibility rating taking into account both detection and identification.

755 Experiment 3 was a follow-up from Experiment 1 in which we changed the order
756 of the tasks (to check for potential task order effects) and included confidence ratings on
757 the two tasks, instead of visibility ratings. The sizes of the stimuli were similar to the ones
758 used in Experiment 2, but they were centered, as in Experiment 1. Participants were first
759 prompted to report whether the target was uppercase or lowercase, and how confident
760 they were about this decision on a 9-point scale. Then they were required to report the
761 target word verbally, and rate their confidence on their word identification response on a
762 9-point scale. This time, the scales were used to measure only the confidence on the
763 previous decision; there was no mention of visibility. Participants were asked to use the

764 lowest level when they were entirely not-confident about their decision i.e. response at
765 random, the first half of the scale if their confidence was low to medium, and the second
766 half if they were fairly to definitely certain of their response.

767 **Multivariate modeling of word and feature response.** We analyzed participants'
768 accuracies at identifying the masked word and reporting its visual features (e.g., case or
769 position) with a multivariate logistic model (Cox, 1972). Given that our dataset is grouped
770 according to 'observational units' (participants) which should be considered as random
771 sample from a population, we adopted a hierarchical Bayesian approach. To correct for
772 these guesses in congruent trials, we counted for each participant and each stimulus
773 duration the percentage of incongruent trials where they produced the matching
774 semantically associated word (resulting in a special type of error). In the context of our
775 fully Bayesian approach, these counts were used as priors to inform participant- and
776 duration-specific Beta priors over the guess rates (see **Supplementary Note 2** for
777 details). Our model estimated the effect of congruency while accounting for word-specific
778 and participant-specific effects on visual feature and word performance, guess rate
779 probability for each target duration, as well as the effect of performance at one task on the
780 other. An association parameter ρ was computed from the model, and indicates whether
781 the two objective responses are more likely to be either both correct or wrong, rather
782 than being one correct and the other wrong (expressed in log-odds units, and reported as
783 probabilities in the main text) . The full description of the model is available in
784 **Supplementary Note 2**.

785 **Frequentist analyses.** Detection sensitivity and confidence sensitivity were derived from
786 the distributions of visibility or confidence ratings respectively using Signal Detection
787 Theory(Hanley & McNeil, 1982; McMillan & Creelman, 2005). For the conditional
788 performance panels, performance at the case or position task was calculated for trials
789 where the word was correctly identified. In the congruent condition this required
790 discarding "correct-guesses" trials without the help of the model. We calculated an
791 individual, per target duration, visibility cut-off below which a correct trial could safely
792 be classified as guess (since the response was correct but the subjective rating was low).
793 If a participant's point-estimate of the guess rate (used as a prior in the Bayesian Model)
794 was n% of congruent trials, we excluded all trials with visibility equal or below that of the
795 n^{th} percentile. This cut-off was most of the time between 0 and 1 (0.27±0.46 for 12ms,
796 0.47±0.64 for 24ms, 0.74±0.88 for 36ms, 0.87±0.64 for 48ms for Experiment 1).

Conditional performance for "retro-perceived" trials, i.e. the additional "word-correct" trials brought about by congruent retro-cues was also computed. The rationale behind this analysis is that if all incongruent trial errors are caused by a variety of factors grouped in X (such as wrong button presses, attentional lapses, or effect of the task excluding congruency), then all errors in congruent trials are due to the sum of errors caused by X and those caused by congruency of the cue. We aimed at dissociating these two phenomena and estimating performance in the congruent condition given that one has retroperceived (e.g. removing errors shared with the incongruent condition), using Bayes' theorem and the definition of retro-perceived trials. The full demonstration is available in **Supplementary Note 3** and **Supplementary Figure 3**. Trial exclusion criteria and word identification task analysis are available in **Supplementary Note 10**.

Data and script availability: See **Supplementary Note 11**.

Acknowledgements. We thank Benjamin Rohaut, Lionel Naccache, Benoît Laslier and Laurent Cohen for their useful comments.

References

Abrams, R. L., & Grinspan, J. (2007). Unconscious semantic priming in the absence of partial awareness. *Consciousness and Cognition*, *16*(4), 942-953.

Baars, B. J. (2005). Global workspace theory of consciousness : Toward a cognitive neuroscience of human experience. In S. Laureys (Éd.), *Progress in Brain Research* (Vol. 150, p. 45-53). Elsevier. https://doi.org/10.1016/S0079-6123(05)50004-9

Bernstein, I. H., Bissonnette, V., Vyas, A., & Barclay, P. (1989). Semantic priming : Subliminal perception or context? *Perception & Psychophysics*, *45*(2), 153-161.

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision : A label advantage in object recognition. *Journal of Neuroscience*, *35*(25), 9329-9335.

Bouvier, S., & Treisman, A. (2010). Visual Feature Binding Requires Reentry. *Psychological science*, *21*(2), 200-204. https://doi.org/10.1177/0956797609357858

Briand, K., Heyer, K. den, & Dannenbring, G. L. (1988). Retroactive semantic priming in a lexical decision task. *The Quarterly Journal of Experimental Psychology*, *40*(2), 341-359.

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annual review of neuroscience*, *21*(1), 47-74.

Chen, J., Chen, Y., & Liu, Y. (2018). Color-motion feature binding occurs unconsciously. *Neuroscience Letters*, *675*, 54-58. https://doi.org/10.1016/j.neulet.2018.03.060

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream : The case for the visual word form area. *Neuroimage*, *22*(1), 466-476.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, *82*(6), 407.

Cox, D. R. (1972). The analysis of multivariate binary data. *Applied statistics*, 113-120.

Dark, V. J. (1988). Semantic priming, prime reportability, and retroactive priming are interdependent. *Memory & Cognition*, *16*(4), 299-308.

Debrenne, M. (2011). Le dictionnaire des associations verbales du français et ses applications. *Variétés, variations and formes du français. Palaiseau: Éditions de l'Ecole polytechnique*, 355-366.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing : A testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204-211. https://doi.org/10.1016/j.tics.2006.03.007

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. https://doi.org/10.1016/j.tics.2005.05.004

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness : Basic evidence and a workspace framework. *Cognition*, *79*(1), 1-37. https://doi.org/10.1016/S0010-0277(00)00123-2

Di Lollo, V. (2018). Attention is a sterile concept; iterative reentry is a fertile substitute. *Consciousness and Cognition*, *64*, 45-49. https://doi.org/10.1016/j.concog.2018.02.005

Ellis, A. W. (2004). Length, formats, neighbours, hemispheres, and the processing of words presented laterally or at fixation. *Brain and language*, *88*(3), 355-366.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345-352. https://doi.org/10.1016/S1364-6613(00)01520-5

Felleman, D. J., & Van, D. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, *1*(1), 1-47.

Ferrand, L. (2001). Normes d'associations verbales pour 260 mots «abstraits». *L'Année psychologique*, *101*(4), 683-721.

Ferrand, L., & Alario, F.-X. (1998). Normes d'associations verbales pour 366 noms d'objets concrets. *L'Année psychologique*, *98*(4), 659-709.

Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., Andelman, F., Neufeld, M. Y., Kramer, U., & Fried, I. (2009). Neural “ignition” : Enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron*, *64*(4), 562-574.

Fowler, C. A., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General*, *110*(3), 341.

Hagenbeek, R. E., & Van Strien, J. W. (2002). Left–right and upper–lower visual field asymmetries for face matching, letter naming, and lexical decision. *Brain and Cognition*, *49*(1), 34-44.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29-36.

Humphreys, G. W., Evett, L. J., & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory & Cognition*, *10*(6), 576-590.

Keizer, A. W., Hommel, B., & Lamme, V. A. F. (2015). Consciousness is not necessary for visual feature binding. *Psychonomic Bulletin & Review*, *22*(2), 453-460. https://doi.org/10.3758/s13423-014-0706-2

Koivisto, M., & Silvanto, J. (2011). Relationship between visual binding, reentry and awareness. *Consciousness and Cognition*, *20*(4), 1293-1303. https://doi.org/10.1016/j.concog.2011.02.008

Kouider, S., De Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in cognitive sciences*, *14*(7), 301-307.

Kouider, S., de Gardelle, V., & Dupoux, E. (2007). Partial awareness and the illusion of phenomenal consciousness. *Behavioral and Brain Sciences*, *30*(5-6), 510-511.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception : A critical review of visual masking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*(1481), 857-875.

Kouider, S., & Dupoux, E. (2004). Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological science*, *15*(2), 75-81.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571-579. https://doi.org/10.1016/S0166-2236(00)01657-X

Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking Interrupts Figure-Ground Signals in V1. *Journal of Cognitive Neuroscience*, *14*(7), 1044-1053. https://doi.org/10.1162/089892902320474490

Lamme, V. A., Super, H., Landman, R., Roelfsema, P. R., & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vision research*, *40*(10-12), 1507-1521.

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, *110*(35), 14196-14201. https://doi.org/10.1073/pnas.1303312110

McMillan, N. A., & Creelman, C. D. (2005). Detection theory. *LEA Pub*.

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, *80*(3), 215-229. https://doi.org/10.1016/S0010-0277(00)00139-6

Neely, J. H. (1991). Semantic priming effects in visual word recognition : A selective review of current findings and theories. *Basic processes in reading: Visual word recognition*, *11*, 264-336.

Pitts, M. A., Padwal, J., Fennelly, D., Martínez, A., & Hillyard, S. A. (2014). Gamma band activity and the P3 reflect post-perceptual processes, not visual awareness. *Neuroimage*, *101*, 337-350.

Rimsky-Robert, D., Störmer, V., Sackur, J., & Sergent, C. (2019). Retrospective auditory cues can improve detection of near-threshold visual targets. *Scientific Reports*, *9*(1), 1-11.

Sergent, C. (2018). The offline stream of conscious representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1755), 20170349.

Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2013). Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. *Current Biology*, *23*(2), 150-155.

Thibault, L., Van den Berg, R., Cavanagh, P., & Sergent, C. (2016). Retrospective attention gates discrete conscious access to past sensory stimuli. *PloS one*, *11*(2), e0148504.

Thomas, M. A., Neely, J. H., & O'Connor, P. (2012). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language*, *66*(4), 623-643.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136. https://doi.org/10.1016/0010-0285(80)90005-5

VanVoorhis, B. A., & Dark, V. J. (1995). Semantic matching, response mode, and response mapping as contributors to retroactive and proactive priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 913.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream : Dissecting the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143-156. https://doi.org/10.1016/j.neuron.2007.05.031

Weiskrantz, L., Barbur, J. L., & Sahraie, A. (1995). Parameters affecting conscious versus unconscious visual discrimination with damage to the visual cortex (V1). *Proceedings of the National Academy of Sciences*, *92*(13), 6122-6126.

Xia, Y., Morimoto, Y., & Noguchi, Y. (2016). Retrospective triggering of conscious perception by an interstimulus interaction. *Journal of vision*, *16*(7), 3-3.