Stepping out of the Chinese Room: Word meaning with and without consciousness

Roberto Bottini
Center for Mind/Brain Sciences,
University of Trento, Italy
bottini.r@gmail.com

Daniel Casasanto
University of Chicago, IL
USA
casasanto@uchicago.edu

Andrea Nadalini
International School for Advanced Studies (SISSA)
Trieste, Italy
anadalini@sissa.it

Davide Crepaldi
Milan Center for Neuroscience
Milan, Italy
davide.crepaldi@sissa.it

Abstract

English. What is the role of consciousness in language processing? Unconscious priming experiments show that words can prime other words with related meanings (cat – dog), and these priming effects are assumed to reflect the activation of conceptual knowledge in semantic memory. Alternatively, however, unconscious priming effects could reflect predictive relationships between the words' forms, since words that are semantically related are also statistically related in language use. Therefore, unconscious "semantic" priming effects could be due to relationships between words' forms mimicking conceptual relationships, as in Searle's Chinese Room thought experiment. To distinguish word-form-based and semantics-based accounts of priming we conducted an experiment in which temporal words (e.g., earlier, later) were preceded by spatial words that were processed either consciously or unconsciously. Time is typically conceptualized as a spatial continuum extending along either the sagittal (front-back) or the lateral (left-right) axis, but only the sagittal space-time mapping is encoded in language (e.g. the future is *ahead*, not *to the right*). Results showed that temporal words were primed both by sagittal words (back, front) and lateral words (left, right) when primes were perceived consciously, as predicted by both wordform-based and semantics-based accounts. Yet, only sagittal words produced an unconscious priming effect, as predicted by the word-form-based account. Unconscious word processing appears to be limited to relationships between words' forms, and consciousness may be needed to activate words' meanings.

Italiano. *Qual è il ruolo della coscienza nell'elaborazione semantica delle parole? Esperimenti di masked priming semantico mostrano che la vista di una parola può facilitare il riconoscimento di un'altra parola dal contenuto semantico simile (gatto – cane). Questo effetto di priming è solitamente interpretato come evidenza che la parola inconscia è processata a livello semantico. Tuttavia, tale effetto può essere spiegato anche sulla base di relazione tra forme lessicali (senza attivazione di informazione nella memoria semantica). Infatti, parole che sono semanticamente legate sono anche legate statisticamente nel linguaggio. Il priming semantico inconscio potrebbe semplicemente emulare relazioni concettuali, come nel famoso esperimento mentale della stanza cinese di Searle. Per distinguere il priming lessicale dal priming semantico abbiamo condotto un esperimento in cui parole temporali (ieri, domani) erano precedute da parole spaziali mostrate sia a livello subliminale che supraliminale. Il tempo è tipicamente concettualizzato attraverso mappe spaziali che si estendono lungo l'asse sagittale (il passato è dietro, e il futuro davanti) e lungo l'asse laterale (il passato è a sinistra, futuro a destra). Solo la mappatura sagittale è però codificata nel linguaggio (il futuro è davanti, non a destra). I risultati mostrano come sia le parole sagittali (dietro, davanti) che quelle laterali (sinistra, destra) facilitano l'elaborazione di parole temporali (ad esempio prima e dopo), quando percepite consciamente. Al contrario, quando i prime sono elaborati al di fuori della coscienza, l'effetto sull'asse laterale viene meno. Il processo inconscio delle parole sembra dunque essere limitato a relazioni tra forme lessicali; la coscienza*

potrebbe essere necessaria per attivarne il significato.

1 Introduction

What role does consciousness play in word meaning's construction? As previous literature has pointed out, lexical items seem to be processed up to the semantic level even when processed out of awareness (Quinn & Kinoshita, 2008; Ansorge, Kiefer, Khalid, Grassl, & König, 2010). Most evidence for this claim comes from masked priming: When two words are sequentially presented, the recognition of the latter is made easier if the two are semantically related (cat-dog), even when the visibility of the former (the prime) is prevented by displaying it very briefly, embedded between visual masks (Forster, 2006; Dehaene et al., 1998). For instance, participants are likely to classify more quickly the word *dog* as referring to a living entity when it is preceded by the semantically related word *cat*, rather than by a semantically unrelated word like *apple*. As similar effects are attested when the prime word is clearly visible, it has been suggested that lexical items can be processed up to the semantic level irrespective of their visibility. We will refer to this perspective as the *semantic-based account* of masked priming, as it assumes that words are processed beyond their surface structure and activate conceptual knowledge about their referents. Such knowledge is thought to be stored within the semantic memory, where concepts and concepts' features are represented in an interconnected network (Tulving, 1972; Masson, 1995). In such a view, *cat* would prime *dog* as both words refer to mammals that have four legs, have a tail, can be pet, and so on; and therefore they are more closely related to each other than to *apple*.

However, there is an aspect of semantic similarity that has been largely overlooked in the priming-related literature, that words with similar meaning tend to have a similar contextual distribution. As corpus-based studies have pointed out, words referring to entities with similar perceptual and conceptual attributes tend be used together (e.g., *dog* and *cat* are more likely to co-occur in the same sentence than *dog* and *apple*), and to be used in similar contexts (e.g., both *cat* and *dog* tend to appear when speaking about pets, whether or not they co-occur within a given utterance; Louwerse, 2011; Landauer & Dumais, 1997). Based on this fact, unconscious priming may be alternatively explained through predictive relationships between words' forms established in language use. According to the *word-form-based account*, *cat* would prime *dog simply* because the two words share a similar contextual distribution. No conceptual representation is involved, as the locus of the unconscious semantic priming would be the lexical system, not semantic memory (Collins & Loftus, 1975).

This latter interpretation of unconscious word processing somehow resembles the Chinese Room thought experiment developed by the philosopher John Searle, where an English-speaking man is closed in a room receiving message written in Chinese characters (Searle, 1980). Due to a set of norms that determine the relationships between those characters (if you see X followed by Y, than reply Z), he is able to provide answers that would look perfectly appropriate to a native speaker. From the outside, it would appear that the man has a good understanding of the language, while instead he is acting on the basis of associations between word forms. Indeed, if he received a message saying that the room is about to explode, he would reply appropriately; but would not leave the room.

For the *semantic-based* and *wordform-based* accounts of masked priming to be distinguished, it is necessary to find concepts that are related in the semantic system, but not in the lexical system. This is the case for the metaphorical relationship linking time to space. The two domains are strictly intertwined in the human mind, in such a way that space is often used to think about time (Lakoff & Johnson, 1980). Time conceptualization involves both the sagittal and the lateral axis (Casasanto & Bottini, 2014; Bonato, Zorzi & Umiltà, 2012). For example, participants are faster in responding to past-related words by providing a leftward response, and to future-related words by providing a rightward response, relative to the opposite pattern. The same holds for the sagittal arrangement, with backward response associated with past-related words and forward response associated with future-related words. Moreover, neurological evidence shows that patients with hemispatial neglect have also impairments in temporal judgments (i.e. if they neglect the left side on space, they also show worst memory for past-related events; Saj, Fuhrman, Vuilleumier, & Boroditsky, 2013). Finally, people have been found to use hand gestures along both the lateral and the sagittal line when speaking about time (Casasanto & Jasmin,

2012). Critically, while the sagittal mapping is linguistically encoded in sentences such as "a bright future *in front of* you", the lateral one is not. The existence of these two mental timelines, and the fact that only one of them is linguistically expressed, offers the ideal test-bed for contrasting the *semantic-based* and *wordform-based* accounts of unconscious word processing. The latter predicts that space–time priming would only emerge along the linguistically encoded sagittal axis when primes are kept outside of awareness; while the former would predict priming to emerge along both axes, both supra- and subliminally.

We tested these predictions in a priming study with spatial words related to the lateral (*left-right*) and the sagittal (*ahead-behind*) axis as primes, and temporal words referring to either the past (*yesterday*) or the future (*tomorrow*) as target stimuli. In the first experiment, primes were clearly visible In the second experiment, prime visibility was prevented by means of a masking procedure.

2 Experiments

2.1 Experiment 1a - Visible primes

Participants: 60 volunteers were recruited for the experiment; all subjects were right-handed, and they all stated being native Italian speakers, with normal or corrected-to-normal vision and no history of neurological disorders. Each subject gave written informed consent for participation.

Stimuli, apparatus and procedure: Primes were 2 spatial words related to the lateral axis ("sinistra", *left*, and "destra", *right*) and 2 spatial words related to the sagittal axis ("davanti", *front*, and "dietro", *back*). Target stimuli were 8 temporal words. Four of them refer to the past ("prima", *earlier*, "ieri", *yesterday*, "passato", *past*, "scorso", *previous*), and four refers to the future ("dopo", *later*, "domani", *tomorrow*, "futuro", *future*, "successivo", *next*).

Each trial consisted of a fixation point (+) displayed for 750 ms. Then a blank screen was shown for 200ms, followed by the prime and by another blank screen, both lasting 50 ms. Finally, the target word was presented for 1500 ms, or until a response was provided.

Participants engaged in a GO/NO_GO task: They had to press a key if the target word referred to the past and do nothing if the target word referred to the future, or vice versa, according to the block instructions.

Results and discussion: analyses were conducted only on "GO" trials. Inaccurate trials (less than 1%) were excluded. In order to reduce the effect of outliers, those individual datapoints standing at more than 2 standard deviations from each participant's mean (~5% of the correct trials) were also removed from the analyses. A 2–by–2 ANOVA on the log-transformed RTs revealed a significant main effect of Congruity, $F(1, 59)= 11.47$, $p= 0.001$, indicating that participants were faster in congruent trials (535 ms) compared to incongruent ones (540 ms). We found no effect of Axis, $F(1, 59)= 0.41$, $p> .250$, and no Axis by Congruity interaction, $F(1, 59)= 0.06$, $p>0.250$. Pairwise comparisons showed that the priming effect was *significant* both in the sagittal (4 ms; $F(1, 59)= 5.79$, $p= 0.02$) and the lateral axis (6 ms; $F(1, 59)= 6.76$, $p= 0.01$).

Thus, significant congruity effects were produced both by sagittal and lateral spatial prime words, consistent with previous studies that provide evidence for sagittal and lateral mental timelines.

2.2 Experiment 1b - Subliminal primes

Participants: 60 volunteers from the same population as in Experiment 1a were recruited into the experiment. None of them took part in the previous experiment.

Stimuli, apparatus and procedure were the same as in Experiment 1a with one exception, i.e the blank screens that were displayed before and after the prime word were replaced with two visual masks in order to make the prime invisible (subjects were not informed of the presence of the primes).

Prime visibility task: after the end of the last part of the experiment, all subjects were informed about the presence of the prime word between the masks. Then, they performed a prime visibility test that included 10 practice and 128 experimental trials. The stimuli to be detected were the same spatial words we used in the previous experiment in half of the trials, and a string of identical lowercase letters (<aaaaaaaa>) in the other half.

Results and discussion: only the "GO" trials, in which participants provided a response, were analyzed. Inaccurate trials (less than 1%) were excluded. In order to reduce the effect of extremely long and short RTs, those individual datapoints standing at more than 2 standard deviations from each participant's mean (~4% of the correct trials) were also removed from the analyses.

A 2–by–2 ANOVA on the log-transformed RTs revealed a significant main effect of Congruity, $F(1, 58)= 27.63$, $p< .001$, and an Axis by Congruity interaction, $F(1, 58)= 14.986$, $p< 0.001$, which we followed up through pairwise comparisons showing that priming was significant for the sagittal axis (9 ms; $F(1, 58)= 40.21$, $p< 0.001$), but not for the lateral axis (2 ms; $F(1, 58)= 1.52$, $p= 0.22$).

No participant reported having noticed the prime. Overall, the average d-prime value was 0.33 (SD= 0.37). Although significantly different from zero, $t(58)= 7.03$, $p< 0.001$, this value is widely taken to indicate that the prime was effectively masked from perceivers' awareness (Kouider & Dupoux, 2005).

Experiment 1b clearly suggests that spatio–temporal masked priming is limited to the sagittal axis, with no apparent effect on the lateral axis. Thus, the pattern of results provides evidence in favor of the *wordform-based* account of unconscious word processing.

3 Conclusion

In this study we looked at the nature of word processing with and without awareness. Using the relationships between space and time, we were able to disentangle the *wordform-based* from the *semantic-based* account of masked priming. When words were clearly visible, we found priming effect on both axes, which reflects the sagittal, linguistically encoded, timeline, as well as the lateral mapping, which relies only on conceptual knowledge. Conversely, subliminal priming was obtained only with the sagittal words, matching the predictions of the *wordform-based* account. Therefore, our data suggests that when people read words unconsciously, activation spreads only between predictively related wordforms.

Unconscious priming between semantically related words may mimic semantic priming, much in the same way as the man inside Searle's Chinese Room mimics knowledge of Chinese, on the basis of "meaningless" wordform-wordform relationships. Unconscious word processing appears to be limited to relationships between words' forms, and consciousness may be needed to activate words' meanings.

Reference

Ansorge, U., Kiefer, M., Khalid, S., Grassl, S., & König, P. (2010). Testing the theory of embodied cognition with subliminal words. *Cognition*, 116(3),303–320.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2257–73.

Casasanto, D., & Bottini, R. (2014). Mirror Reading Can Reverse the Flow of Time. *Journal of Experimental Psychology. General*, *143*(2).

Casasanto, D., & Jasmin, K. (2012). The Hands of Time: Temporal gestures in English speakers. *Cognitive Linguistics*, *23*(4), 643 – 674.

Collins, A. M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing, *Psych Review*, 82(6), 407–428.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., Van de Mortele, P-F, & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*(6702), 597–600.

Fodor, J. (1983). *The modularity of mind*. Cambridge. MIT Press.

Forster, K. I. (2006). Early activation of category information in visual word recognition More on the turple effect, *The Mental Lexicon*, 1, 35–58.

Kouider, S., & Dupoux, E. (2005). Subliminal speech priming. *Psychological Science*, 16(8), 617–25.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato ' s problem : The Latent Semantic Analysis Theory of Acquisition , Induction , and Representation of Knowledge. *Psychological Review*, 104(2), 211–240.

Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, 3(2), 273–302.

Masson, M. E. J. (1995). A Distributed Memory Model for Semantic Priming. Journal of Experimental Psychology.

Quinn, W. M., & Kinoshita, S. (2008). Congruence effect in semantic categorization with masked primes with narrow and broad categories. *Journal of Memory and Language*, 58(2), 286–306.

Saj, A., Fuhrman, O., Vuilleumier, P., & Boroditsky, L. (2013). Patients With Left Spatial Neglect Also Neglect the "Left Side" of Time. *Psychological Science*, (November).

Searle, J. R. (1980). *Minds , brains , and programs*, *BBS*, 3, 1–19.

Tulving, E. (1972). *Episodic and semantic memory*. In Organization of memory (Vol. 381). London: Academic.