THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2015
http://dx.doi.org/10.1080/17470218.2015.1018836

Subliminal repetition primes help detection of phonemes in a picture: Evidence for a phonological level of the priming effects

Laura Manoiloff[1], Juan Segui[2,4], and Pierre Hallé[2,3,4]

[1]Equipo de Investigación de Psicología Cognitiva del Lenguaje y Psicolingüística, Laboratorio de Psicologia Cognitiva, Universidad Nacional de Cordoba, Córdoba, Argentina
[2]Laboratoire Mémoire et Cognition (INSERM - Paris 5) and CNRS, Paris, France
[3]Laboratoire de Phonétique et Phonologie (CNRS-Paris 3) Paris, France
[4]Labex EFL, Paris, France

(Received 13 March 2014; accepted 9 February 2015)

In this research, we combine a cross-form word–picture visual masked priming procedure with an internal phoneme monitoring task to examine repetition priming effects. In this paradigm, participants have to respond to pictures whose names begin with a prespecified target phoneme. This task unambiguously requires retrieving the word-form of the target picture's name and implicitly orients participants' attention towards a phonological level of representation. The experiments were conducted within Spanish, whose highly transparent orthography presumably promotes fast and automatic phonological recoding of subliminal, masked visual word primes. Experiments 1 and 2 show that repetition primes speed up internal phoneme monitoring in the target, compared to primes beginning with a different phoneme from the target, or sharing only their first phoneme with the target. This suggests that repetition primes preactivate the phonological code of the entire target picture's name, hereby speeding up internal monitoring, which is necessarily based on such a code. To further qualify the nature of the phonological code underlying internal phoneme monitoring, a concurrent articulation task was used in Experiment 3. This task did not affect the repetition priming effect. We propose that internal phoneme monitoring is based on an abstract phonological code, prior to its translation into articulation.

Keywords: Masked priming; Word–picture cross-form priming; Internal phoneme monitoring; Phonological representation; Articulatory suppression.

Models of speech production assume that the final articulatory process is preceded by other more abstract processing levels: conceptual activation, lexical access, and word-form encoding (Dell, 1986; Garrett, 1975; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; etc.). To examine these hypothetical processes, researchers have generally studied the impact of different factors on explicit speech production tasks such as the picture naming and the word naming tasks. In these tasks, the main

Correspondence should be addressed to Laura Manoiloff, Equipo de Investigación de Psicología Cognitiva del Lenguaje y Psicolingüística, Laboratorio de Psicología Cognitiva, Universidad Nacional de Córdoba (UNC), Enfermera Gordillo esquina Enrique Barros, Ciudad Universitaria, Córdoba, Argentina. E-mail: lmanoiloff@psyche.unc.edu.ar

This work was supported by the Secretaria de Ciencia y Técnica (SECyT) of the Universidad Nacional de Córdoba (UNC), Argentina.

behavioural diagnostic for speech production processing is the time interval that elapses between the presentation of a stimulus and the subject's naming latency. Word frequency has been one of the main factors examined in speech production. Oldfield and Wingfield (1965) were the first to find that naming times are faster for pictures with a frequent name than for pictures with an infrequent name. Since then, a large number of studies have been conducted to characterize the nature and functional locus of this effect. Even if there is still some disagreement on this point, it is generally assumed that frequency effects in picture naming are a signature of lexical access.

In a recent paper (Manoiloff, Segui, & Hallé, 2013), we examined the presence of a word frequency effect using an internal phoneme monitoring task rather than the standard picture naming task. In the internal phoneme monitoring paradigm, subjects are presented with pictures and are asked to respond, as quickly as possible, to pictures whose name begins with a prespecified target phoneme (e.g., to respond if and only if the name of the presented picture begin with the phoneme /t/). This task clearly requires access to some phonological representation of the picture's name without necessarily engaging the articulatory processes involved in picture naming. In one experiment, we demonstrated that phoneme monitoring times were shorter for pictures with a high- than for those with a low-frequency name. These results are in agreement with recent findings by Hutson, Damian, and Spalek (2013), who replicated a frequency effect previously observed in picture naming with picture–word interference, using an internal phoneme monitoring instead of a picture-naming task: Detection times of the initial phoneme of the target picture's name were shorter when the distractor word was a low- than when it was a high-frequency word. In Manoiloff et al. (2013), no frequency effect was obtained in two additional experiments using the same pictures but a different task: a visual object recognition and a semantic categorization task. The absence of a frequency effect with these tasks, bearing respectively on the visual and conceptual level of representation, and its presence with the internal phoneme monitoring task, support the hypothesis that the observed frequency effect using the latter task is related to lexical access and, more precisely, to the retrieval of phonological information through lexical access. Assuming that internal phoneme monitoring times reflect the access to the picture's name word-form, this procedure should be sensitive to the presence of word primes phonologically related to the picture's name. In particular, repetition priming effects should be observed using subliminal word primes. This point is examined in Experiments 1 and 2 of the present paper.

Most of the previous studies on subliminal masked priming effects have been conducted using printed prime–target pairs to explore the early stages of word recognition. In visual masked priming conditions, phonological priming effects are rather elusive and vary significantly according to experimental task and according to language (Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009; Kouider & Dehaene, 2007; Rastle & Brysbaert, 2006). In clear contrast, robust phonological priming effects have been obtained using a cross-form word–picture masked priming paradigm. Using this paradigm, Ferrand, Grainger, and Segui (1994) reported that the prior presentation as a prime of the target picture's name, or of a pseudohomophone of this name, facilitates target picture naming. From these results, Ferrand et al. (1994) concluded that the priming effects they obtained were mainly due to the preactivation by the prime of the whole-word phonological representation of the target picture's name. According to this interpretation, the observed effect would be phonological in nature and related to the processes that mediate between the conceptual and articulatory levels of representation activated in picture naming. The cross-form masked priming paradigm thus seems appropriate to reveal phonological effects. The main goal of the present study is to determine to what extent phonological effects can be obtained using the internal phoneme monitoring procedure instead of picture naming. For this purpose, we designed experiments combining the cross-form word–picture masked priming paradigm and the internal phoneme monitoring task. A second goal of the present study is to qualify the phonological level of representation underlying internal phoneme monitoring responses. Two main levels of representation have been proposed: (a) an abstract phonological code computed from the phonological code retrieved upon lexical access, and (b) a more concrete phonetic–articulatory code specifying actual articulation. One way to gauge the plausibility of these two propositions is to test whether performing a concurrent articulatory task, or "articulatory suppression", affects phonological priming effects in internal phoneme monitoring. Indeed, if the critical level of representation for internal phoneme monitoring were closely linked to articulation, either specifying articulation or produced by articulation, articulatory suppression would logically alter phoneme monitoring and cancel facilitation effects. We examine this point in Experiment 3.

The experiments reported in this paper are conducted with Spanish, a language with a highly transparent orthography. This point is important to note given that languages with highly transparent ("shallow") writing system may promote fast and automatic phonological recoding of printed words in subliminal conditions of presentation.

EXPERIMENT 1

This experiment used a cross-form word–picture masked priming procedure. The participants had to monitor for word-initial phonemes in the names of pictures, which were preceded by a visual masked word prime. We tested for a repetition effect in comparing phoneme-monitoring times for pictures preceded by their name versus an unrelated name.

Method

Participants

Fifty-two psychology students at the National University of Cordoba, Argentina, aged 18–26 years, participated voluntarily in the experiment. All were Argentinean native speakers of Spanish, with normal or corrected-to-normal vision and no known language disorder.

Materials

Twenty-nine black-on-white simple line drawings of common objects served as test picture stimuli. The drawings were selected from the set described in Cycowicz, Friedman, Rothstein, and Snodgrass (1997). Naming agreement for the selected experimental pictures was above 80% according to recently published Argentinean norms (Manoiloff, Artstein, Canavoso, Fernández, & Segui, 2010). The average frequency of the 29 test picture names was of 87 occurrences per million (opm) according to the LEXESP Spanish lexical database provided by Sebastián-Gallés, Martí, Cuetos, and Carreiras (2000). To check whether these lexical database frequencies computed for Spanish in Spain are relevant for the Argentinean subjects we tested, we collected subjective frequencies from 20 Argentinean native speakers of Spanish (none of them participated in Experiments 1–3), who rated the subjective frequency of the test picture names on a 1–5 scale (see Appendix). The overall correlation between the database and subjective frequencies was significantly positive, $r(27) = .398$, $p < .05$, but rather modest. We therefore retained the subjective frequencies to dispatch the materials evenly across sublists (see the *Design* section). Each test picture was paired with two possible word primes: one repetition word prime—the target picture's name—and one unrelated word prime (e.g., "luna", "moon", and "paso", "step", respectively, for the picture of the moon). For each picture, the unrelated and repetition word primes had approximately the same length and frequency. Ninety additional pictures, whose characteristics were similar to those of the test pictures, were chosen as filler pictures. These filler pictures were all paired with an unrelated word prime.

Design

The experimental phase consisted of 29 test and 90 filler trials, hence a total of 119 trials. The experimental phase was preceded by a training phase of six repetition and six unrelated trials. The prime–target experimental pairs were counterbalanced across the two priming conditions (repetition and unrelated) and two groups of participants, so that each participant was presented only once with any single picture. We therefore constructed one list of 14 and another list of 15 test pictures. One group of participants received the first list in the repetition condition and the second list in the unrelated priming condition. The other group received the same lists in the opposite conditions. Picture name frequency and length were balanced as well as possible across the two lists (see Appendix). The 90 filler word–picture pairs, all unrelated, were the same for both groups of participants. Each participant thus only received 14 or 15 repetition prime–target trials (out of 119). The 119 trials were blocked by target phoneme. There were six word-initial target phonemes (/p, t, k, m, l, r/),[1] hence six blocks of variable size, depending of the number of target pictures matching the target phoneme for the block: There were 7, 6, 5, 4, 5, and 2 target pictures whose name began with /p, t, k, m, l, r/, respectively (see Appendix). Each block contained about three times more filler than test trials and at least two test trials.

Procedure

Each trial consisted of the following sequence of four visual stimuli, appearing at the centre of a computer screen (60-Hz refresh rate), and immediately following each other: a 500-ms forward mask, a 29-ms word prime, a 14-ms backward mask, and a target picture, which was displayed until the participant had responded, or for a maximum two seconds (the time-out value for participants' responses). The next trial was launched one second after participant's response (or after response time-out). The word prime, forward mask, and backward mask were displayed as white characters on a black background, in 48-point Times New Roman font. The backward as well as forward masks were quasirandom strings of upper-case consonants (e.g., "ZRQWYSXTV").[2] The word primes were also presented in upper-case letters. The target pictures were drawn in black on white within a 7 × 7-cm white square. Participants were tested individually. They were instructed to orient their gaze to the middle of the forward mask at the beginning of each trial (i.e., at the centre of the screen, where the prime then the target would be displayed). They were told they would be presented with six series of object drawings, each associated with a target phoneme, and had to respond with a button press, as quickly and accurately as possible, whenever the name of the object was beginning with the target phoneme specified for the current series. Each series was therefore introduced with the oral specification of a target phoneme. For example, target /t/ was specified as follows: "*Pulsar el botón de respuesta, lo más rápido que pueda,* si y solo si *el nombre de la imagen comienza con el sonido* [te], *como en* 'templo', 'taxi' *o* 'túnel'." ["*Press the response button, as quickly as possible,* if and only if *the name of the picture begins with the sound* [te], *as in* 'temple', 'taxi', *or* 'tunnel'."]. Participants were also told not to respond when the name of the picture did not begin with the currently specified target sound. That printed word primes were briefly displayed was not revealed to the participants. The experiment was run on a personal computer, using the DMDX experimentation software (Forster & Forster, 2003). After the experiment was completed, participants were asked whether they noticed that a printed word briefly appeared before at least some pictures. Most of the participants reported that they did not notice such words. The data of four participants who reported having noticed word primes were not retained: This left 48 participants.

Results and discussion

We first computed the miss rates averaged by subjects across all the test picture items and by items across all the subjects. We discarded participants

[1]In Spanish, grapheme–phoneme correspondences are largely consistent. This was the case for the word-initial letter in the name of the target pictures we used, except "c", which is pronounced /θ/ (or /s/) before "i" or "e" (/i/ or /e/) and /k/ otherwise. In our material, "c" was only followed by either "a" or "u" (/a/ or /u/), hence was always pronounced /k/.

[2]Strings of consonants are believed to be more effective masks than, for example, strings of "#"s or of "%"s, as used in a number of previous studies (e.g., Ferrand & Grainger, 1992, 1994).

Table 1. *Mean reaction times and miss rate according to the repetition versus unrelated priming condition in Experiment 1*

	Priming condition		
Measure	*Repetition*	*Unrelated*	*Priming effect*
RT	823 (140)	870 (125)	+47
Miss rate	3.30 (5.3)	3.22 (4.8)	−0.08

Note: Reaction time (RT) in ms. Miss rate in percentages. Standard deviations in parentheses.

and picture items with more than 17% miss rate: five participants (three in a group, two in the other group) and two items (one in each list): "caja", "box", and "cuchillo", "knife". This left 43 participants and 27 picture items. For the reaction time (RT) data, we applied a 1500-ms cut-off value, thereby discarding 3.9% of the RT data. The miss rate data were tabulated from all the responses. The miss rate and RT data are shown in Table 1.

We conducted by-subject and by-item analyses of variance on the RT and miss rate data, with group (by-subject) or list (by-item) as a between factor and priming condition (repetition vs. unrelated) as a within factor. The structural factors group or list did not reach significance for either the RT or the miss rate data. We do not discuss them further. Priming had a significant effect in the RT data, with shorter RTs in the repetition than in the unrelated condition (823 < 870 ms) [$F_1(1, 41) = 19.80$, $p < .0001$; $F_2(1, 25) = 11.50$, $p < .005$]. Priming had no effect on the miss rate data, Fs < 1.

Our results thus show that word-initial phoneme detection from the presentation of a picture is facilitated by the picture's name, presented in print as a visual masked prime. If we use the "activation" metaphor, such facilitation suggests that the phonological code activated by the prime "preactivates" the phonological code that subjects retrieve from the picture's name to perform phoneme monitoring. That is, the facilitation observed in the repetition priming condition would be due to preactivation at the phonological level. However, before proposing this account of the repetition effect, it is necessary to examine an alternative interpretation. In Experiment 1, the first letter of the masked prime corresponded to the specified target phoneme in the repetition condition but never in the unrelated condition (see Appendix). For example, the repetition and unrelated primes for the picture of a moon (Spanish "luna") were "luna" and "paso" ("step"), respectively. Possibly then, the sole leftmost phoneme, not the entire phonemic code of the prime, might have been sufficient to trigger a (nonconscious) response preparation. By this account, the observed facilitation effect would not be a whole-word repetition effect, but a response preparation effect due to the congruence of prime and target as to the response to be made (cf. Dehaene et al., 1998). To test for this possibility, we conducted a second experiment, in which repetition and unrelated word primes shared their first letter (as in Ferrand et al., 1994). In both priming conditions, the leftmost letter of the prime corresponded to the target phoneme. With this design, an advantage of the repetition over the unrelated condition could only be interpreted as a genuine repetition effect rather than a preparation effect induced by the leftmost letter of the repetition prime.

EXPERIMENT 2

In this experiment, both unrelated or repetition primes shared their first letter with the associated target picture's name. If the priming effect found in Experiment 1 was simply due to prime word and target picture's name sharing their first letter (i.e., phoneme), unrelated and repetition primes should not differ in their priming effect.

Method

Participants

Twenty-eight psychology students at the National University of Cordoba, Argentina, aged 18–29 years, participated voluntarily in the experiment. All were Argentinean native speakers of Spanish, with normal or corrected-to-normal vision and no known language disorder. None of the subjects had participated in Experiment 1.

Materials

The target pictures were the same as those in Experiment 1. In the repetition condition, the associated primes were the same as those in Experiment 1; in the unrelated condition, the primes were replaced with words beginning with the same letter (hence, phoneme) as in the associated picture's name. For example, the two possible primes for the picture of the moon were "luna", "moon" (repetition) and "locro", a classic Argentinean dish (unrelated). As in Experiment 1, the unrelated primes had approximately the same length and frequency as the associated repetition primes.

Design and procedure

The design and the procedure were the same as those in Experiment 1. As in Experiment 1, participants were asked whether they noticed that printed words appeared before the pictures. The data of two participants who reported having noticed word primes were not retained: This left 26 participants.

Results

The miss rates were quite low in this experiment compared to Experiment 1: 1.87% in either the repetition or the unrelated priming condition. No participant and no item approached the 17% miss rate criterion used in Experiment 1, except the picture for "caja", "box", whose miss rate was 15.4%. We therefore retained the data of all the 26 participants and 29 items. As for Experiment 1, we applied a 1500-ms cut-off value on the RT data, thereby discarding 2.6% of the RT data. The miss rate data were, again, tabulated from all the responses. The resulting miss rate and RT data are shown in Table 2.

We conducted the same by-subject and by-item analyses of variance as for Experiment 1 on the miss rate and RT data. The structural factors group (by-subject analyses) or list (by-item analyses) did not reach significance, $F\text{s} < 1$, and are not discussed further. Priming again had a significant effect on the RT data, with shorter RTs in the repetition than in the unrelated condition (851 < 885 ms) [$F_1(1, 24) = 8.05$, $p < .01$; $F_2(1, 27) = 9.71$, $p < .005$]. Priming had no effect on the miss rate data, $F\text{s} < 1$.

Table 2. *Mean reaction times and miss rate according to the repetition versus unrelated priming condition in Experiment 2*

	Priming condition		
Measure	*Repetition*	*Unrelated*	*Priming effect*
RT	851 (110)	885 (106)	+35
Miss rate	1.87 (4.7)	1.87 (3.7)	0.0

Note: Reaction time (RT) in ms. Miss rate in percentages. Standard deviations in parentheses.

The repetition effect obtained in Experiment 2 was thus similar to the effect obtained in Experiment 1, though slightly smaller numerically (35 ms vs. 47 ms, respectively). We ran a by-subject analysis of the RT data across the two experiments, with Experiment (Experiment 1 vs. 2) as a between-subject factor and priming condition (repetition vs. unrelated) as a within-subject factor. Neither experiment nor the Experiment × Priming interaction reached significance, $F_1\text{s} < 1$. We also ran a by-item analysis restricted to the items common to the two experiments—that is, excluding the RT data for "caja" and "cuchillo" (see Experiment 1). In this analysis, both experiment and priming were within-item factors. Experiment did not reach significance, $F_2(1, 25) = 2.61$, $p = .12$, and did not interact with priming, $F_2 < 1$. Thus, the repetition priming effect did not significantly differ across experiments. In particular, the repetition effect found in Experiment 2 allows us to dismiss the interpretation of the repetition effect in Experiment 1 as simply due to prime word and target picture's name sharing their initial phoneme.

DISCUSSION OF EXPERIMENTS 1 AND 2

In Experiment 1, we found that the detection of the initial phoneme of the target picture's name was faster when the picture was preceded by the brief masked presentation of the picture's name than by that of an unrelated name with a different first letter/phoneme from the picture's name. Experiment 2 replicated this repetition effect with unrelated word primes beginning with the same letter/phoneme as the picture's name. The shorter phoneme monitoring times in the repetition than in the unrelated condition thus cannot be explained by the prime word and target picture's name sharing their initial letter. Therefore, a phonological priming account of the observed priming effects is quite plausible.

An alternative account of these effects at the level of conceptual representation seems quite unlikely. Although semantic priming effects have been found using subliminal picture primes (Dell'Acqua & Grainger, 1999; Van den Bussche, Notebaert, & Reynvoet, 2009), such effects are at least controversial when using subliminal *word* primes instead of picture primes (for contradictory results, see, for example, Chauncey, Holcomb, & Grainger, 2009; Finkbeiner & Caramazza, 2006; Quinn & Kinoshita, 2008; Spalek & Damian, 2013). We nevertheless conducted a semantic categorization experiment with picture targets and repetition versus unrelated word primes, using the same cross-form masked priming settings as those in Experiments 1–2. Participants had to categorize the target pictures into an artefactual or a natural category.[3] No sign of repetition priming was observed[4], thus supporting our hypothesis that the functional locus of the repetition priming effects obtained with the internal phoneme monitoring task is phonological rather than conceptual. We therefore assume that repetition priming facilitates internal phoneme monitoring through preactivation of a phonological code for the target picture's name, in line with Ferrand et al.'s (1994) conclusions for primed naming. We further assume, within the framework of Levelt's (1989; Levelt et al., 1999) model of speech production, that this phonological code receives preactivation at several levels, starting from the word-form initially retrieved at lexical access (the "lexeme"), down to the various levels of phonological code elaborated during phonological encoding towards an articulatory plan, including the critical level, which internal monitoring scans. The issue now arises of better qualifying this critical level of representation.

As noted previously, two main levels of representation have been proposed: a phonetic–articulatory level of coding, specifying actual articulation, and a more abstract phonological level. Internal phoneme monitoring should logically be based on a more abstract level of representation than a level closely linked to articulation since it does not engage "explicit" articulation processes. However, we cannot dismiss the possibility that the critical level for internal monitoring is articulatory. If this were the case, a concurrent task of articulatory suppression would seriously impede retrieval of the critical code that is monitored: It would be available, as it were, only during the brief pauses between consecutive words or syllables produced during "articulatory suppression", blurring out any systematic facilitation effect. Wheeldon and Levelt (1995, Experiment 1b), used articulatory suppression with a similar motivation to test the abstract versus articulatory nature of the code allowing internal phoneme monitoring in their Experiment 1a. In Experiment 3, we follow these authors and examine whether the repetition priming effect obtained in Experiments 1–2 is affected by articulatory suppression. The presence of a phonological priming effect in this experimental condition would confirm the prearticulatory, abstract nature of the phonological code that internal monitoring is based on in Experiments 1–2.

[3]We used 20 target pictures with frequent names for each category. Pictures were primed by their own name or by an unrelated name (sharing the first letter/phoneme of the picture's name). The participants ($N = 27$) were drawn from the same population as that in Experiments 1–2 but did not participate in those experiments.

[4]Mean response times for the repetition versus unrelated condition were 741 versus 734 ms (*ns*) for the artefactual category and 719 versus 709 ms (*ns*) for the natural category, respectively.

EXPERIMENT 3: REPLICATION OF EXPERIMENT 2 WITH ARTICULATORY SUPPRESSION

This experiment used the same materials and methods as those in Experiment 2 with the addition of an articulatory suppression task. Following the logic of Wheeldon and Levelt (1995), articulatory suppression should interfere with internal phoneme monitoring, were internal phoneme monitoring performed on the "phonetic code" (in Wheeldon & Levelt's, 1995, terminology) that serves as input to the articulatory programme constructed to name the picture. Absence of interference would support the claim that internal phoneme monitoring is performed on a more abstract phonological code rather than this "phonetic code".

Method

Participants

Thirty-nine psychology students at the National University of Cordoba, Argentina, aged 17–32 years (mean age 21.4 years), participated voluntarily in the experiment. All were Argentinean native speakers of Spanish, with normal or corrected-to-normal vision and no known language disorder. None of them had participated in Experiment 1 or 2.

Materials

The target pictures and the associated primes in the repetition and unrelated conditions were identical to those in Experiment 2.

Design and procedure

The design and the procedure were the same as those in Experiments 1–2, except that participants had to repeatedly say aloud the syllable /sa/ during each trial. This /s/-initial syllable was chosen so that it would minimally interfere with the six target phonemes /p, t, k, m, l, r/; /s/ indeed differs from all six target phonemes at least in manner of articulation and for all of them except /t/ in voicing and/or place of articulation. Participants were instructed to repeat /sa/ at a comfortable, yet not too slow, speaking rate (see Wheeldon & Levelt, 1995, p. 320). In each block (corresponding to a specified target-phoneme), after target-phoneme specification and before the first trial, participants practised repeating aloud /sa/ for about one minute on average, under the supervision of the experimenter. The experiment then proceeded trial by trial. Participants were instructed to repeat the syllable /sa/ throughout the entire current block from the first to the last trial (the end of the last trial was signalled by a message displayed on the screen). The data of two participants were not retained because they stopped repeating /sa/ during some trials. As in Experiments 1 and 2, participants were asked whether they noticed that printed words appeared before the pictures. The data of one participant who reported seeing the primes were not retained. This left 36 participants.

Results and discussion

We first applied the same criteria as those in Experiment 1 to discard participants and picture items with more than 17% miss rate. Three participants and three items ("caja", "cuchillo", and "media") were discarded by these criteria. This left 33 participants (17 in one group and 16 in the other) and 26 picture items (13 in each list). For the RT data, we applied a 1600-ms cut-off value, thereby discarding 2.9% of the RT data. This value was chosen, instead of the 1500-ms cut-off value in Experiments 1–2, because the mean (unfiltered) RT in Experiment 3 was about 100 ms longer than that in Experiments 1 or 2 (970 ms compared to 847 ms in Experiment 1 and 868 ms in Experiment 2). The miss rate data were tabulated from all the responses. The miss rate and RT data are shown in Table 3.

We conducted the same by-subject and by-item analyses of variance as for Experiments 1 and 2 on the data. The structural factors group (by-subject analyses) or list (by-item analyses) did not reach significance (Fs < 1 for list, ps $> .1$ for group) and are not discussed further. Priming had a significant effect in the RT data, with shorter RTs

Table 3. *Mean reaction times and miss rate according to the repetition versus unrelated priming condition in Experiment 3*

	Priming condition		
Measure	*Repetition*	*Unrelated*	*Priming effect*
RT	900 (116)	929 (119)	+29
Miss rate	5.4 (6.5)	3.7 (4.8)	−1.7

Note: Reaction time (RT) in ms. Miss rate in percentages. Standard deviations in parentheses.

in the repetition than in the unrelated condition (900 < 929 ms) [$F_1(1, 31) = 6.01$, $p < .05$; $F_2(1, 24) = 6.75$, $p < .05$]. Priming had no effect on the miss rate data [$F_1(1, 31) = 1.46$, $p = .24$; $F_2(1, 24) = 1.74$, $p = .20$].

The repetition effect obtained in Experiment 3 was thus quite similar to the effect obtained in Experiment 2, though slightly smaller numerically (29 ms vs. 35 ms). We ran a by-subject analysis of the RT data across the two experiments, with experiment (Experiment 2 vs. 3) as a between-subject factor and priming condition (repetition vs. unrelated) as a within-subject factor. Experiment was not significant, $F_1(1, 55) = 2.64$, $p = .11$, and did not interact with priming, $F_1 < 1$. We also ran a by-item analysis restricted to the items common to Experiments 2 and 3—that is, excluding the data for "caja", "media", and "cuchillo". In this analysis, both experiment and priming were within-item factors. Experiment was found to be significant, $F_2(1, 24) = 41.02$, $p < .0001$, reflecting longer RTs overall in Experiment 3 than in Experiment 2 (914 vs. 868 ms). More importantly, the Experiment × Priming interaction was far from significance, $F_2 < 1$, showing that the repetition priming effect did not significantly differ across experiments.

To summarize, although the additional task of repeating a syllable entailed an increased difficulty at performing the internal phoneme monitoring task, as suggested by the longer RTs as well as the higher error rates in Experiment 3 than in Experiment 2 (see Wheeldon & Levelt, 1995, for a similar pattern), articulatory suppression did not affect the repetition effect found in Experiments 1–2. This result supports our interpretation that the representations monitored for phoneme detection, whose preactivation by repetition priming facilitates phoneme detection, are more abstract than those in the "phonetic plan" (see Levelt, 1989) that feeds articulation. We return to this interpretation in the General Discussion.

GENERAL DISCUSSION

In this study, we asked whether phoneme monitoring on internally generated picture names shows the same sensitivity to repetition priming as the direct task of naming pictures. Experiments 1 and 2 provide a positive answer to this question: Monitoring for word-initial /p/ from the picture of a door (Spanish "puerta") was faster when the picture was preceded by the printed word "puerta" than "rosa", "rose" (Experiment 1) as well as "postre", "dessert" (Experiment 2). We thus obtained a repetition priming effect similar indeed to the effect obtained in Ferrand et al. (1994), with the same cross-form masked priming format of presentation but with an internal phoneme monitoring task. We concluded that the priming effects we obtained must reflect preactivation by the prime of a phonological code of the target picture's name, given that the effects were not due to prime word and target picture's name sharing their initial letter (Experiment 2), and probably not due, either, to semantic priming (as suggested by the results of an additional test of semantic categorization). We thus successfully replicated the main result of Ferrand et al. (1994) using internal phoneme monitoring instead of picture naming.

Experiment 3 explored further the nature of the critical phonological code in these experiments. Because repetition priming effects were not affected by a concurrent task of articulatory suppression, the phonological code that the internal monitor scans to detect phonemes, and which we assume is preactivated by repetition primes, cannot be the articulatory or motor code sent to the articulation system: It must be a more abstract code. We thus reach a conclusion similar to that drawn by Wheeldon and Levelt (1995) from their Experiments 1a–1b. We now return to the Wheeldon and Levelt study within the framework of Levelt's model of speech production, trying to provide an account of the internal monitoring issue and to qualify our results more accurately.

The main motivation of the Wheeldon and Levelt (1995) study was testing the hypothesis that phonological encoding proceeds incrementally from left to right. As it turned out, this also led to an interesting speculation about internal monitoring. Left-to-right encoding may seem a logical constraint for producing speech smoothly without waiting for the planning of the current utterance to be completed. Experimental data from the preparation paradigm (Meyer, 1990, 1991) also suggest that phonological encoding is incremental. Wheeldon and Levelt (1995) reasoned that left-to-right elaboration of the speech code for production would entail left-to-right availability to an "internal monitor" of the successive segments to be produced. Wheeldon and Levelt (1995; also see Morgan & Wheeldon, 2003; Wheeldon & Morgan, 2002) indeed showed that internal phoneme monitoring latencies varied as a function of within-word syllable position and within-syllable phoneme position: Target phonemes were in particular monitored faster in the first than in the second syllable onset position. The first syllable advantage survived articulatory suppression, precluding the possibility that subjects monitored a subvocal articulation of the carrier target word. Wheeldon and Levelt (1995) concluded that internal monitoring operates on a prearticulatory phonological representation specified for syllable structure.

But what exactly could be an articulatory representation with which articulatory suppression can interfere? It must be the "internal speech" of Levelt's model. In the original version of Levelt's model of speech production (Levelt, 1989), "internal speech" is identified as corresponding to a "phonetic code", which is the input to the articulation mechanisms. This phonetic code or internal speech thus constitutes the material that would be stored temporarily in an "articulatory buffer" before it is sent to articulation. In other words, it constitutes the ultimate output of the "phonological encoding" stage in Levelt's model of speech production. The articulatory buffer metaphor, although perhaps simplistic, is quite handy to understand how articulatory suppression can interfere with internal phoneme monitoring. Repeating the syllable /sa/, as in our Experiment 3, would supposedly fill the articulatory buffer, perhaps only partially, and render the monitoring of target pictures' names more difficult, and possible only at random moments between successive /sa/s. Therefore, any systematic pattern of monitoring times based on the code in the articulatory buffer would probably be blurred by articulatory suppression. Wheeldon and Levelt (1995) found that articulatory suppression left intact the pattern of increasing monitoring times with left-to-right phoneme position. They concluded that internal phoneme monitoring was scanning a more abstract code than the "phonetic code". Likewise, in our Experiment 3, articulatory suppression does not alter the repetition priming effect found in Experiments 1–2. We therefore conclude that the code on which phoneme monitoring is based, and which is preactivated by repetition primes, must be prearticulatory and more abstract than the phonetic code.

Because Wheeldon and Levelt's (1995) found clear syllable rank effects, they proposed that internal phoneme monitoring operates on an abstract phonological code in which segments have been assigned to syllabic frames. More recently, Schiller, Jansma, Peters, and Levelt (2006) obtained left-to-right effects in internal monitoring for stressed syllables in a given position. They proposed that internal monitoring in general applies to a syllabified and prosodified phonological code. This seems to set the critical level of phonological encoding, appropriate for monitoring phonological properties in internally generated speech, at a rather high level of elaboration within phonological encoding. However, the real difficulty for qualifying the phonological representations on which internal monitoring operates does not lie so much on the level of elaboration as on the functional locus within the speech production system at which they can be monitored. Schiller et al. (2006) have suggested two kinds of internal monitor: a "production monitor", which scans directly phonological encoding as it builds up, "moving in parallel with phonological word encoding" (Schiller et al., 2006, p. 135), and a "perception monitor", which scans the end-result of phonological encoding. This may sound a bit vague. In our understanding, the perception monitor could rather be the "language comprehension system" in Levelt's speech production model: This system, in Levelt's view, can respond to (or monitor for) either "internal speech" (the output of phonological encoding) via an "inner loop" or to external, overt speech via an "outer loop". In other words, the same device would be able to monitor for either overt or internally generated speech. This would be possible because for both overt and internal speech; the information to be processed is made available sequentially. A strong argument supporting the view of a speech perception system monitoring internal speech just like it was external, overt speech is offered by Özdemir, Roelofs, and Levelt (2007) who obtained uniqueness point effects in internal phoneme monitoring—that is, effects logically linked to lexical access from external speech (also see Morgan & Wheeldon, 2003, and Wheeldon & Morgan, 2002, for more mitigated results).

These results of course need to be replicated in some ways. The general idea to be tested is that the effects typically found with overt speech are also found with internally generated speech. Several robust effects found in perception, other than uniqueness point effects, could be tested: effects of syllable structure complexity (Segui, Dupoux, & Mehler, 1990); advantage for syllable over phoneme detection (syllables are faster to detect than their phoneme constituents; Segui, Frauenfelder, & Mehler, 1981); and so on. For the latter effect, a pilot experiment we ran recently suggests that the advantage for syllable over phoneme also holds when monitoring for internally generated speech.

REFERENCES

Braun, M., Hutzler, F., Ziegler, J., Dambacher, M., & Jacobs, A. (2009). Pseudohomophone effects provides evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, *30*, 1977–1989. doi:10.1002/hbm.20643

Chauncey, K., Holcomb, P., & Grainger, J. (2009). Primed picture naming within and across languages: An ERP investigation. *Cognitive, Affective, and Behavioral Neuroscience*, *9*, 286–303. doi:10.3758/CABN.9.3.286

Cycowicz, Y., Friedman, D., Rothstein, M., & Snodgrass, J. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*, 171–237. doi:10.1006/jecp.1996.2356

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., … Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597–600. doi:10.1038/26967

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321. doi:http://dx.doi.org/10.1037/0033-295X.93.3.283

Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition*, *73*, B1–B15. doi:10.1016/S0010-0277(99)00049-9

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology*, *45A*, 353–372. doi:10.1080/02724989208250619

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked from priming. *The Quarterly Journal of Experimental Psychology*, *47*, 365–382. doi:10.1080/14640749408401116

Ferrand, L., Grainger, J., & Segui, J. (1994). A study of masked form priming in picture and word naming. *Memory and Cognition*, *22*, 431–441. doi:10.3758/BF03200868

Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex*, *42*, 790–796. doi:10.1016/S0010-9452(08)70419-2

Forster, K., & Forster, J. (2003). DMDX: A windows display program with millisecond accuracy. *Behaviour Research Methods, Instruments, and Computers*, *35*, 116–124. doi:10.3758/BF03195503

Garrett, M. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–177). New York: Academic Press.

Hutson, J., Damian, M., & Spalek, K. (2013). Distractor frequency effects in picture word-interference tasks with vocal and manual responses. *Language and Cognitive Processes*, *28*, 615–632. doi:http://dx.doi.org/10.1080/01690965.2011.605599

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 857–875. doi:10.1098/rstb.2007.2093

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences*, *22*, 1–75. doi:10.1017/S0140525X99001776

Manoiloff, L., Arstein, M., Canavoso, M., Fernandez, L., & Segui, J. (2010). Expanded norms for 400 experimental pictures in an Argentinean Spanish-speaking population. *Behavior Research Methods*, *42* (2), 452–460. doi:10.3758/BRM.42.2.452

Manoiloff, L., Segui, J., & Hallé, P. (2013). L'effet de fréquence dans l'accès aux propriétés phonologiques des noms d'objets. *L'Année Psychologique*, *113*, 335–348. doi:http://dx.doi.org/10.4074/S0003503313003023

Meyer, A. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524–545. doi:10.1016/0749-596X(90)90050-A

Meyer, A. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, *30*, 69–89. doi:10.1016/0749-596X(91)90011-8

Morgan, J., & Wheeldon, L. (2003). Syllable monitoring in internally and externally generated English words. *Journal of Psycholinguistic Research*, *32*, 269–296. doi:10.1023/A:1023591518131

Oldfield, R., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273–281. doi:10.1080/17470216508416445

Özdemir, R., Roelofs, A., & Levelt, W. J. M. (2007). Perceptual uniqueness point effects in monitoring internal speech. *Cognition*, *105*, 457–465. doi:10.1016/j.cognition.2006.10.006

Quinn, W., & Kinoshita, S. (2008). Congruence effect in semantic categorization with masked primes with narrow and broad categories. *Journal of Memory and Language*, *58*, 286–306. doi:10.1016/j.jml.2007.03.004

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, *53*, 97–145. doi:10.1016/j.cogpsych.2006.01.002

Schiller, N., Jansma, B., Peters, J., & Levelt, W. (2006). Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes*, *21*, 112–140. doi:10.1080/01690960400001861

Sebastián-Gallés, N., Marti, M., Cuetos, F., & Carreiras, M. (2000). *Lexesp: una base de datos informatizada del español*. Barcelona: Universidad de Barcelona.

Segui, J., Dupoux, E., & Mehler, J. (1990). The role of the syllable in speech segmentation, phoneme identification, and lexical access. In G. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 263–280). Cambridge, MA: MIT Press.

Segui, J., Frauenfelder, U., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, *72*, 471–477. doi:10.1111/j.2044-8295.1981.tb01776.x

Spalek, K., & Damian, M. (2013, July). *Picture-word interference with masked and visible distractors: Different types of semantic relatedness inhibit lexical selection*. Presentation at the Thirty-Fifth Annual Conference of the Cognitive Science Society, Berlin.

Van den Bussche, E., Notebaert, K., & Reynvoet, B. (2009). Masked primes can be genuinely semantically processed: A picture prime study. *Experimental psychology*, *56*, 295. doi:10.1027/1618-3169.56.5.295.

Wheeldon, L., & Levelt, W. (1995). Monitoring the time-course of phonological encoding. *Journal of Memory and Language*, *34*, 311–334. doi:10.1006/jmla.1995.1014

Wheeldon, L., & Morgan, J. (2002). Phoneme monitoring in internal and external speech. *Language and Cognitive Processes*, *17*, 503–535. doi:10.1080/01690960143000308

APPENDIX

Materials of Experiments 1–2

Target pictures' names with LEXESP frequencies and frequency ratings. Differences between Lists 1 and 2 are all nonsignificant ($ts < 1$).

Picture's name	*No. characters*	*First consonant*	*LEXESP frequency*	*Frequency rating*
List 1				
cama	4	/k/	136.4	4.30
casa	4	/k/	629.8	4.57
cuchillo	8	/k/	15.4	4.33
limón	5	/l/	6.6	3.13
luna	4	/l/	52.5	3.40
media	5	/m/	162.0	3.23
mesa	4	/m/	172.1	4.40
pan	3	/p/	54.6	3.70
pantalón	8	/p/	18.2	4.80
papa	4	/p/	35.9	3.44
puerta	6	/p/	278.0	4.40
regla	5	/r/	26.2	3.30
taza	4	/t/	16.8	4.10
televisor	9	/t/	20.5	3.22
tenedor	7	/t/	3.7	4.10
Mean	5.3		108.6	3.89
(*SD*)	(1.8)		(164.9)	(0.57)
List 2				
caja	4	/k/	45.0	3.50
cuchara	7	/k/	3.7	4.10
libro	5	/l/	194.3	4.60
lápiz	5	/l/	7.0	4.30
lapicera	8	/l/	0.0	3.52
mano	4	/m/	387.1	4.20
manzana	7	/m/	11.1	3.00
peine	5	/p/	5.0	4.20
perro	5	/p/	60.5	4.40
pierna	6	/p/	24.6	3.33
reloj	5	/r/	50.7	4.20
tomate	6	/t/	6.8	3.03
teléfono	8	/t/	80.2	4.70
torta	5	/t/	3.0	3.03
Mean	5.7		62.8	3.87
(*SD*)	(1.3)		(106.7)	(0.61)