Subliminal speech perception and auditory streaming

Emmanuel Dupoux*, Vincent de Gardelle, Sid Kouider

Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS/EHESS/DEC-ENS, 29 rue d’Ulm, 75005 Paris, France

Article history:
Received 16 November 2007
Revised 6 June 2008
Accepted 22 June 2008

Keywords:
Spoken word recognition
Subliminal speech priming
Interaural time delay
Consciousness

Abstract

Current theories of consciousness assume a qualitative dissociation between conscious and unconscious processing: while subliminal stimuli only elicit a transient activity, supraliminal stimuli have long-lasting influences. Nevertheless, the existence of this qualitative distinction remains controversial, as past studies confounded awareness and stimulus strength (energy, duration). Here, we used a masked speech priming method in conjunction with a submillisecond interaural delay manipulation to contrast subliminal and supraliminal processing at constant prime, mask and target strength. This delay induced a perceptual streaming effect, with the prime popping out in the supraliminal condition. By manipulating the prime-target interval (ISI), we show a qualitatively distinct profile of priming longevity as a function of prime awareness. While subliminal priming disappeared after half a second, supraliminal priming was independent of ISI. This shows that the distinction between conscious and unconscious processing depends on high-level perceptual streaming factors rather than low-level features (energy, duration).

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The characterization of the difference between conscious and unconscious processing has become an active area of research over the last decade. It remains debated whether, apart from the phenomenology associated with consciousness, one can find measurable processing correlates of the presence or absence of consciousness (Baars, 1988, 2002; Edelman, 2003; Koch & Crick, 2004; Kouider & Dehaene, 2007). In addition, provided such measures exist, it remains unclear whether they would reveal a qualitative rather than a quantitative distinction (Cleeremans, 2005; Dehaene & Naccache, 2001).

A survey of the literature on consciousness reveals four distinct theoretical positions (Kouider & Dehaene, 2007). Two of them simply deny the validity of a conscious/unconscious distinction. The “no-unconscious-representation” view argues that the notion of unconscious representations is not demonstrated, or is even inconsistent, and that one can only talk about conscious representations (Dulany, 1997; Holender & Duscherer, 2004; Perruchet & Vinter, 2002). The “no-functional-distinction” view argues that all forms of computation can be performed regardless of consciousness (Chalmers, 1996; Marcel, 1983) and that conscious perception is only an epiphenomenal experience superimposed on unconscious processes. According to these two extreme views, conscious and unconscious perceptions cannot be functionally dissociated. Two more moderate and widespread classes of models explicitly acknowledge a functional dissociation between conscious and unconscious processes. According to the “quantitative-dissociation” view, there is a continuum from unconscious to conscious perception corresponding with an increase in depth of processing (Cleeremans, 2005, 2006; Farah, 1994, 2000; Overgaard, Rote, Mouridsen, & Ramsoy, 2006). The “qualitative-dissociation” view, in contrast, posits a discrete, all-or-none distinction (e.g., Baars, 1988; Di Lollo, Enns, & Rensink, 2000; Lamme, 2003; Sergent & Dehaene, 2004).
The “qualitative-dissociation” models, contrary to all others, propose that consciousness corresponds to a qualitatively special processing mode: unconsciously perceived stimuli are processed in a bottom-up fashion and can only have a small, local, and transient effect on the perceptual system. Conscious stimuli, in contrast, are amplified by top-down connections, are encoded in working memory, and have long-lasting influences at several levels of processing. This predicts that under masking conditions, subliminal stimuli can have an effect for a limited duration (less than a second, see Ferrand, 1996; Forster & Davis, 1984; Greenwald, Draine, & Abrams, 1996; Humphreys, Besner, & Quinlan, 1988), whereas supraliminal stimuli can have an effect for a much longer time. In such models, longevity is not determined by physical properties per se, but rather by whether the stimulus is represented as a conscious event (Baars, 2002; Dehaene & Naccache, 2001). Thus, the longevity of a stimulus’ influences constitutes a good marker for the distinction between conscious and unconscious processes.

The three other models described above make different predictions regarding longevity. First, “no-unconscious-representation” models simply deny the possibility of subliminal influences, not to speak about their potential longevity. Second, “no-functional-dissociation” models expect the same influences, and thus the same longevity, for subliminal and supraliminal stimuli, all other things being equal. Third, in “quantitative-dissociation” models, because the conscious/unconscious dissociation is grounded on continuous variables like the strength of the stimulus (e.g., duration, energy, etc.), longevity should primarily depend on prime strength. Such models can account for past studies because consciousness was always confounded with energetic or low-level changes. Indeed, the contrast between subliminal vs. supraliminal priming was always associated with a change in the physical properties of the prime stimulus (e.g., increase in the duration or the intensity for supraliminal primes, Dehaene et al., 2001; Greenwald et al., 1996; Kouider, Dehaene, Jobert, & Le Bihan, 2007).

In order to distinguish qualitative and quantitative models, we used the masked speech priming paradigm (Kouider & Dupoux, 2005) and manipulated the prime-target ISI (inter stimulus interval, i.e., the time between the offset of the prime and the onset of the target). Importantly, contrary to past research, we kept identical the low-level properties (energy, duration) of all the stimuli (prime, target and masks). In order to render the primes conscious while keeping the energy constant, we relied on perceptual grouping (Bregman, 1990) using a manipulation of the interaural time delays (ITD). When we hear an external sound source, the difference in arrival time of the sound waves between the two ears (i.e., the ITD) is used by the brain to reconstruct the spatial location of the source (Sach & Bailey, 2004). Therefore, by artificially manipulating the ITD of stimuli presented in headphones, one can manipulate the subjective location of the source(s) of the stimuli. Thus, in both the subliminal and supraliminal conditions, the prime and targets were identical but the stream of masks was presented with different ITDs, as depicted in Fig. 1. In the supraliminal condition, due to the introduction of a 750 µs ITD, the masks appear to belong to a lateralized stream, while the prime and target remains on a central stream. In the subliminal condition, the prime, target and masks are binaurally presented without ITD. Hence they appear to belong to the same central stream. Thus, at the level of each ear, all the stimuli in the subliminal and supraliminal conditions are presented with exactly the same physical properties. It is only by listening to the two ears simultaneously, in the supraliminal condition, that the perceptual organization in two streams emerges, one central for the prime and the target, one lateral for the masks.

The prediction of qualitative-dissociation accounts is that priming is long-lasting only when the prime is segregated from the mask and becomes conscious. When the prime is streamed together with the masks, it remains unconscious, and hence produces only short-lived priming. By contrast, quantitative-dissociation models that rely only on the physical properties of the primes predict identical priming in both conditions.

2. Method

2.1. Participants

The participants were 96 students recruited from Paris universities. They reported no hearing impairment, were native French speakers, and were paid for their participation. Sixty-four participants (mean age = 22) participated in Experiment 1 (subliminal priming) and 32 participants (mean age = 22) participated in Experiment 2 (supraliminal priming).

2.2. Stimuli

In both experiments, the stimuli consist in a list of 224 experimental pairs of items (half words and half pseudowords; half monosyllabic and half disyllabic) and 30 supplementary training pairs. Within each pair, items were matched in frequency (for words), length, phonological structure (consonant-vowel), but did not share any obvious formal or semantic relation. One item of the pair is used as a target and repeated prime, while the other is used as unrelated prime. All items were recorded only once by a male French speaker and sampled at 16 kHz.

In the subliminal experiment, we used the same speech masking method as Kouider & Dupoux (2005). Target items were played at a normal level, primes were time-compressed to 35% of their original duration, attenuated (−15 dB). Each mask was made out of a randomly selected time-reversed prime. The same signal was presented to both ears and in the absence of any ITD. The target was always preceded by a sequence of nine masks, one of which being substituted by the prime (i.e., the prime replaced a mask but was not superimposed over a mask). Hence, the ISI was manipulated in eight discrete steps through the position of the prime relative to the target (the first mask was never replaced by a prime).
While the subliminal experiment used an ITD of 0 μs for masks, primes and targets, the supraliminal experiment introduced an ITD of 750 μs for the masks only. This was obtained by shifting the stream of masks in one of the channels by 12 samples (the left for half of the participants, the right for the other half). The protocol was programmed using the audio mixing table function of the Expe software package (Pallier, Dupoux, & Jeannin, 1997).

2.3. Procedure

Prime-target relation (repetition vs. unrelated) and ISI were counterbalanced across items and participants by using a latin square design. Every participant received every target item only once. Relation and ISI factors were randomized across trials. Participants were asked to perform a speeded lexical decision on the target and to ignore other auditory events. They received first a 30 trials training, then four experimental blocks of 56 trials each.

Following the priming experiment, we performed two prime awareness tests: a lexical decision (LD) and a speech/non-speech decision (SD) (Kouider & Dupoux, 2005). Half of the target stimuli were presented either under the same situation as in the priming experiment (with both related and unrelated primes from the priming conditions), and half were preceded by a prime with a different lexicality (for LD), or a backward-speech prime (for SD). Half of the participants received an SD block followed by a LD block, and vice versa for the other participants. Each block included 56 trials and was preceded by 20 training trials where participants were instructed to focus on the prime which was presented in an unmasked way (played at a higher level, and separated from the masks by 100 ms of silence) and received feedback on their errors.

3. Results

3.1. Priming effects

We ran one ANOVA restricted on responses to words and a separate one for pseudowords, using Experiment as a between-participant factor, and ISI and Relation as within-participant factors. Additionally, we included two counterbalancing group factors corresponding to the subject assignment in the latin square described above. Error trials were excluded, and reaction times more or less than 2 standard deviations from the mean for a particular participant and condition were replaced by the relevant cut-off.

For words, there was an effect of Relation ($F(1,64) = 162, p < 0.001$) and Experiment ($F(1,64) = 11.3, p = 0.001$). In addition, there was an interaction between Relation and Experiment ($F(1,64) = 171.4, p < 0.001$), an interaction between ISI and Experiment ($F(7,448) = 3.3, p = 0.002$), and, crucially, a triple interaction between ISI, Relation, and Experiment ($F(1,47) = 2.6, p = 0.012$). To break down the interactions, we analysed separately each experiment in planned contrasts (see Table 1). In both experiment there was an effect of Relation (subliminal: $F(1,48) = 12.0, p = 0.001$, supraliminal: $F(1,16) = 164.7, p < 0.001$). In the subliminal experiment, there was a marginal interaction between ISI and Relation ($F(7,336) = 2.0, p = 0.057$).

For pseudowords, there was an effect of Experiment ($F(1,64) = 14.2, p < 0.001$) and of ISI ($F(7,448) = 3.0$,...
Table 1

Reaction time (RT) and standard error of the mean (SE) for the lexical decision tasks on words, for the eight ISI and subliminal and supraliminal experiments

<table>
<thead>
<tr>
<th>ISI (ms)</th>
<th>0</th>
<th>152</th>
<th>304</th>
<th>456</th>
<th>608</th>
<th>760</th>
<th>912</th>
<th>1064</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subliminal condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT (ms)</td>
<td>686</td>
<td>696</td>
<td>689</td>
<td>702</td>
<td>699</td>
<td>708</td>
<td>703</td>
<td>708</td>
</tr>
<tr>
<td>SE (ms)</td>
<td>18</td>
<td>17</td>
<td>14</td>
<td>15</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>Unrelated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT (ms)</td>
<td>712</td>
<td>707</td>
<td>715</td>
<td>726</td>
<td>704</td>
<td>710</td>
<td>694</td>
<td>711</td>
</tr>
<tr>
<td>SE (ms)</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Priming (ms)</td>
<td>26**</td>
<td>11</td>
<td>25*</td>
<td>24*</td>
<td>53</td>
<td>26**</td>
<td>11</td>
<td>25*</td>
</tr>
<tr>
<td>Supraliminal condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT (ms)</td>
<td>796</td>
<td>734</td>
<td>728</td>
<td>729</td>
<td>719</td>
<td>746</td>
<td>742</td>
<td>743</td>
</tr>
<tr>
<td>SE (ms)</td>
<td>25</td>
<td>26</td>
<td>24</td>
<td>27</td>
<td>36</td>
<td>30</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>Unrelated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT (ms)</td>
<td>850</td>
<td>854</td>
<td>832</td>
<td>838</td>
<td>854</td>
<td>831</td>
<td>829</td>
<td>830</td>
</tr>
<tr>
<td>SE (ms)</td>
<td>24</td>
<td>30</td>
<td>19</td>
<td>23</td>
<td>20</td>
<td>28</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>Priming (ms)</td>
<td>54**</td>
<td>120***</td>
<td>103***</td>
<td>109***</td>
<td>134***</td>
<td>84***</td>
<td>87***</td>
<td>87***</td>
</tr>
</tbody>
</table>

*p < 0.05.
**p < 0.01.
***p < 0.005.

There was an interaction between Relation and Experiment (F(1,64) = 4.5, p = 0.037) and an interaction between ISI and Experiment (F(7,448) = 4.7, p < 0.001). Separately planned analyses in each experiment revealed an effect of Relation (F(1,16) = 4.6, p = 0.046), and an effect of ISI (F(7,112) = 4.1, p < 0.001), in the supraliminal experiment only. No other effects or interactions with the experimental factors were significant. This result is congruent with past work showing that pseudowords do not yield masked priming in lexical decision unless the primes become conscious (Forster & Davis, 1984; Kouider & Dupoux, 2001, 2005).

3.2. Prime longevity

To assess the longevity of priming for words, we ran at each ISI post-hoc tests using Relation and the two counterbalancing factors. In the supraliminal experiment, the Relation effect was significant for each ISIs, whereas in the subliminal experiment, it was significant in three of the four shortest ISIs, while no effect was found for all the four longest ISIs (see Table 1). A linear regression ran on the average priming as a function of ISI revealed a significant correlation (R = 0.78, F(1,6) = 9.7, p = 0.020), for the subliminal but not for the supraliminal experiment (R = 0.04, F < 1, p > 0.1).

3.3. Prime awareness

The average error rate across the two prime awareness tasks as a function of ISI is presented in Fig. 2. The participants were very poor for all ISIs in the subliminal condition (average error rate: 46.1%), and well above chance in the supraliminal condition at all ISIs (28.3%). We computed d’ values across all ISIs for each participant by treating speech as signal in the SD task, and words as signal in the LD task. The d’ values for the supraliminal experiment were significantly larger than for the subliminal experiments (F(1,94) = 159.6, for LD, and F(1,94) = 75.3, for SP, both p < 0.001), confirming a difference in prime audibility. In the supraliminal experiment, the d’ values were significantly higher than zero in both tasks (d’ = 1.56, t(31) = 14.8 for LD, and d’ = 0.99, t(31) = 11.5 for SP, both p < 0.001). In the subliminal experiment, although the d’ values were lower, they were significantly different from zero (d’ = 0.25, t(63) = 4.92 for LD, and d’ = 0.19, t(63) = 3.95 for SP, both p < 0.001). To verify that priming could obtain under genuine subliminal conditions, we performed two further analyses. First, we assessed whether there was a still a significant residual priming effect for a null d’ score was assessed using the regression of priming against d’ (Greenwald et al., 1996; Kouider & Dupoux, 2005). We analysed only the four shortest ISIs since there was no significant priming for the longest ISIs. We observed that the priming effect extrapolated to a null d’ was significant (15 ms, t(62) = 2.05, p < 0.05), as shown in Fig. 3, and also that the slope of the regression was not significant (t(62) = 1.58, p = 0.12). These two results show that priming was still significant in the absence of prime audibility and that although some participants might have had a small awareness of the primes, it did not correlate with priming. Second, to confirm that priming at the four shortest ISIs indeed obtains in the absence of audibility, we restricted our analysis by excluding the twenty participants with the highest d’ score. The remaining subjects (N = 44) had an average d’ score of 0.06 which did not differ from null performance (t > 1), yet they still showed a significant priming effect (16 ms, p < 0.05).

4. Discussion

We used a masked priming paradigm with speech stimuli and compared its longevity as a function of prime awareness. We manipulated prime awareness while
keeping the strength of the auditory events constant by varying the ITD on the stream of masking stimuli. This manipulation leads to the subjective perception that the auditory source of the masks was located on a lateral channel, while the prime and target were in a central channel. This percept was obtained despite the fact that the masks, prime and target were presented with the same energy to both ears. Yet, the ITD difference induces the prime to pop out from the steam of masks and hence to become consciously accessible.

Under such conditions, we found that priming longevity has a qualitatively different profile as a function of awareness: in the subliminal condition, priming was maximal at an ISI of zero and declined with longer ISIs until its disappearance around 600 ms. In contrast, in the supraliminal condition, priming did not depend on ISI, and was maintained up to 1000 ms after the presentation of the prime. This study constitutes the first experimental assessment of the longevity of subliminal influences in the auditory modality. It is noteworthy that although subliminal priming was short-lived here, it appears to last somewhat longer than in previous studies in the visual modality where priming did not extend beyond 100 or 200 ms (Ferrand, 1996; Greenwald et al., 1996). One reason for this discrepancy may be due to the fact that iconic memory decays faster than its echoic equivalent (Darwin, Turvey, & Crowder, 1972; Sperling, 1960).

Our results reinforce the idea that the distinction between conscious and unconscious stimuli cannot be reduced to the strength of the stimulus, measured in terms of energy or duration (e.g., Cleeremans, 2006). Indeed,
the physical characteristics of the primes are strictly identical in the supraliminal and subliminal conditions. In addition, they are surrounded by the same acoustic materials (compressed primes) in both conditions, so that the low level auditory masking remains identical. What is different though is that the ITD is used as a cue to segregate the auditory scene into separate streams, and hence in the supraliminal condition, the prime can be perceptually separated from the surrounding masks. The distinction between conscious and unconscious is hence crucially linked not to the strength of the stimuli, but rather to the capacity of participants to perceptually segregate the primes from the masks (Bregman, 1990). One could object that ITD differences are encoded at a rather early stage in the auditory pathway (midbrain inferior colliculus and brainstem medial superior olive, see Joris & Yin, 2007), and that hence our subliminal and supraliminal stimuli differ in a low level property. What is important though, is not the level at which ITD is calculated, but rather the level at which it is being used to segregate auditory stimuli. Indeed, psycho-physical experiments have shown that ITD is remarkably ineffective as cue for the perceptual grouping of spectral components in concurrent signals (Culling & Summerfield, 1995). Rather, ITD seems to be used to track the spatial localization of objects that have already being perceptually grouped (Darwin & Hukin, 1999; Sach & Bailey, 2004). As such, the distinction between subliminal and supraliminal stimuli in our experiments is arguably taking place at a rather advanced level of processing (i.e., after perceptual streaming).

Finally, our results are difficult to square with models that deny the existence of a distinction between conscious and unconscious processing, or state that this distinction is merely epiphenomenal. They are consistent with qualitative models of consciousness, whereby consciousness corresponds to the all-or-none transfer of a transient activation into working memory. Reformulated from a neurobiological perspective, supraliminal stimuli induce global broadcasting i.e., long distance reinforcement and reverberation loops (Baars, 2002; Dehaene & Naccache, 2001). Such sustained activity allow for the continuing activation of the representation even after the stimulus has disappeared, which becomes available for a variety of non-automatic tasks. In contrast, subliminal activation only propagates in a stimulus-driven feed-forward way in chains of networks that have been setup for highly automatized tasks (Kouider & Dehaene, 2007). Such a bottom-up activity quickly fades when the stimulus disappears.

A first glance, these results are difficult to handle by quantitative/gradual accounts of consciousness (Cleermans, 2006; Farah, 2000; Overgaard et al., 2006). Indeed, in such models, the quantity and longevity of activation in neural networks is proportional to the strength of the stimulation. However, they could accommodate these data by considering stimulus strength not only in terms of low-level properties (duration, energy) characterizing the prime, but also in terms of properties related to perceptual segregation of the prime from the background masks. Further research is needed to further specify the conditions leading to the dissociation between conscious and unconscious processes.

Acknowledgements

We thank Stanislas Dehaene and Christian Lorenzi for very useful discussion. This work was supported by a doctoral fellowship from the Direction Générale de l’Arme- ment (France), a grant from the French Agence Nationale pour la Recherche (“Early Language Acquisition: Experi- ments and Computational Approaches”), and by the Euro-opean Commission FP6 Neurocom STREP project.

References


