

L'accès à la conscience

Stanislas Dehaene  
Chaire de Psychologie Cognitive Expérimentale

Cours

**Mécanismes cérébraux du masquage visuel**

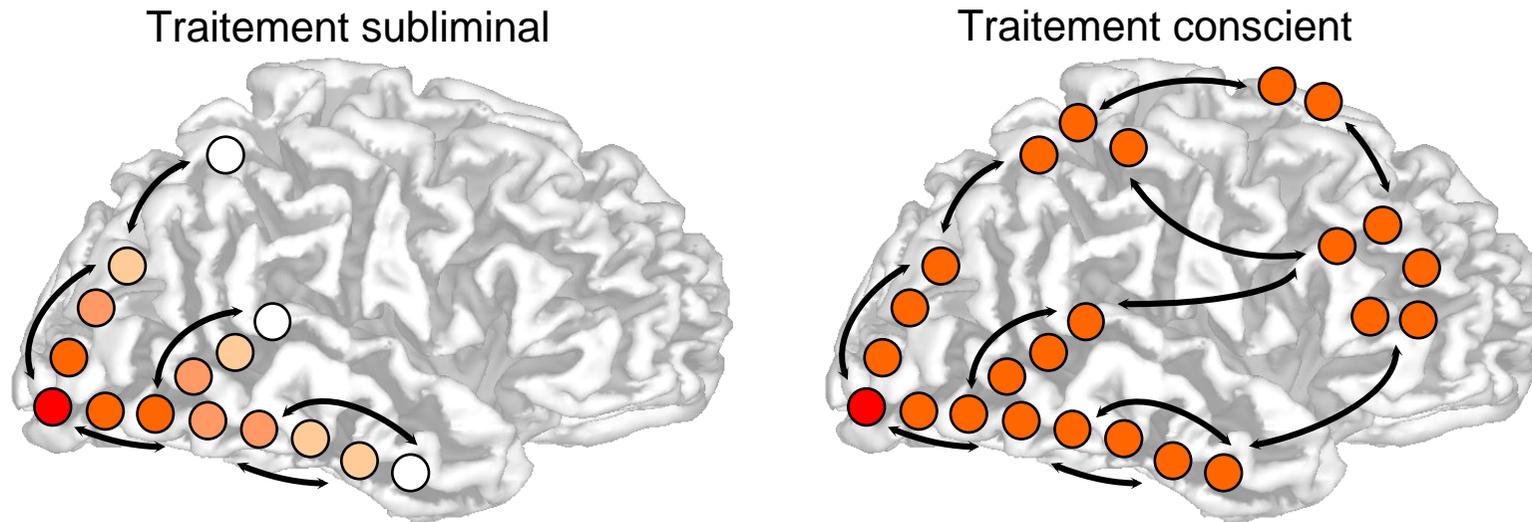
## La synthèse moderne des idées sur l'accès à la conscience

Il existe un début de convergence, sinon vers un « modèle standard », du moins en faveur d'un ensemble d'idées proposées depuis les années 1950-1960, et de mieux en mieux acceptées:

La conscience serait associée à

- Un système de **supervision centrale**
- A **capacité limitée**
- Nécessitant des **boucles lentes**, réentrantes et descendantes (« top-down ») par opposition au traitement ascendant (« bottom-up ») rapide et non-conscient des données sensorielles
- Un **espace interne** de synthèse, de maintien et de partage des données (« théâtre », « tableau noir », « espace de travail neuronal global »)

# L'accès à la conscience selon le modèle GNW



- Ce que nous appelons “l'accès à la conscience” est la **disponibilité globale d'une information**, qui la rend **accessible** à tous les processeurs et donc **rapportable** (à nous même ou aux autres)
- L'accès à la conscience correspond à l'**ignition soudaine** de neurones à axones longs, distribués mais particulièrement denses dans les cortex préfrontaux et pariétaux (*global neuronal workspace*, GNW)
- Durant la perception conscience, le GNW se stabilise, pour une durée minimale, dans un **état réverbérant et métastable** et distribue cette information au reste du cortex
- Le même processeur peut participer à des traitements conscients et non-conscients, le mode conscient se caractérisant par une **amplification** descendante et une **mise en synchronie** avec les neurones du GNW.
- Les neurones du GNW sont le siège d'une **activité spontanée permanente**, une série temporelle d'états métastables, modulée par (1) des signaux ascendants d'éveil; (2) des signaux sélectifs de renforcement

# Un modèle neuronal détaillé de l'ignition du GNW

Dehaene, Sergent & Changeux, 2003; Dehaene & Changeux, 2005

Simulations neuro-informatiques détaillées d'un réseau de neurones

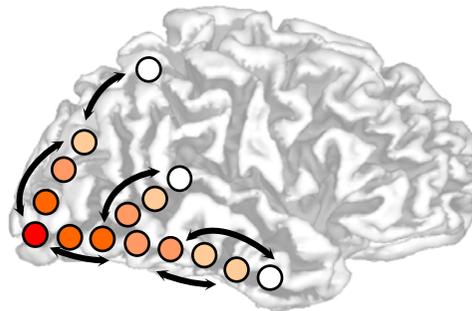
- Neurones « integrate-and-fire » avec dynamique réaliste des potentiels post-synaptiques (AMPA, NMDA, GABA)

- Intégrés dans une architecture reproduisant une partie du GNW:

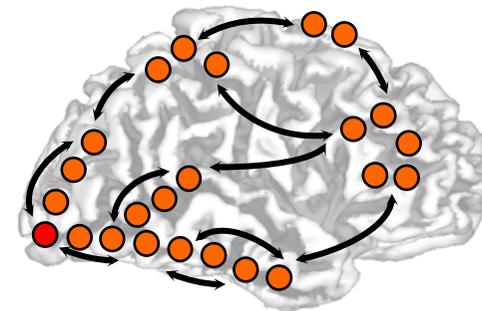
- 4 représentations thalamo-corticales successives, liées par des connections ascendantes (AMPA) et descendantes globales (NMDA)

- Hypothèse: un stimulus représenté dans les aires C et D serait conscient, car diffusé à l'ensemble du *workspace*

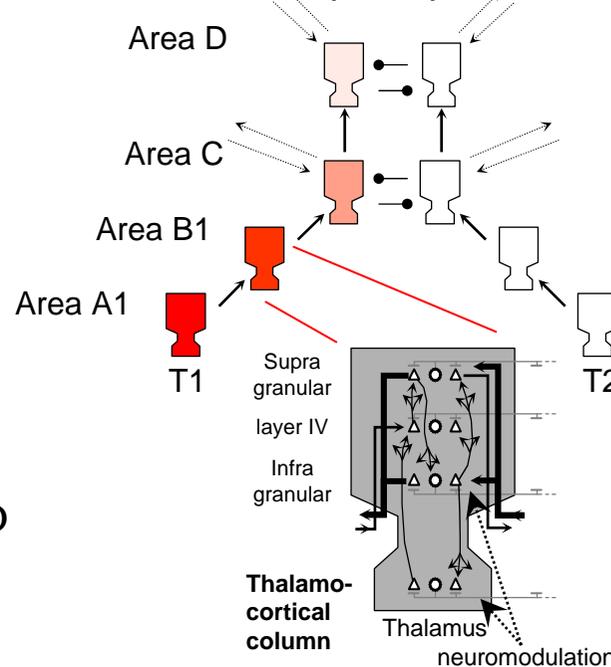
**Feed-forward propagation (subliminal processing)**



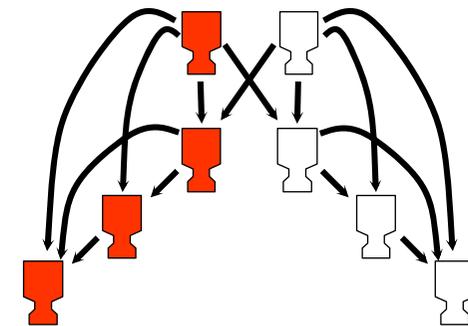
**Reverberating global neuronal workspace (conscious representation)**



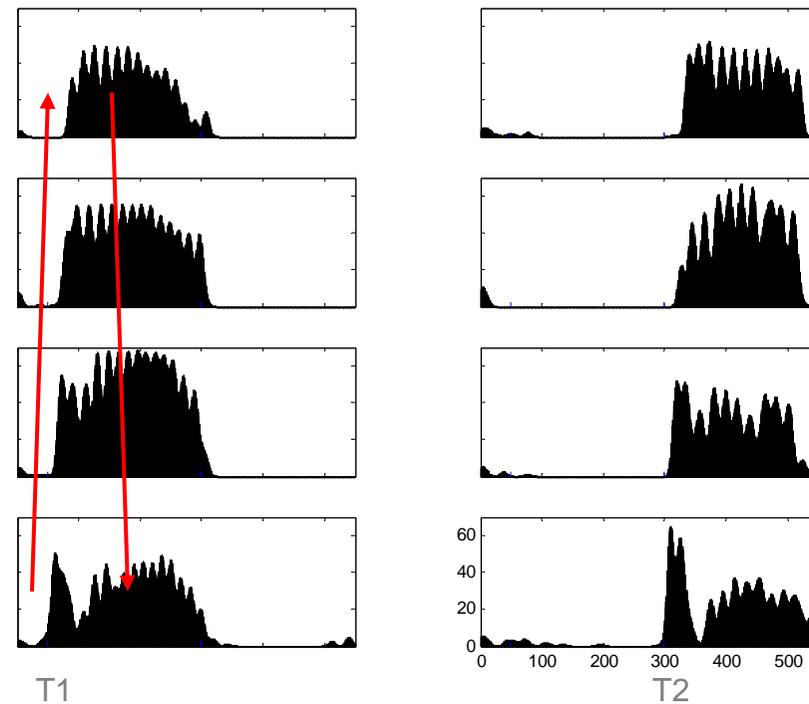
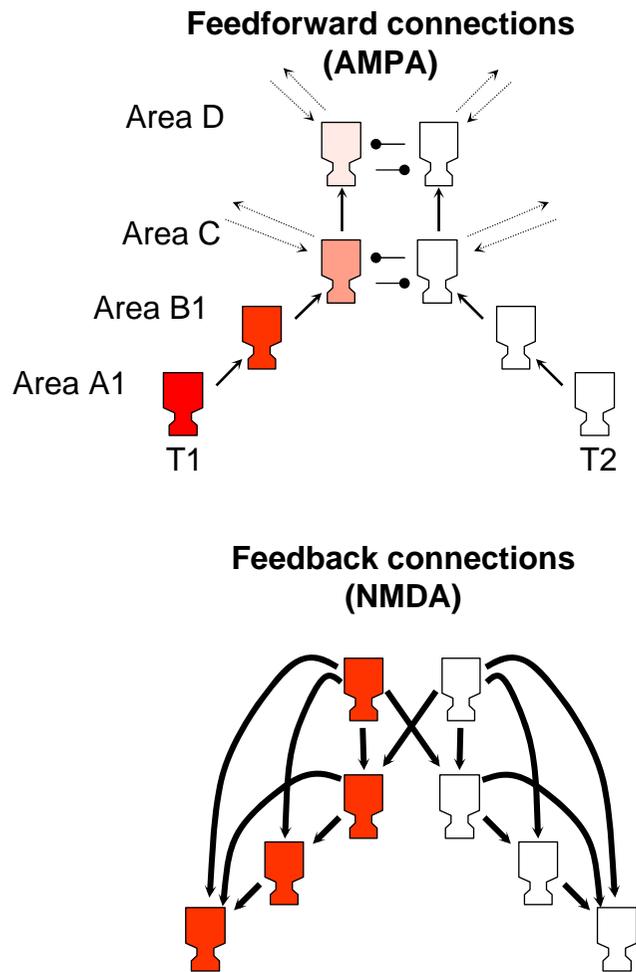
**Feedforward connections (AMPA)**



**Feedback connections (NMDA)**



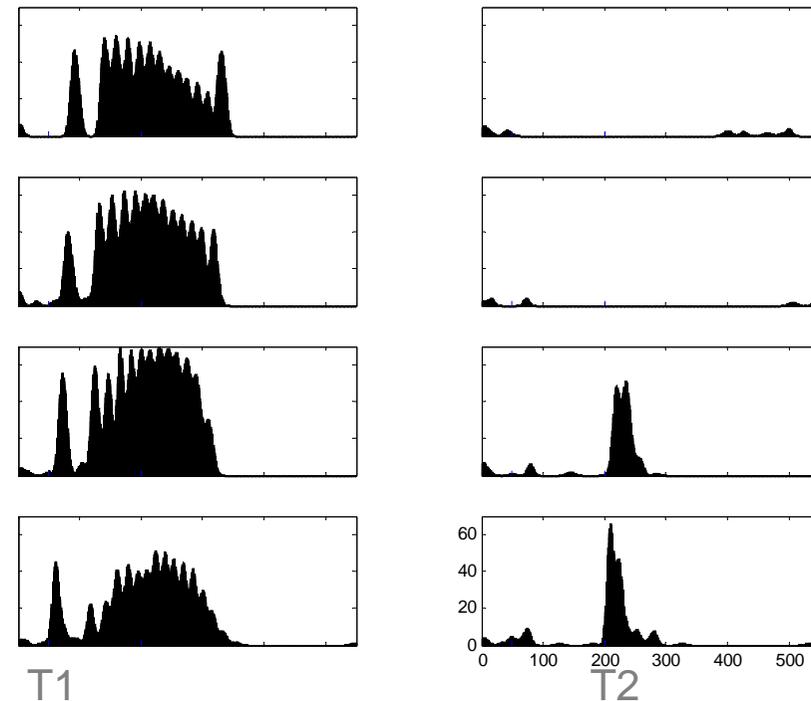
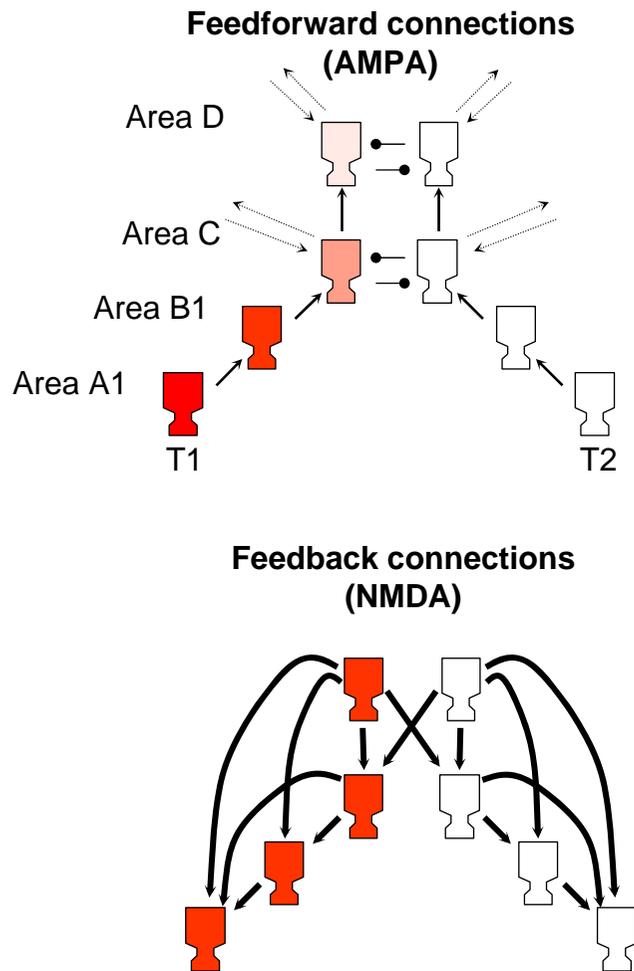
- L'accès au *workspace* se produit en deux étapes:
- propagation ascendante
  - amplification descendante et réverbération (ignition)



'Seen' trial (250 ms lag)

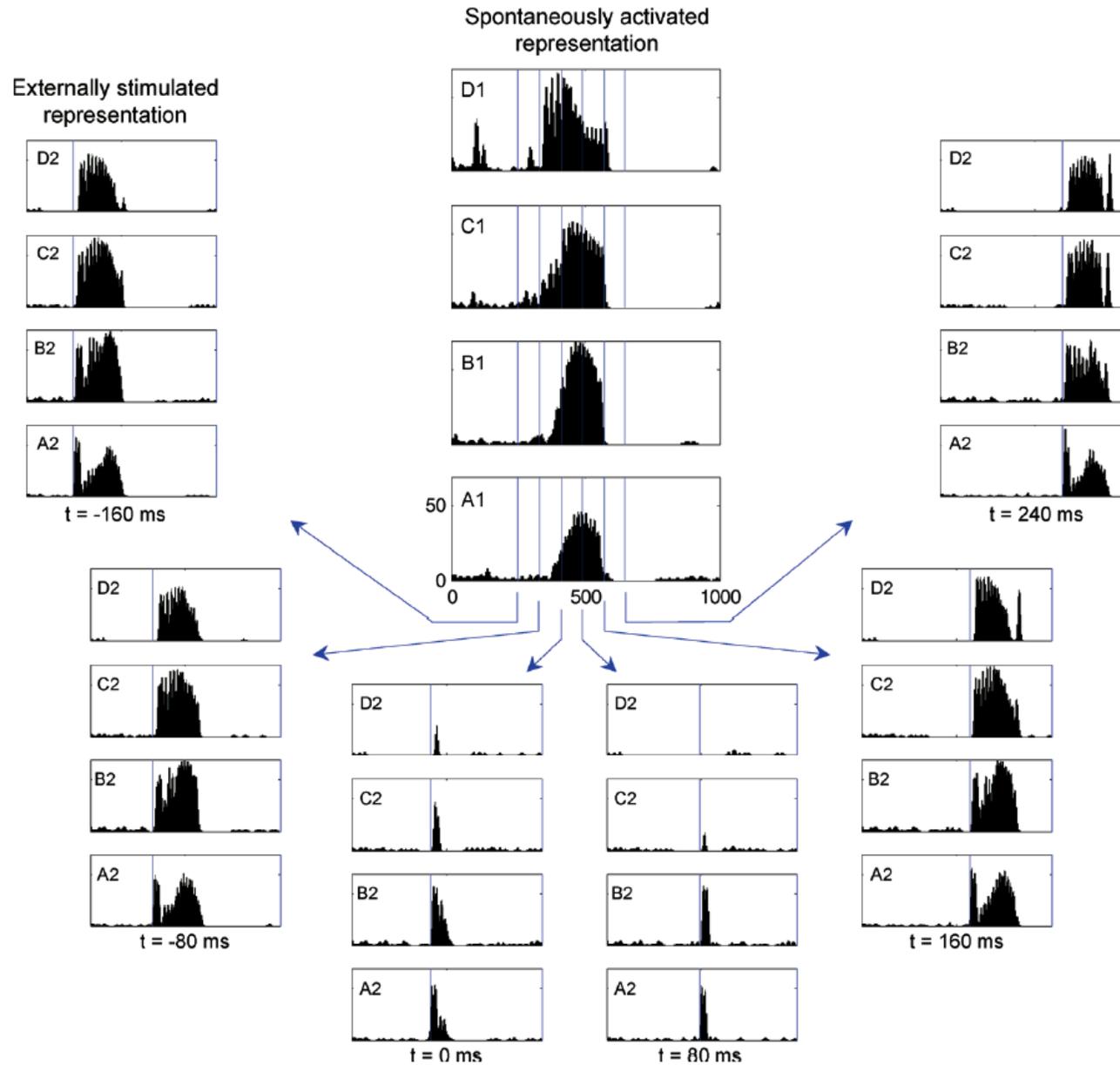
- L'accès au *workspace* se produit en deux étapes:
- propagation ascendante
  - amplification descendante et réverbération (ignition)

Si un second stimulus est présenté pendant l'ignition causé par le premier, il entre en compétition avec lui et peut ne pas conduire à une ignition consciente (*attentional blink*)

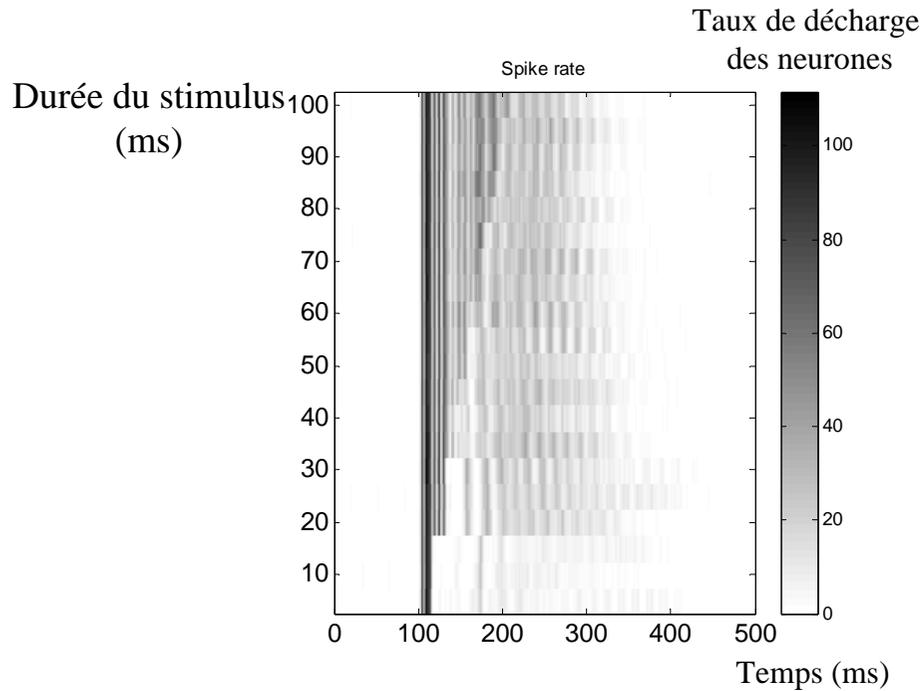


'Blinked' trial (150 ms lag)

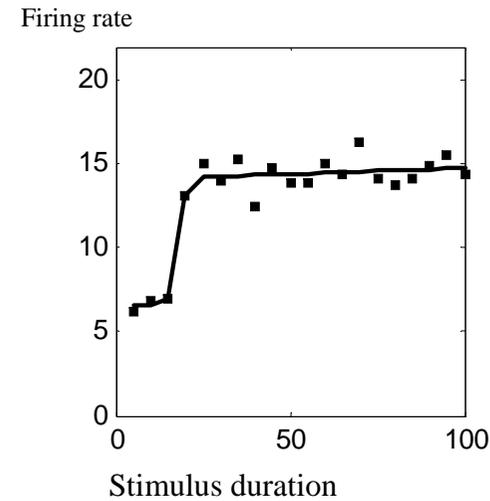
L'activité du réseau s'organise en séquences d'ignitions spontanées qui entrent en compétition avec les stimuli du monde extérieur.



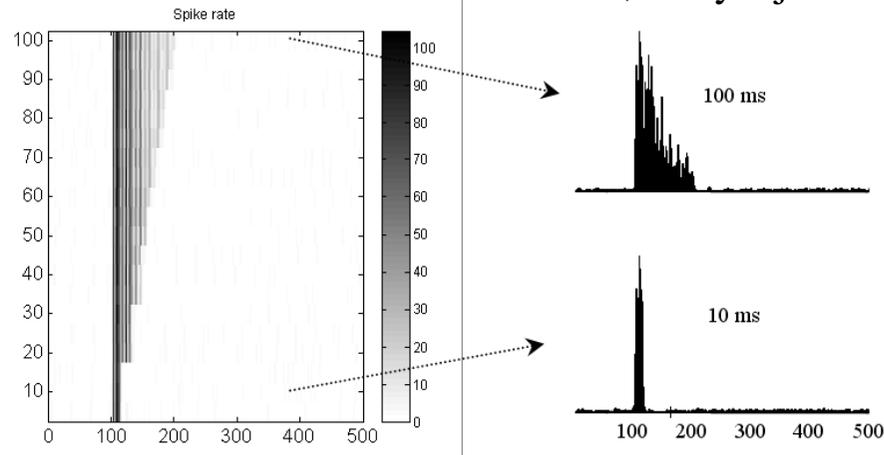
# Existence d'un seuil d'ignition globale en réponse à la présentation d'un seul stimulus



L'ignition tardive suit une fonction de seuil

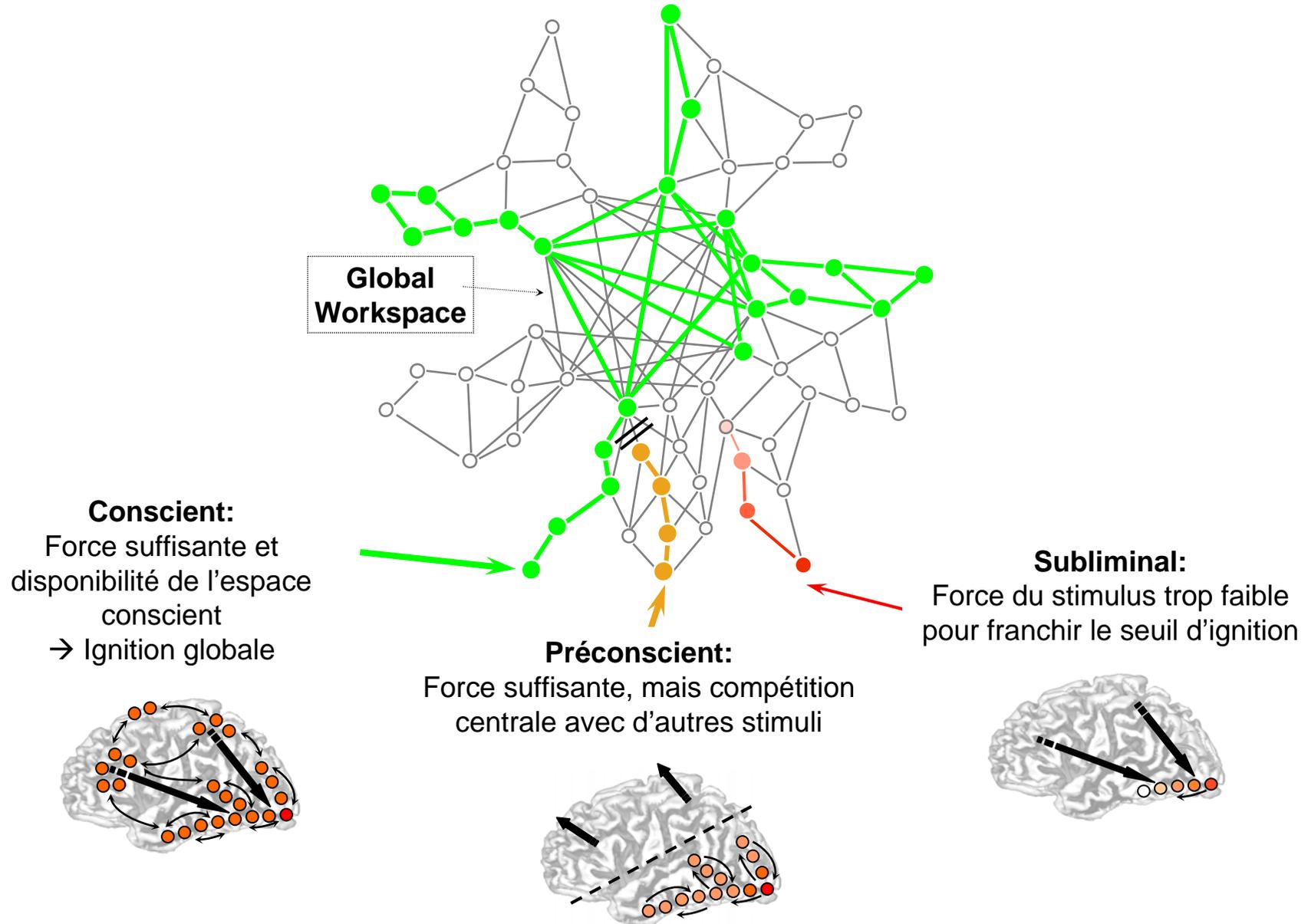


En l'absence de connexions descendantes, il n'y a jamais d'ignition



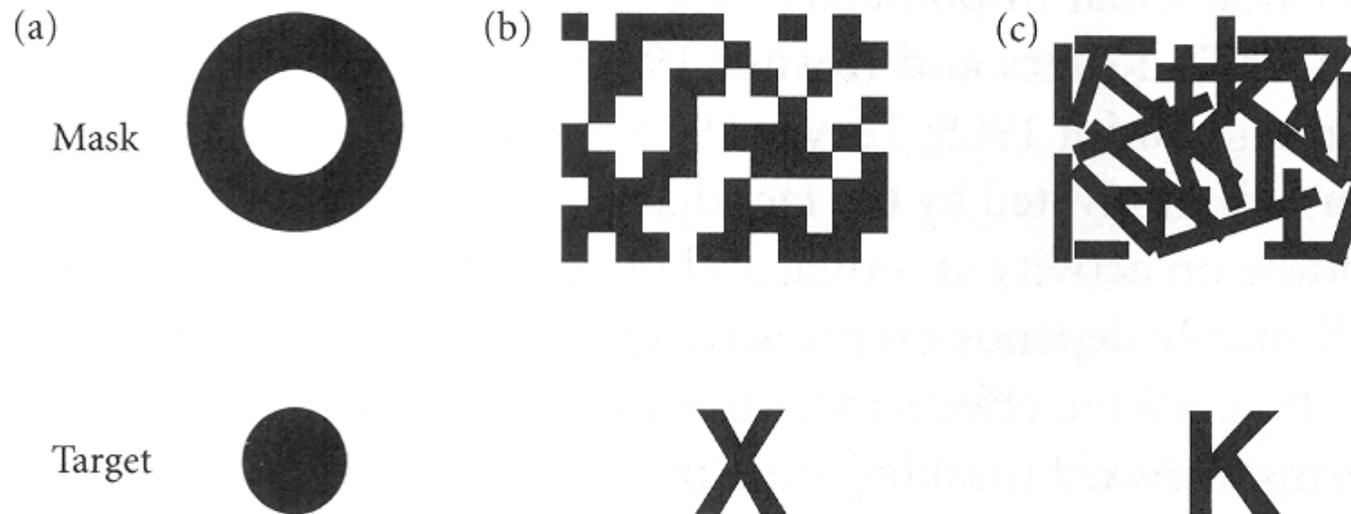
# Le modèle conduit à une taxonomie simple: traitement **subliminal**, **préconscient** ou **conscient**

Dehaene, Changeux, Naccache, Sackur, & Sergent, TICS, 2006



# Le masquage

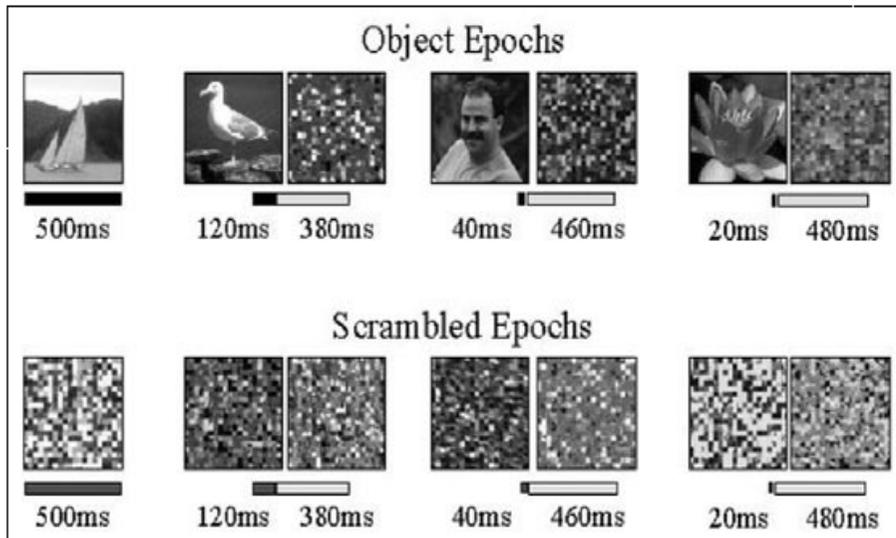
- Définition (Breitmeyer et Ogmen, 2006): « Réduction de la visibilité d'un stimulus présenté brièvement (la *cible*) par un second stimulus, également bref et adjacent dans l'espace et dans le temps »
- Différentes méthodes de masquage:



**Fig. 2.1** Examples of target and mask stimuli typically used in (a) paracontrast and metacontrast, (b) pattern masking by noise, and (c) pattern masking by structure. (Reproduced from Breitmeyer and Ganz 1976.)

# Première utilisation de l'IRM fonctionnelle pour étudier les bases cérébrales du masquage

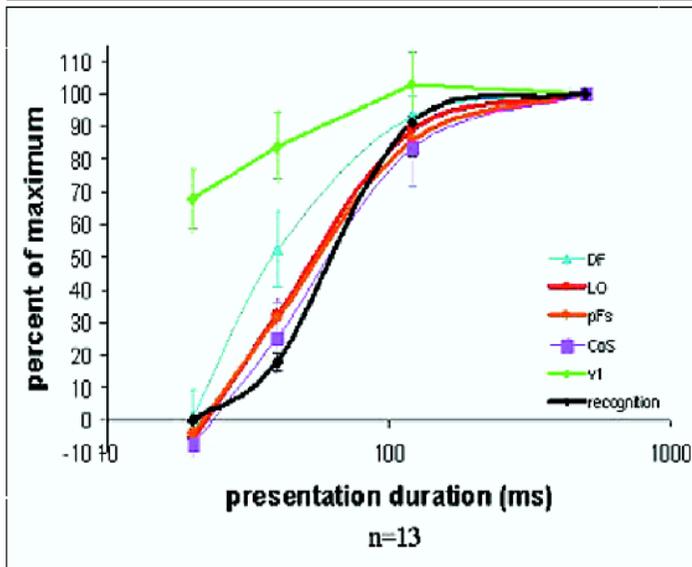
K. Grill-Spector, T. Kushnir, T. Hendler, R. Malach, *Nature Neurosci.* 3, 837-843 (2000).



Présentation, pour une durée variable, d'images masquées

Limitations:

- Etude restreinte aux régions occipitales
  - Paradigme en blocs (présentation de 28 images de durée fixe pendant 14 secondes)
- Pas de mesure essai par essai

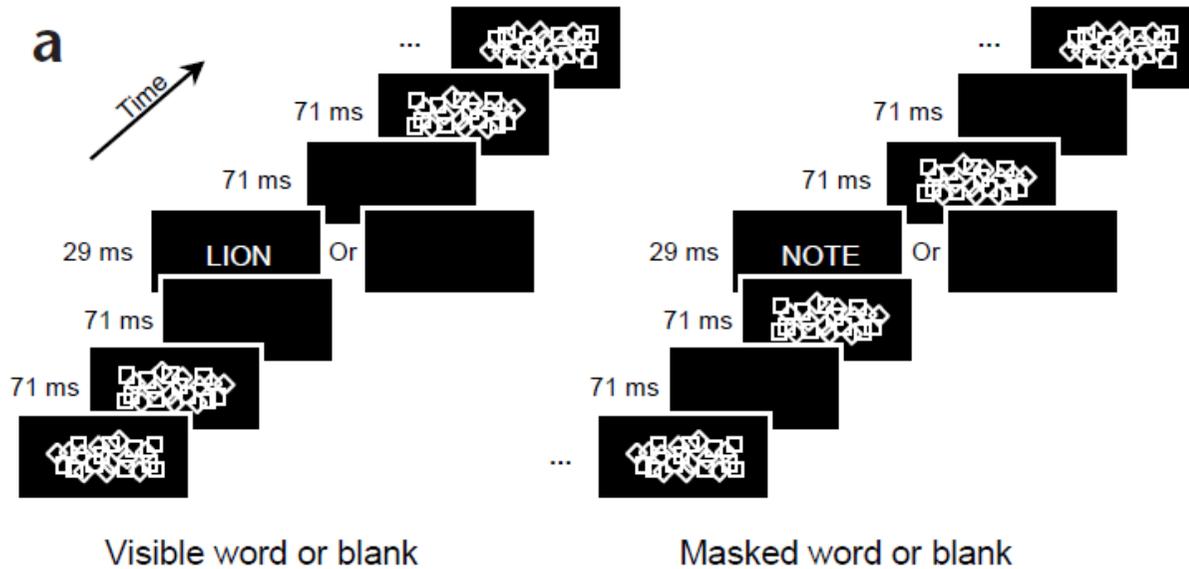


Résultats:

- Dans les aires les plus précoces, le signal reste élevé même aux durées les plus brèves, qui ne permettent pas de percevoir l'image
- Dans la région occipitale latérale, le signal suit étroitement la courbe de perception.

# L'activité cérébrale évoquée par des mots masqués ou démasqués

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nat Neurosci*, 4(7), 752-758.



IRM fonctionnelle en cerveau entier, en réponse à des événements uniques:

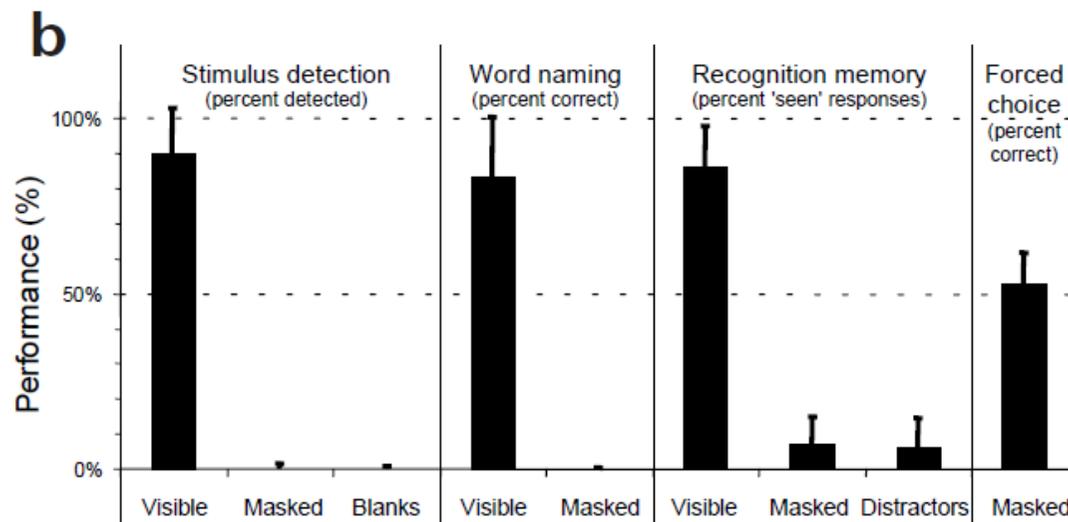
- Mots masqués ou démasqués
- Conditions de contrôle correspondantes (masques seuls)

Critiques:

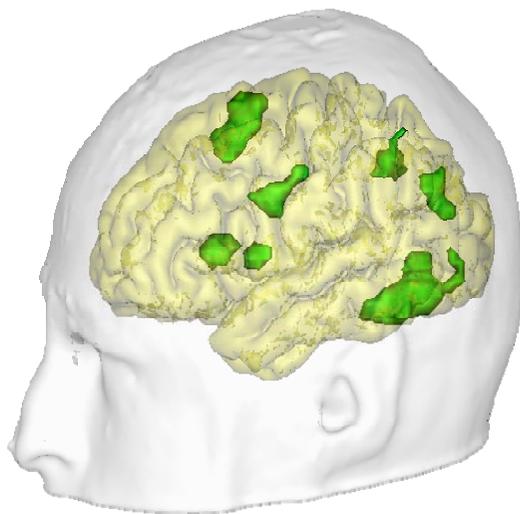
- Différence importante entre les deux conditions (partiellement contrôlé par la soustraction des masques)

- Pas de contrôle essai par essai de la perception, mais une vérification en fin d'expérience:

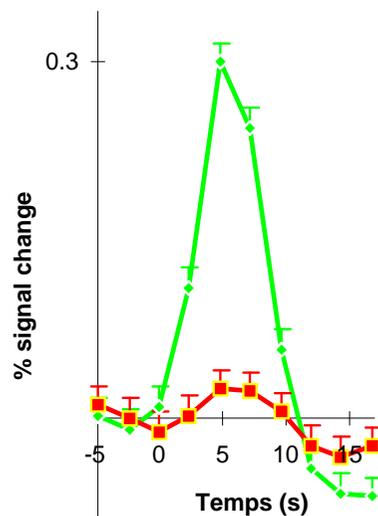
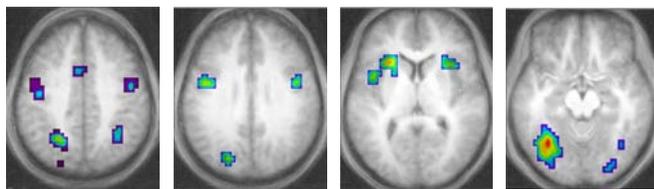
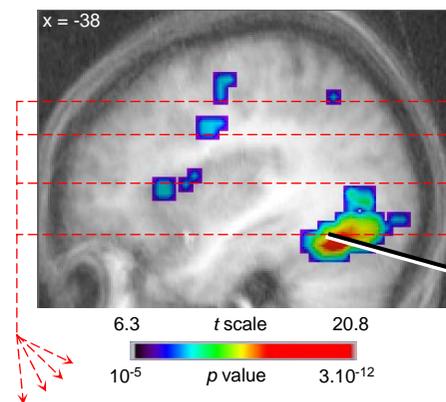
Les sujets ne sont pas capables de détecter, nommer, reconnaître ou sélectionner les mots masqués parmi plusieurs possibilités.



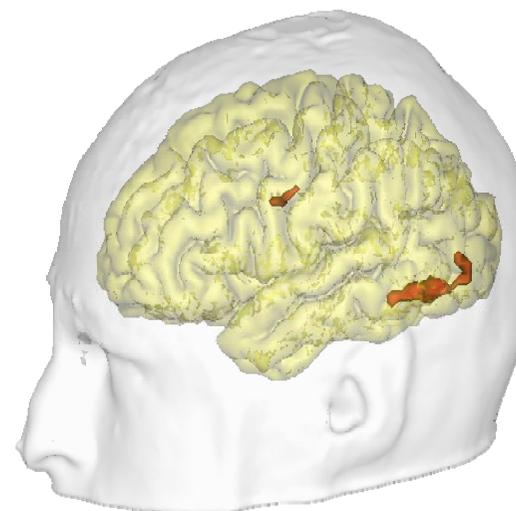
- Deux signatures hypothétiques de la conscience:
- L'amplification de l'activité des processeurs pertinents
  - L'entrée en activité d'un réseau pariéto-frontal global



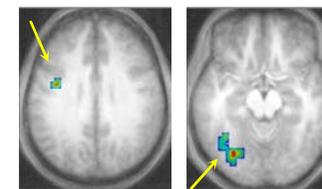
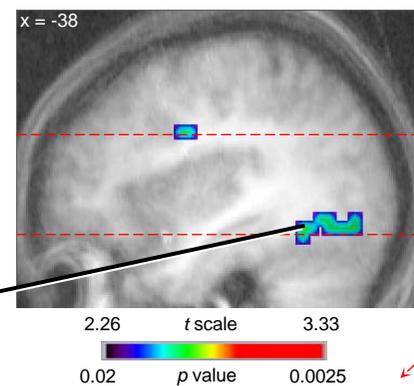
Mots visibles



left fusiform gyrus  
(-48, -60, -12)

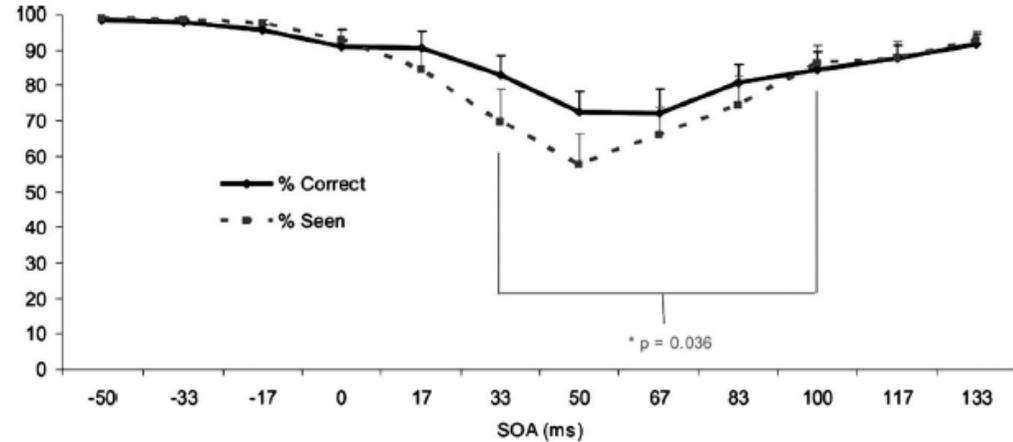
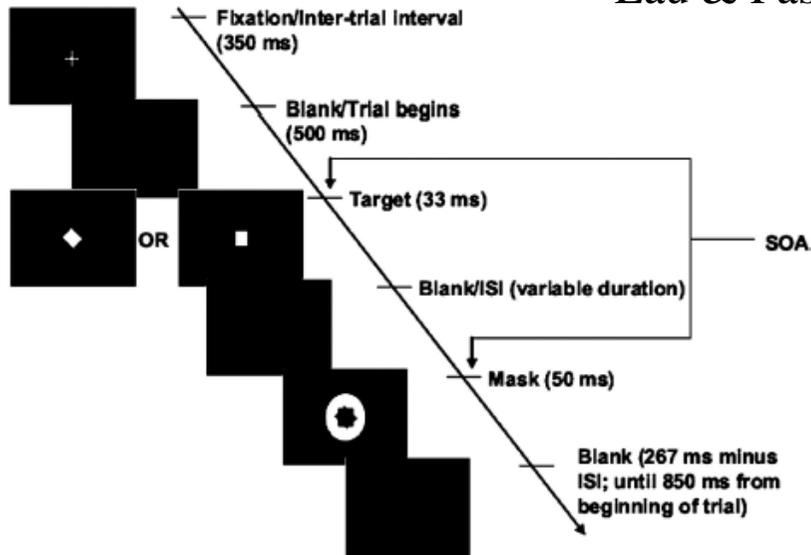


Mots masqués



# Une expérience où les performances sont égalisées

Lau & Passingham (*PNAS* 2006)



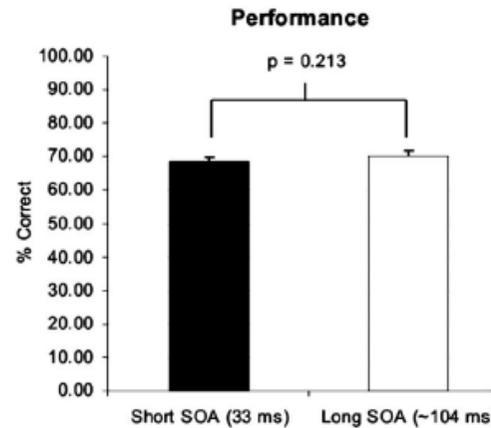
Stimulus visuel masqué

Après chaque essai, deux questions:

- Carré ou diamant?
- Vu ou deviné?

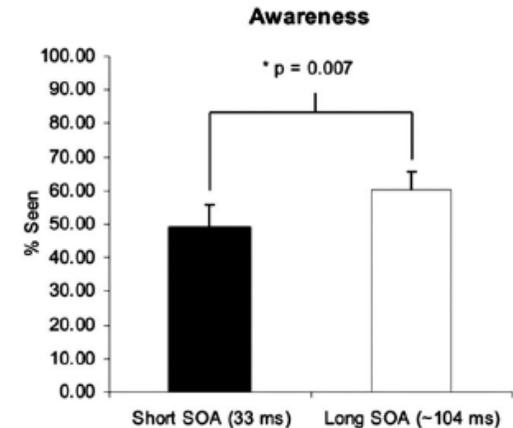
Les stimuli sont présentés à deux valeurs du SOA, choisies pour différer dans leur visibilité

Le SOA long et la taille des stimuli à SOA courts sont ajustés en permanence en sorte que la performance reste autour de 71%.



Jugement à choix forcé: carré ou diamant?

→ Performance identique

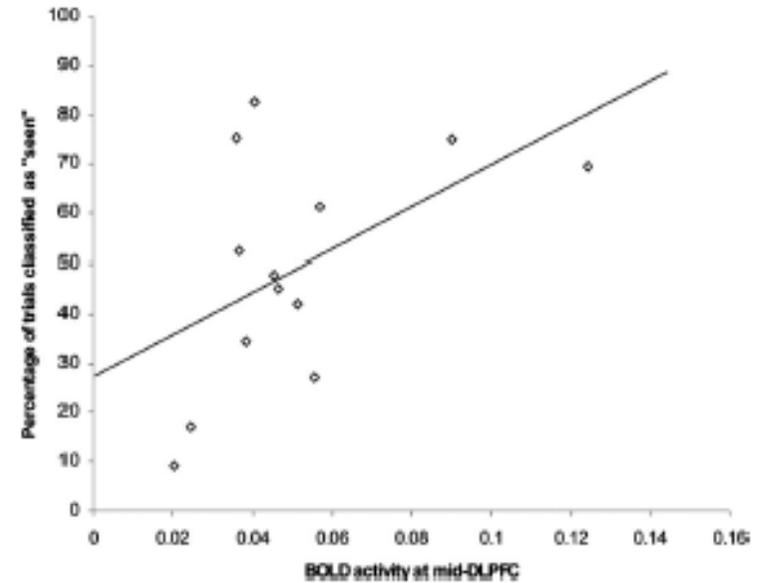
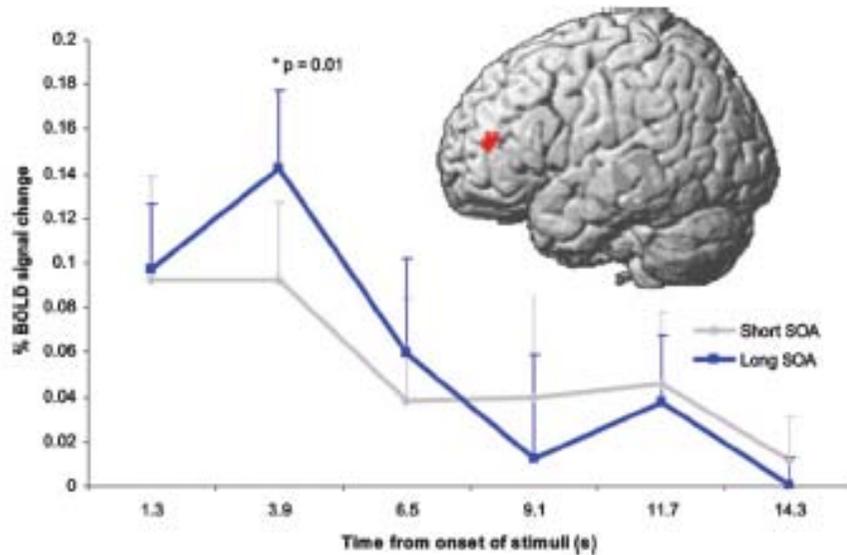


Jugement vu ou deviné:

→ Les stimuli sont plus visibles à SOA long

# Une expérience où les performances sont égalisées

Lau & Passingham (PNAS 2006)



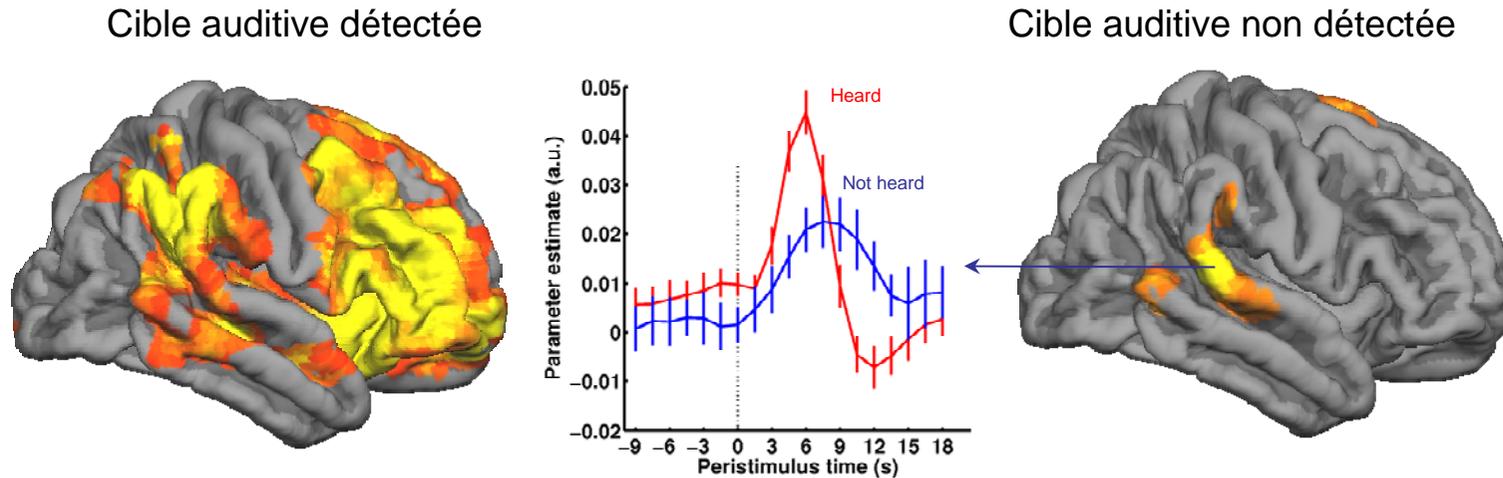
## Résultats:

Seul le cortex préfrontal dorsolatéral gauche montre un surcroît d'activité au SOA long, c'est-à-dire pour les essais où la visibilité est meilleure.

Dans cette région, l'activité est corrélée avec le taux d'essais classifiés come vus au SOA court.

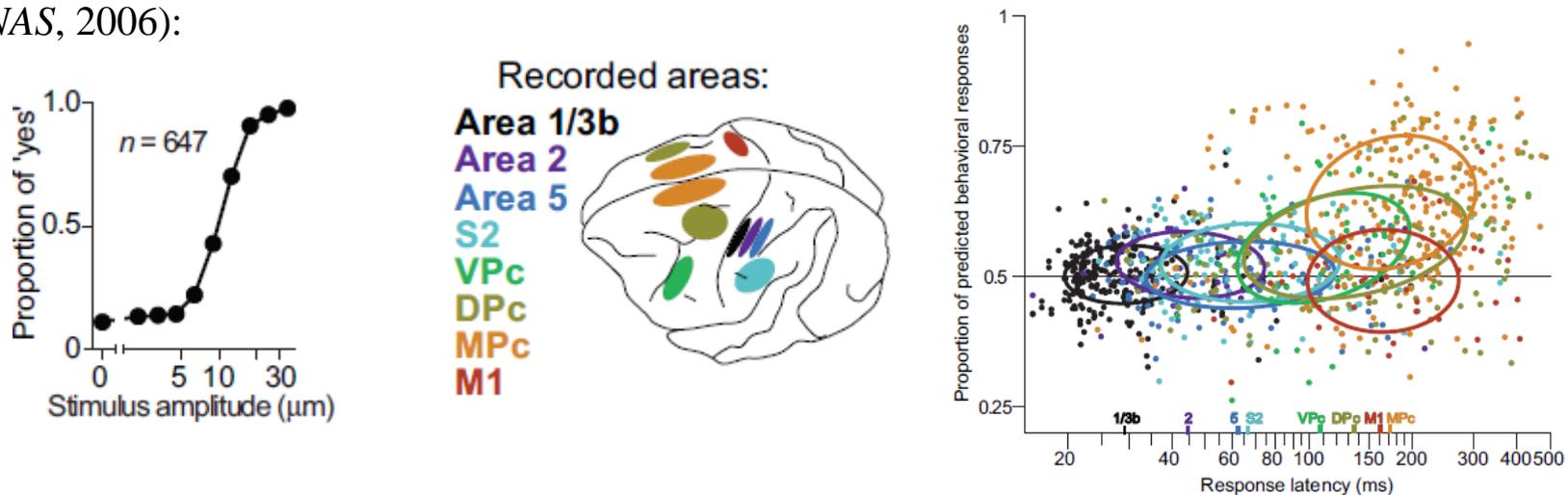
Ainsi, le résultat ne peut pas être dû aux (petites) différences entre les deux types de stimuli.

# L'accès à la conscience dans d'autres modalités



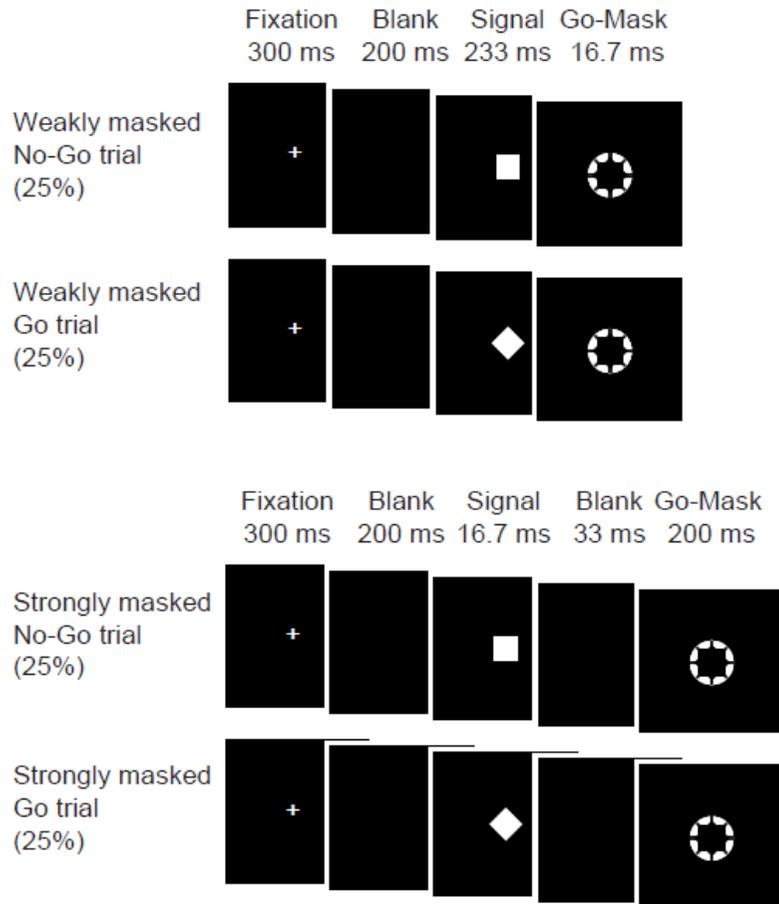
Sadaghiani et al. (*J. Neuroscience*, 2009): La détection de rares cibles auditives s'accompagne d'une très intense activation d'un réseau préfronto-pariétal bilatérale, avec amplification des aires auditives temporales. Celles-ci s'activent pour les stimuli non-détectés.

Résultats similaires dans la modalité tactile: Boly et al. (*PNAS*, 2007); de Lafuente, Romo et al (*PNAS*, 2006):



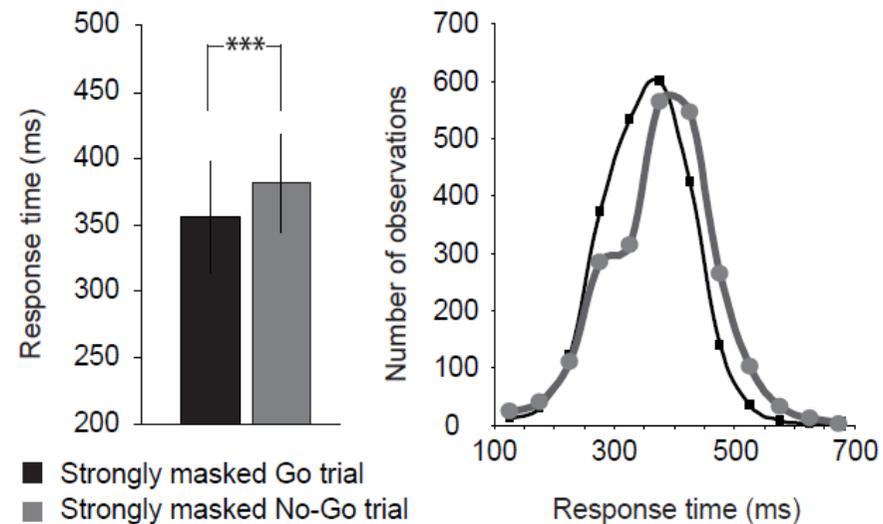
# Le contrôle exécutif conscient ou non-conscient

Van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. (2010). Unconscious activation of the prefrontal no-go network. *submitted*.



Le carré implique une réponse motrice (*go*), tandis que le diamant (*no-go*) demande d'inhiber cette réponse prépondérante.

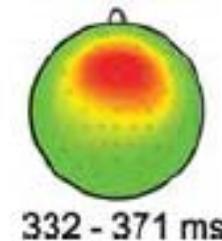
L'inhibition de la réponse est présente même lorsque les indices sont masqués et rendus subliminaux.



Dans le cours de l'an dernier, nous avons vu qu'un indice no-go, même subliminal, ralentit les réponses et induit une onde frontale P3a

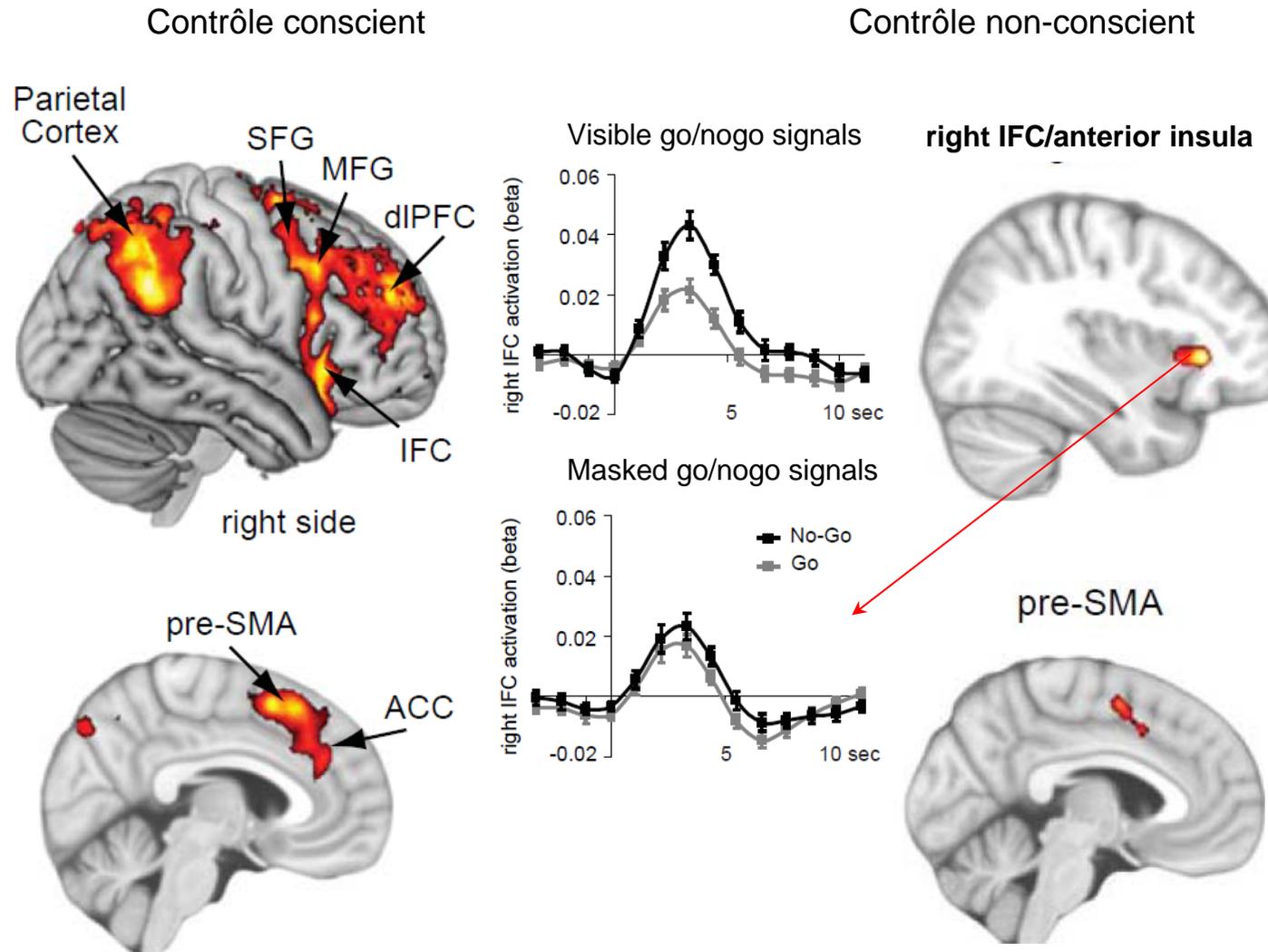
van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2008).

Frontal cortex mediates unconsciously triggered inhibitory control. *J Neurosci*, 28(32), 8053-8062.



# Le contrôle exécutif conscient ou non-conscient

Van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. (2010). Unconscious activation of the prefrontal no-go network. *submitted*.



## Conclusions:

Certaines régions préfrontales (pré-SMA, insula antérieure) sont susceptibles de s'activer non-consciemment.

Toutefois, à nouveau, la prise de conscience s'accompagne d'une amplification de ces réseaux et l'extension de l'activation à de vastes régions préfrontales et pariétales inférieures bilatérales

(remerciements à Simon Van Gaal)

Une exigence supplémentaire  
pour un corrélat neuronal de la conscience:  
la capacité de décoder précisément le contenu conscient

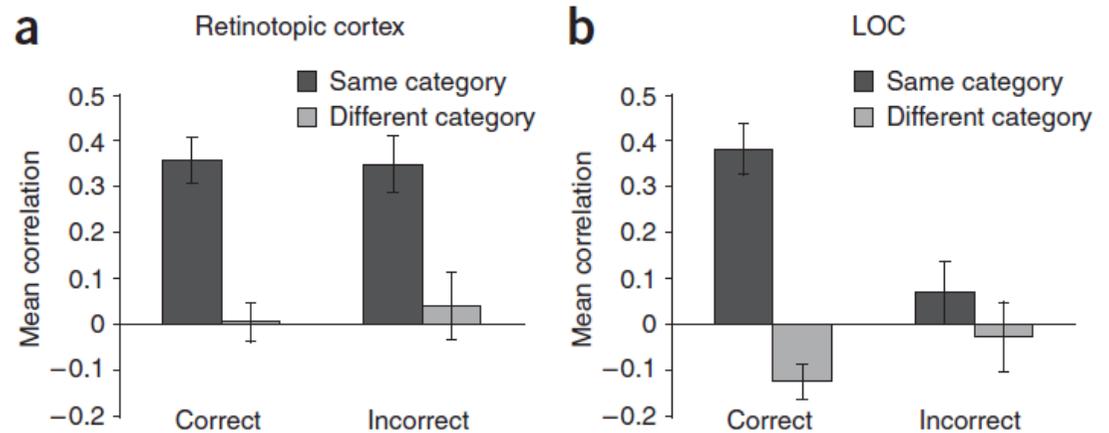
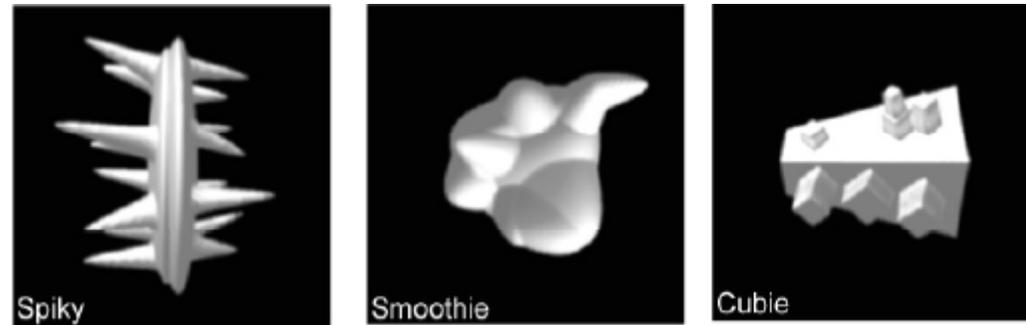
Haynes, J. D. (2009). Decoding visual consciousness from human brain signals. *Trends Cogn Sci*, 13, 194-202.

- Selon Haynes (2009), il ne suffit pas de détecter la présence d'un signal lors des essais conscients, qui ne serait pas présent lors des essais non-conscients
- Il faut également montrer que ce signal constitue un *code neural* de l'état de conscience du sujet
  - Un code **total** et **inversible**, c'est-à-dire que l'on peut trouver un algorithme de décodage qui, partant de l'activation observée, reconstitue précisément l'état de conscience du sujet, et rien que celui-ci
- Le code observé doit
  - Présenter le même degré d'invariance que le contenu conscient du sujet
  - Préserver les relations de similarité subjective entre les contenus conscients
  - Ne pas coder des différences physiques que le sujet ne perçoit pas
  - Coder l'objet perçu, même lorsqu'il s'agit d'une fausse perception

# Le décodage de l'objet perçu

Williams, M. A., Dang, S., & Kanwisher, N. G. (2007). Only some spatial patterns of fMRI response are read out in task performance. *Nat Neurosci*, 10(6), 685-686.

- Présentation brève (33 ms) et masquée, pendant l'IRMf, d'objets de trois catégories différentes
- Tâche = identifier la catégorie
- Corrélation entre la configuration d'activation observée lors des essais pairs et impairs → index de la présence d'un code neural de l'objet présenté.
- Les aires rétinotopiques du cortex occipital codent pour l'identité objective de l'objet présenté, même quand le sujet se trompe.
- Le cortex occipital latéral (LOC) ne code pour l'objet que lorsque que celui est correctement identifié.
- Le LOC est un bon candidat pour un code neural de la conscience des formes

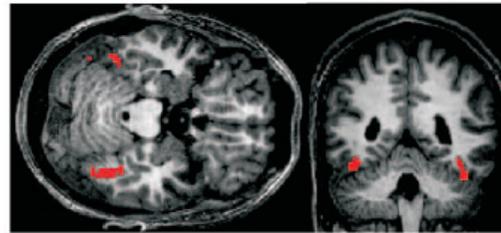


# Le décodage de l'objet perçu

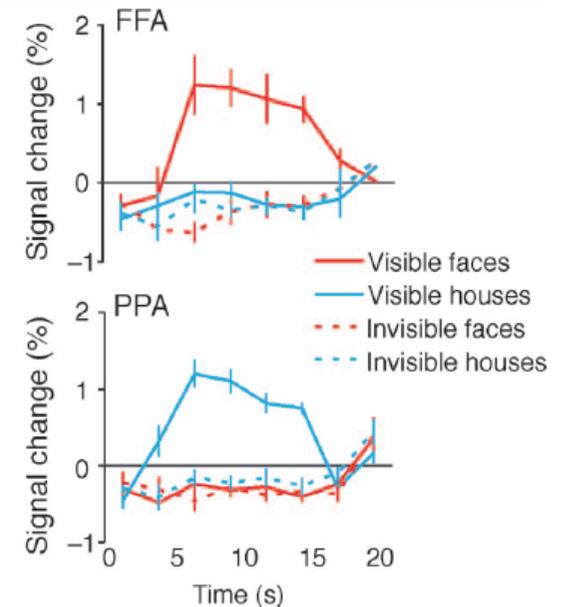
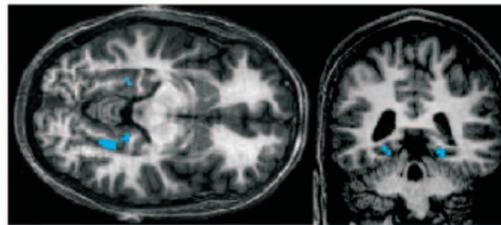
Sterzer, Haynes & Rees (Journal of Vision, 2008)

- Classification visage / maison, visibles ou rendus invisibles par suppression binoculaire par des flashes
- Activation nette de la FFA et de la PPA, seulement en réponse aux images conscientes
- Présence de configurations d'activation qui permettent de classifier les essais visibles, mais également les essais invisibles
- Généralisation, dans les deux sens
  - A peine significative, mais néanmoins totale, lors de l'entraînement avec des stimuli invisibles et test avec stimuli visibles
- Résultats compatibles avec l'existence d'un code partiellement activé par les stimuli invisibles, mais amplifié et étendu lors de l'accès à la conscience
- Pas d'analyse en cerveau entier, ni de corrélation fonctionnelle avec le reste du cerveau.

FFA = Faces > Houses



PPA = Houses > Faces

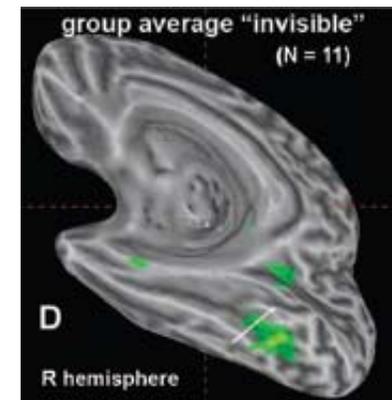
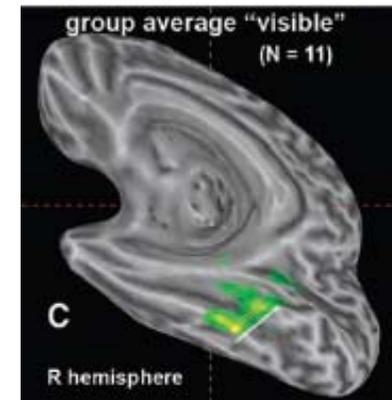
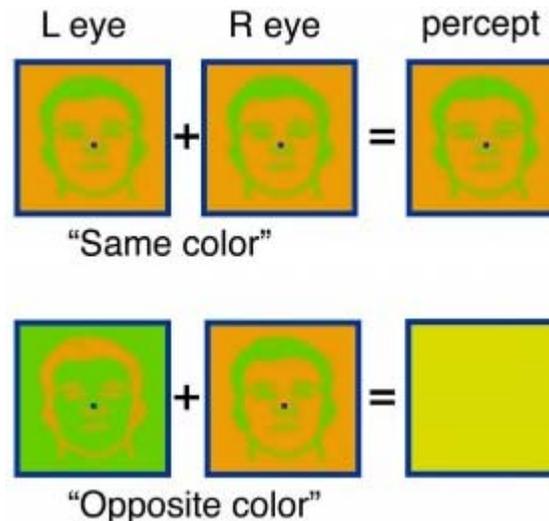


% de classification (à partir des deux régions)	Test avec essais VISIBLES	Test avec essais INVISIBLES
Entraînement avec essais VISIBLES	88.7% $p < 0.001$	71.3% $P = 0.008$
Entraînement avec essais INVISIBLES	63.1% ( $p = 0.10$ )	63.5% $p = 0.02$

# Le décodage de l'objet perçu

Schurger, Pereira, Treisman & Cohen (*Science* 2009)

- Classification visage / maison, visibles ou rendus invisibles par fusion binoculaire de couleurs
- Activation de la PPA, et plus faiblement de la FFA, seulement en réponse aux images conscientes (communication personnelle)
- Présence de configurations d'activation qui permettent de classifier les essais visibles, mais également les essais invisibles
- Les régions pertinentes sont plus antérieures lors des essais visibles que lors des essais invisibles
- Pas de généralisation croisée
- Variabilité plus grande, dans les voxels postérieurs, pour les essais invisibles
- Conclusion: code partiellement activé par les stimuli invisibles; l'accès à la conscience rend les activations (1) plus fortes (2) moins variables (3) plus antérieures dans la région ventrale



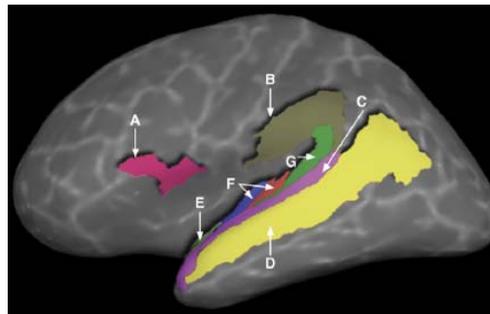
% de classification (à partir des deux régions)	Test avec essais VISIBLES	Test avec essais INVISIBLES
Entraînement avec essais VISIBLES	63% p<0.002	48% n.s.
Entraînement avec essais INVISIBLES	52% n.s.	58% p=0.02

## Exemple 2: le code neural de la perception audio-visuelle

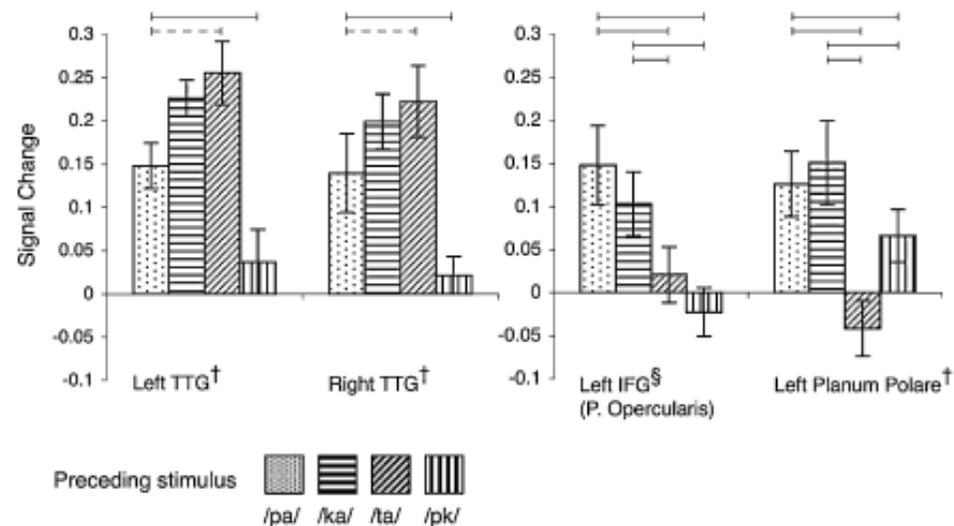
Hasson, U., Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2007).

Abstract coding of audiovisual speech: beyond sensory representation. *Neuron*, 56(6), 1116-1126.

- L'illusion de McGurk: un son « pa », présenté en même qu'une vidéo de la bouche prononçant « ka », est perçu comme « ta »
- Cette situation permet de dissocier le stimulus objectif de sa perception subjective
- Expérience de répétition en IRMf: cible pa-ka précédée d'un stimulus objectivement ou subjectivement répété



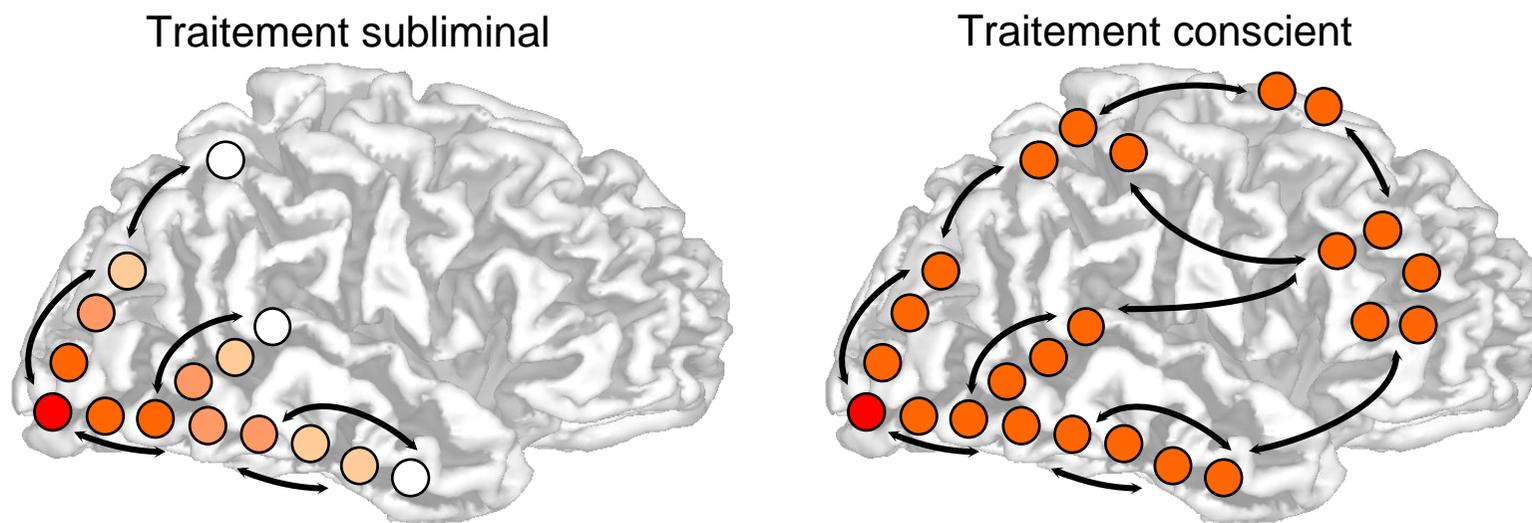
Preceding Stimulus	Dimension of Overlap with Target /PK/		
	Auditory	Visual	Perceptual
/PA/ (PA <sub>Aud</sub> PA <sub>Vid</sub> )	✓	×	×
/KA/ (KA <sub>Aud</sub> KA <sub>Vid</sub> )	×	✓	×
/TA/ (TA <sub>Aud</sub> TA <sub>Vid</sub> )	×	×	✓
/PK/ (PA <sub>Aud</sub> KA <sub>Vid</sub> )	✓	✓	✓



Les régions temporales bilatérales (F) répondent à la répétition du stimulus objectif, tandis que les régions frontale inférieure gauche (A) et du pôle temporal (E) répondent à la répétition subjective

→ la conscience perceptive *auditive* n'est pas issue des régions auditives précoces, mais d'un traitement plus avancé et cross-modal, impliquant conjointement des régions temporales et préfrontales

## Conclusion: Corrélats de la conscience en IRMf



- Amplification de l'activation sensorielle
- Progression de l'activation vers des régions sensorielles plus antérieures
- Activation additionnelle de régions très distribuées du cortex préfrontal et pariétal inférieur
- Mise en corrélation de ces régions à longue distance

La semaine prochaine, nous examinerons le déroulement temporel de ces activations, et la présence de “signatures” de l'accès conscient