Task-dependent masked priming effects in visual word recognition

Sachiko Kinoshita[1]* and Dennis Norris[2]

[1] Department of Psychology, ARC Centre of Excellence in Cognition and its Disorders (CCD), Macquarie University, Sydney, NSW, Australia
[2] MRC Cognition and Brain Sciences Unit, Cambridge, Cambridgeshire, UK

A method used widely to study the first 250 ms of visual word recognition is masked priming: These studies have yielded a rich set of data concerning the processes involved in recognizing letters and words. In these studies, there is an implicit assumption that the early processes in word recognition tapped by masked priming are automatic, and masked priming effects should therefore be invariant across tasks. Contrary to this assumption, masked priming effects are modulated by the task goal: For example, only word targets show priming in the lexical decision task, but both words and non-words do in the same-different task; semantic priming effects are generally weak in the lexical decision task but are robust in the semantic categorization task. We explain how such task dependence arises within the Bayesian Reader account of masked priming (Norris and Kinoshita, 2008), and how the task dissociations can be used to understand the early processes in lexical access.

Keywords: masked priming, visual word recognition, Bayesian reader

Human readers are remarkably efficient at recognizing words: As noted in the introduction to this special issue, the time window in which a letter string passes from being a mere sequence of printed curves and strokes to being perceived as a word takes no longer than one-third of a second. Speed and minimal effort are hallmarks of automatic processes, and a procedure that has been valuable in studying the automatic aspects of visual word recognition is masked priming.

Forster and Davis (1984) pioneered the masked priming procedure that has come to be the standard in studies of visual word recognition. In this procedure, a trial consists of a sequence of three events: a forward mask (typically a series of # symbols) presented for 500 ms, a prime presented briefly (usually about 50 ms), followed immediately by the target to which a response is required – usually lexical decision. The target is presented either for a fixed duration (e.g., 500 ms) or until subject's response. Typically the prime is presented in lowercase letters and the target in uppercase, so that the prime related in form does not overlap the target physically, and the target functions as a backward mask for the prime. Despite the prime being presented so briefly that subjects have little phenomenological awareness of it; a prime related to the target in some way – for example, by identity (e.g., chair-CHAIR) or form (e.g., cheir-CHAIR) facilitates the response to the target, relative to an unrelated control.

The standard interpretation of masked priming has been driven by two consistent findings. First, as noted above, the prime and target are presented in different cases. Combined with the results of experiments that have specifically manipulated the perceptual overlap between primes and targets (Bowers et al., 1998), this implies that priming is driven by representations at the level of abstract letter identity rather than visual form. Second, in the lexical decision task, identity priming is found consistently for

words, but not for non-words. At first blush this seems to be clear and convincing evidence that priming is purely lexical. If priming were not lexical, why would it be observed only for words? Indeed, the idea that masked priming is lexically mediated has become the "conventional wisdom." For example, Forster (2004) suggested that masked priming is an "index of lexical access" (p. 277).

However, despite the consistency of these findings when using lexical decision, a completely different picture emerges when the task is changed. With careful choice of task, priming can be obtained for non-words, and even made to disappear for words (Forster, 1985; Norris and Kinoshita, 2008). The same pairing of primes and targets can produce different patterns of priming in different tasks. Priming is not an automatic function of the relation between prime and target but depends critically on the nature of the experimental task. Note that we are not the first to make this claim: Earlier, Dehaene et al. (1998) had proposed that subliminal processing can be found under conditions where "subjects unconsciously apply the task instruction to the prime" (p. 598)[1]. This would seem to imply that masked priming effects are "strategic," in which case much of the appeal of the procedure is lost. However, the fact that priming is task-dependent is not a cause

---

[1]Dehaene et al. (1998) put forward the view in the context of subliminal perception. To date, there has been relatively little contact between researchers who use masked priming to study subliminal perception, and those who use it to study the visual word recognition. In part, this is because the former have typically used a small set of stimuli (e.g., arrows pointing left and right, single digits) that are presented repeatedly, and there has been a concern that the mechanism supporting masked priming in this case is a simple stimulus-response mapping process which has little to do with visual word recognition (e.g., Damian, 2001). We will discuss this issue later, with respect to whether masked priming effects can be semantic.

for gloom and despondency, nor does it make a case for abandoning the paradigm. On the contrary, we will suggest that the lawful way in which the influence of a prime presented for only 50 ms can be modulated by the task provides significant insights into the process involved in the first 250 ms of word recognition.

In this paper, we first explain why it is that masked priming should be influenced by the task. We will do this in the context of the framework of the Bayesian Reader model Norris, 2006, 2009; Norris and Kinoshita, 2008). We will then illustrate this by reviewing data[2] that show that almost all of the main phenomena that have been studied with masked priming are modulated by task demands. In fact, some patterns of data are turned on their heads completely when the task is changed. The data challenge all of the following common assumptions about masked priming:

- Masked priming is lexically mediated.
- Orthographic processing is modulated by morphology.
- Orthographic processing is language-specific.
- Transposed-letter priming effects reflect the orthographic representations in the lexicon.
- There is no semantic priming.
- Mapping of orthography to phonology occurs serially across a letter string.

We conclude by noting how the task dependence in masked priming does not imply that it is "strategic."

DECISION PROCESS IN THE BAYESIAN READER

An assumption common to almost all models of reading, whether verbal or computational, is that there is a fixed processing architecture. Word recognition always proceeds in the same automatic fashion and all that varies between tasks is how the output of the system is used. According to this view priming takes place within the fixed component of the system and should therefore be unaffected by task. Even in models that explicitly recognize task differences (e.g., Bimodal Interactive Activation model, BIAM; Grainger and Ferrand, 1994; Grainger and Ziegler, 2011), these differences are explained in terms of the different representations and pathways involved in different tasks (or different memory systems – lexical vs. episodic – in the case of Forster, 1985), that is, in terms of different architectures supporting different tasks.

Here we present a very different view: we suggest that it is the task that drives perception. More specifically, we suggest that perception can be characterized as embodying a process of optimal Bayesian decision making (cf. Knill et al., 1996). According to this view, all perception involves making decisions. A necessary implication of this is that behavior will vary with task demands, i.e., the nature of the decision required. This admits the possibility that the pattern of priming may vary quite radically as a function of the subject's task, even if they involve the same representations. This possibility has huge ramifications for the kind of inferences that can be drawn from masked priming data.

TASK DEPENDENCE IN MASKED PRIMING

We now turn to the review of findings demonstrating task dependence in masked priming. The review is necessarily selective. Several different tasks have been used with the masked priming procedure. In studies of visual word recognition, the most popular task is lexical decision, in which subjects are asked to classify a letter string as either a word, or a non-word. The read-aloud task (also called the naming task or the pronunciation task) is another task that is frequently used by visual word recognition researchers. The task here is to read aloud the word as quickly and as accurately as possible. In a semantic categorization task, subjects are asked to decide whether a word is an exemplar of a category (e.g., "animals"). In addition to these tasks, more recently, Norris and Kinoshita (2008) adapted the masked priming procedure to be used with the same-different task, which we describe below. In all of these tasks, response latency (reaction time, RT) is the main dependent variable, as the tasks are generally designed to produce a high level of accuracy and it is less sensitive to masked priming[3].

IS PRIMING LEXICAL?

As noted earlier, the dominant view of masked priming is that masked priming is lexical (Forster and Davis, 1984; Forster, 1998; Forster et al., 2003). The main support for this view comes from the fact that in the lexical decision task word targets show robust masked priming effects but non-word targets do not. Forster (1998) reported that across 40 lexical decision experiments, the mean size of identity priming effect for non-word targets was 8.7 ms, and only in three cases it was statistically significant at the 0.05 level. This effect of lexical status is easily explained within the lexical view: A masked prime automatically activates its corresponding representation (or within Forster's "entry-opening" account, "opens the entry") in the lexicon, but non-words have no representations to activate (or have no entries to open), hence non-words do not show priming.

Norris and Kinoshita (2008) developed an account of masked priming based on the Bayesian Reader model of word recognition (Norris, 2006). The Bayesian Reader is a stimulus sampling model. The model accumulates samples of evidence from the perceptual input and makes near-optimal decisions as dictated by the experimental task. This simple assumption correctly accounts for a wide range of phenomena in visual word recognition including the logarithmic function relating ease of recognition to word frequency, how neighborhood effects are modulated by task, and how reaction-time distributions change as a function of both frequency and the type of non-words used in lexical decision (Norris, 2009).

---

[2]Our review is limited to behavioral data. For a review of ERP data in masked priming studies, see, e.g., Grainger and Holcombe (2009).

[3]In addition, the perceptual (tachistoscopic) identification task has also been used with the masked priming procedure (e.g., Evett and Humphreys, 1981; Humphreys et al., 1988). Here, a trial sequence consists of a forward mask, prime, target, and a backward mask. The prime and the target are both presented briefly, and the subject's task is to identify the target ("the item presented in uppercase letters"), with accuracy being the dependent measure. Unlike the RT tasks in which the target is presented clearly, subjects sometimes report the prime instead of the target, and intrude the letters from the prime. Furthermore, Davis and Forster (1994) have shown that priming in this task may be entirely attributable to target legibility which results from the physical fusion between the prime and target. As these pattern of priming effects are quite different from the other tasks in which the target is presented undegraded, we will limit our coverage of masked priming to RT tasks in which the target perception is accurate, and RT is the main dependent variable.

In order to extend the model to simulate masked priming Norris and Kinoshita made one additional assumption: they assumed that the perceptual system does not treat the prime and target as separate perceptual events, and therefore evidence from both the prime and targets are integrated in reaching a decision.

As we will explain below, this accounts for the fact that priming in lexical decision is seen only for words but not non-words, but it also makes an interesting and rather counter-intuitive prediction: If the task is changed then it should be possible to observe priming for non-words, and to eliminate priming for words. The task chosen was a same-different task. This has the same basic procedure as masked priming in lexical decision, but now an additional word or non-word is presented for 1000 ms before the prime. The subject's task is to decide whether the target is the same as or different from this referent stimulus. In contrast to the lexical decision task, the same-different task shows equally robust priming effects for word and non-word targets requiring a Same decision, but no priming for either words or non-words for *Different* decision (Norris and Kinoshita, 2008, Experiment 1; Kinoshita and Norris, 2009, Experiment 1, 4). That is, targets requiring a *Same* response pattern like *words* in lexical decision regardless of their lexical status, whereas targets requiring a *Different* response pattern like *non-words* in lexical decision regardless of their lexical status. These results are exactly as predicted by the Bayesian Reader.

In order to understand how the model predicts this pattern of data we need to remind ourselves exactly what the decision is that we ask subjects to make. In a typical lexical decision experiment subjects are told to respond "Yes" if the stimulus is a word, and "No" if it is not a word. They are not told to respond only when they know exactly what the word is, or only when they know exactly what the non-word is. Similarly, in the same-different task, subjects are not required to uniquely identify a stimulus that is *Different*, they just need to know that it is not the same as the referent. The significance of this rather pedantic analysis is that it highlights an important parallel between the two tasks. In both cases subjects have a set of stimuli in memory. In lexical decision this set corresponds to the entire lexicon. In the same-different task the set contains only the referent. In both cases the task is to determine whether the target is a member of the specified set. The data have a simple pattern: targets in the set show priming, targets not in the set show no priming.

But how is this pattern predicted by the Bayesian Reader? Although we explain this more formally and back it up with simulations in Norris and Kinoshita (2008; see also Norris et al., 2010) here we will try to give a more intuitive account of how this pattern emerges from the combination of stimulus sampling and optimal decision making. We begin by examining why in lexical decision only the words show priming. We will use the standard statistical metaphor of drawing balls from an urn. On each trial the urn contains balls which are a mixture of colors. On half of the trials the balls are mainly black, and on the other half they are some other color. The task is to draw balls from the urn to determine whether the balls are mainly black or not. We can think of black balls as words, and colored balls as non-words. The assumption that the balls in the urn are never purely one color reflects the stimulus sampling component of the Bayesian Reader. The mix of colors corresponds to noise in the sampling process and means

that many samples are required to make a confident decision. See Norris (2006) for full details of the mathematics of the decision process.

Consider what happens when we add a "prime" where some extra balls are sampled before the balls representing the target. We assume that these "prime" balls are mainly a single color (black, blue, red, etc.). In the case where the prime and target balls are both mainly black, you will obviously need that many fewer balls to reach a decision that most of the balls are black. If you prime the black balls with a sample of balls of a different color you will clearly need to sample more balls form the target to appreciate that most of the balls are in fact black. There will therefore be an overall priming effect with identity primes producing faster responses (needing fewer samples) than unrelated primes. Now consider what happens when the target balls are some other color. A prime of the same color will provide exactly the same kind of evidence as the samples from the target, so this will provide a head start equal to the number of balls in the prime. But, a prime of a different color will provide exactly the same head start. Any ball that is not black contributes in exactly the same way to the decision that the target balls are not black. As noted above, when performing lexical decision it does not matter exactly what the non-word is, all that matters is that it is not a word. Similarly, the exact color of the balls is immaterial because there is no need to know what color the balls are, so long as they are not black. All that counts in reaching the decision is whether or not the balls are black. There will therefore be a "priming effect" for black balls, but no priming effect for balls of any other color. Balls of any other color are simply balls that are not black.

As with all analogies, the urn analogy fails to capture the more subtle details of the full model. For example, it might seem to imply that priming a colored-ball target with a black-ball prime should bias the decision process toward "black" and give rise to inhibition. However, in lexical decision, RTs to non-words are unaffected by the lexical status of the prime. This is because the main effect of the prime is to generate evidence that the target is in a particular area of orthographic/lexical space, but is insufficient to provide specific evidence that the target is a word or not.

Below we describe examples where this framework has been useful in understanding the nature of task dependence. Before doing so, we put to rest an alternative account of why there is no priming for non-words in lexical decision and *Different* decisions in the same-different task.

FAMILIARITY BIAS

It has sometimes been suggested that masked priming effects reflect a combination of lexical activation and "familiarity bias." If this were true it would not only undermine our explanation of masked priming, but also undermine the value of the task as a tool for providing insights into the first 250 ms of reading. It is worth emphasizing that the changes in pattern of priming with task reported by Norris and Kinoshita (2008) were exactly as predicted by the Bayesian Reader. No other account of masked priming would have lead one to expect that priming would be completely different in lexical decision and the same-different task. Nevertheless, Bowers (2010) suggested that the data can be

explained in terms of a "familiarity bias," originally suggested by Bodner and Masson (1997). The idea is that a masked non-word prime facilitates the identification of a repeated non-word target by "preactivating the relevant sublexical representations," but a repeated non-word target is "perceived as more familiar (due to its improved perception)" (Bowers, 2010, p. 786), and this familiarity bias counteracts the benefit due to the preactivation of sublexical representation for non-word targets. Similarly, in the same-different task, the absence of priming effect for *Different* responses is explained in terms of the familiarity bias due to increased fluency of perceiving the target producing a bias toward responding "*Same*" (because "the increased fluency can be taken as evidence that the target has been repeated," Bowers, 2010, p. 787), which inhibits a *Different* response.

The problem with the familiarity bias hypothesis is that it is *ad hoc*, and it would seem possible to rationalize any conceivable pattern of data within this loosely formulated view. Furthermore, attempts to test it empirically have not succeeded in producing support for it. Kinoshita and Norris (2011) attempted to replicate a finding by Bodner and Masson (1997), which has been taken as providing evidence for the familiarity bias hypothesis. The finding concerns the emergence of priming for non-word targets in a lexical decision task presented in a cAsE-AlTeRnAtEd format. According to the familiarity bias hypothesis, case-alternated targets are visually unfamiliar, and this should have the effect of reducing reliance on "perceived fluency" which generates a bias toward responding "Word" in the lexical decision task. Consequently, when targets are presented in a visually unfamiliar, case-alternated format, masked priming effects should emerge for non-word targets, reflecting the preactivation of sublexical representations, and this is what Bodner and Masson (1997) found. Kinoshita and Norris' (2011) replication of Bodner and Masson's lexical decision experiment, using the stimuli used by Bodner and Masson, showed that the priming for case-alternated non-word targets did emerge. However, further analysis showed that this was limited to non-words containing letters that were ambiguous when presented in a case-alternated format, namely, a lowercase "l" or an uppercase "I" (e.g., lOvInk, lOmIT); non-words that did not contain ambiguous letters (e.g., jAsEnT, nOrBaT) did not show any priming. These results suggest that the priming of non-words emerged with case-alternated targets because of priming of letters made ambiguous by case-alternation, not because subjects abandoned familiarity bias when all targets were visually unfamiliar. Follow-up experiments using case-alternated targets that did not contain ambiguous letters showed the standard pattern of priming, namely, robust masked priming effects for words but not non-words in lexical decision, and masked priming effects for the *Same* decision but no priming for the *Different* decision in the same-different task. These results provide little support for the familiarity bias hypothesis, but are as expected from the Bayesian Reader account[4].

DO TRANSPOSED-LETTER PRIMING EFFECTS DEPEND ON LEXICAL REPRESENTATIONS?

One of the most widespread uses of masked priming has been to investigate the nature of orthographic representations. The standard procedure is to manipulate the degree of orthographic overlap between prime and target with the aim of "cracking the orthographic code" (Grainger, 2008). The paradigmatic example of this enterprise is the transposed-letter (hereafter TL) priming effect. This effect refers to the finding that a prime generated by transposing the positions of letters in the target (usually adjacent, internal letters, e.g., *jugde-JUDGE*) facilitates the recognition of the target (often almost as much as the identity prime) more than a prime generated by replacing the corresponding letters with letters not contained in the target (e.g., *junpe-JUDGE*). First reported by Forster et al. (1987), this effect has been replicated many times in the lexical decision task across different languages (e.g., Perea and Lupker, 2003; Lupker et al., 2008, in English; Perea and Lupker, 2004 in Spanish; Schoonbaert and Grainger, 2004, in French; Perea and Perez, 2009, in Japanese kana)[5].

As with other priming effects, in the lexical decision task, TL priming is readily observed with words but not with non-words. In studies examining TL priming with non-words (e.g., Perea and Lupker, 2003; Schoonbaert and Grainger, 2004; Perea and Carreiras, 2008) the effect is absent or unreliable at best. This invites the inference that TL priming is telling us about specifically lexical representations rather than representations at a purely orthographic level (Grainger and van Heuven, 2003; Whitney and Cornelissen, 2005, 2008). However, in the light of the task differences we have already described, we need to ask whether these effects are indeed truly lexical. That is, might we see TL effects for non-words in another task? It should now be clear that the obvious way to investigate orthographic effects in non-word processing is to use the same-different task. Kinoshita and Norris (2009) did exactly this and found that TL priming effects were equally robust for words and non-words, indicating that the effect is pre-lexical in origin. Also using the same-different task, García-Orza et al. (2010) extended the finding of TL priming effects to digit- and symbol-strings, demonstrating that the effect is not even limited to letter stimuli. These results are consistent with the assumption of the Overlap model (Gomez et al., 2008) and the noisy-position Bayesian Reader model (Norris et al., 2010). Both of these models suggest that TL priming effects arise from perceptual uncertainty in the location of visual objects during the brief period in which masked primes are presented.

The comparison between the lexical decision task and the same-different task has also been useful in further elucidating how TL priming effects interact with linguistic factors. In lexical decision, TL priming effects have been reported to be modulated by morphological structure. Duñabeitia et al. (2007) used the lexical decision task and reported finding robust TL priming effects in both Basque and Spanish if the letters in the prime are transposed within a morpheme (e.g., spekaer) but not if the letters were transposed across a morphemic boundary (e.g., speaekr). (See however

---

[4]To be complete, Kinoshita and Norris (2011) noted that there are conditions under which familiarity bias does seem to operate, but these are not the conditions under which masked priming is used to study early processes in visual word recognition. Readers are referred to Kinoshita and Norris (2011) for details.

[5]But not in Hebrew (e.g., Velan and Frost, 2009, 2011). We will turn to this finding shortly.

Rueckl and Rimzhim, 2011, for a failure to replicate these "morphological boundary effects" in English.) Previous studies (e.g., Rastle et al., 2004) have shown that priming is observed with masked primes that merely appear to be morphologically complex (e.g., "corner," which appears to contain the suffix "er") primes the stem target (e.g., CORN) and that this effect is not due to mere orthographic overlap ("brothel," where the ending "el" is not a suffix does not prime "BROTH"). This has been taken as evidence for an orthographically driven morphological decomposition process. Within this context, Duñabeitia et al. interpreted their own results as suggesting that morphologically complex words are decomposed into morpheme constituents at the same stage as when letter position coding takes place. Duñabeitia et al. (2010) investigated whether morphological decomposition is an obligatory part of orthographic processing, or occurs only in the service of lexical access. They used the Spanish stimuli used by Duñabeitia et al. (2007) and found that in this task, unlike the lexical decision task, TL priming effects were unaffected by the morphological structure: TL priming effects were equally robust when the letter transposition was within a morpheme, or when it straddled across a morpheme boundary. From these results, Duñabeitia et al. (2010) concluded that the presence of an orthographically defined morpheme (prefix or suffix) is not sufficient to drive the morphological decomposition process (as would be expected, for example, from the view that orthographic representations become structured as a result of learning of structural regularities such as low bigram frequency associated with morpheme boundaries); the morpho-orthographic segmentation process only comes into play when lexical access is attempted.

We (Kinoshita et al., 2012) recently reported a similar task dissociation with Hebrew, a language for which morphology is believed to play a more important role. In contrast to the Indo-European languages with linear concatenative morphology (where prefix/suffix is simply appended to the stem), Semitic morphology is comprised of tri-consonantal *roots* which are embedded in *phonological word patterns*. For example, the Hebrew word TIZ-MORET ("orchestra") consists of the root ZMR, which alludes to the concept of singing, and the phonological word pattern TI-O-ET, which is used to form feminine nouns. Frost et al. (1997) showed that in lexical decision roots but not word patterns prime the whole word and argued that in Hebrew roots are "lexical units." Frost et al. (2005) further showed that unlike in English, primes that are one-letter-different from the target do not produce priming in Hebrew. Based on these results, Frost (2009) argued that whereas in English and other Indo-European languages the lexical space is structured in terms of the constituent letters and their positions, Hebrew lexical space is structured according to the morphological roots. Velan and Frost (2009, 2011) further pointed out that in many Hebrew words, transposing letters in a root produces another root, and showed that in these words, TL priming effects are not found. With this as background, Kinoshita et al. (2012) tested whether Hebrew morphology also modulates TL priming in the same-different task. The results were clear: Robust TL priming effects were found with Hebrew words and non-words, irrespective of morphological structure, even for the words for which Velan and Frost (2009, 2011) did not find TL priming effects in the lexical decision task. Norris and Kinoshita (2012) took these results

to argue that the basic perceptual processes supporting the identification of written symbols are universals: They are governed by exactly the same principles as all other forms of visual object recognition, and that it is what the reader does with those symbols that depend on the properties of the language. These dynamic, task-dependent patterns of TL priming effects would be hard to explain within models of word recognition which assume that orthographic representation with fixed properties – properties that are built in to the orthographic representations to reflect the structure of the language – get activated automatically whenever the word is presented.

The way that TL priming is modulated by morphology might appear to suggest that orthographic processing is different in the two tasks. For example, one might assume that in lexical decision there is some kind of feedback from morphology that alters orthographic processing. However, there is a much simpler explanation that is in line with our suggestion that the primary difference between the tasks is in the way information is used in making decisions.

A system performing optimal decisions should obviously make use of all of the information available. Lexical decision could be based on whole-word forms, but any morphological representations that become available in during the access process should also be taken into account. Whereas transposing letters within a word-sized unit, whether in lexical decision or in the same-different task, may produce a representation that is still a close match to the target, transposing letters between the much smaller morphemic units in a word may well cause much greater disruption, simply because the units are shorter. Indeed, letter transpositions are more apparent with short words (compare for example ALBE/ABLE vs. TRANLSATE/TRANSLATE). So, to the extent that word recognition and lexical decision take advantage of morphology, TL priming effects will be modulated by morphology in tasks that require lexical access, even if morphology has no direct influence on orthographic or letter-level processing. This is another advantage of comparing data from different tasks. One tasks informs us about the importance of morphology, the other tells us that lower level orthographic processing can be completely independent of morphology.

IS PRIMING SEMANTIC? SEMANTIC CATEGORIZATION

The semantic priming effects found with masked primes in the early studies of subliminal perception (e.g., Marcel, 1983) were treated with a great deal of skepticism, and the results generally did not stand up to close methodological scrutiny (e.g., Holender, 1986, see Kouider and Dehaene, 2007, for a historical review of the literature). In visual word recognition studies also, semantic priming effects with masked primes are generally weak and unreliable (e.g., Frost et al., 1997; Rastle et al., 2000). These studies used tasks such as lexical decision and perceptual identification that do not require semantic processing. In contrast, the semantic categorization task necessarily requires semantic processing, as the decision is whether the target word has the semantic features of a category exemplar (e.g., McRae and Boisvert, 1998; Grondin et al., 2009). Thus, masked primes that share semantic features with the target ought to produce priming in a semantic categorization task, and indeed this has now been shown in many studies

(e.g., Dehaene et al., 1998; Bueno and Frenck-Mestre, 1999, 2002; Kunde et al., 2003; Quinn and Kinoshita, 2008). Bueno and Frenck-Mestre (2008) further showed that semantic priming effects with masked primes can be demonstrated at a shorter prime-target SOA in semantic categorization than in lexical decision.

It is important to consider what sort of decision process is involved in semantic categorization when using categories like "animals," and "living things" typically used in studies of word recognition. Natural categories like these are characterized by family resemblance: While category exemplars generally resemble each other to some extent and share features, there is no necessary and sufficient set of features that all exemplars possess (Rosch and Mervis, 1975). That is, contrary to the suggestion made by some researchers (e.g., Carreiras et al., 1997; Forster and Hector, 2002) in semantic categorization, there is no single dimension (e.g., "animalness") that can be monitored to make a decision. One implication of this is that the prime need not be a category member to produce priming; sharing many semantic features with the exemplars should be sufficient to produce priming. Quinn and Kinoshita (2008, Experiment 3, 4) demonstrated this with what they called "impostor" priming. Impostors were non-members of a category which nevertheless shared many semantic features with the exemplars, e.g., for the category "Planets," "comet," "asteroid"; for the category "Human body parts," "mind," "claw." Although subjects correctly rejected these items as non-members when the items were presented as targets in a non-speeded condition, they were slower to reject them in a speeded condition, indicating that these items were similar to the exemplars. When used as primes, these impostors facilitated categorization of targets relative to unrelated primes.

Another point to note with regards natural categories is that the degree of family resemblance varies from category to category. In general, small categories (categories with a small number of members) such as "single-digit numbers" (e.g., one, seven, four), "precious stones" (e.g., diamond, ruby, sapphire), or "planets" (e.g., Mars, Venus, Jupiter) are homogeneous, whereas categories containing a large number of exemplars like "animals" tend to be a superordinate category that comprise heterogeneous subcategories such as birds, mammals, fish, etc. Quinn and Kinoshita (2008) noted that consequently, the prime-target pairs drawn from a small category are more likely to share semantic features relevant to the category classification (e.g., category – "planet," mars-VENUS) than when the exemplars are drawn from a large category (e.g., category – "animal," parrot-RABBIT). Quinn and Kinoshita attributed the failures to find facilitation of target categorization by category-congruent primes with large categories (e.g., Forster et al., 2003; Forster, 2004, Experiment 3) to the lack of semantic feature overlap, and showed that provided that feature overlap is high (e.g., hawk-EAGLE, frog-TOAD)[6], category congruence effects with masked primes are also found with large categories.

---

[6]Quinn and Kinoshita (2008) quantified the amount of feature overlap using the feature production norm of McRae et al. (2005). Other measures of semantic overlap using co-occurrence statistics of words based on text corpus analysis [e.g., Latent Semantic Analysis (LSA): Landauer and Dumais, 1997; Correlated Occurrence Analog to Lexical Semantics (COALS): Rohde et al., 2004] are also possible.

RESPONSE CONGRUENCE EFFECT AND STIMULUS-RESPONSE MAPPING

One of the issues that has been of debated vigorously in research on subliminal perception (see Kouider and Dehaene, 2007, for a review) is whether the category congruence effects observed with masked primes are semantic, or reflect a response conflict that has a different origin. In semantic categorization tasks requiring a binary categorization decision (e.g., is the target bigger/smaller than 5?) and a key-press response to indicate the decision, category congruence is confounded with decision congruence and response congruence. Consequently, these terms are often used interchangeably even though they are conceptually distinct. As noted above, our view is that it is the congruence in the information used to make the decision required by the task that produces priming, not merely response congruence. Nevertheless, when a small set of stimuli are responded to repeatedly in the same task, repetition benefits not only the semantic classification process but also other levels of response representation such as action (which finger is used; see Horner and Henson, 2009). This is a point to consider when interpreting masked priming effects in categorization tasks that used a small set of items repeatedly.

Dehaene et al. (1998) reported a highly influential study using single digits (Arabic numerals, e.g., 1, 8) and number words (e.g., ONE, EIGHT) as primes and targets in a "bigger-than-5?" task. They showed that primes that belonged to the same category as the target (e.g., prime = 3, target = ONE) facilitated the response to the target relative to primes that belonged to the opposite category (e.g., prime = 7, target = ONE). In addition to the behavioral data, they showed congruence effects in the hemodynamic (fMRI) data and the electrophysiological (ERP) measures of brain activity related to the preparation of motor responses, and took the results to argue that masked primes were semantically categorized and then processed all the way to the level of a motor response.

Damian (2001) questioned this conclusion on the basis that Dehaene et al. (1998) used a small set of stimuli repeatedly as both primes and targets and it was therefore possible that the primes were activating a motor response directly, on the basis of a learned stimulus-response mapping. In support of this claim, Damian showed that in a categorization task that required size judgment against an arbitrary reference ("Is the real-world object corresponding to the word larger or smaller than 20 cm × 20 cm?") using a small set of Dutch words (e.g., appel/apple, huis/house), the congruence effect emerged only from the second block, after the prime had been used as a target. Moreover, when the primes were used in a task that did not require the same categorization decision and key-press response (a read-aloud task), they did not produce a congruence effect. Damian concluded from these results that Dehaene et al.'s (1998) findings also reflected stimulus-response mappings rather than congruence in semantic category.

Damian's (2001) claim has in turn been challenged by the finding that with other stimuli (e.g., numbers), primes that have not been responded to ("novel" primes) do produce congruence effects (Naccache and Dehaene, 2001; Kunde et al., 2003; Forster, 2004; Reynvoet et al., 2005; Kinoshita and Hunt, 2008; Quinn and Kinoshita, 2008). Naccache and Dehaene (2001) have further reported that with their number stimuli in the "bigger-than-5" task, the priming effect was greater the closer in numerical distance

the prime and target were (e.g., for the target 1, the prime 2 produced more facilitation than the prime "3"), indicating the semantic nature of the effect. The semantic categorization studies reviewed earlier (e.g., Forster et al., 2003; Forster, 2004, Experiment 3; Quinn and Kinoshita, 2008) used novel primes and a large set of targets presented only once and obtained robust category congruence effects. Stimulus-response mapping alone therefore cannot explain the masked priming effects found in these studies.

Kinoshita and Hunt (2008) used RT distribution analysis to tease apart the contribution of these two levels of congruence – stimulus-response mapping and semantic features – in the "bigger-than-5" task, using single digits as stimuli. They found different RT distributions for novel primes and used primes, with a disproportionate slowdown of congruent trials (e.g., 3-1; 6-9) in the slow RT bins when the prime was a used prime (see Ansorge et al., 2010, for a similar finding). Borrowing ideas developed in the context of response conflict literature (e.g., De Jong et al., 1994; Hommel, 1994), Kinoshita and Hunt suggested that the slowdown specific to the used primes reflects the decay over time, or active suppression of the response code activated directly by the prime. The congruence effect for novel primes, in contrast, was suggested to be semantic in origin, and time locked to the processing of the target. More specifically, the congruence effect for the novel primes was suggested to be a semantic priming effect, and is based on the overlap in semantic features (quantity information) between the prime and the target that are relevant to the categorization.

In sum, when categorizing a small set of stimuli repeatedly, the priming effect could be semantic in origin, or could be due to stimulus-response mapping (see Finkbeiner and Friedman, 2011, for converging evidence based on the trajectory measure using a reaching response). Because visual word recognition researchers have typically used a large set of stimuli presented only once, this issue has not been a concern, however, it is a factor to consider when the set of potential stimuli is small as in single digits or letters of the alphabet, as we will see below.

STIMULUS-RESPONSE MAPPING IN LETTER PROCESSING

As noted in the introduction, one of the most consistent findings in masked priming studies using words as stimuli is that the perceptual overlap between the prime and target does not modulate the size of priming, indicating that priming is driven by representations at the level of abstract letter identity, that is letters that are abstract with regards font, size, and case (i.e., $_A = A = A = a$). Bowers et al. (1998) made an important observation that in contrast to studies using word stimuli, studies using single letters as stimuli failed to find evidence for masked priming of abstract letter identities. Bowers et al. reasoned that if the letter representations are abstract, then the size of identity priming effect should not differ whether the prime and target differing in case are visually similar (e.g., c/C, k/K) or dissimilar (e.g., a/A, b/B). While this was the pattern found with the lexical decision task and noun-verb decision task for words made up of visually similar prime-target pairs (e.g., kiss–KISS) and dissimilar pairs (e.g., edge–EDGE), for individual letter stimuli used in an alphabet decision task and consonant-vowel decision task, priming effects were found only for visually similar pairs. A similar interaction between priming and visual

similarity of prime-target pairs was reported by Arguin and Bub (1995) and Ziegler et al. (2000) in an alphabet decision task. Bowers et al. (1998) concluded that from these data that "abstract letter codes and abstract word codes exist in the orthographic system but for some reason, only orthographic word codes support priming" (p. 1718).

Kinoshita and Kaplan (2008) noted that in studies using single letters as stimuli, the stimulus set was necessarily small (Bowers et al., 1998, used eight visually similar pairs and eight dissimilar pairs), and the letters were used repeatedly both as primes and targets. Thus, just as Damian (2001) suggested with regards the "bigger-than-5?" task with numbers as stimuli, the priming effect observed in these studies may have reflected stimulus-response mapping, based on a partial analysis of the visual features of the prime. To circumvent this response strategy, Kinoshita and Kaplan used the cross-case same-different task. Here, the referent is always in the opposite case to the target, and subjects are asked to decide whether the target is the same letter as the referent, ignoring case. Kinoshita and Kaplan reasoned that here because a letter can be used both in the SAME and DIFFERENT conditions (e.g., SAME: referent = A, target = a; DIFFERENT: referent = B, target = a) equally often, stimulus-response mapping cannot be learned. Because the decision requires abstract letter identity and not physical identity, the decision supporting priming in this task was assumed to be based on abstract letter representations (e.g., one that corresponds to both uppercase A and lowercase a). In line with this assumption, the results showed robust priming effects which were equal in magnitude for the prime-target letter pairs which were visually similar (e.g., c-C, x-X) and visually dissimilar pairs (e.g., a-A, b-B).

To sum up, as summarized in a review of letter perception by Grainger et al. (2008), in the binary categorization tasks (alphabet decision, consonant-vowel decision), letter priming effects are largely driven by the visual similarity of prime-target pairs (for priming effects found with the letter naming task, see Is the Assembly of Phonology Serial? Onset Priming Effect in Reading Aloud). This pattern could not been taken as evidence for priming of abstract letter identities. Only in the cross-case same-different match task, robust priming effects which were insensitive to the visual similarity could be demonstrated. These results highlight the usefulness of task analysis in guiding the design of masked priming experiment.

IS THE ASSEMBLY OF PHONOLOGY SERIAL? ONSET PRIMING EFFECT IN READING ALOUD

Although the Bayesian Reader is not a model of reading aloud, the pattern of data seen in reading aloud should still be modulated by the goal of the task. In reading aloud, the goal is to generate a speech response, and to initiate articulation as quickly as possible while minimizing errors. One feature of the masked priming effect in this task is that it is highly sensitive to the overlap in phonemic onset. Forster and Davis (1991) were the first to note that relative to the all-letter-different unrelated control (e.g., fame-SINK), overlap in the onset alone (e.g., same-SINK) facilitates the naming of the target. This onset priming effect has been replicated in a number of languages that use alphabetic scripts

(in English, Kinoshita, 2000; Kinoshita and Woollams, 2002; in Dutch, Schiller, 2004; in French, Grainger and Ferrand, 1996; in Spanish, Dimitropoulou et al., 2010; in Korean alphabetic Hangul, Kim and Davis, 2002) but not in syllabic scripts like Korean Hanja (e.g., Kim and Davis, 2002) or in mora-based Japanese (Verdonschot et al., 2011). The onset priming effect has been found with the naming of single letter targets (Bowers et al., 1998) and picture targets (Schiller, 2008), so it is not specific to reading. Taken together with the fact that it is absent in the lexical decision task (Forster and Davis, 1991, Experiment 5; Carreiras et al., 2005, Experiment 1; Grainger and Ferrand, 1996, Experiment 4; Kim and Davis, 2002, Experiment 1b), the results indicate that the onset priming effect reflects the task goal of the need to generate a speech output for the target (Grainger and Ferrand, 1996; Kinoshita and Woollams, 2002; Kinoshita, 2003; Carreiras et al., 2005).

The onset priming effect has important methodological implications. The facilitation due to the mere overlap of onset in the naming task can be sizeable (up to about 30 ms in Grainger and Ferrand, 1996), and it can complicate the interpretation of priming effects in the read-aloud task. An example of this is seen in studies that investigated identity priming of abstract letter identities using the letter naming task (Arguin and Bub, 1995; Bowers et al., 1998). In these studies the size of identity priming effect for prime-target pairs that were visually similar across case (e.g., c-C, k/K) and dissimilar (e.g., a-A, b/B), which is consistent with the idea that there is an abstract letter identity which is invariant across shape and case. Bowers et al. however noted that this pattern was not found in other tasks such as the alphabet decision task and consonant-vowel decision task (as discussed above), and followed up the locus of the identity priming effect in the letter naming task by comparing it to priming produced by word homophone primes (e.g., sea-C, cue-Q), and phonologically similar letter primes (e.g., i-Y). Homophone primes produced facilitation that was as large as the identity primes, whereas the phonological primes produced little facilitation. These results are readily interpretable from the perspective that the priming effects in the naming task are mainly driven by the overlap in the phonemic onset between the prime and the target.

The task dependence of the onset priming effect also has important implications for theory development. Because reading aloud necessarily requires the generation of phonology, the naming task has been a task of choice for researchers interested in the role of phonology in word recognition (e.g., Plaut et al., 1996; Coltheart et al., 2001; Perry et al., 2007, 2010). To date, two computational models have provided accounts of the onset priming effect. According to the DRC model (Coltheart et al., 2001, for the most recent version with simulations of the onset priming effect, see Mousikou et al., 2010), the onset priming effect reflects a serial, left-to-right letter-to-phoneme mapping process implemented in the non-lexical route. The CDP+ (and CDP++) model (Perry et al., 2007, 2010), on the other hand, suggests the locus of the effect is the graphemic parsing process within the sublexical route. A sequence of letters is segmented into graphemes (which can include multi-letter graphemes such as SH, TCH) which are then placed into the Onset, Vowel, and Coda slots, and this parsing process is assumed to occur from left to right across the letter

sequence[7]. Both models account well for the serial, left-to-right nature of the effect (see Kinoshita, 2000; Montant and Ziegler, 2001; Malouf and Kinoshita, 2007): The phonemic overlap has little benefit if it is in the latter part of the stimuli but the onset differs (e.g., suf-SIB < muf-SIB but mub-SIB = muf-SIB; noon-MOON = need-MOON). However, both the DRC and CDP+ models lack an independent motivation (other than to account for the data) as to why the sublexical letter-to-phoneme mapping process or the sublexical graphemic parsing process operates in the left-to-right fashion. The models also have a difficulty accommodating the fact that in a task that does not require a speech output, the effect of sublexical phonology is not necessarily left-to-right. Recently, Kinoshita and Norris (2012) reported that pseudohomophone primes (e.g., cymptom-SYMPTOM, frajile-FRAGILE) produced greater priming effects than orthographic control primes (e.g., lymptom-SYMPTOM, franile-FRAGILE) and that the benefit did not differ for initial (e.g., cymptom) and medial (e.g., frajile) positions. Such a finding suggests that the phonology for the pseudohomophone was generated across the letter string in parallel, rather than serially, from left to right.

Recognizing the task-dependent nature of onset priming effect (that it occurs only when speech output needs to be generated) provides a rationale for proposing a different locus for a left-to-right process. In the speech production literature, there is much evidence for a serial, segment-to-frame association process (e.g., Levelt et al., 1999). In picture naming, the word-form phonology is retrieved as a whole, then (re)syllabified into metrical frames with stress markings. The phoneme segments are then slotted into the metrical frames, and this process is assumed to be left-to-right. The masked onset priming effect could have its origin in this frame-to-segment association process when preparing a speech response to the target. In line with the view, Roelofs (2004) showed that the serial, left-to-right process is observed both when naming objects and reading their printed names, and suggested that "the observed seriality is due to phonological encoding mechanisms shared by naming and reading rather than a grapheme-to-phoneme conversion in oral reading" (p. 221).

IS MASKED PRIMING STRATEGIC?

Before concluding this review, one comment is in order regarding the question of whether masked priming effect is "strategic." One of the major appeals of the masked priming procedure is that because subjects have no awareness of the prime, any priming effects should be immune to the influence of conscious strategies (e.g., Forster et al., 2003)[8]. This related to the belief that an "unconscious" prime will tap into automatic processes and can therefore be used to identify the obligatory representations and processes that support reading. We note that our claim that masked priming effects are driven by the goal of the task does not entail the

---

[7]The fact that unpronounceable consonant string primes (e.g., cdkm-CARO) do not produce onset priming effect (Dimitropoulou et al., 2010) is problematic for the DRC, but can be accounted for by the CDP+ model by assuming that such strings cannot be parsed into graphemes.

[8]An example can be seen in De Groot's (1983) investigation of semantic priming effects in a lexical decision task. By masking the prime, she successfully precluded subjects from using the "post-access coherence check," which creates a bias toward responding "non-word" when the prime and target are semantically unrelated.

view that masked priming is strategic, in the sense that subjects can choose how to make use of the prime-target relationship to facilitate responding to the target in a way that best suits the specific experimental context. Subjects "apply the task instruction to the prime" when the prime is masked because they are unaware that the prime and target are separate perceptual events. On this view, the relationship between the prime and target is necessarily veiled from awareness, and hence cannot be used strategically to facilitate responding to the target.

In apparent contradiction of this view, Bodner and Masson (2001, 2003, 2004) have reported finding in a number of studies that the proportion of related trials with masked primes modulated the size of priming effects: The size of masked priming effect is larger in a block containing a high proportion (0.8) of related trials and a low proportion (0.2) of unrelated trials relative to a block containing the opposite mix. This was found with the lexical decision task (Bodner and Masson, 2001; Bodner and Stalinski, 2008; note however that the effect of relatedness proportion in this task was variable across individual experiments), the read-aloud task (Bodner and Masson, 2004), and the parity judgment task (deciding whether a number is odd or even, Bodner and Dypvik, 2005). The interpretation offered by Bodner and Masson of these effects of relatedness proportion – which they referred to as the "prime validity effect" – is that "the processing operations applied to the prime to identity and interpret it form a new memory representation," and this memory episode can be recruited, without awareness, to assist with processing of a subsequent target. The degree of recruitment is modulated as a function of list context, with "a context containing a high proportion of task-useful primes" cuing the cognitive system to increase prime recruitment to facilitate target processing (Bodner et al., 2006, p. 1299). In other words, Bodner and Masson suggest that the recruitment of the prime, even when it is masked and not consciously available, is under strategic control, and is a function of the list-wide utility of the prime-target relationship.

While we agree broadly with the view that masked priming reflects the overlap of processing operations applied to the prime and the target, we do not believe that this implies that a new memory episode is established for masked primes. In fact, our view is that when the prime is masked, subjects are unaware that the prime and target are distinct perceptual events[9]. Consequently, we do not believe that the cognitive system is able to modulate the impact of the masked prime strategically as a function of the list-wide utility of the prime-target relationship – note that relatedness proportion is not necessarily the same as list-wide prime utility. For example, in a parity (odd-even) decision task, in a block containing a low proportion of category-congruent trials, the prime has high utility, as it predicts the opposite response to the prime (e.g., if the prime is odd, the target is likely to be even, and vice versa). Consistent with our view, the list-wide predictability of the target response from the prime modulates the size of priming when the prime is visible, but not when the prime is masked. In a parity decision task, when the relatedness proportion is 0.5 [i.e., when the

prime is an odd (even) number, the target is also odd (even) on half of the trials and even (odd) on the other half], the list-wide prime utility is zero, as the target parity cannot be predicted from the prime parity: Kinoshita et al. (2011) showed that in this condition there is no effect of parity congruence if the prime was visible, but there is a positive effect of congruence if the prime was masked. In a similar vein, in the same-different task, Kinoshita and Norris (2009) showed that the prime-target response contingency had a large impact on the size of priming when the prime was visible, but little impact when the prime was masked. With visible primes, the priming effect was large when the response to the target was the same as the prime on 75% of the trials (predictable contingency), and reduced to a negligible level when the they were the same on 50% of the trials (zero-contingency); in contrast, with masked primes, the priming effects was equally large and robust in the predictable- and zero-contingency conditions.

Given these dissociations between effects of relatedness proportion with the visible and masked primes, with only the former being a function of list-wide prime utility, how could the effects observed by Bodner and Masson with the masked primes be explained? Kinoshita et al. (2008, 2011) have presented an alternative account – termed the Adaptation to the statistics of the environment (ASE) – which explains the effect in terms of the adaptation of the response initiation process to the history of trial difficulty. The account is based on the assumption that in RT tasks, subjects attempt to meet the instruction to "respond as quickly as possible without making too many errors" by estimating the optimal point to initiate responding that minimizes the total cost of responding too early and risking an error, and delaying the response unnecessarily. The history of trial difficulty, in particular, that of immediately preceding trials, is used in conjunction with the evidence accumulated in the current trial to estimate the optimal point to respond. Consistent with this, RT of previous trial is positively correlated with the current trial RT (independent of post-error slowdown). When the condition is such that the easy trials show greater sensitivity to the previous trial RT than the hard trials, the relatedness proportion effect with masked primes falls out of the ASE model from the fact that related trials are "easy" trials (and unrelated trials are "hard" trials). Although the ASE model is silent with regards why easy trials should be more sensitive to the previous trial RT, this pattern is expected to hold in general from the assumption that when trials slow down by a fixed amount of time, the benefit in the reduction in error rate would be greater for the easy trials than for the hard trials due to its greater rate of evidence accumulation. Kinoshita et al. (2011) reanalyzed the masked priming experiment showing an apparent relatedness proportion effect (Experiment 3) as a function of previous trial RT, and found greater sensitivity of related trials than the unrelated trials to the previous RT, consistent with the ASE. Furthermore, they showed that the adaptation to the list-wide difficulty (as determined by the proportion of easy vs. hard trials) is a noisy process requiring many trials (over 300 trials) – many more than are standard in masked priming experiments.

In sum, the pattern of relatedness proportion effects are different with visible and masked primes, and the latter – referred to

---

[9]For evidence that an explicit episodic record of the prime capable of supporting long-term priming (priming spanning several intervening trials) is formed only with visible primes, see Humphreys et al. (1988).

as the prime validity effect and interpreted by Bodner and Masson as the main evidence for the view that masked priming is strategic – has an alternative explanation. We take this to argue that the concern that masked priming effects are strategic – in the sense that the impact of the prime on the target processing can be modulated as a function of the list-wide prime utility – is unwarranted.

CONCLUSION

From the perspective that different tasks should all tap into the output of a fixed lexical processing system, the diverse pattern of results found with different masked priming tasks makes little sense. Indeed, one might be tempted to conclude that because different tasks behave differently, none of them is particularly useful. After all, which task provides the true measure of lexical processing? In contrast, the view that we have presented is much more optimistic about the value of masked priming data. In fact, it implies that the different tasks all have something valuable to tell us. Different tasks tap into different facets of the word recognition process in ways that are quite systematic and lawful. We have argued that the variation in the pattern of priming is what should be expected if perception can be characterized as approximating optimal Bayesian decision making operating by sampling evidence accumulated from the perceptual input. As we explained earlier, we also need to assume that in masked priming the evidence from the prime and target is integrated in reaching a decision. That is, the prime and target are not treated as separate perceptual events. We illustrated this with the urn analogy. In order to make optimal

decisions there is no need to know which samples or balls come from the "prime" and which from the target. Each sample, or ball, provides an independent piece of evidence that can be used in making the decision. This verbal description of the optimal decision making process is supported by simulations (see, e.g., Norris and Kinoshita, 2008; Norris et al., 2010).

Much of the data we have reviewed here focuses on the differences between lexical decision and the same-different task. This contrast is particularly illuminating because it allows us to investigate the source of these effects – whether it is in the orthographic representations, or in the lexicon. For example, the fact that TL priming effects are found with non-words (or even symbols) in the same-different task has told us that the effect reflects the perceptual uncertainty of positions of individual objects in a string. It is not surprising therefore that TL priming effects are found universally across languages in the same-different task. In contrast, in the lexical decision task, TL priming effects are modulated by language, and by morphological structure. From the perspective that priming effects reflect automatic activation of representations with a fixed property, that masked priming effects are task-dependent may be puzzling. However, from the view that masked priming reflects the accumulation of task-specific evidence contributed by the prime, that the pattern of task-dependent masked priming effects mirrors the task goal is exactly what is expected, given that the evidence needs to be accumulated is determined by the task. This framework provides a useful guide to interpreting what information is available to the readers in the first 250 ms of visual word recognition.

REFERENCES

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

Arguin, M., and Bub, D. (1995). Priming and response selection processes in letter classification and identification tasks. *J. Exp. Psychol. Hum. Mem. Percept. Perform.* 21, 1199–1219.

Bodner, G. E., and Dypvik, A. T. (2005). Masked priming of number judgments depends on prime validity and task. *Mem. Cognit.* 33, 29–47.

Bodner, G. E., and Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: evidence for a nonlexical basis for priming. *J. Mem. Lang.* 37, 268–293.

Bodner, G. E., and Masson, M. E. J. (2001). Prime validity affects masked repetition priming: evidence for an episodic resource account of priming. *J. Mem. Lang.* 45, 616–647.

Bodner, G. E., and Masson, M. E. J. (2003). Beyond spreading activation: an influence of relatedness proportion on masked semantic priming. *Psychon. Bull. Rev.* 10, 645–652.

Bodner, G. E., and Masson, M. E. J. (2004). Beyond binary judgments: prime validity modulates masked repetition priming in the naming task. *Mem. Cognit.* 32, 1–11.

Bodner, G. E., Masson, M. E. J., and Richard, N. T. (2006). Repetition proportion biases masked priming of lexical decisions. *Mem. Cognit.* 34, 1298–1311.

Bodner, G. E., and Stalinski, S. M. (2008). Masked repetition priming and proportion effects under cognitive load. *Can. J. Exp. Psychol.* 62, 127–131.

Bowers, J. S. (2010). Does masked and unmasked priming reflect Bayesian inference as implemented in the Bayesian Reader? *Eur. J. Cogn. Psychol.* 22, 779–797.

Bowers, J. S., Vigliocco, G., and Haan, R. (1998). Orthographic, phonological, and articulatory contributions to masked letter and word priming. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1705–1719.

Bueno, S., and Frenck-Mestre, C. (1999). Semantic features and semantic categories: differences in rapid activation of the lexicon. *Brain Lang.* 68, 199–204.

Bueno, S., and Frenck-Mestre, C. (2002). Rapid activation of the lexicon: further investigation with

behavioral and computational results. *Brain Lang.* 81, 120–130.

Bueno, S., and Frenck-Mestre, C. (2008). The activation of semantic memory: effects of prime exposure, prime-target relationship and task demands. *Mem. Cognit.* 36, 882–898.

Carreiras, M., Ferrand, L., Grainger, J., and Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychol. Sci.* 16, 585–589.

Carreiras, M., Perea, M., and Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: cross-task comparisons. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 857–871.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. C. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256.

Damian, M. (2001). Congruity effects evoked by subliminally presented primes: automaticity rather than semantic processing. *J. Exp. Psychol. Hum. Mem. Percept. Perform.* 27, 154–165.

Davis, C., and Forster, K. I. (1994). Masked orthographic priming: the effect of prime-target legibility. *Q. J. Exp. Psychol. A* 47, 673–697.

De Groot, A. M. B. (1983). The range of automatic spreading activation in word priming. *J. Verbal Learn. Verbal Behav.* 22, 417–436.

De Jong, R., Liang, C.-C., and Lauber, E. (1994). Conditional and unconditional automaticity: a dual-process model of effects of spatial stimulus-response correspondence. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 731–750.

Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van der Moortele, P. F., and Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature* 395, 597–600.

Dimitropoulou, M., Dunabeitia, J. A., and Carreiras, M. (2010). Influence of prime lexicality, frequency, and pronounceability on the masked onset priming effect. *Q. J. Exp. Psychol.* 63, 1813–1837.

Duñabeitia, J. A, Kinoshita, S., Carreiras, M., and Norris, D. (2010). Is morph-orthographic decomposition purely orthographic? Evidence from masked priming in the same-different task. *Lang. Cogn. Process.* 26, 509–529.

Duñabeitia, J. A., Perea, M., and Carreiras, M. (2007). Do transposed-letter similarity effects

occur at a morpheme level? Evidence for ortho-morphological decomposition. *Cognition* 105, 691–703.

Evett, L. J., and Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Q. J. Exp. Psychol.* 33, 325–350.

Finkbeiner, M., and Friedman, J. (2011). The flexibility of nonconsciously (deployed) cognitive processes: evidence from masked congruence priming. *PLoS ONE* 6, e17095. doi:10.1371/journal.pone.0017095

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Lang. Cogn. Process.* 1, 87–108.

Forster, K. I. (1998). The pros and cons of masked priming. *J. Psycholinguist. Res.* 27, 203–233.

Forster, K. I. (2004). Category size effects revisited: frequency and masked priming effects in semantic categorization. *Brain Lang.* 90, 276–286.

Forster, K. I., and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 680–698.

Forster, K. I., and Davis, C. (1991). The density constraint on form-priming in the naming task: interference effects from a masked prime. *J. Mem. Lang.* 30, 1–25.

Forster, K. I., Davis, C., Schoknecht, C., and Carter, R. (1987). Masked priming with graphemically related forms: repetition or partial activation? *Q. J. Exp. Psychol.* 39, 211–251.

Forster, K. I., and Hector, J. (2002). Cascaded and noncascaded models of lexical and semantic processing: the turple effect. *Mem. Cognit.* 30, 1106–1117.

Forster, K. I., Mohan, K., and Hector, J. (2003). "The mechanics of masked priming," in *Masked Priming: The State of the Art*, eds S. Kinoshita and S. J. Lupker (Hove: Psychology Press), 3–37.

Frost, R. (2009). "Reading in Hebrew vs. reading in English: is there a qualitative difference?" in *How Children Learn to Read: Current Issues and New Directions in the Integration of Cognition, Neurobiology, and Genetics of Reading and Dyslexia Research and Practice*, eds K. Pugh and P. McCradle (New York: Psychology Press), 235–254.

Frost, R., Forster, K. I., and Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 829–856.

Frost, R., Kugler, T., Deutsch, A., and Forster, K. I. (2005). Orthographic structure versus morphological structure: principles of lexical organization in a given language. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1293–1326.

García-Orza, J., Perea, M., and Muñoz, S. (2010). Are transposition effects specific to letters? *Q. J. Exp. Psychol.* 63, 1603–1618.

Gomez, P., Ratcliff, R., and Perea, M. (2008). The overlap model: a model of letter position coding. *Psychol. Rev.* 115, 577–600.

Grainger, J. (2008). Cracking the orthographic code: an introduction. *Lang. Cogn. Process.* 23, 1–35.

Grainger, J., and Ferrand, L. (1994). Phonology and orthography in visual word recognition: effects of masked homophone primes. *J. Mem. Lang.* 33, 218–233.

Grainger, J., and Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: cross-task comparisons. *J. Mem. Lang.* 35, 623–647.

Grainger, J., and Holcombe, P. J. (2009). Watching the word go by: on the time course of component processes in visual word recognition. *Lang. Linguist. Compass* 3/1, 128–156.

Grainger, J., Rey, A., and Dufau, S. (2008). Letter perception: from pixels to pandemonium. *Trends Cogn. Sci. (Regul. Ed.)* 12, 381–387.

Grainger, J., and van Heuven, W. (2003). "Modeling letter position coding in printed word perception," in *The Mental Lexicon*, ed. P. Bonin (New York: Nova Science), 1–23.

Grainger, J., and Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Front. Psychol.* 2:54. doi:10.3389/fpsyg.2011.00054

Grondin, R., Lupker, S. J., and McRae, K. (2009). Shared features dominate semantic richness effects for concrete concepts. *J. Mem. Lang.* 60, 1–19.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: a survey and appraisal. *Behav. Brain Sci.* 9, 1–23.

Hommel, B. (1994). Spontaneous decay of response-code activation. *Psychol. Res.* 56, 261–268.

Horner, A. J., and Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 757–779.

Humphreys, G. W., Besner, D., and Quinlan, P. T. (1988). Event perception and the word repetition effect. *J. Exp. Psychol. Gen.* 117, 51–67.

Kim, J., and Davis, C. (2002). Using Korean to investigate phonological priming effects without the influence of orthography. *Lang. Cogn. Process.* 17, 569–591.

Kinoshita, S. (2000). The left-to-right nature of the masked onset effect in naming. *Psychon. Bull. Rev.* 7, 133–141.

Kinoshita, S. (2003). "The nature of masked onset priming effects in naming," in *Masked Priming: The State of the Art*, eds S. Kinoshita and S. J. Lupker (Hove: Psychology Press), 223–238.

Kinoshita, S., and Hunt, L. (2008). RT distribution analysis of category congruence effects with masked primes. *Mem. Cognit.* 36, 1324–1334.

Kinoshita, S., Forster, K. I., and Mozer, M. C. (2008). Unconscious cognition isn't that smart: modulation of masked repetition effect in the naming task. *Cognition* 107, 623–649.

Kinoshita, S., Mozer, M. C., and Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *J. Exp. Psychol. Gen.* 140, 622–636.

Kinoshita, S., and Norris, D. (2009). Transposed-letter priming of pre-lexical orthographic representations. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1–18.

Kinoshita, S., and Norris, D. (2011). Does the familiarity bias hypothesis explain why there is no masked priming for "NO" decisions? *Mem. Cognit.* 39, 319–334.

Kinoshita, S., and Norris, D. (2012). Pseudohomophone priming in lexical decision is not fragile in a sparse lexical neighborhood. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 764–775.

Kinoshita, S., Norris, D., and Siegelman, N. (2012). Transposed-letter priming effect in Hebrew in the same-different task. *Q. J. Exp. Psychol.* doi:10.1080/17470218.2012.655749

Kinoshita, S., and Woollams, A. M. (2002). The masked onset priming effect in naming: computation of phonology or speech-planning? *Mem. Cognit.* 30, 237–245.

Kinoshita, S., and Kaplan, L. (2008). Priming of abstract letter identities in Hebrew in the letter match task. *Q. J. Exp. Psychol.* 61, 1873–1885.

Knill, D. C., Kersten, D., and Yuille, A. (1996). "A Bayesian formulation of visual perception," in *Perception as Bayesian Inference*, eds D. C. Knill and W. Richards (Cambridge: Cambridge University Press), 1–21.

Kouider, S., and Dehaene, S. (2007). Levels of processing during nonconscious perception: a critical review of visual masking. *Philos. Trans. R. Soc. Lond. B.* 362, 857–875.

Kunde, W., Kiesel, A., and Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition* 88, 223–242.

Landauer, T. K., and Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychol. Rev.* 104, 211–240.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38.

Lupker, S. J., Perea, M., and Davis, C. J. (2008). Transposed-letter effects: consonants, vowels, and letter frequency. *Lang. Cogn. Process.* 23, 93–116.

Malouf, T., and Kinoshita, S. (2007). Masked onset priming effect for high-frequency words: further support for the speech-planning account. *Q. J. Exp. Psychol.* 60, 1155–1167.

Marcel, A. J. (1983). Conscious and unconscious perception: experiments on visual masking and word recognition. *Cogn. Psychol.* 15, 197–237.

McRae, K., and Boisvert, S. (1998). Automatic semantic similarity priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 558–572.

McRae, K., Cree, G. S., Seidenberg, M. S., and McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behav. Res. Methods Instrum. Comput.* 37, 547–559.

Montant, M., and Ziegler, J. (2001). Can orthographic rimes facilitate naming? *Psychol. Bull. Rev.* 8, 351–356.

Mousikou, P., Coltheart, M., Finkbeiner, M., and Saunders, S. (2010). Can the DRC computational model of reading offer a valid account of the masked onset priming effect? *Q. J. Exp. Psychol.* 63, 984–1003.

Naccache, L., and Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition* 80, 223–237.

Norris, D. (2006). The Bayesian Reader: explaining word recognition as an

optimal Bayesian decision process. *Psychol. Rev.* 113, 327–357.

Norris, D. (2009). Putting it all together: a unified account of word recognition and reaction-time distributions. *Psychol. Rev.* 116, 207–219.

Norris, D., and Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: insights from masked priming. *J. Exp. Psychol. Gen.* 137, 433–455.

Norris, D., and Kinoshita, S. (2012). Orthographic processing is universal; it's what you do with it that's different: comment on Frost. *Behav. Brain Sci.* (in press).

Norris, D., Kinoshita, S., and van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *J. Mem. Lang.* 62, 254–271.

Perea, M., and Carreiras, M. (2008). Do orthotactics and phonology constrain the transposed-letter effect? *Lang. Cogn. Process.* 23, 69–92.

Perea, M., and Lupker, S. J. (2003). "Transposed-letter confusability effects in masked form priming," in *Masked Priming: The State of the Art*, eds S. Kinoshita and S. J. Lupker (New York: Psychology Press), 97–120.

Perea, M., and Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *J. Mem. Lang.* 51, 231–246.

Perea, M., and Perez, E. (2009). Beyond alphabetic orthographies: the role of form and phonology in transposition effects in Katakana. *Lang. Cogn. Process.* 24, 67–88.

Perry, C., Ziegler, J. C., and Zorzi, M. (2007). Nested modeling and strong inference testing in the development of computational theories: the CDP+ model of reading aloud. *Psychol. Rev.* 27, 301–333.

Perry, C., Ziegler, J. C., and Zorzi, M. (2010). Beyond single syllables: large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cogn. Psychol.* 61, 106–151.

Plaut, D., McClelland, J. L., Seidenberg, M. S., and Patterson, K. E. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* 103, 56–115.

Quinn, W. M., and Kinoshita, S. (2008). Category congruence effect in semantic categorization with masked primes with narrow and broad categories. *J. Mem. Lang.* 58, 286–306.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., and Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: a time-course study. *Lang. Cogn. Process.* 15, 507–537.

Rastle, K., Davis, M. H., and New, B. (2004). The broth in my brother's brothel: morpho-orthographic segmentation in visual word recognition. *Psychon. Bull. Rev.* 11, 1090–1098.

Reynvoet, B., Gevers, W., and Caessens, B. (2005). Unconscious primes activate motor codes through semantics. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 991–1000.

Roelofs, A. (2004). Seriality of phonological encoding in naming objects and reading their names. *Mem. Cognit.* 32, 212–222.

Rohde, D. L. T., Gonnerman, L. M., and Plaut, D. C. (2004). *An Improved Method for Deriving Word Meanings from Lexical Co-occurrence Statistics.* Available at: http://dlt4.mit.edu/%7Edr/COALS/Coals.pdf

Rosch, E., and Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cogn. Psychol.* 7, 573–605.

Rueckl, J. G., and Rimzhim, A. (2011). On the interaction of letter transpositions and morpheme boundaries. *Lang. Cogn. Process.* 26, 482–508.

Schiller, N. O. (2004). The onset effect in word naming. *J. Mem. Lang.* 50, 477–490.

Schiller, N. O. (2008). The masked onset priming effect in picture naming. *Cognition* 106, 952–962.

Schoonbaert, S., and Grainger, J. (2004). Letter position coding in printed word perception: effects of repeated and transposed letters. *Lang. Cogn. Process.* 19, 333–367.

Velan, H., and Frost, R. (2009). Letter-transposition effects are not universal: the impact of transposing letters in Hebrew. *J. Mem. Lang.* 61, 285–302.

Velan, H., and Frost, R. (2011). Words with and without internal structure: what determines the nature of orthographic and morphological processing? *Cognition* 118, 141–156.

Verdonschot, R. G., La Heij, W., Kiyama, S., Tamaoka, K., Kinoshita, S., and Schiller, N. O. (2011). The functional unit of Japanese word naming: evidence from masked priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1458–1471.

Whitney, C., and Cornelissen, P. (2005). Letter-position coding and dyslexia. *J. Res. Read.* 28, 274–301.

Whitney, C., and Cornelissen, P. (2008). SERIOL reading. *Lang. Cogn. Process.* 23, 143–164.

Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., and Grainger, J. (2000). Visual and phonological codes in letter and word recognition: evidence from incremental priming. *Q. J. Exp. Psychol.* 53A, 671–692.