The broth may *not* be in my brother's brothel: a meta-analysis of visual morphological decomposition

Roberto Petrosino[1], Diogo Almeida[1]

[1]New York University Abu Dhabi

1 Introduction

One of the pivotal findings in visual word recognition dates to back to the early 2000s, when both truly bimorphemic (e.g., *driver*) and opaquely bimorphemic, but actually monomorphemic words (e.g., *brother*) were first shown to facilitate the recognition of the contained stem target word (*DRIVE* and *BROTH*, respectively) when being briefly presented as primes [i.e., for no more than 60 ms; Rastle et al. (2000), Rastle, Davis, and New (2004)]. That these effects are independent of semantic and orthographic influence is commonly agreed upon on the basis of two independent, though complementary pieces of evidence. First, both semantically-related pairs [e.g., *cello-VIOLIN*: Rastle et al. (2000); but see *Feldman literature* for controversial results] and orthographically-related pairs (e.g., *typhoid-TYPHOON*: Rastle et al. 2000; *brothel-BROTH*: Rastle, Davis, and New 2004) usually show numerically small (< 5 ms) to null, and statistically non-significant, masked priming effects. Second, the size of the masked priming response to morphologically transparent pairs (i.e., *driver* is legitimately drived from the word *drive*, since a driver is someone who drives) and morphologically opaque pairs (i.e., *brother* may not possibly derived from the word *broth*, as a brother is not someone who broths) was found to be numerically and statistically similar Rastle and Davis (2008). Taken together, this evidence suggests that masked priming is exclusively sensitive to letter clusters that correspond to orthographic instantiations of abstract morphemes, but crucially may not actually be as such: e.g., the letter clusters $\$drive\$$, $\$broth\$$ are orthographic units corresponding to the actual lexical units associated to the meaning of driving and broth; but $\$er\$$ may or may not correspond to the actual lexical unit associated with agency, with no reliable way to predict it before knowing the meaning of the whole word.[1] For this reason, such early, semantically-blind effects are usually explained as triggered by early activation of "morpho-orthographic units", i.e. purely orthographic

[1]In this paper, we use dollar signs \$ flaking orthographic units to distinguish their orthographic nature from the lexical, more abstract nature of the corresponding morphemes.

(thus, devoid of any semantic meaning) letter clusters that are found occurring together statistically more often than not. Formation of such units – also called “islands of regularity” (Rastle and Davis 2008) – had been argued to be due to placement of a morpheme boundary occurring whenever a bigram has a lower transitional probability than the flanking bigrams (the “trough pattern”: among others, Seidenberg 1987). For example, in words like *driver, brother*, the bigram cluster such as $\$er\$$ has a higher transition probability (TP) as compared to the linearly preceding bigrams $\$ve\$$, $\$he\$$, respectively. Thus, a morpheme boundary is placed between the monograms involving the latter bigrams: $\$driv\|er\$$, $\$broth\|er\$$. Automatic identification of morpho-orthographic units within words is commonly known as *morphological decomposition*, and believed to be responsible for the dissociation between the morphologically transparent and opaque masked priming response on the one hand, and the semantically or orthographically masked priming response on the other.

Morphological masked priming effects have consistently reported across languages (*English, French, Italian, etc.: add more languages and references*) and modalities (*visual, auditory, and visual/auditory: add references*), and, as such, are considered to be uncontroversial evidence for the separation between morphological and non-morphological effects at early stages of visual word processing. However, we believe that some crucial results reported in the literature have been overlooked, and would actually call for a thorough reconsideration. Here, we mention two important contributions in this sense. Morris et al. (2011) found similar masked priming effects in response to bare target words (*FLEX*) preceded by either an actually derived prime word (*flexible*), a prime pseudo-word consisting of the corresponding monomorphemic target suffixed with a real, but syntactically illicit morpheme (*flexity*), or a prime pseudo-word consisting of the corresponding monomorphemic target suffixed with a phonologically licit, non-suffixal ending (*flexire*). Similar results are also reported in compound priming study by Fiorentino et al. (2015). In this study, similar masked priming effects were found for pairs involving compounds made of two extant English words (e.g, *drugrack-RACK*) and compounds made of an English pseudo-word and an extant English word (e.g., *slegrack-RACK*). These two sets of results are admittedly at stake with the results reported above. On the one hand, Rastle, Davis, and New (2004) morphologically transparent and opaque words such as *driver, brother* decompose because they are fully decomposable: $\$drive\$$, $\$broth\$$, $\$er\$$ are all extant morphemes of English; monomorphemic words such as *brothel* do not decompose instead because they are not fully decomposable: *broth* is a morpheme, but $\$el\$$ is not. We will call this pattern of results “brothel non-effect,” and seems to suggest that decomposition is contingent on the constraint that all the putative constituent of a word must be extant morphemes of the language. On the other hand, Morris et al. (2011) and Fiorentino et al. (2015) report that both pseudo-suffixed pseudo-words (*flexity, flexire*) and pseudo-compounds (*drugrack, slegrack*) prime the contained stem target (*FLEX, RACK*). We conflate these results and label them as the “slegrack/flexire effect,” and seem to suggest that decomposition does occur for pseudo-words, thus circumventing the constraint just mentioned. The most obvious way to explain the two effects would be to make reference to the different lexicality status of the prime stimuli tested: *brothel* is a word and does not prime *BROTH*, since $\$el\$$ is not a morpheme; *slegrack, flexire* are pseudo-words, and they prime *RACK, FLEX* respectively, regardless of the fact that $\$sleg\$$, $\$ire\$$ are not morphemes. However, accessing the lexicality status of a given visual stimulus essentially entails contacting the lexicon, which is usually argued not to impinge on early stages of word processing in the light of the absence of semantic masked priming.

In this paper, we try to solve the impasse in two ways. First, we entertain the hypothesis that visual decomposition relies on a phono-orthographic procedure of syllabification, which breaks down the visual stimulus into syllable-like chunks (similar to Taft 2003's BOSS units). Occurring at early stages, this procedure is rather crude and clusters together letters from left to right. This procedure first identifies an orthographic nucleus (that is, a vocoid string) and groups with it an orthographic coda (that is, a consonantal cluster string) according to the phono-orthographic syllabic constraints of the language. Under this hypothesis, the *brothel* and the *slegrack/flexire* effects are explained as follows. On one hand, the slegrack/flexire effect is a direct result of the syllabification procedure just described. Stimuli such as *slegrack, flexire* are loosely syllabified as $\$sleg.rack\$$, $\$flex.ire\$$ (where the dot "." signals a loose syllabic boundary). These syllabic chunks then activate the lexical entries {rack} and {flex}, which in turn trigger priming onto the target words *RACK, FLEX*. On the other hand, the brothel non-effect is a confound resulting from the diverse syllabic nature of the word pairs tested in the same condition. In the literature on morphological priming, the brothel non-effect has been reported in orthographic conditions in which the primes consisted of the primed root followed by a syllabic (i.e., *brothel-BROTH*) or by a consonantal (*against-AGAIN*) non-suffixal ending (as also suggested by Taft and Nguyen-Hoan 2010). Thus, under the syllabification hypothesis we are entertaining here, a word with a syllabic, non-suffixal ending like brothel is loosely syllabified as $\$broth.el\$$; the syllabic chunk $\$broth\$$ activates the lexical entry {broth}, which in turn triggers priming onto the target word (*BROTH*). Conversely, a word with a non-syllabic, non-suffixal ending like against is loosely syllabified as $\$against\$$, which activates the lexical entry {against}, thus inhibiting/suppressing priming onto the target word (*AGAIN*). When averaged together as part of the same condition, the opposite priming magnitudes cancel each other out, thus leading to the brothel non-effect. The syllabification driven decomposition procedure seems like a promising solution for the impasse we are in, since (i) it would explain the difference between the brothel non-effect and the slegrack/flexire effect by providing (ii) a reasonable mechanistic explanation that is relatively easy to implement in models of lexical access. To test this hypothesis, we elicited priming effects to the transparent and opaque conditions, akin to Rastle, Davis, and New (2004); additionally, we tested priming effects to two different orthographic conditions: a syllabic condition, in which primes were words made of a monosyllabic real word and a syllabic non-suffixal ending (e.g., *can.vas-CAN*), and a non-syllabic condition, in which the primes were words made of a real word and a consonantal non-suffixal ending (e.g., *starch-STAR*).

Second, we entertained the hypothesis that the results reported in the studies mentioned above may suffer from low statistical power, as it has been claimed for most published research in the field (Brysbaert and Stevens 2018). As a way to circumvent this problem, we ran four different online replications of the same experiment with increasing sample size ($140 \geq N \leq 400$). Two different stimulus delivery web-apps were used: *PsychoJS* (J. Peirce et al. 2019) and *Labvanced* (Finger et al. 2017). Even though online data collection has already been shown to elicit a comparable masked priming response to in-lab data collection (Angele et al. 2022; Petrosino, Sprouse, and Almeida submitted), we additionally ran a fifth replication of the experiment in a typical in-lab environment with a greater than normal sample size ($N = 84$). We finally ran a meta-analysis over the whole dataset to estimate effect sizes for all conditions tested across the five replications, and directly assess replicability of the brothel non-effect as reported in the literature . Our results seem to contrast with the commonly-accepted divide between the morphologically transparent and opaque priming on the one side, and orthographic and semantic priming on the other. Rather, they offer a more complex pattern, in which the effects for morphological opacity are more similar to the effects for orthographic relatedness than to the effects for morphological transparency. We discuss the implications of these results and offer a new model of word processing that is able to explain the results, while admittedly being at odds with the traditional view of masked priming and lexical access.

2 Methods

2.1 Materials

Similarly to Rastle, Davis, and New (2004), we elicited the masked priming response to morphologically transparent (e.g., *driver-DRIVE*) and morphologically opaque (*brother-BROTH*). All replications reported here present two additions: an identity condition to compare the morphological priming response to the identity priming response; and the two separate orthographic conditions to catch potential differences in the masked orthographic priming response triggered by the hypothesized syllabification mechanism described above. One-hundred and sixty pairs were selected from the English Lexicon Project corpus (ELP: Balota et al. 2007), 32 in each of the following five conditions.

1. The *identity* condition included monomorphemic words that were presented both as prime and target (e.g., *fuss-FUSS*). Although not discussed in the literature review above, the identity condition represents the upper bound reference for any masked priming experiment, as it is expected to trigger priming at its maximal size (see Forster, Mohan, and Hector 2003 for a review).
2. The *transparent* condition included pairs involving a semantically and morphologically transparent relationship between a bimorphemic prime word and their corresponding target stem (e.g., *boneless-BONE*).
3. The *opaque* condition included pairs involving an apparent morphological relationship between a seemingly bimorphemic, but actually monomorphemic prime word and the therein-contained target stem (e.g., *belly-BELL*).

In lieu of Rastle, Davis, and New (2004)'s orthographic condition (*brothel-BROTH*), we constructed two separate conditions:

4. The *syllabic* condition consisted of prime words (e.g., *banjo*) that were made of the corresponding target word (*BAN*) plus an additional syllabic pseudo-suffix (*-jo*).
5. The *non-syllabic* condition consisted of prime words (*starch*) that were made of the corresponding target word (*STAR*) and an additional consonantal, non-syllabic pseudo-suffix (*-ch*).

In compiling the lists of these two conditions, we were particularly careful so that both of them adhered to the constraints listed below. Since syllabification (namely, the process whereby words are supposedly chunked into syllabic units) is part of phonology, these constraints guaranteed that prime and target words consistently carried the same orthographic and phonological information, without mismatches (e.g., different spellings in homophonic pairs) or conflicts (e.g., different pronunciations in homographic pairs).

i. the target word was phonologically and orthographically identical to the left most portion of the prime (e.g., *heaven-HEAVE* was excluded because of the root vowel change)
ii. the target word did not contain extra letters with respect to the corresponding prime word (e.g., *sort-SORE* was excluded because of the silent *e* present in the target word but absent in the prime);
iii. the prime word could not possibly be segmented differently from the way the target word is extracted (e.g., *restore-REST* was excluded since restore may be possibly paired up with *REST, STORE*, or *TORE* as target words);
iv. the target word was not an abbreviation or a proper name (e.g., *caroline-CAROL* was excluded because both prime and target word words can be proper names);
v. the pseudo-suffix in the prime was not a real suffix from either the orthographic and phonological point of view (e.g., *covenant-COVEN* was excluded since covenant may possibly segmented as affixed with the *-ant* suffix; similarly, *brothel-BROTH* was excluded since the pseudo-suffix *-el* in brothel is phonologically identical to the suffix *-al*);
vi. prime and target word pairs were not morphologically, semantically and/or (pseudo-)etymologically related (e.g., _arcade-A_RC was excluded);
vii. the pseudo-suffix did not consist of more than one syllable (e.g., *armada-ARM* was excluded since armada is trisyllabic);
viii. the target word was not an inflected word form (e.g., *rancho-RAN* was excluded since ran is the past tense form of the verb *run*);
ix. in the syllabic condition, the pseudo-suffix was always a legitimate syllable (e.g, *galaxy-GALA* was excluded since [ksi] is not a legitimate syllabic unit in English).

To ensure best comparability across conditions, none of the primes in the transparent, opaque, or syllabic conditions were monosyllabic bimorphemic words (e.g., a word like *dipped* was excluded); each of these conditions had only bimorphemic prime word with more than one syllable.

Primes and targets across the five conditions were matched separately as closely as possible for frequency (HAL) and length. The mean frequency and length values across all conditions are shown in Table 1 below. Because of the substantial differences between the non-syllabic condition and the other conditions, we were unable to control the prime words for orthographic length. The non-syllabic condition was included in the experiment regardless, since the primary goal of the experiment was to trigger priming *within* each condition and across the five replications of the same experiment. The stimuli of the identity condition were matched in length and frequency only with the targets of the other five conditions. Thirty-two unrelated prime words were selected for all target words; these words were orthographically, morphologically and semantically unrelated to the corresponding targets and were matched as closely as possible on frequency and length (HAL: $t(159)$=-0.16, p=.87; length: $t(159)$=0.03, p=.97). Thirty-two pairs of unrelated words were added to the final item set to reduce the prime-TARGET relatedness proportion to 40%. The filler targets were matched on frequency ($F(5,184)$=1.82, p=.1) and length ($F(5, 186)$=1.07, p=.37) to the targets of the experimental conditions; they were preceded by unrelated suffixed word primes.

	property	condition: transparent	opaque	syllabic	non-syllabic	identity	statistics
prime	Log HAL	7.25 (0.88)	7.83 (1.73)	7.17 (1.64)	7.19 (1.64)	NA	F(3, 122)=1.382, p=0.2
	length	6.69 (1.03)	6.19 (1.15)	6.8 (1.56)	4.78 (0.79)	NA	F(3, 122)=20.242, p=1.
target	Log HAL	9.39 (0.73)	8.79 (1.3)	9.75 (1.71)	9.12 (1.95)	9.22 (1.05)	F(4, 154)=1.992, p=0.0
	length	3.56 (0.67)	3.81 (0.79)	3.47 (0.95)	3.41 (0.5)	3.5 (0.67)	F(4, 154)=1.399, p=0.2

Table 1: Lexical properties of the main five conditions used in the experiment across the five replications. For each property, condition, and item type, the mean value is reported together with the relative standard deviation (in parenthesis).

An additional set of 192 pseudo-word targets were chosen from English Lexicon Project corpus. All pseudo-word targets complied with English spelling and phonotactics, and were generated by changing one or two letters in a corresponding target word. All pseudo-word targets were matched to the word targets used in the experiment in length ($F(6, 377)$=.9, p=.49).

As a way to further compare the masked priming response across conditions, in the second Labvanced replication we removed the unrelated word condition and added two more related word conditions (for a total of 7 conditions): an orthographic condition, in which prime and target words were neighbors (i.e., they differed in exactly one letter: *basil-BASIN*); and a semantic condition, in which prime and target words were not orthographically, but only semantically related (e.g., *poem-RHYME*). These two additional conditions were treated as a set of conditions separate from the others, and therefore did not share the same length and frequency with the other conditons (see Table 2 below for further information about these two additional conditions). As such, an additional set of 64 non-words and unrelated words were selected with the same length, similarly to above. All word items used in our replications may be found in the appendix below.

	property	condition: orthographic	semantic	statistics
prime	Log HAL	7.92 (0.6)	7.71 (0.48)	F(1, 62)=2.365, p=0.129
	length	5.22 (0.49)	5.09 (0.96)	F(1, 62)=0.428, p=0.515
target	Log HAL	8.13 (0.56)	8.18 (0.58)	F(1, 62)=0.112, p=0.739
	length	5.22 (0.49)	4.91 (1.15)	F(1, 62)=2.01, p=0.161

Table 2: Lexical properties of the additional conditions tested in the fifth replication of the experiment. For each property, condition, and item type, the mean value is reported together with the relative standard deviation (in parenthesis).

2.2 Procedure

For the in-lab replication, eighty-four native English speaking students of the University of Connecticut participated in the experiment (68 females; mean age: 19.25, s.d.: 0.82), none reporting atypical vision and/or other cognitive impairment (e.g., dyslexia). For each of the online replications on PsychoJS, one hundred and forty native speakers (*ms version*: 55 females, mean age: 36.43, s.d.: 11.07; *frame version*: 49 females, mean age: 35.06, s.d.: 11.51) of English recruited from Amazon Mechanical Turk (MT). English nativeness was assessed by checking their local IP address and time zone. For the two online replications on Labvanced, three-hundred and twenty-five (144 females; mean age: 40.28, s.d.: 12.25) and four hundred (200 females; mean age: 40.02, s.d.: 13.83) participants were recruited on Prolific, respectively. For these batches, we took advanced of a few criteria Prolific offers to ensure that all participants were actual native speakers of English. All participants had to be born in the Unites States of America, speaking English as their first and only language, and have no language-related disorder. In all online batches, participants were also asked to response to two open-ended questions (e.g., "*Your best friend is getting married in one week. You just saw your friend's partner with someone else at a restaurant and you suspect he/she is having an affair. What would you do?*") interspersed through the experiment during breaks as a way to further assess their nativeness and to screen for potential bots.

In the in-lab replication, stimulus presentation and data recording were performed with a python script run through the PsychoPy software (J. W. Peirce 2009). In the two PsychoJS versions, stimulus presentation and data recording were performed with a python script run through the PsychoJS program (J. Peirce et al. 2019). In the two Labvanced versions, stimulus presentation and data recording were performed through the Labvanced web-app (Finger et al. 2017). In all replications, prime words were presented in lowercase and preceded by a 500ms-long forward mask (#######) and immediately followed by a target word in uppercase, which remained on the screen until a response was made. The order of pairs was chosen randomly across participants. The replications varied in the way the prime duration was set up. In the in-lab replication, the prime duration was set at 2 frames (which is roughly 34 ms on our 60Hz monitor). To monitor prime duration, the code also reported prime duration both in frames and in milliseconds for each trial 34 ms (*ms version*) or 2 frames (*frames version*), respectively. In the two Labvanced replications, the prime duration was set at 28 ms to ensure primes to actually last 34 ms, as suggested by the developers of the program.

Each participant was assigned one between two different word lists, which differed only in the relatedness of the prime with respect to the target; other than that, the two lists presented the same set of target words and non-words. In one list, teh word conditions had 16 target items being preceded by themselves (the *related* primetype condition) and the remaining 16 target items being preceded by one of the unrelated primes belonging to the same frequency bin (the *unrelated* primetype condition). In the other list, the order was reversed. Participants were asked to perform a lexical decision task by pressing either the ‘J’ (for word) or ‘F’ (for non-word) keys on their keyboard. Participants were given 10 practice pairs before the actual experiment began. A total of 384 pairs were presented to each participant during the actual experiment. Participants were also given the possibility to take a few breaks throughout the experiment to avoid fatigue.

In all replications, we encouraged subjects to avoid any sort of distraction throughout the experiment. In particular, in the online replications, we asked participants to close any program that may be running in the background during the experiment, as a way to boost a decent performance of the stimulus presentation tool as much as possible. Besides, subjects could not be monitored in any way during data collection. Finally, to reduce the variability across devices, we restricted the experiment to be run on the Chrome browser only, since other browsers appears to be perform worse in previous pilot runs (possibly as a byproduct of the different engines used).

All replications consisted of a total of 384 items, except for the second Labvanced replication, which consisted of 512 items (because of the addition of two more conditions). The median time to finish the experiment was between 12 and 15 minutes. For the in-lab replication, participants were compensated in the form of course credit; for the online replications, participants were paid with the standard rate of $15/hour. A summary of the five replications are reported in Table 3. For the sake of clarity, in the remainder of the paper, we will refer to each dataset with in the corresponding number assigned in the column # in the table below.

#	setting	prime duration setting	platform	recruitment pool	no. of conditions	no. of items	sample size
1	in-lab	frame	PsychoPy	University	5	384	84
2	online	milliseconds	PsychoJS	MT	5	384	140
3	online	frame	PsychoJS	MT	5	384	140
4	online	milliseconds	Labvanced	Prolific	5	384	325
5	online	milliseconds	Labvanced	Prolific	7	512	400

Table 3: Summary of the main properties of the replications.

3 Analysis

3.1 Within-replication priming analyses

The data for all replications are available on OSF (https://osf.io/XXXX). In each dataset, we first looked at the distribution of the actual duration of prime words for each trial for each subject, to ensure that all trials included in further analyses had the corresponding prime last for the wanted duration (33 ms), regardless of the setting (in-lab/on-line), platform (PsychoPy/PsychoJS/Labvanced), and recruitment pool (university pool/MT/Prolific). Trials with the prime duration below 25 ms and above 60 ms were removed from analysis. Then, we looked at the by-subject and by-item performance. We removed subjects and items (words) whose overall error score was higher than 30%. We also made sure that all subjects included in the analysis had at least half of the trials that were presented with for each condition*primetype subdesign (i.e., 16 trials). Finally, we looked at the RT distribution. Individual trials were excluded if the corresponding RT was below 200 ms and above 1800 ms (Ratcliff 1993). Table Table 4 below provides the total number of observations and sample size for each replication after the aforementioned analysis pipeline and before statistical analyses.

#	setting	version	platform	recruitment	observations	sample size
1	in-lab	frames	PsychoPy	university	9598	73
2	online	ms	PsychoJS	MT	12327	107
3	online	frames	PsychoJS	MT	11658	103
4	online	ms	Labvanced	Prolific	22677	163
5	online	ms	Labvanced	Prolific	37955	196

Table 4: Number of observations for each replication after the analysis pipeline

3.2 Meta-analyses

In order to make sense of the large number of results reported in the five replications reported above, we performed a meta-analysis, along the lines of what has been previously done in our lab (Tucker, Idrissi, and Almeida 2021; but see also: Rosenthal and DiMatteo 2001; Schmidt and Hunter 2015; Cumming 2013; Cooper, Hedges, and Valentine 2019). To do this, we combined the results of the five replications into a unified dataset so to provide a less biased and better statistically-grounded summary of the cumulative evidence for a given pattern of results. We conducted a fixed effects meta-analysis, given that the five datasets reported here are replications of the exact same experiment, and the population being tested was similar across the five replications. To ensure further comparability, we excluded the two additional conditions tested in the second Labvanced batch (i.e., the orthographic and semantic conditions) from the meta-analysis. We therefore conducted five separate meta-analyses, each on the masked priming response to one of the conditions being tested, with the ultimate goal of comparing the priming effects for each of the conditions tested, and assessing the variance of the corresponding effect sizes across replications. In each analysis, the priming response was weighed by the inverse of the corresponding variance. All analyses were performed using the `metafor` R package (Viechtbauer 2010).

4 Results

4.1 Within-replication priming analyses

For each dataset, priming effects of each condition were calculated for each subject by subtracting the by-subject mean RT to the related sub-condition from the by-subject mean RT to the unrelated sub-condition. Standardized effect sizes (i.e., Cohen's *d*) were then calculated for each condition. Finally, by-subject mean priming effects were grand-averaged across subjects for each condition. Figure 1 and Table 5 below report the descriptive statistics of each replication. Overall, the identity and the transparent conditions systematically had the biggest effects across the five replications, and the syllabic orthographic condition had the smallest effects. The priming effects for the opaque and the non-syllabic conditions varied from being equally closer to the identity and transparent conditions in the in-lab replication (dataset # 1); to being closer to the syllabic condition in the all of the online replications (datasets # 2-5). The small effects for identity in the dataset # 2 (online, PsychoJS, ms, MT) is the odd one out as compared to the identity effects in the other datasets, and may be just accidental.

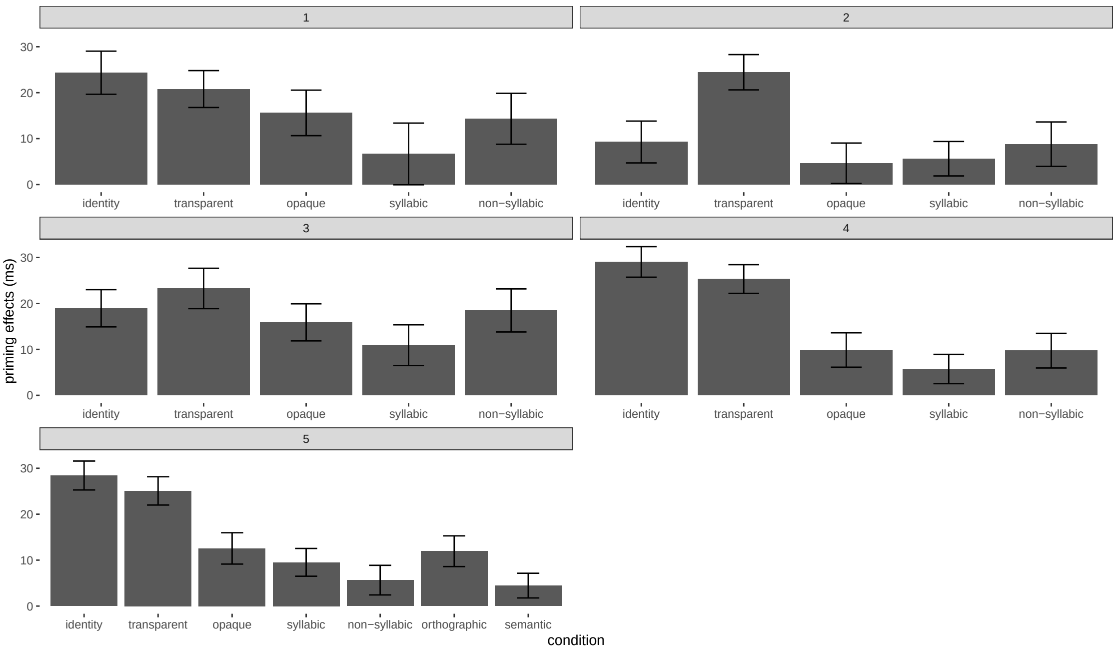


Figure 1: Barplot of the priming effects for each replication. Error bars represent one standard error in either direction

	#	setting	version	platform	recruitment	priming	sd	cor	ES
identity	1	in-lab	frames	PsychoPy	university	24.34 [15 33.68]	41.98	0.74	0.58
	2	online	ms	PsychoJS	MT	9.26 [0.24 18.29]	46.87	0.83	0.20
	3	online	frames	PsychoJS	MT	18.95 [10.91 26.98]	40.08	0.90	0.47

	4	online	ms	Labvanced	Prolific	29.03 [22.47 35.6]	42.46	0.85	0.68
	5	online	ms	Labvanced	Prolific	28.42 [22.22 34.62]	44.03	0.86	0.65
transparent	1	in-lab	frames	PsychoPy	university	20.79 [12.79 28.78]	35.93	0.79	0.58
	2	online	ms	PsychoJS	MT	24.45 [16.83 32.07]	39.17	0.88	0.62
	3	online	frames	PsychoJS	MT	23.26 [14.55 31.98]	43.93	0.88	0.53
	4	online	ms	Labvanced	Prolific	25.31 [19.15 31.48]	39.85	0.87	0.64
	5	online	ms	Labvanced	Prolific	25.07 [19 31.15]	43.13	0.86	0.58
opaque	1	in-lab	frames	PsychoPy	university	15.6 [5.74 25.46]	44.30	0.78	0.35
	2	online	ms	PsychoJS	MT	4.64 [-4.09 13.36]	44.41	0.86	0.10
	3	online	frames	PsychoJS	MT	15.88 [7.88 23.89]	40.13	0.91	0.40
	4	online	ms	Labvanced	Prolific	9.86 [2.47 17.26]	47.81	0.83	0.21
	5	online	ms	Labvanced	Prolific	12.55 [5.83 19.28]	47.73	0.83	0.26
syllabic	1	in-lab	frames	PsychoPy	university	6.67 [-6.7 20.03]	60.05	0.59	0.11
	2	online	ms	PsychoJS	MT	5.63 [-1.81 13.06]	38.04	0.90	0.15
	3	online	frames	PsychoJS	MT	10.91 [2.11 19.71]	43.20	0.89	0.25
	4	online	ms	Labvanced	Prolific	5.73 [-0.57 12.02]	40.70	0.87	0.14
	5	online	ms	Labvanced	Prolific	9.52 [3.56 15.48]	42.31	0.88	0.22
non-syllabic	1	in-lab	frames	PsychoPy	university	14.32 [3.28 25.35]	49.58	0.66	0.29
	2	online	ms	PsychoJS	MT	8.79 [-0.79 18.38]	48.81	0.84	0.18
	3	online	frames	PsychoJS	MT	18.47 [9.17 27.78]	45.93	0.88	0.40
	4	online	ms	Labvanced	Prolific	9.72 [2.27 17.17]	48.14	0.85	0.20
	5	online	ms	Labvanced	Prolific	5.65 [-0.71 12]	45.12	0.85	0.13
orthographic	5	online	ms	Labvanced	Prolific	11.94 [5.35 18.54]	46.80	0.86	0.26
semantic	5	online	ms	Labvanced	Prolific	4.46 [-0.84 9.76]	37.61	0.90	0.12

Table 5: Mean magnitude (and 95% confidence intervals in square brackets) and standard deviation (SD) pf priming effects, Pearson's r between the related and unrelated conditions, and Cohen's d (ES) for all conditions and replications.

We ran one t-test for each condition. The estimates, statistic values, and p-values of both analyses are reported in Table 6. The identity and transparent priming effects were found significant across all replications. The statistical results for the remaining conditions were not as uncontroversial, and varied across datasets. The opaque effects were found significant in all replication but one (dataset #2). The syllabic effects were found significant only in the last replication (dataset #5). The nonsyllabic effects were found significant in datasets #3 and #4.

	#	setting	platform	recruitment	t	df	p	signif.
identity	1	in-lab	PsychoPy	university	-4.42	2071.01	1.05e-05	**
	2	online	PsychoJS	MT	-2.87	2744.00	4.16e-03	*
	3	online	PsychoJS	MT	-3.89	2460.81	1.04e-04	**

	4	online	Labvanced	Prolific	-7.79	4747.60	7.95e-15	**
	5	online	Labvanced	Prolific	-8.08	5696.37	7.56e-16	**
transparent	1	in-lab	PsychoPy	university	-4.20	2160.86	2.78e-05	**
	2	online	PsychoJS	MT	-6.25	2733.52	4.65e-10	**
	3	online	PsychoJS	MT	-5.15	2596.02	2.76e-07	**
	4	online	Labvanced	Prolific	-7.16	4912.00	8.97e-13	**
	5	online	Labvanced	Prolific	-7.25	5816.74	4.74e-13	**
opaque	1	in-lab	PsychoPy	university	-2.93	1843.16	3.46e-03	*
	2	online	PsychoJS	MT	-0.68	2297.86	4.97e-01	n.s.
	3	online	PsychoJS	MT	-4.20	2239.95	2.82e-05	**
	4	online	Labvanced	Prolific	-2.80	4447.26	5.10e-03	*
	5	online	Labvanced	Prolific	-3.92	5246.99	9.04e-05	**
syllabic	1	in-lab	PsychoPy	university	-1.76	1795.78	7.90e-02	n.s.
	2	online	PsychoJS	MT	-1.26	2299.35	2.07e-01	n.s.
	3	online	PsychoJS	MT	-1.44	2232.94	1.50e-01	n.s.
	4	online	Labvanced	Prolific	-1.55	4417.66	1.22e-01	n.s.
	5	online	Labvanced	Prolific	-2.87	5222.47	4.13e-03	*
non-syllabic	1	in-lab	PsychoPy	university	-1.60	1691.99	1.09e-01	n.s.
	2	online	PsychoJS	MT	-1.16	2176.06	2.45e-01	n.s.
	3	online	PsychoJS	MT	-3.38	2095.44	7.35e-04	**
	4	online	Labvanced	Prolific	-2.68	4101.00	7.46e-03	*
	5	online	Labvanced	Prolific	-1.84	4771.34	6.56e-02	n.s.
orthographic	5	online	Labvanced	Prolific	-2.98	5354.33	2.93e-03	*
semantic	5	online	Labvanced	Prolific	-1.42	5797.97	1.57e-01	n.s.

Table 6: Summary of the statistical results

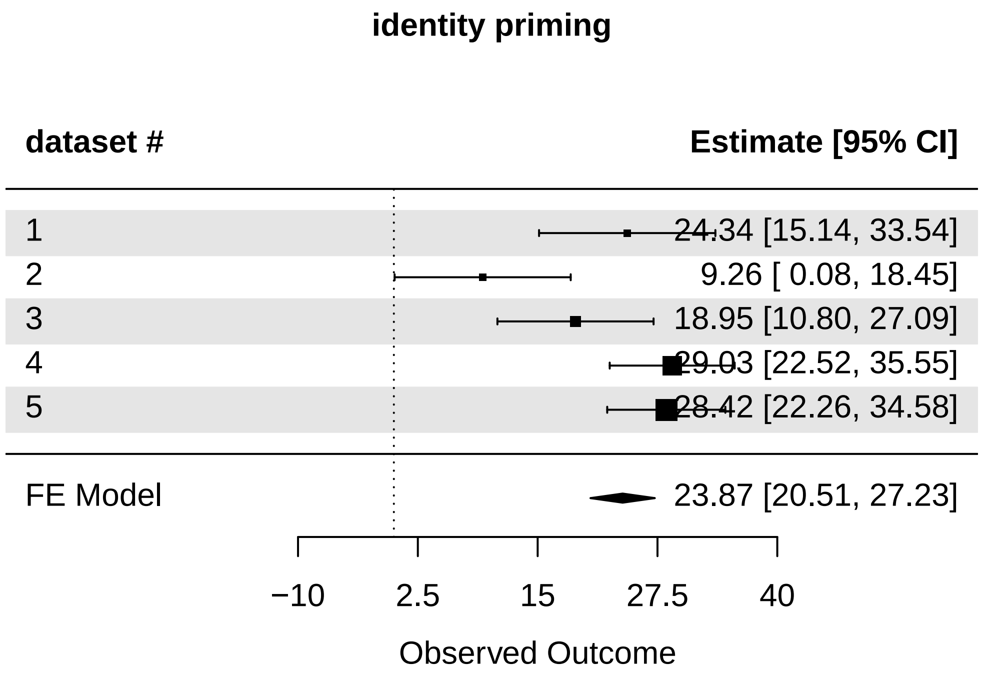
4.2 Meta-analysis

The statistical results reported above revealed on the one hand a substantial consistency in the priming effects to identity and transparent conditions across different setting (in-lab vs. online), platform (PsychoPy/PsychoJS vs Labvanced), version (ms vs. frames), and recruitment pool (university pool vs. MT vs. Prolific). This is very comforting, as it substantiates the results reported in the literature. On the other hand, the results above also revealed an evident variance in the distribution of the priming effects to the opaque, syllabic, and non-syllabic condition. We deem this variance worth further investigation, since the divide between the opaque condition (which is somewhat morphological), and the syllabic and the non-syllabic conditions (which are instead not morphological at all) is the crucial argument for a level of processing specifically devoted to morphological analysis. Should this divide be revealed not to be as clear as previously believed, it would necessarily call for a reconsideration of the models of word processing currently accepted. The meta-analytical results reported here may be apt for this purpose.

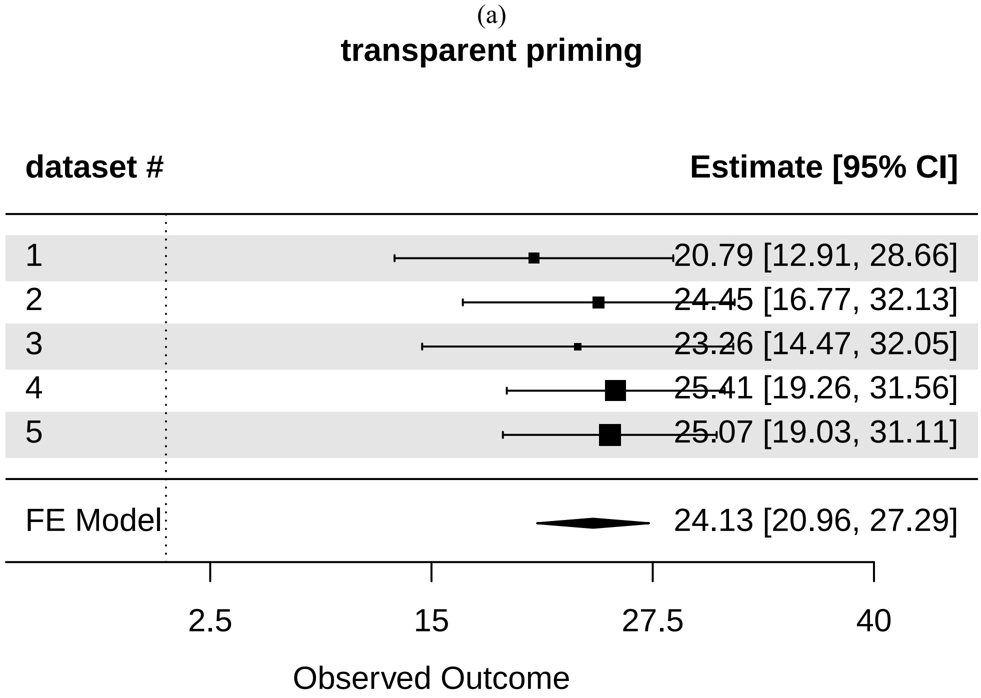
Figure 2, Figure 3 display the forest plots of the meta-analyses conducted for the five conditions tested across the five replications. The results straightforwardly suggest priming effects higher than 0 for all conditions. The largest effects were found in the identity and transparent conditions, with a rounded point-estimate effect size of 24 ms for each condition (Figure 2a, Figure 2b). The three remaining conditions were found to elicit smaller, yet positive, priming effects, with rounded point-estimates of 12 and 10 ms for the opaque and non-syllabic conditions (Figure 3a, Figure 3c), and 8 ms for the the syllabic condition (Figure 3b).

5 General discussion

In the past, a great deal of evidence has been reported suggesting a seemingly consistent bipartite pattern in the masked priming response to words: (a) similar, positive effects to identical (e.g., *dog-DOG*), morphologically related (e.g., *driver-DRIVE*, where the prime is morphologically derived from the target), and morpho-orthographically related word pairs (e.g., *brother-BROTH*, where the prime only seems to be morphological related to target since *-er* is a real English morpheme, but the two words are actually unrelated to one another); and (b) null or close-to-null effects to orthographically similar word pairs, where the target is orthographically contained in the prime but is not related to it (e.g., *canvas-CAN*; *starch-STAR*). This seemingly clear-cut dichotomy suggested that morphological decomposition occurs onto the incoming visual stimulus at early stages of visual word processing, i.e. within the first tens of milliseconds post stimulus onset. This procedure has since been thought to be "automatic", i.e. occurring with no regards to lexical (i.e., semantic) relatedness between the prime and target stimuli, but exclusively on the basis of long-term "morpho-orthographic" units, i.e. letter clusters memorized as a result of orthographic statistical predictions ("islands of regularity": Rastle and Davis 2008). The meta-analysis of the five replications reported here suggests instead that such a divide is actually not as reliable as previously thought – or, at least, not within the same purview. Across the five replications of the same experiment, we did find a dichotomic pattern in the masked priming response to the five conditions tested, but different from the one above in at least two respects (see Figure 5). First, regarding (a), the magnitude of the priming effects to the morphologically opaque condition is numerically closer to the magnitude of the priming effects to the stem-containing orthographic conditions (i.e., the syllabic and non-syllabic conditions), rather than to the magnitude of the identity and morphologically transparent conditions. Second, regarding (b), orthographic priming –and, specifically, priming to word pairs in which the target is only orthographically contained in the prime, but not morphologically related to it– is not-null, nor near-to-null. As further shown in the forest plot below Figure 4, the point-estimate of the masked priming response to the two stem-containing orthographic conditions merged as a single condition was around 9 ms – indeed very close to the 12-ms estimate of the opaque condition.

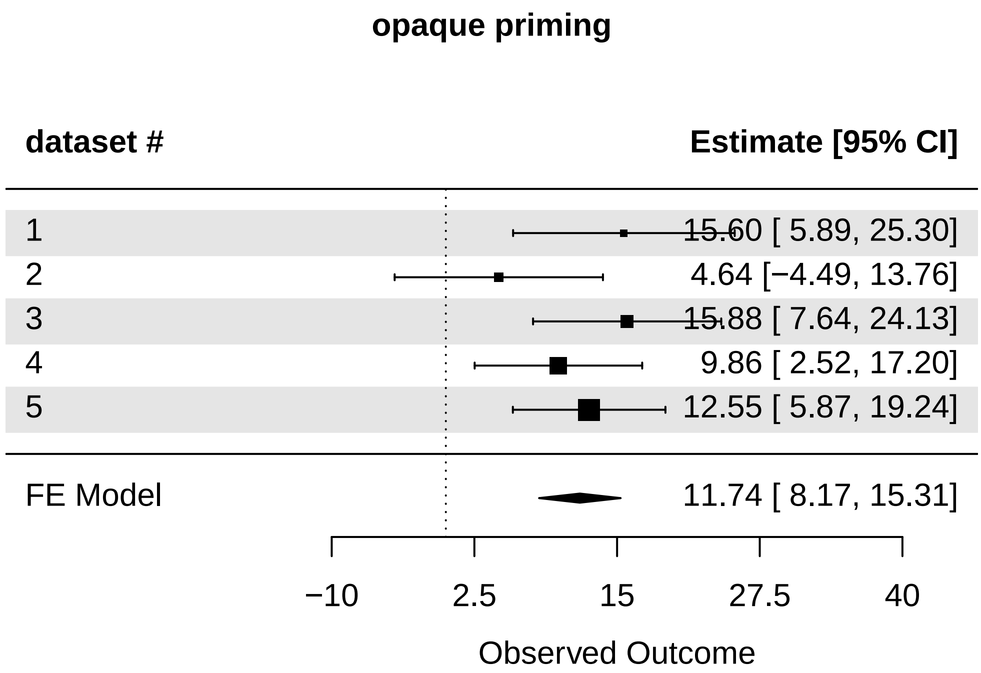


(a)

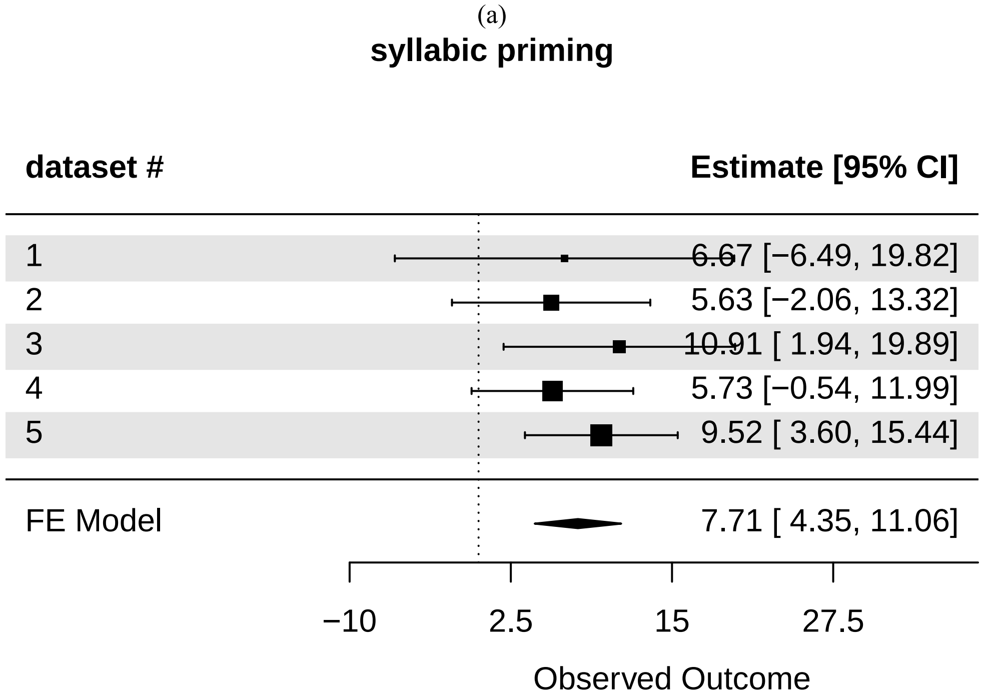


(b)

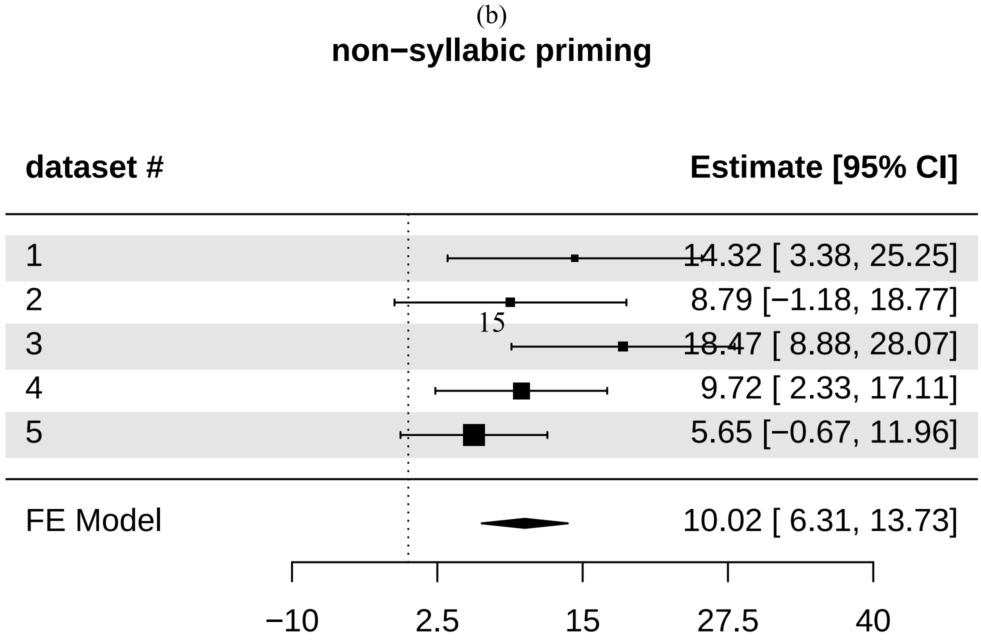
Figure 2: Forest plots of meta-analyses conducted for the identity (a) and transparent (b) conditions. The size of the square representing each replication is proportional to its weight in the meta-analysis. The estimated effect size of the meta-analysis is represented by the black diamond at the bottom of each graph. The width of the diamond represents the estimated 95% CI of the meta-analysis effect size. Experiments are plotted in chronological order from top to bottom.


opaque priming
dataset #
Estimate [95% CI]
1
15.60 [5.89, 25.30]
2
4.64 [−4.49, 13.76]
3
15.88 [7.64, 24.13]
4
9.86 [2.52, 17.20]
5
12.55 [5.87, 19.24]
FE Model
11.74 [8.17, 15.31]
−10
2.5
15
27.5
40
Observed Outcome

(a)


syllabic priming
dataset #
Estimate [95% CI]
1
6.67 [−6.49, 19.82]
2
5.63 [−2.06, 13.32]
3
10.91 [1.94, 19.89]
4
5.73 [−0.54, 11.99]
5
9.52 [3.60, 15.44]
FE Model
7.71 [4.35, 11.06]
−10
2.5
15
27.5
Observed Outcome

(b)


non−syllabic priming
dataset #
Estimate [95% CI]
1
14.32 [3.38, 25.25]
2
8.79 [−1.18, 18.77]
3
18.47 [8.88, 28.07]
4
9.72 [2.33, 17.11]
5
5.65 [−0.67, 11.96]
FE Model
10.02 [6.31, 13.73]
−10
2.5
15
27.5
40

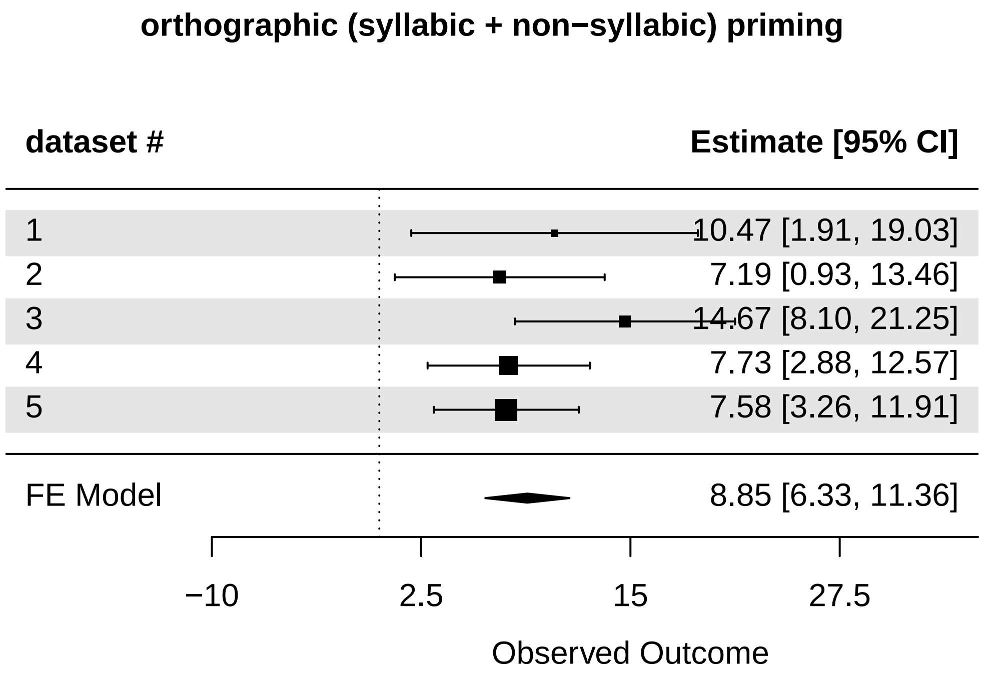


Figure 4: Forest plots of meta-analyses conducted for the two orthographic conditions (syllabic and non-syllabic) conditions merged together. The size of the square representing each replication is proportional to its weight in the meta-analysis. The estimated effect size of the meta-analysis is represented by the black diamond at the bottom of each graph. The width of the diamond represents the estimated 95% CI of the meta-analysis effect size. Experiments are plotted in chronological order from top to bottom.

5.1 Implications for morphological decomposition and priming

The large total sample size from which these estimates were calculated ($N_{total} = $ 642) makes these results highly reliable, and call for a general reconsideration of the mechanisms underpinning morphological decomposition as indexed by visual masked priming. In what follows, we abstract away from the statistical results reported above, or reported in the literature (i.e., the *F*s, *t*s, and *p*s), and focus on the actual raw magnitude estimates. The reason for this is two-fold. The first reason is that the little consensus regarding the right statistical methodology to measure the statistical significance of priming makes comparison across studies very hard, or even impossible at times. Most studies indeed report just a subset of the information necessary to run a meta-analysis similar to the one we have reported here (e.g., standard deviations, standardized effect size, correlation between conditions are usually never reported in papers), or to replicate the experiment altogether (e.g., the pseudo-word list is usually never published). The second reason is that statistical significance is strictly correlated with statistical power. It has been recently acknowledged that research in psychology (and sub-fields) suffers from a fundamental problem of low statistical power, due to the insufficiently low size of the sample recruitable (Brysbaert and Stevens 2018). The risks that comes with low statistical power are not negligible, as scientific theoretical progress fundamentally hinges on the reliability of statistical results. Under these circumstances, the absence of statistical significance reported in the relevant literature for a given condition turns out to be moot at best, as it may not be indicative of the actual absence of an effect, rather amere byproduct of low statistical power. It is primarily for this reason that four of the five replications reported here were conducted online, which seems to be the fastest way to recruit a sufficient sample size that would guarantee a minimum power of 80%, while guaranteeing experimental replicability (e.g., Petrosino and Almeida submitted).

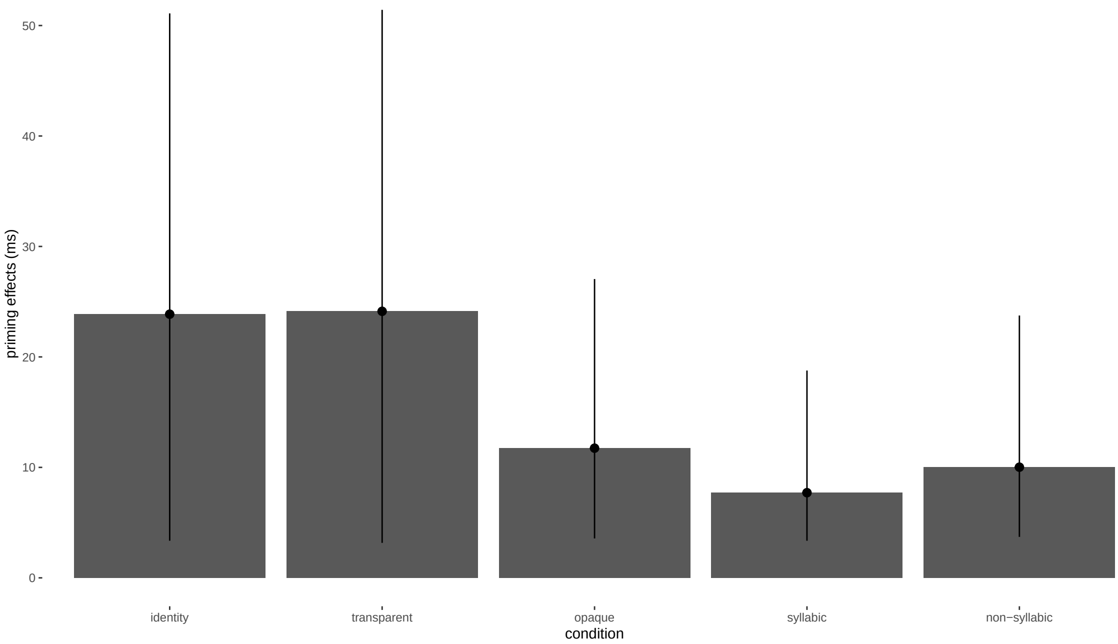


Figure 5: Barplot of the priming effects (with the 95% confidence intervals) as calculated in the meta-analysis described above

In general, the comparison that morphological (i.e., what the literature has traditionally called “transparent”) priming has with semantic priming and orthographic priming is quite different in nature. On the one hand, our results seems in line with previous studies in suggest a pretty clear distinction between morphological and semantic priming, with the magnitude semantic priming being very close to null (95% CIs: [-0.84 9.76]), even with the fairly large sample size of dataset #5 ($N = 196$). On the other hand, the comparison of morphological priming with orthographic priming seems more complex than thought of before. In our study, we triggered priming effects to two types of orthographic relatedness. The first type of orthographic relatedness involves an existent English target stem (e.g., *can, star*) orthographically contained inside a monomorphemic prime word which was either one syllable (*canvas*) or one consonantal cluster (*starch*) longer than the target, but with no evident morphological and semantic relationship with it. While this type of orthographic relatedness has been believed to trigger null, or occasionally small (< 5 ms), priming effects (Rastle and Davis 2008), our results suggest tht it may still elicit small to medium (cumulative 95% CIs: [6.33 11.36]) priming effects – i.e., typically around half of the size of identity/transparent priming. The second type of orthographic relatedness we have tested in this study involves monomorphemic English orthographic neighboring words, i.e. words of equal orthographic length and differing in only one letter (e.g., *basil-BASIN*). In this case, we reported effects that were larger than the root-contained orthographic conditions (with a mean priming effect of around 12 ms, 95% CI: [5.35 18.53]), and still smaller than the identity/transparent conditions. We should be kept mind that, however, that this estimate refers to dataset #5 only, and therefore to a much smaller sample size than the other conditions ($N = 196$, just around a third of the sample size for the other conditions). To further complicate the picture, our meta-analytical results also showed a substantial difference in size between the transparent and opaque priming response, with the former being nearly twice as big as the latter ($M = 11.74, 95$), in contrast to what was previously reported. Since Rastle, Davis, and New (2004), the opaque priming response has since been assumed to be equal in size to the transparent priming response. The fact that the two responses were statistically equal has been routinely put forth as the main argument for a model in which morphological decomposition must occur blindly to any semantic, and therefore *before* lexical access, so to apply in both morphologically transparent bimorphemic words (such as *driver*) as well as morphologically opaque monomorphemic words (such as *brother*). Our meta-analysis of our five replications of the same experiment suggests instead that the size of opaque priming is much smaller than the size of transparent priming, and comparable to the size of orthographic priming.

The complex relationship of morphological priming with opaque, orthographic, and semantic priming suggests that morphological priming has a special status in early visual word recognition, which is yet to be completely understood. The fact the size of morphological priming is large, or in any case comparable to the the size of identity priming, possibly suggests that stems are activated quite early in visual word recognition. Once activated, the stem presented as target has been already activated (or "open" in the parlance of Forster 1999's model of priming), and is recognized faster, thus ultimately eliciting a facilitation (($M = 24.13, 95$) that is numerically very similar to cases of identity priming ($M = 23.87, 95$), in which the exact same word is repeated as prime and target. Admittedly, the reason of this numerical equality is yet unclear, and relatively understudied. Here, we attempt a theoretical account that might explain all the effects, but further research will need to address each testing hypothesis therein. The account we propose here rests upon two main assumptions. The first assumption comes from what we have just said regarding early stem activation, which strictly entails that lexical access occurs earlier than commonly assumed. To the best of our knowledge, the idea of early lexical access in word processing has never been considered as an actual possibility, since it is usually equated to semantic access (which is commonly thought of as occurring at later stages of word processing). Actually, lexical access has a rather vague nature, and never had a proper and commonly accepted definition. In our view, lexical access might be seen as a gradual unfolding of word properties, rather than a one-time step in processing: in other words, we entertain the possibility that lexical access may *start* after stimulus presentation, thus allowing the parser to gradually access bits of lexical information of the stimulus, rather than accessing it all in one go. Further research will be needed to clarify the exact unfolding of properties, but here we assume that a prime duration smaller than 60 ms will forbid the parser to access higher-level information such as meaning. The second assumption we make is that concerns the relationship between the processing of the prime and the processing of the target. Most research on masked priming hinges on the assumption that the short duration of the prime stops the processing of the prime as soon as the prime it is visually replaced by the target on the screen. Under this assumption, modulating the prime duration allows us to probe the set of properties available to the parser during the decomposition of the prime stimulus. A plausible alternative, just hinted at in a small number of contributions (e.g., Forster, Mohan, and Hector 2003; Petrosino 2020), contends that the processing of the prime continues even after target presentation, and runs in parallel with it. In the model we present here, we entertain this possibility, thus allowing any property of the prime being accessed to influence processing of the target (and ultimately its recognition). Once the prime is presented, its processing starts right away: its low-level orthographic features are identified first, and, by the target is presented, the letters of the prime are clustered together to form more complex units ("morpho-orthographic units": Rastle and Davis 2008). After the target presentation, the processing of the prime continues along with the processing of the target. The properties of the prime that the parser has accessed therefore influence the processing of the target. If there is no match between the orthographic code of the target and the orthographic code of the prime (e.g., in semantically related pairs: *poem-RHYME*), the latter cannot elicit any facilitation onto the recognition of the former. If there is, some facilitation (i.e., priming) may occur, the extent of which is strictly modulated by two different levels of analysis of the prime: (i) the size of the orthographic overlap between the prime and the target; and (ii) the successful activation of the morpheme (in our cases, the stem) contained in the prime. In identical pairs (*fuss-FUSS*), the size of the overlap is at its maximum and the corresponding stem is successfully activated; thus, the target stimulus received the maximal facilitation [arguably, numerically contingent on the actual duration of the prime; see Forster, Mohan, and Hector (2003)]. In morphologically transparent pairs (*driver-DRIVE*), the size of the overlap, though not maximal, is substantial (since the non-overlapping string portion only minimally concerns the left side of the prime, where the suffix surfaces). The derivational information of the prime is accessed, and the corresponding stem (*drive*) gets activated. As perfectly matching with the target, it elicits priming effects that are comparable to the effects triggered in identity pairs. In morphologically opaque pairs (*belly-BELL*), and stem-containing orthographically related pairs (syllabic: *canvas_CAN*; non-syllabic: *starch-STAR*), the stem contained in the prime is initially set to trigger priming onto the corresponding target; however, the facilitation effect gets partially inhibited when the derivational information of the prime is accessed and the purported stem is not found. As a result, the priming effects to these pairs is penalized, and its size reduced to half of the size of identity and transparent priming. The extent of the inhibition may also vary depending on the size of the non-overlapping portion, which was roughly the same in the opaque and non-syllabic conditions (consisting of one to two letters), but was smaller in the syllabic condition (as consisting of a whole CVC syllable, i.e. three letters). Indeed, the magnitude of the priming of the two former conditions was slightly larger than the magnitude of the priming of the latter condition. Finally, in our orthographically neighboring pairs (*basil-BASIN*), the size of orthographic overlap is almost maximal (as the non-overlap size being one letter by definition). As a result, the prime stimulus facilitates the recognition of the target stimulus at the orthographic level, but not at the lexical level, where no morpheme activation occurs so to facilitate eventual target recognition. Hence, the size of priming cannot reach its maximum, and remains at the similar magnitude to the other orthographically related pairs. Albeit in sharp contrast with the traditional view of lexical access (at the theoretical/computational level) and of the mechanics of priming (at the empirical/algorithmic level), this account seems to offer a plausible explanation for the differential priming effects we obtained in the meta-analysis reported above. Furthermore, it maintains the commonly accepted view of morphological procedure as an automatic operation occurring at early stages of word processing, and represents a new fertile source for further research on the topic.

References

Angele, Bernhard, Ana Baciero, Pablo Gómez, and Manuel Perea. 2022. "Does Online Masked Priming Pass the Test? The Effects of Prime Exposure Duration on Masked Identity Priming." *Behavior Research Methods* 55 (1): 151–67. https://doi.org/10.3758/s13428-021-01742-y.

Balota, David A., Melvin J. Yap, Keith A. Hutchison, Michael J. Cortese, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson, and Rebecca Treiman. 2007. "The English Lexicon Project." *Behavior Research Methods* 39 (3): 445–59.

Brysbaert, Marc, and Michaël Stevens. 2018. "Power Analysis and Effect Size in Mixed Effects Models: A Tutorial." *Journal of Cognition* 1 (1). https://doi.org/10.5334/joc.10.

Cooper, Harris, Larry V. Hedges, and Jeffrey C. Valentine, eds. 2019. "The Handbook of Research Synthesis and Meta-Analysis," June. https://doi.org/10.7758/9781610448864.

Cumming, Geoff. 2013. "The New Statistics." *Psychological Science* 25 (1): 7–29. https://doi.org/10.1177/0956797613504966.

Finger, Holger, Caspar Goeke, Dorena Diekamp, Kai Standvoß, and Peter König. 2017. "LabVanced: A Unified JavaScript Framework for Online Studies." In *2017 International Conference on Computational Social Science*. Cologne, Germany.

Fiorentino, Robert, Stephen Politzer-Ahles, Natalie S Pak, María Teresa Martínez-García, and Caitlin Coughlin. 2015. "Dissociating Morphological and Form Priming with Novel Complex Word Primes: Evidence from Masked Priming, Overt Priming, and Event-Related Potentials." *The Mental Lexicon* 10 (3): 413–34.

Forster, Kenneth I. 1999. "Microgenesis of Priming Effects in Lexical Access." *Brain and Language* 68: 5–15.

Forster, Kenneth I., Kathleen Mohan, and Jo Hector. 2003. "The Mechanics of Masked Priming." In *Masked Priming: The State of the Art*, edited by Sachiko Kinoshita and Stephen J. Lupker, 3–37. New York, NY/Hove, UK: Psychology Press.

Morris, Joanna, James H Porter, Jonathan Grainger, and Phillip J Holcomb. 2011. "Effects of Lexical Status and Morphological Complexity in Masked Priming: An ERP Study." *Language and Cognitive Processes* 26 (4-6): 558–99.

Peirce, Jonathan W. 2009. "Generating Stimuli for Neuroscience Using PsychoPy." *Frontiers in Neuroinformatics* 2: 10.

Peirce, Jonathan, Jeremy R Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. "PsychoPy2: Experiments in Behavior Made Easy." *Behavior Research Methods* 51: 195–203.

Petrosino, Roberto. 2020. "More Than Islands of Regularity: An Investigation of the Sensitivity of Morphological Decomposition to Higher-Level Linguistic Properties." PhD thesis, University of Connecticut.

Petrosino, Roberto, and Diogo Almeida. submitted. "The Frequency Attenuation Effect in the Masked Priming Response: A Large-Scale Online Study." *Journal of Cognitive Neuroscience*, submitted.

Petrosino, Roberto, Jon Sprouse, and Diogo Almeida. submitted. "Asymmetries in the Stem and Suffix Masked Priming Response in a Large-Scale Online Study." In *Quaderni Di Linguistica e Studi Orientali*.

Rastle, Kathleen, Matt H. Davis, William D. Marslen-Wilson, and Lorraine K. Tyler. 2000. "Morphological and Semantic Effects in Visual Word Recognition: A Time-Course Study." *Language and Cognitive Processes* 15 (4-5): 507–37.

Rastle, Kathleen, and Matthew H. Davis. 2008. "Morphological Decomposition Based on the Analysis of Orthography." *Language and Cognitive Processes* 23 (7-8): 942–71.

Rastle, Kathleen, Matthew H. Davis, and Boris New. 2004. "The Broth in My Brother's Brothel: Morpho-Orthographic Segmentation in Visual Word Recognition." *Psychonomic Bulletin & Review* 11 (6): 1090–98.

Ratcliff, Roger. 1993. "Methods for Dealing with Reaction Time Outliers." *Psychological Bulletin* 114 (3): 510–32. https://doi.org/10.1037/0033-2909.114.3.510.

Rosenthal, R., and M. R. DiMatteo. 2001. "Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews." *Annual Review of Psychology* 52 (1): 59–82. https://doi.org/10.1146/annurev.psych.52.1.59.

Schmidt, Frank L., and John E. Hunter. 2015. "Methods of Meta-Analysis: Correcting Error and Bias in Research Findings." https://doi.org/10.4135/9781483398105.

Seidenberg, Mark S. 1987. "Sublexical Structures in Visual Word Recognition: Access Units or Orthographic Redundancy?" In *Attention and Performance 12: The Psychology of Reading*, edited by M. Coltheart. Hillsdale, NJ: Erlbaum.

Taft, Marcus. 2003. "Morphological Representation as a Correlation Between Form and Meaning." In *Reading Complex Words*, edited by E. Assink and D. Sandra, pp. 113–137. Amsterdam: Kluwer.

Taft, Marcus, and Minh Nguyen-Hoan. 2010. "A Sticky Stick? The Locus of Morphological Representation in the Lexicon." *Language and Cognitive Processes* 25 (2): 277–96. https://doi.org/10.1080/01690960903043261.

Tucker, Matthew A., Ali Idrissi, and Diogo Almeida. 2021. "Attraction Effects for Verbal Gender and Number Are Similar but Not Identical: Self-Paced Reading Evidence from Modern Standard Arabic." *Frontiers in Psychology* 11 (January). https://doi.org/10.3389/fpsyg.2020.586464.

Viechtbauer, Wolfgang. 2010. "Conducting Meta-Analyses in*R*with the**metafor**Package." *Journal of Statistical Software* 36 (3). https://doi.org/10.18637/jss.v036.i03.