The detection and accurate estimation of frequency attenuation effects in masked repetition priming: A large scale web browser-based study

Roberto Petrosino[1], Diogo Almeida[1]

ARTICLE HISTORY
Compiled July 11, 2024

[1] Psychology Program, Division of Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

ABSTRACT
This study investigates the controversy surrounding the sensitivity of masked repetition priming to word frequency: while unmasked priming exhibits a frequency attenuation effect, wherein high frequency words yield smaller repetition effects, this phenomenon has been inconsistently reported in masked priming. We conducted two large online experiments with rigorously validated frequency databases to reconcile past discrepancies. The first experiment confirmed the viability of conducting masked priming experiments in web browser-based settings. The pre-registered second study, designed for high statistical power and precision, identified a 10-ms attenuation effect under masked priming. This result suggests that the repetition effect in masked priming is less qualitatively distinct from unmasked priming than previously assumed. This finding has implications for masked priming experimental design and theoretical consequences for models of priming. Crucially, models that predict either the presence or absence of frequency attenuation under masked conditions need to account for a small but reliable effect.

KEYWORDS
masked repetition priming; frequency attenuation effect; online browser-based experiment; power analysis

1. Introduction

The masked priming technique has been an invaluable tool in visual word recognition research. It has allowed researchers to study the conditions under which orthographic, phonological, morphological, and semantic information impact access to visual word forms while mitigating strategic effects and minimizing the influence of controlled processes (Forster 1998). First introduced in its traditional form by Forster and Davis (1984; see also Evett and Humphreys 1981), this technique involves a forward mask (i.e., usually a string of hashes, #####), followed by a prime string presented for very short time ($SOA < 60$ ms), and a target string presented immediately after. Because the

CONTACT: Roberto Petrosino. Email: roberto.petrosino@nyu.edu. Diogo Almeida. Email: diogo@nyu.edu.

prime presentation is so brief and masked by preceding and subsequent stimuli, most participants report not being aware that a prime string has been presented, and can at most report a screen flicker just before the target presentation (Forster, Mohan, and Hector 2003).

Among possible manipulations of prime-target relatedness, masked repetition priming (in which the same word is presented as both the prime and target within the same trial: e.g., *love-LOVE*) has been well studied, because its response seems to be qualitatively different from the unmasked counterpart ($SOA > 60ms$): while high-frequency words benefit less from repetition than low-frequency words in the unmasked design (*frequency attenuation effect*, henceforth FAE; Scarborough, Cortese, and Scarborough 1977), this does not seem to be the case when the prime is masked (Forster and Davis 1984; Forster et al. 1987; Segui and Grainger 1990; Sereno 1991; Forster and Davis 1991; Rajaram and Neely 1992; Bodner and Masson 1997; Forster, Mohan, and Hector 2003; Nievas 2010).

This asymmetry in sensitivity to lexical frequency between the masked and unmasked repetition priming responses has been important in distinguishing among different models of priming in visual word recognition. More specifically, *interactive activation models* (McClelland and Rumelhart 1981; Grainger and Jacobs 1996; Coltheart et al. 2001) conceive of priming as a "head start" in processing due to the pre-activation of the target word due to the presentation of the prime. Thus, according to *interactive activation* models, priming is ultimately caused by a single mechanism, making the qualitatively different profiles for repetition priming in masked and unmasked conditions a difficult empirical finding to explain.

Similarly, episodic models (e.g., Jacoby and Dallas 1981; Jacoby 1983) posit a different single mechanism for priming effects: the activation/retrieval of the episodic memory trace of the encounter with the prime word. These models therefore encounter the same type of difficulty in accounting for qualitatively different patterns of repetition priming effects in masked and unmasked conditions. A similar type of model, called the *memory recruitment model* makes very similar predictions to the episodic memory models, positing a non-lexical source for priming effects (Bodner and Masson 1997; Masson and Bodner 2003; Bodner and Masson 2014). Repetition priming effects under this view stem from the exploitation, strategically or automatically, of a memory resource created by the encounter with the prime word. The frequency attenuation effect, under episodic and memory recruitment models alike, is predicted on the basis that low frequency primes, being more distinctive stimuli, create a more potent and effective memory resource compared to high frequency primes.

In contrast, other models appear to successfuly sidestep the problem posed by the qualitatively different repetition priming profiles observed in masked and unmasked conditions. One such model is the *entry-opening model* (also known as the *bin model*; Forster and Davis 1984). According to this model, when the visual stimulus is presented, lexical entries are assigned to specific bins based on orthographic similarity. In the first stage (fast search stage), a fast, frequency-ordered search goes through the entries within a given bin, and compares each one with the the input stimulus, assigning to each entry a goodness-of-fit score. This comparison is fast and crude, and sorts entries into (a) perfect (i.e., no difference is detected between the input and the entry), (b) close (i.e., small differences are detected), and (c) irrelevant matches (i.e., substantial difference are detected). Any entry of type (a) or (b) is opened, so that the entry can be further analyzed and compared to the input in the subsequent verification stage.

Under a masked presentation, the entry of the prime word is opened at the fast search stage, but the short duration of the stimulus prevents it from reaching the evaluation stage. Crucially, the entry is nonetheless left open. Upon the presentation of the target stimulus, the access procedure will follow its two stage course, with a frequency-sensitive fast search and a subsequent entry opening for evaluation/verification. In this view, the fast search for the target word proceeds normally, but the evaluation/verification procedure starts and ends sooner than it otherwise would, because the target entry has already been left open after the brief processing of the prime. Thus, the *entry-opening model* explains the masked repetition priming as the benefit from having the entry of the target word already open by the time the second stage of recognition starts. Crucially, this occurs *after* the target word is initially accessed, which happens in order of frequency. Put differently, according to the *entry-opening model*, masked repetition priming occurs because of the time savings from not having to open the entry, which is a frequency-insensitive process (i.e., every entry takes the same time to be opened), but *after* the frequency-sensitive first access stage. As a consequence, the *entry-opening model* predicts a frequency-insensitive masked repetition priming effect, which is what has been traditionally reported in the literature (see Table 1). In addition, it also (correctly) predicts that pseudowords should not benefit from masked repetition priming, as they have no entries in the mental lexicon to be left open after the brief processing of the prime.

However, as Table 1 shows, there are nonetheless a few studies that do report significant FAEs in masked repetition priming Nievas (2010). Bodner and Masson (2001) argues that when stimuli are presented in alternating case (e.g., *pHoNe*), this increases the lexical decision difficulty and therefore generates an extra incentive to draw on the memory resource created by the brief processing of the prime. Under such conditions, they were able to observe a statistically significant FAE.

In the same vein, Kinoshita (2006) noticed that in earlier studies the low frequency words often had very high error rates, and suggested that perhaps many participants did not know them. If participants treated a substantial number of low frequency words as nonwords, and nonwords do not exhibit repetition priming under masked conditions, it could artificially depress the repetition priming effect for the low frequency condition alone, which could make any existing FAE harder to detect. In two separate experiments, Kinoshita (2006) showed that larger repetition priming effects for low frequency words were only obtained when the low frequency words were vetted to make sure the participants knew them prior to the experiment. Following up on Kinoshita (2006), Norris and Kinoshita (2008) were also able to find an interaction between lexical frequency and repetition in masked repetition priming, as was Nievas (2010) in Spanish (exp. 1B).

Finally, as Table 1 shows, it is noteworthy that 15 out of 18 previous studies showed numerically larger masked priming effects for low frequency words as opposed to high frequency words, irrespective of statistical significance. Similarly, the average repetition effect for low frequency words in the studies reviewed in Table 1 is 13 ms larger when compared to that of high frequency words. These results are not in line with the predictions dictated by the *entry opening model*, and seem to align better with the predictions made by *interactive activation models* and *memory recruitment* models.

Table 1. Summary of the masked repetition priming effects as a function of word frequency reported in the literature. The statistical power range estimates were calculated by simulation with the corresponding sample size (N) and for two representative FAE magnitudes. Simulations were performed across a range of correlation values between conditions (from 0.6 to 0.9, in increments of 0.1) as well as plausible standard deviations per conditions (from 60 ms to 180 ms, in increments of 10 ms), with 10,000 simulated datasets for each combination of parameters.

				MOP (ms)		FAE (ms)		Power range [min max]	
Study	Language	N	SOA	HF	LF	ES	$p<.05$?	FAE=15ms	FAE=30ms
Forster, Davis, Schoknecht, & Carter (1987), exp. 1	English	16	60	61	66	5		[0.02 0.24]	[0.04 0.84]
Norris, Kinoshita, Hall, & Henson (2018)	English	16	50	38	51	13		[0.02 0.24]	[0.04 0.84]
Sereno (1991), exp. 1	English	20	60	40	64	24		[0.02 0.33]	[0.04 0.92]
Forster & Davis (1991), exp. 5	English	24	60	54	72	18		[0.02 0.4]	[0.05 0.96]
Bodner & Masson (1997), exp. 1	English	24	60	29	45	16		[0.02 0.4]	[0.05 0.96]
Bodner & Masson (1997), exp. 3	English	24	60	36	50	14		[0.02 0.4]	[0.05 0.96]
Forster, Mohan, & Hector (2003), exp. 1	English	24	60	63	60	-3		[0.02 0.4]	[0.05 0.96]
Kinoshita (2006), exp. 1	English	24	53	32	38	6		[0.02 0.4]	[0.05 0.96]
Kinoshita (2006), exp. 2	English	24	53	29	59	30	*	[0.02 0.4]	[0.05 0.96]
Norris & Kinoshita (2008), exp. 1	English	24	53	35	66	31	*	[0.02 0.4]	[0.05 0.96]
Forster, Davis, Schoknecht, & Carter (1987), exp. 4	English	27	60	34	25	-9		[0.03 0.46]	[0.05 0.98]
Forster & Davis (1984), exp. 1	English	28	60	45	38	-7		[0.03 0.48]	[0.06 0.98]
Nievas (2010), exp. 1b	Spanish	30	50	44	65	21	*	[0.03 0.52]	[0.06 0.99]
Nievas (2010), exp. 2a	Spanish	30	50 or 33[1]	51	58	7		[0.03 0.52]	[0.06 0.99]
Segui & Grainger (1990), exp. 4	French	36	60	42	45	3		[0.03 0.63]	[0.07 1]
Bodner & Masson (2001), exps. 2A, 2B, 3, & 6 (average)[2]	English	40	60	37	69	32	*	[0.03 0.68]	[0.08 1]
Rajaram & Neely (1992), exp. 1	English	48	50	30	37	7		[0.04 0.76]	[0.09 1]
Rajaram & Neely (1992), exp. 2	English	48	50	45	78	33		[0.04 0.76]	[0.09 1]
Mean				41	55	13			
SD				10	14	13			
Correlation						0.46			

[1]SOA for each subject determined by pre-test
[2]Reported in Masson & Bodner (2003)

2. The present study

It is somewhat surprising that the status of the FAE in masked priming remains largely unresolved in the literature, given its non-negligible average magnitude across studies and its theoretical significance in elucidating the underlying cognitive processes of masked priming.

One possible interpretation of the conflicting past findings revolves around the fact that only 4 out of 18 studies demonstrate a statistically significant FAE. Notably, this number potentially diminishes further when considering that, among these four studies, the FAE is detected only through the pooling of data across multiple studies employing a unique alternating-case stimulus presentation (Bodner and Masson 2001; Masson and Bodner 2003). This line of reasoning suggests a qualitatively distinct profile between masked and unmasked repetition priming, with the FAE more firmly established in the latter.

Conversely, one could argue that 15 out of 18 studies exhibit numerically larger repetition effect sizes for low-frequency words compared to high-frequency words —- a pattern that is challenging to reconcile with a genuine absence of interaction between frequency and masked repetition. Additionally, the average FAE across all studies stands at 13 ms, a modest yet non-negligible effect size. In fact, the naïve assumption that the two conditions are similar enough across experiments could justify the use of a *t*-test with statistically significant results: $M\_FAE = 13$, CI_95% = [7, 20]), $t(17) = 4.24$, $p = .0005$. These considerations suggest that a genuine FAE may exist in masked priming but might be smaller than the magnitudes that are statistically detectable in most previous experiments. This interpretation is supported by the results from Adelman et al. (2014) in a large scale, multi-site lab-based study on orthographic priming. They report a small but reliable FAE, but caution this effect could simply be an orthographic neighborhood effect masquerading as a frequency effect, due to the high correlations between the two variables.

In addition, another potential contributor to past discrepancies is the reliance on the dated Kučera and Francis (1967) word frequency database, which 15 out of 18 studies have depended on. This poses a potential problem, as this frequency database has consistently demonstrated inferior predictive performance in psycholinguistic experiments, particularly with low-frequency words, compared to more contemporary databases (Burgess and Livesay 1998; Zevin and Seidenberg 2002; Balota et al. 2004; Brysbaert and New 2009; Yap and Balota 2009; Brysbaert and Cortese 2011; Gimenes and New 2016; Herdağdelen and Marelli 2017; Brysbaert, Mandera, and Keuleers 2018). Both of these issues are addressed in the subsequent sections.

2.1. *Issues with frequency databases*

Due to the well-documented concerns over the reliability of the Kučera and Francis (1967) frequency database for psycholinguistic experiments (Burgess and Livesay 1998; Zevin and Seidenberg 2002; Balota et al. 2004; Brysbaert and New 2009; Yap and Balota 2009; Brysbaert and Cortese 2011; Gimenes and New 2016; Herdağdelen and Marelli 2017; Brysbaert, Mandera, and Keuleers 2018), our studies exclusively sourced materials from the HAL (Lund and Burgess 1996) and SUBTLEX$_{US}$ (Brysbaert and New 2009) databases, which reflect more recent linguistic usage and offer better validation in behavioral experiments (e.g., Balota et al. 2004; Brysbaert and New 2009;

Yap and Balota 2009; Brysbaert and Cortese 2011; Gimenes and New 2016; Herdağdelen and Marelli 2017). While these databases outperform Kučera and Francis (1967) in predicting psycholinguistic task outcomes, it is important to note potential discrepancies in individual frequency counts, particularly in the low and mid-frequency ranges. It is possible that this variation, attributable to the primary genre of their sources (USENET groups for HAL and movie subtitles for SUBTLEX$_{US}$),[1] may not have an oversized impact on megastudies with very large word samples (e.g., Balota et al. 2004; Brysbaert and New 2009; Yap and Balota 2009; Brysbaert and Cortese 2011; Gimenes and New 2016; Herdağdelen and Marelli 2017). However, corpus-specific frequency skew can become significant when dealing with smaller samples of words, as is the case in most masked priming studies (cf. Adelman et al. (2014)). Table 2 illustrates the potential discrepancy in considering words as high or low frequency based on the different aforementioned databases.

Table 2. Example of frequency count imbalances (in occurrences per million) across the frequency norms of Kucera & Francis (KF), HAL and SUBTLEX$_{\mathrm{US}}$ for 4 to 6 letter words.

Word	KF	HAL	SUBTLEX$_{\mathrm{US}}$
Skew in KF			
negro	104	3	5
poet	99	9	9
mercer	71	4	2
swung	48	3	2
mantle	48	8	2
Skew in HAL			
web	6	351	9
user	4	297	2
mint	7	211	5
format	9	198	1
warp	4	125	5
Skew in SUBTLEX$_{US}$			
daddy	4	16	185
bitch	6	24	169
cute	5	28	88
pardon	8	12	65
steal	5	28	53

2.2. *Issues with statistical power*

The inconsistency of past findings regarding the FAE in masked priming has been linked to a potential lack of statistical power in previous research (Bodner and Masson 1997, 2001; Masson and Bodner 2003; Adelman et al. 2014). This is a reasonable concern,

[1]A separate, though relevant issue which cannot be addressed here is to how to mitigate the discrepancies across the databases available, but see Yap and Balota (2009), and Brysbaert and Cortese (2011) for proposals about combining the frequency counts from different corpora.

as interactions like the FAE often require larger sample sizes for statistical detection (Potvin and Schutz 2000; Brysbaert and Stevens 2018) compared to main effects. We outline below three ways in which neglecting statistical power might frustrate our understanding of FAE in masked repetition priming.

First, our literature review revealed crucial gaps in the reporting of relevant statistical information, which impedes the assessment of the statistical power attained by past experiments. The inconsistent reporting of each conditions' standard deviations (in only 7 out of 18 studies) and the complete absence of reporting of the correlation structure between conditions complicates power assessments. Researchers are thus forced to explore a range of plausible values for standard deviations and correlation structures on their own.

Table 1 details our attempt to conduct power simulations for two hypothesized frequency attenuation effect sizes: 15 ms (close to the averaged FAE of 13 ms) and 30 ms (close to the only three observed statistically significant FAE in English). Standard deviations (ranging between 60 ms and 180 ms, in 10 ms increments) and correlation between conditions (uniformly set to range between 0.6 and 0.9, with 0.1 unit increments) were simulated for each study's sample size, with 10,000 replications for each simulation. These range of values were derived from our literature review and previous in lab and online experiments (Petrosino 2020; Petrosino, Sprouse, and Almeida 2023). For each simulated dataset, a paired *t*-test was performed comparing the repetition effect for high frequency words and low frequency words. This calculation is mathematically identical to the interaction term in a 2x2 factorial repeated-measures design^[the resulting t value, when squared, is equal to the F value for the interaction calculated in the 2x2 repeated-measures ANOVA), but it is less computationally expensive to perform in large scale simulations. Power to detect this interaction was then calculated as the proportion of statistically significant tests (= 5%) obtained across replications. All else being equal, standard deviations and correlations between conditions have opposite effects on statistical power: increases in standard deviations lead to less power, while increases in correlation between conditions lead to more power.

The results reported in Table 1 reveal a wide range of possible statistical power attained by previous studies, depending solely on the combination of plausible standard deviation and correlation across conditions. For instance, the study with the smallest sample size (Forster et al. 1987, N=16) had a 2% to 24% chance of detecting a 15 ms frequency attenuation effect and a 4% to 84% chance to detect a 30 ms effect. Similarly, the study with the largest sample size (Rajaram and Neely 1992, N=48) exhibited a range of 4% to 76% for a 15 ms frequency attenuation effect and 9% to 100% for a 30 ms effect. As a consequence of the limited reporting of relevant statistical information in past studies, it is nearly impossible to determine if any of them were adequately powered to detect the effect of interest.

A second concern arising from the ambiguity surrounding statistical power in the literature is the potential impact of a prevalence of low-powered experiments on the scientific record. An excess of such experiments increases the risk of observed statistically significant effects being spurious (Button et al. 2013). As highlighted in Table 1, only 4 out of 18 studies demonstrate a statistically significant FAE. The absence of clarity regarding the statistical power of previous research poses challenges in assessing the likelihood of these significant findings being spurious.

Finally, it is widely acknowledged that experiments with approximately 50% power are akin to a coin toss in their ability to detect a true effect (Cohen 1992). A less-appreciated fact is that, in the presence of even lower power (<25%), statistically significant results can substantially overestimate the effect size – a type-M error (Gelman and Carlin 2014). When power drops to levels below 10%, a statistically significant result may occur even when the observed effect goes in the opposite direction of the true effect – a type-S error (Gelman and Carlin 2014). Our power simulations for within-subjects data revealed a similar relationship between statistical power, type-M, and type-S errors in line with the observations detailed by Gelman and Carlin (2014) for the independent samples t-test. For instance, at 10% power (a possibility for virtually all previous studies, as indicated in Table 1), a statistically significant result could indicate an overestimation of the magnitude of the frequency attenuation effect by a factor between 2 and 5, with up to a 5% chance of incorrectly determining the direction of the effect.

The two studies reported here were designed to mitigate these two confounding issues: the overreliance on the Kučera and Francis (1967) frequency data as well as a potential lack of statistical power observed in previous research. As a large increase in statistical power requires a large sample size, Experiment 1 aimed to assess the suitability of using *Labvanced* (Finger et al. 2017), an online platform for running web browser-based experiments, for running masked priming studies online.

3. Experiment 1

As evident in Table 1, conducting a properly powered experiment for a FAE close to the averaged value calculated from previous studies requires a sample size that would be impractical to pursue in standard university research settings, i.e. typically quiet lab rooms with a small number of research computers. In response to this challenge, the present study was exclusively conducted online, leveraging the growing trend in online behavioral research facilitated by HTML5 capabilities and the availability of advanced web software such as *jsPsych* (de Leeuw 2014), *PsychoJS* (the JavaScript counterpart of *PsychoPy*, Peirce et al. (2019)), *Gorilla* (Anwyl-Irvine et al. 2020), and *Labvanced* (Finger et al. 2017). Notably, three recent studies have already demonstrated the viability of conducting masked priming experiments online, employing different software tools: Angele et al. (2023) with *PsychoJS*, Cayado, Wray, and Stockall (2023) with *Gorilla* and Petrosino, Sprouse, and Almeida (2023) with *Labvanced*. In this study, we opted for *Labvanced* (Finger et al. 2017), given our previous successful experience with it (Petrosino, Sprouse, and Almeida 2023). Similar to *Gorilla*, *Labvanced* eliminates local installation issues, ensuring cross-platform consistency and simplifying experimental design without necessitating proficiency in additional programming languages.

We determined the sample size large enough to guarantee an acceptable statistical power (> 80) by a full-fledged power analysis specifically targeting what we construed as the smallest theoretically interesting FAE (i.e., 5ms). The details of the power analysis is available as supplemental material of the paper. Similarly, the code used for the power simulations, along with the simulated datasets are available online (https://osf.io/r7d2q/). Our analysis identified a sample size of 1,250 participants as optimal, ensuring robust statistical power especially for a raw FAE equal to or exceeding 10 ms —- a value closely aligned with the average FAE calculated from previous studies (cf. Table 1). In light of the limitations in the temporal accuracy and precision of current online stimulus delivery programs (observed in several pilots and previous published studies conducted in our lab), we aimed for an intended sample size of 2,600. This decision was made to enhance the likelihood of obtaining a sample size of at least 1,250 participants after applying all the necessary exclusion criteria to the data. In addition, sample sizes exceeding 1,250 can only help increase the precision of the estimated effect size.

In this experiment, the prime duration was set at 33 ms. The motivation for the choice of such a short prime duration (as compared to the literature, in which it is usually between 50 and 60 ms; see Table 1) is threefold. First, previous experiments on *Labvanced* (Petrosino, Sprouse, and Almeida 2023) showed that, due to the inherent difficulties in presenting stimuli for very short set durations in the browser, a longer set duration would increase the number of trials in which the prime duration would rise above the subliminal threshold (usually thought to be around 60 ms) due to timing inaccuracies and missing screen refreshes, which could trigger the adoption of experiment-wide strategies in the task, and ultimately contaminate the masked priming response (Zimmerman and Gomez 2012). Second, Angele et al. (2023), Cayado, Wray, and Stockall (2023) and Petrosino, Sprouse, and Almeida (2023) have demonstrated that a 33 ms priming duration is sufficient to elicit repetition priming effects in online experiments. Finally, setting such a short prime duration prevents virtually everyone from consciously perceiving the prime word Nievas (2010), and thus presents a less contaminated estimate of early putatively automatic processes in word recognition.

3.1. *Methods*

3.1.1. Preregistration

We preregistered the results of the power analysis, the goals, the design and analysis plan for this experiment prior to data collection. The preregistration, detailing the experimental hypotheses, the desired sample size as well as the planned analyses is available online (https://doi.org/10.17605/OSF.IO/3NFQP).

3.1.2. Participants

Two thousand and six hundred participants (1,445 females; *mean age* = 42, *sd age* = 14) were recruited on Prolific (https://www.prolific.com). Several criteria were selected to ensure recruitment of native speakers of U.S. English. Participants had to be born in the Unites States of America, speak English as their first and only language, and have no self-reported language-related disorder. We encouraged participants to avoid any sort of distraction throughout the experiment, and to close any program that may be running in the background. Because the experiment was run online, participants could not be monitored during data collection. Finally, to further reduce variability across participants' devices, we restricted the experiment to be run on Google Chrome only, which is the most used browser worldwide (W3 Counter 2023), and reportedly performs better than any other across operating systems (likely thanks to the *Blink* engine; see Lukács and Gartus 2023).

3.1.3. Design

The masked priming procedure relied on a lexical decision task (LDT), in which a 2 (frequency: *high* vs *low*) x 2 (prime type: *repetition* vs *unrelated*) factorial design was used. Both factors were manipulated within-subjects. The dependent variables were lexical decision latency (RT, in milliseconds) and error rate (in percentages).

3.1.4. *Materials*

One-hundred and four five-letter words, half of low frequency (between 7 and 24 in the SUBTLEX$_{US}$ frequency per million) and half of high frequency (between 57 and 2,961 in the SUBTLEX$_{US}$ frequency per million) were sampled from ELP (Balota et al. 2007), but this time based on the SUBTLEX$_{US}$ frequency counts rather than HAL. Table 3 shows that although the SUBTLEX$_{US}$ frequency ranges of the two conditions were very far from one another they still show some overlap when HAL frequencies are used. As mentioned before, this seems to be a general problem when jointly considering different frequency databases for a smaller set of stimuli that need to be manipulated and controlled in different ways (see also fn. 1 and Adelman et al. (2014)). From each condition, 26 words were selected to be presented as targets and related primes (the *repetition* condition), and the remaining 26 were presented as unrelated primes (the *unrelated* condition). All word items were also controlled for orthographic neighborhood (i.e., Coltheart's N): $t \approx 0$. All words used were monomorphemic nouns, adjectives, or verbs, thus excluding particles, prepositions, and derived or inflected forms.

Table 3. Experiment 1. Descriptive statistics of the word items used. For both frequency databases, the word frequencies were converted to per-million count to ensure cross-comparison.

		HAL				**SUBTLEX$_{US}$**				**Orthographic *N***			
frequency	N	min	max	mean	SD	min	max	mean	SD	min	max	mean	SD
high	52	45	4984	573	808	57	2691	210	388	0	10	3.98	2.60
low	52	6	570	64	93	7	24	13	5	0	11	3.92	2.79

One-hundred and four five-letter, phono-orthographically legal non-words were randomly selected from the ELP database as well. Half of them (i.e., 52) were randomly selected to be presented as targets; the other half was instead used as unrelated non-word primes. None of the non-words contained any existing English morpheme. Both the words and non-words used in the experiments are reported in the appendix below. In addition, all items had a reported error rate smaller than 10%, so to ensure that they were all clearly distinguishable by participants.

3.1.5. *Procedure*

Each recruited participant was assigned one of two word lists, which differed only in the relatedness of the prime with respect to the target; otherwise, the two lists presented the same set of target words and nonwords (i.e., 104 pairs for each list). In one list, the three conditions (the high- and low-frequency word conditions, and the non-word condition) had half of the target items being preceded by themselves (the *repetition* condition) and half of the target items being preceded by one of the unrelated primes belonging to the same frequency bin (the *unrelated* condition). In the other list, these assignments were reversed. The order of stimulus presentation was randomized for each participant.

After being recruited in the *Prolific* online platform, participants were asked to click on a link redirecting them to the Labvanced online service. During the experiment, they were asked to perform a lexical decision task by pressing either the 'J' (for word) or 'F' (for non-word) keys on their keyboard. Each trial consisted of three different stimuli appearing at the center of the screen: a series of five hashes (#####) presented for 500 ms, followed by a prime word presented for 33 ms, and finally the target word; the target word disappeared from the screen as soon as a decision was made.

Participants were given 5 breaks throughout the experiment. When the experiment was over, the participants were then redirected to Prolific in order to validate their submission. The median time to finish the experiment was 6 minutes. Each participant was paid with a standard rate of GBP 9/hour.

3.2. *Data analysis*

Analysis scripts and an abridged version of the data collected can be found online (https://osf.io/vn3r2), and consisted of 297,598 observations in total. We performed the same three steps of analysis described for experiment 1 (Section 3.2).

3.2.1. *Step 1: subject and item performance*

Item and subject error rates were calculated. The item error rate was never above 14%, so no item was excluded from analysis. 19 subjects were removed because their error rate was above 30%. Thus, a total of 269,652 observations and 2,593 participants were included in further analyses.

3.2.2. *Step 2: prime durations*

Prime fluctuations were dealt with in the same way as in experiment 1 (Section 3.2.2). The mean (mean = 32.32 ms, sd = 15) and the median (median = 33 ms) prime durations were closer to the intended value (33 ms). The same prime duration cut-off set for experiment 1 (i.e., any trial whose prime duration was out of the 25-60ms range) removed 13 % of the trials. No participant was excluded, for a total of 237,287 observations.

3.2.3. *Step 3: RT analysis*

After removing the incorrect responses, similarly to experiment 1 (Section 3.2.3), 0.51% of the trials were excluded if their corresponding RT was below 200 ms or above 1800 ms. Finally, 249 subjects were removed because the number of trials within the same condition was less than 7 (i.e., about half of the total number of trials being presented within the same condition, i.e. 13). A total of 210,889 observations and 2,341 subjects were included in the statistical analysis below.

3.3. *Results*

Table 4 below report the descriptive statistics of the experiment. For each frequency condition, priming effects were calculated by subtracting the mean RT to the related condition to the mean RT from the unrelated condition. We ran two different analyses, for both RT and error data. First, we ran a 2x2 repeated-measures ANOVA (condition, 3 levels: high vs. low vs. non-word; primetype, 2 levels: unrelated vs. repetition) on the by-subject averaged data, paired with planned comparisons between related and unrelated primetype for each condition. In addition, a Generalized Linear Mixed Model (GLMM) analysis on the raw RT and the raw accuracy data (rather the by-subject means or error rates used in the ANOVA analyses). Unlike linear mixed-effect models, GLMMs do not assume normal distribution of the data, and are therefore particularly useful for non-normally distributed data such as RTs and accuracy. For the RT data, a Gamma distribution was used; for the accuracy data, a binomial distribution was used instead, both with an identity link between fixed effects and the dependent variable (Lo and Andrews 2015). To prevent converge failure, the model was kept as simple as possibly, with condition, primetype, and their interaction as fixed effects, and subjects and items as random intercepts. Before fitting the model, the contrasts were also set to sum-to-zero contrasts (i.e., by using the R function `contr.sum()`) to facilitation interpretation of main and interaction effects. The fitting was performed by using the `lme4` R-package (Bates et al. 2015) with the Laplace approximation technique, using 1 million iterations and the BOBYQA optimizer to help convergence. Then, we used the function `Anova()` from the `car` R-package (Fox and Weisberg 2019) to obtain estimates and probability values for fixed effects calculated for Type-III sums of squares.

In the word analysis, the ANOVA analysis showed significant main effects (condition: $F(1, 2340)=1572$, $p<.0001$; primetype: $F(1, 2340)=1113$, $p<.0001$) and their interaction $F(1, 2340)=52.48$, $p<.0001$). Planned comparisons confirmed statistically significant repetition priming effects for both word conditions (*MOP_HF*=18 ms, *CI_95%*=[16 20], $t(2340)=19.7$, $p<.0001$; *MOP_LF* = 28 ms, *CI_95%*=[26 30], $t(2340)=27.8$, $p<.0001$), with the low-frequency word repetition priming effect being 10 ms larger than the high-frequency word repetition priming effect. This FAE effect was statistically significant (*M_FAE*=10 ms, *CI_95%*=[7 13]), $t(2340)=7.24$, $p<.0001$l). A very small but statistically significant inhibitory priming effect was observed in the non-word condition (*MOP_NW*=-2 ms, *CI_95%*=[-4 0], $t(2340)=-2.33$, $p<.0001$). Similarly, the GLMM analysis confirmed that both main effects (condition: $\chi^2 = 2978.35, p < .0001$; primetype: $\chi^2 = 1531.12, p < .0001$) and their interaction ($\chi^2 = 870.27, p < .0001$) were significant.

In the error analysis, the ANOVA analysis revealed significant effects for both main effects (condition: $F(1, 2340)=392.5$, $p<.0001$; primetype: $F(1, 2340)=380.5$, $p<.0001$) and interaction $F(1, 2340)=55.47$, $p<.0001$). Planned comparison revealed significant priming effects in the form of fewer errors in the repeated compared to unrelated trials in the all conditions (high: $t(2340)=9.95$, $p<.0001$; low: $t(2340)=16.9$, $p<.0001$; non-word: $t(2340)=-3.27$, $p=.001$). As in the ANOVA error analysis, both main effects (condition: $\chi^2 = 30.45, p < .0001$; primetype: $\chi^2 = 108.88, p < .0001$) and their interaction ($\chi^2 = 307.57, p < .0001$) were significant.

Table 4. Experiment 1. Summary of the word priming results. *Legend.* MOP: magnitude of priming.

	unrelated RT			repetition RT				priming effects				t-test		
factor	mean	SD	Error (%)	mean	SD	Error (%)	cor	MOP	95% CI	SD_p	ES	t	df	p
high	573	83	3	555	85	2	0.860	18	[16 20]	45	0.41	19.7	2340	2.88e-80
low	605	88	6	577	88	3	0.850	28	[26 30]	49	0.58	27.8	2340	1.52e-147
non-word	623	103	4	625	103	4	0.910	-2	[-4 0]	43	-0.05	-2.33	2340	0.0197
frequency:primetype							0.029	10	[7 13]	66	0.15	7.24	2340	5.86e-13

3.4. *Discussion*

Experiment 1 was designed to investigate whether Frequency Attenuation Effects (FAE) can be detected under masked priming conditions (with SOA = 33 ms). We managed to recruit a very large sample size ($N = 2341$) to ensure adequate statistical power for detecting even small effect sizes. Our results not only showed statistically significant main effects of repetition for high and low frequency words alike, but also detected a statistically significant interaction: the low-frequency condition yielded priming effects that were 10 ms larger than the high-frequency condition with a margin of error as narrow as 5 ms. The absence of a robust non-word masked priming response has been used as an additional piece of evidence supporting the view that the masked priming response stems from lexical memory and is devoid of episodic influences (e.g., Forster 1999). The results of this experiment align with the previous evidence in showing at best very small inhibitory masked repetition priming for non-words, with very high precision: the 95% CI indicates the plausible range for the masked repetition priming effect for non-words to be between -4 and 0 ms when prime duration is 33 ms.

4. Experiment 2

The results of experiment 1 showed that with the right sample size, the repetition priming response may indeed be modulated by word frequency. However, the experiment was carried out with a unusually short prime duration (i.e., 33 ms). In previous pilot experiments carried out at longer durations (e.g., 48 ms), we observed that the distribution of the prime duration tended to be more positively skewed, with a substantial increase of the number of trials with a prime duration above the subliminal threshold of 60 ms. While being aware of these methodological limitations, we carried out a follow-up experiment that elicited the repetition priming response to the same materials used in experiment 1, with a longer prime duration (i.e., 48 ms), which is more in line with the literature on the topic. The aim for this experiment was three-fold. From the methodological point of view, it may show that a shorter prime duration in online masked priming experiment may prevent unnecessary data loss, while still providing reliable results. From the theoretical point of view, it may provide further evidence on two separate issues. First, it may inform on the interaction between priming and prime duration. In particular, we expect the longer prime duration to elicit a larger facilitation, and therefore a larger magnitude of priming. Second, and more importantly for the question being asked here, it may further inform on the size of the interaction between priming and word frequency.

4.1. *Methods*

4.1.1. Participants

Two thousand and six hundred participants (1,551 females; *mean age* = 39, *sd age* = 12) were recruited on Prolific (https://www.prolific.com) with the same criteria specified for experiment 1 (Section 3.1.2). Prior to recruitment, the participants participating in experiment 1 were excluded from the pool, so that the participants recruited for experiment 2 was completely different from the participants recruited for experiment 1.

4.1.2. Design

The experimental design was identical to experiment 1.

4.1.3. Materials

The experimental items were the same as experiment 1.

4.1.4. Procedure

Experiment 2 followed the same procedures as experiment 1 (see Section 3.1.5). The median time to finish the experiment was the same as experiment 1 (i.e., about 6 minutes).

4.2. *Data analysis*

Analysis scripts and an abridged version of the data collected can be found online (https://osf.io/k3gpc), and consisted of 295,940 observations in total. We performed the same three steps of analysis described for experiment 1 (Section 3.2).

4.2.1. Step 1: subject and item performance

Item and subject error rates were calculated. Similarly to experiment 1, the item error rate was never above 12%, so no item was excluded from analysis. 39 subjects were removed because their error rate was above 30%. Thus, a total of 265,982 observations and 2,558 participants were included in further analyses.

4.2.2. Step 2: prime durations

Prime fluctuations were dealt with in the same way as in experiment 1 (Section 3.2.2). The mean (mean = 50.11 ms, sd = 11.13) and the median (median = 50 ms) prime durations were closer to the intended value (50 ms). The same prime duration cut-off set for experiment 1 (i.e., any trial whose prime duration was out of the 25-60ms range) removed 20 % of the trials. As compared to experiment 1, experiment 2 had therefore a 7% larger percentage of removed out-of-range trials, the majority of which (18%) were above the subliminal threshold. This is in sharp contrast with the distribution of the prime durations in experiment 1, where the trials above the range amounted to only 0% of the dataset. As observed in previous studies and pilot conducted in our lab, this distribution suggests that setting the prime duration closer to either limit of a given range has the side effect of allowing for more fluctuations around either limit, thus potentially leading to greater data loss (see also further below). No participant was excluded, for a total of 213,078 observations.

4.2.3. Step 3: RT analysis

After removing the incorrect responses, similarly to experiment 1 (Section 3.2.3), 0.59% of the trials were excluded if their corresponding RT was below 200 ms or above 1800 ms. Finally, 634 subjects were removed because the number of trials within the same condition was less than 7 (i.e., about half of the total number of trials being presented within the same condition, i.e. 13). A total of 168,195 observations and 1,924 subjects were included in the statistical analysis below. The substantial number of subjects being removed at this final stage was the ultimate side effect of the increased number of out-of-range trials that were removed during the previous step of analysis, and provides further evidence of the risks with setting the prime duration closer to the upper bound of the subliminal range.

4.3. *Results*

Table 5 below report the descriptive statistics of the experiment. For experiment 2, we ran the same statistical analyses as experiment 1. In the word analysis, a 2x2 repeated-measures ANOVA revealed significant main effects (condition: $F(1,\ 1923)$=987.4, p<.0001; primetype: $F(1,\ 1923)$=1447, p<.0001) and interaction $F(1,\ 1923)$=36.82, p<.0001). Planned comparisons confirmed statistically significant repetition priming effects for both word conditions (*MOP_HF*=26 ms, *CI_95%*=[24 28], t(1923)=24.6, p<.0001; *MOP_LF* = 35 ms, *CI_95%*=[33 37], t(1923)=30.2, p<.0001), with the low-frequency word repetition priming effect being 10 ms larger than the high-frequency word repetition priming effect. This FAE effect was statistically significant (*M_FAE*=9 ms, *CI_95%*=[6 12]), t(1920)=6.07, p<.0001). A very small but statistically significant inhibitory priming effect was observed in the non-word condition (*MOP_NW*=-4 ms, *CI_95%*=[-6 -2], t(1923)=-3.53, p=.0004). The GLMM analysis confirmed significant effects for condition ($\chi^2 = 2613.3, p < .0001$), primetype ($\chi^2 = 1700.5, p < .0001$), and their interaction ($\chi^2 = 1158.9, p < .0001$).

In the error analysis, the ANOVA showed significant effects for both main effects (condition: $F(1,\ 2340)$=392.5, p<.0001; primetype: $F(1,\ 2340)$=380.5, p<.0001) and interaction $F(1,\ 2340)$=55.47, p<.0001). Planned comparisons confirmed significant priming effects in the form of fewer errors in repeated compared to unrelated trials in the all conditions (high: t(1923)=9.30, p<.0001; low: t(1923)=15.8, p<.0001; non-word: t(1923)=-2.32, p=.002). Similar results were obtained in the GLMM error analysis (condition: $\chi^2 = 51.50, p < .0001$; primetype: $\chi^2 = 97.42, p < .0001$; interaction: $\chi^2 = 283.77, p < .0001$.

Table 5. Experiment 2. Summary of the word priming results. *Legend.* MOP: magnitude of priming.

	unrelated RT			repetition RT				priming effects				*t*-test		
factor	mean	SD	Error (%)	mean	SD	Error (%)	cor	MOP	95% CI	SD_p	ES	t	df	p
high	574	83	3	548	83	2	0.850	26	[24 28]	46	0.56	24.595576	1923	2.27e-116
low	605	89	6	570	89	3	0.840	35	[33 37]	51	0.69	30.185474	1923	3.51e-164
non-word	629	108	4	633	110	5	0.900	-4	[-6 -2]	50	-0.08	-3.525726	1923	4.32e-04
frequency:primetype							0.024	9	[6 12]	68	0.13	6.068254	1923	1.55e-09

4.4. ***Discussion***

Experiment 2 was designed to investigate whether a frequency attenuation effect akin to the one detected in experiment 1 at SOA of 33 ms can be detected at the longer, and most commonly used SOA of 48 ms. To this end, we used the same stimuli and recruited the same sample size as experiment 1, and only set the prime duration accordingly. First, we found that the masked repetition priming to both high and low-frequency words were respectively about 8 ms larger in experiment 2 than in experiment 1, resulting from the longer prime duration (Forster, Mohan, and Hector 2003). Nevertheless, the significant interaction effect size amounted to 9 ms, which is only 1 ms away from the estimate of the interaction in experiment 1 (10 ms), but well within its 95% CI (which ranged from 7 ms to 13 ms). Finally, similarly to experiment 1, the magnitude of non-word masked repetition priming response in experiment 2 was inhibitory, though slightly larger (-4 ms, as compared to -2 ms of experiment 1).

At the face value, experiment 2 is very similar to experiment 1 on several respects. First, they involved the exact the same stimuli and the same sample size from the same pool (i.e., the Prolific pool; crucially though, different participants were recruited for each experiment). Second, they were both analyzed with the exact same analysis pipeline, and, in particular, with the same criterion to detect trials with outlying prime durations. Third, the estimates (means, SDs, error percentages, and correlations) of both experiments are numerically very close to one another, with the maximal differences almost exclusively present in the repetition conditions. For these reasons, one may therefore argue against its methodological and theoretical validity, and may deem the differences of the magnitude of priming across the two experiments might have just been due to sampling error having two major sources. The first source of sampling error is the inherent imprecision in the presentation of the prime. Regardless of the preset duration of the prime, there was an inherent inaccuracy in the actual presentation of the prime in both experiments. The analysis pipeline (and, in particular, with the criterion to detect trials with outlying prime durations) described in detail above might have therefore led to two datasets with a similar distribution of the prime durations. Figure 1 shows the distributions of the prime durations for both datasets after the last analysis step. The two distributions minimally overlap, with most of the trials peaking at about the relative preset duration (i.e., 33 and 48 ms). The difference between the two distribution was also statistically significant: $t(394431) = -868.47, p < .0001$.

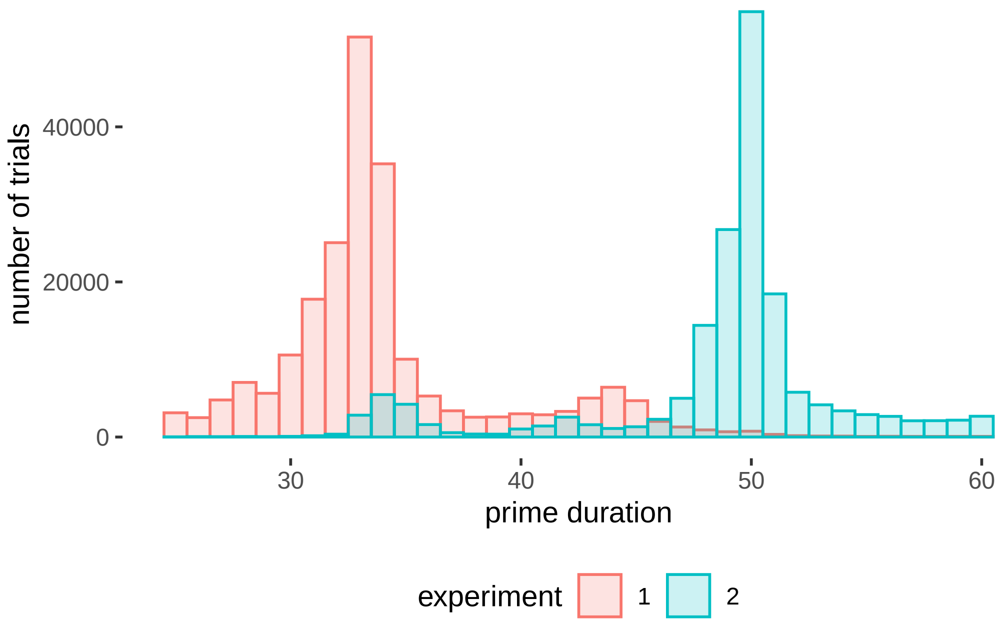

Figure 1. Distribution of the prime durations across the two experiments.

The second source of sampling error is equally unavoidable and present in every study: participant selection, and is ensured by the fact that recruiting sample sizes approaching infinity is virtually impossible. One way to gauge the extent to which such source of sampling error might have affected the differences in the effect sizes of the two experiments is to calculate the *prediction intervals* on the results of experiment 1, and check if the results of experiment 2 fall within them. Unlike confidence intervals, the lesser-known prediction intervals are calculated around the original study's mean (rather than the population mean), and express the amount of deviation that a future replication of the study may allow for due to sampling error (Spence and Stanley 2016). In our case, if the results of experiment 2 falls within the prediction interval above, it would suggest that the deviation between the two experiment was just due to sampling error, and therefore experiment 2 is a replication of experiment 1. Conversely, if the results of experiment 2 falls within the prediction interval above, it would suggest that the deviation between the two studies could not be due to sampling error, and therefore experiment 2 is not a replication of experiment 1. The prediction intervals for the means of the three conditions and the interaction between the two word condition (i.e., FAE) of experiment 1 were calculated by using the function `pi.m()` of the R package `predictionInterval` (available in the CRAN repository), and are reported in Table 6 below. In experiment 2, while the estimates of the non-word condition and the FAE fall within the relative prediction interval (*MOP_NW* = -4 ms, *PI_95%* = [-5 1]; *FAE* = 9 ms, *PI_95%* = [6 14]), the estimates of both word conditions are respectively 5 and 4 ms above the corresponding PIs (*MOP_HF* = 26 ms, *PI_95%* = [15 21]; *MOP_LF* = 35 ms, *PI_95%* = [25 31]). Our interpretation of these results is three-fold. First, the fact that the the word priming effects of experiment 2 (for both the high and low-frequency conditions) are outside the prediction intervals of experiment 1 suggest that the effects elicited in experiment 2 may not be considered as mere replications of the effects elicited in experiment 1. Rather, they corroborate the view that priming may significantly benefit from a longer prime duration. Second, the masked repetition non-word priming effect size being replicated in both experiments further corroborates the commonly-accepted view that the masked priming response taps into lexical memory, rather than pure orthographic decoding. Finally, and more importantly for the main purpose of the paper, the estimate of the FAE of experiment falling within the relative prediction interval further confirms the presence of an interaction effect between priming and frequency, while not being contingent on the duration of the prime.

Table 6. Prediction intervals calculated on the means and standard deviations (SD) of the conditions tested in experiment 1.

			sample size		
factor	mean	SD	experiment 1	experiment 2	95% PI
high	18	45	2341	1924	[15 21]
low	28	49	2341	1924	[25 31]
non-word	-2	43	2341	1924	[-5 1]
frequency:primetype	10	66	2341	1924	[6 14]

5. General discussion

The repetition priming response stands as a cornerstone in psycholinguistic investigations, offering insights into the mechanisms governing word recognition. An ongoing debate surrounds the interpretation of these effects, particularly concerning their source in the memory system. On the one hand, *interactive activation models* (McClelland and Rumelhart 1981; Grainger and Jacobs 1996; Coltheart et al. 2001) posit a lexical source for repetition priming effects, either in terms of temporarily raised resting activation levels for lexical nodes in unmasked priming, or as a head start in the retrieval process in masked priming. *Episodic* and *memory recruitment models* (Jacoby and Dallas 1981; Jacoby 1983; Bodner and Masson 1997; Masson and Bodner 2003; Bodner and Masson 2014) on the other hand, invoke a non-lexical source for the repetition effect, namely an episodic or episodic-like memory resource formed upon brief exposure to the prime word that can be recruited during the processing of the target item. Crucially, both models predict a single mechanism underlying masked and unmasked priming. Differential mechanisms between unmasked and masked repetition priming, however, are predicted by the *entry-opening model* (Forster and Davis 1984), which propose both lexical and episodic sources of priming effects.

Thus, the existence of qualitatively distinct outcomes in masked and unmasked priming presented a direct challenge to some, but not all of these models. One such finding is the *Frequency Attenuation Effect* (FAE), in which higher frequency words exhibit smaller repetition effects compared to lower frequency words. The FAE has been described as observable only in unmasked priming since the work of Forster and Davis (1984), who demonstrated that when the prime word is presented very briefly (SOA $<$ 60 ms), it becomes masked by the target word, and this is hypothesized to prevent the conscious encoding of the prime. Under such conditions, the FAE purportedly disappears. Forster and Davis (1984) argued that this potentially shows that the FAE is subserved by a different type of memory source (perhaps episodic) than the masked repetition priming response. This conclusion, however, is the source of ongoing debates (see Table 1 for review of past findings), which the two experiments reported here were meant to address.

Within this research landscape, our experiments targeting the frequency sensitivity of the repetition effect under masked conditions contribute methodological and theoretical insights. Methodologically, our results help establish the viability and reliability of online data collection for the masked priming paradigm, building on the work of Angele et al. (2023), Cayado, Wray, and Stockall (2023) and Petrosino, Sprouse, and Almeida (2023).

In the same vein, the FAEs observed in experiments 1 and 2 have important theoretical ramifications. The historical belief that FAE fails to obtain in masked priming arose from a lack of statistically significant results. These were possibly rooted in the reliance of outdated frequency corpora by earlier experiments or inadequate statistical power to detect plausible effect sizes. Our design addressed these concerns, yielding statistically significant FAE results aligning with the literature's average effect (see Table 1; the 95% CI implies that the FAE is unlikely to be larger than 13 ms with a 33 ms prime duration). These results challenge the supposed qualitative distinction between masked and unmasked repetition priming cleaved by the FAE, complicating the rejection of single-mechanism theories, and suggesting that *interactive-activation models* and *memory recruitment models* may yet offer unifying explanations for masked and unmasked priming.

Similarly, our results also challenge the entry-opening model's prediction of the absence of FAE in masked priming. One potential way of dealing with this in the *entry opening model* is to claim that masked priming severely reduces, but does not entirely eliminate, the use of sources other than lexical memory (see Forster 1998; Forster, Mohan, and Hector 2003, for proposals along this line). Alternatively, within the entry-opening model, the results of this study may be explained by the frequency-based mechanism occurring in the fast search stage. A potential mechanism in this direction was already hinted at by Forster and Davis (1984) themselves, and consists of a procedure, whereby during the fast search stage, the entry of a prime word is promoted to the top position of the search list. As a consequence, low-frequency words (which are fairly low in the search list) will benefit from such promotion procedure more than high-frequency words (which are instead already in higher positions), thus ultimately giving rise to the FAE.

While our findings present a compelling case for the presence of FAE in masked priming that is seemingly parallel to the unmasked case, questions about potential mechanistic differences persist. The larger sample size needed for masked FAEs raises intriguing considerations about the influence of memory sources and warrants further investigation. For example, there is independent evidence for different mechanisms in masked and unmasked repetition priming from RT distributional analyses (cf. Gomez, Perea, and Ratcliff (2013)) that suggests that repetition priming under masked conditions affect primarily the encoding stage of the stimulus. Given that frequency is often associated with facilitation of encoding, our results could help support this view. Additionally, the trivially small inhibitory effect sizes of non-word masked repetition priming in experiments 1 and 2 align with the trend (overwhelmingly shown in the literature) that facilitatory effect may be exclusive to unmasked designs (Forster 1998; Forster, Mohan, and Hector 2003; but see Masson and Bodner 2003), and suggests avenues for future exploration.

Finally, the finding that the FAE occurs under masked priming conditions may impact our understanding of masked morphological priming. In this literature, there is a unresolved question about the ability of affixes to elicit masked morphological priming results (for a review, Amenta and Crepaldi 2012). In English, the evidence seems to indicate that only stems, but not affixes, have the ability to prime entries across the lexicon. This finding can and has been used to support models in which affixes are initially stripped before stems are accessed in the lexicon (Taft and Forster 1975; Forster and Azuma 2000; Stockall and Marantz 2006). However, stems and affixes do also have a large frequency imbalance, with most affixes being substantially more frequent that most stems. The observation of FAE under masked priming can provide an alternative reason for why masked stem morphological priming is well attested but masked affix morphological priming is not: the latter could be due to a ceiling frequency attenuation effect. This is an intriguing possibility that must be left for future work to explore.

In summary, our study successfully replicated and expanded upon the work of Angele et al. (2023), Cayado, Wray, and Stockall (2023) and Petrosino, Sprouse, and Almeida (2023), confirming the viability of observing repetition priming effects in masked priming experiments conducted online with a brief SOA of 33 ms. Notably, we addressed a lingering question in the literature by establishing the presence of the Frequency Attenuation Effect (FAE) under masked conditions. The use of large online samples proved instrumental in overcoming the longstanding challenge of insufficient statistical power to detect interactions in factorial designs, which we believe had impeded previous investigations into detecting the FAE in masked priming.

References

Adelman, James S, Rebecca L Johnson, Samantha F McCormick, Meredith McKague, Sachiko Kinoshita, Jeffrey S Bowers, Jason R Perry, et al. 2014. "A Behavioral Database for Masked Form Priming." *Behavior Research Methods* 46: 1052–67.

Amenta, Simona, and Davide Crepaldi. 2012. "Morphological Processing as We Know It: An Analytical Review of Morphological Effects in Visual Word Identification." *Frontiers in Psychology* 3: 232.

Angele, Bernhard, Ana Baciero, Pablo Gómez, and Manuel Perea. 2023. "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.

Anwyl-Irvine, Alexander L, Jessica Massonnié, Adam Flitton, Natasha Kirkham, and Jo K Evershed. 2020. "Gorilla in Our Midst: An Online Behavioral Experiment Builder." *Behavior Research Methods* 52: 388–407.

Balota, David A., Michael J. Cortese, Susan D. Sergent-Marshall, Daniel H. Spieler, and Melvin J. Yap. 2004. "Visual Word Recognition of Single-Syllable Words." *Journal of Experimental Psychology: General* 133 (2): 283.

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. https://doi.org/10.3758/bf03193014.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth, and Harald Baayen. 2015. "Parsimonious Mixed Models." https://arxiv.org/abs/1506.04967.

Bodner, Glen E., and Michael E. J Masson. 2014. "Memory Recruitment: A Backward Idea about Masked Priming." In *Psychology of Learning and Motivation*, 61:179–213. Elsevier.

Bodner, Glen E., and Michael E. J. Masson. 1997. "Masked Repetition Priming of Words and Nonwords: Evidence for a Nonlexical Basis for Priming." *Journal of Memory and Language* 37 (2): 268–93. https://doi.org/10.1006/jmla.1996.2507.

———. 2001. "Prime Validity Affects Masked Repetition Priming: Evidence for an Episodic Resource Account of Priming." *Journal of Memory and Language* 45 (4): 616–47. https://doi.org/10.1006/jmla.2001.2791.

Brysbaert, Marc, and Michael J Cortese. 2011. "Do the Effects of Subjective Frequency and Age of Acquisition Survive Better Word Frequency Norms?" *Quarterly Journal of Experimental Psychology* 64 (3): 545–59.

Brysbaert, Marc, Paweł Mandera, and Emmanuel Keuleers. 2018. "The Word Frequency Effect in Word Processing: An Updated Review." *Current Directions in Psychological Science* 27 (1): 45–50.

Brysbaert, Marc, and Boris New. 2009. "Moving Beyond Kučera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English." *Behavior Research Methods* 41 (4): 977–90. https://doi.org/10.3758/brm.41.4.977.

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.

Burgess, Curt, and Kay Livesay. 1998. "The Effect of Corpus Size in Predicting Reaction Time in a Basic Word Recognition Task: Moving on from Kučera and Francis." *Behavior Research Methods, Instruments, & Computers* 30 (2): 272–77.

Button, Katherine S, John PA Ioannidis, Claire Mokrysz, Brian A Nosek, Jonathan

Flint, Emma SJ Robinson, and Marcus R Munafò. 2013. "Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience." *Nature Reviews Neuroscience* 14 (5): 365–76.

Cayado, Dave Kenneth Tayao, Samantha Wray, and Linnaea Stockall. 2023. "Does Linear Position Matter for Morphological Processing? Evidence from a Tagalog Masked Priming Experiment." *Language, Cognition and Neuroscience*, 1–16.

Cohen, Jacob. 1992. "A Power Primer." *Psychological Bulletin* 112 (1): 155–59. https://doi.org/10.1037/0033-2909.112.1.155.

Coltheart, Max, Kathleen Rastle, Conrad Perry, Robyn Langdon, and Johannes Ziegler. 2001. "DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud." *Psychological Review* 108 (1): 204–56.

de Leeuw, Joshua R. 2014. "jsPsych: A JavaScript Library for Creating Behavioral Experiments in a Web Browser." *Behavior Research Methods* 47 (1): 1–12. https://doi.org/10.3758/s13428-014-0458-y.

Evett, Lindsay J., and Glyn W. Humphreys. 1981. "The Use of Abstract Graphemic Information in Lexical Access." *The Quarterly Journal of Experimental Psychology* 33 (4): 325–50.

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.

Forster, Kenneth I. 1998. "The Pros and Cons of Masked Priming." *Journal of Psycholinguistic Research* 27 (2): 203–33.

———. 1999. "Microgenesis of Priming Effects in Lexical Access." *Brain and Language* 68: 5–15.

Forster, Kenneth I., and Tamiko Azuma. 2000. "Masked Priming for Prefixed Words with Bound Stems: Does Submit Prime Permit?" *Language and Cognitive Processes* 15 (4-5): 539–61.

Forster, Kenneth I., and Chris Davis. 1984. "Repetition Priming and Frequency Attenuation in Lexical Access." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10 (4): 680.

Forster, Kenneth I., and Christopher Davis. 1991. "The Density Constraint on Form-Priming in the Naming Task: Interference Effects from a Masked Prime." *Journal of Memory and Language* 30 (1): 1–25. https://doi.org/10.1016/0749-596x(91)90008-8.

Forster, Kenneth I., C. Davis, C. Schoknecht, and R. Carter. 1987. "Masked Priming with Graphemically Related Forms: Repetition or Partial Activation?" *The Quarterly Journal of Experimental Psychology Section A* 39 (2): 211–51. https://doi.org/10.1080/14640748708401785.

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.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*. Third. Thousand Oaks CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Gelman, Andrew, and John Carlin. 2014. "Beyond Power Calculations." *Perspectives on Psychological Science* 9 (6): 641–51. https://doi.org/10.1177/1745691614551642.

Gimenes, Manuel, and Boris New. 2016. "Worldlex: Twitter and Blog Word Frequencies for 66 Languages." *Behavior Research Methods* 48: 963–72.

Gomez, Pablo, Manuel Perea, and Roger Ratcliff. 2013. "A Diffusion Model Account of Masked Versus Unmasked Priming: Are They Qualitatively Different?" *Journal of Experimental Psychology: Human Perception and Performance* 39 (6): 1731.

Grainger, Jonathan, and Arthur M. Jacobs. 1996. "Orthographic Processing in Visual Word Recognition: A Multiple Read-Out Model." *Psychological Review* 103 (3): 518.

Herdağdelen, Amaç, and Marco Marelli. 2017. "Social Media and Language Processing: How Facebook and Twitter Provide the Best Frequency Estimates for Studying Word Recognition." *Cognitive Science* 41 (4): 976–95.

Jacoby, Larry L. 1983. "Remembering the Data: Analyzing Interactive Processes in Reading." *Journal of Verbal Learning and Verbal Behavior* 22 (5): 485–508.

Jacoby, Larry L, and Mark Dallas. 1981. "On the Relationship Between Autobiographical Memory and Perceptual Learning." *Journal of Experimental Psychology: General* 110 (3): 306.

Kinoshita, Sachiko. 2006. "Additive and Interactive Effects of Word Frequency and Masked Repetition in the Lexical Decision Task." *Psychonomic Bulletin & Review* 13 (4): 668–73. https://doi.org/10.3758/bf03193979.

Kučera, J., and W. N. Francis. 1967. *Computational Analysis of Present Day American English*. Providence, RI: Brown University Press.

Lo, Steson, and Sally Andrews. 2015. "To Transform or Not to Transform: Using Generalized Linear Mixed Models to Analyse Reaction Time Data." *Frontiers in Psychology* 6: 1171.

Lukács, Gáspár, and Andreas Gartus. 2023. "Precise Display Time Measurement in JavaScript for Web-Based Experiments." *Behavior Research Methods* 55 (3): 1079–93. https://doi.org/10.3758/s13428-022-01835-2.

Lund, Kevin, and Curt Burgess. 1996. "Producing High-Dimensional Semantic Spaces from Lexical Co-Occurrence." *Behavior Research Methods, Instruments, & Computers* 28 (2): 203–8.

Masson, Michael E. J., and Glen E. Bodner. 2003. "A Retrospective View of Masked Priming: Toward a Unified Account of Masked and Long-Term Repetition Priming." *Masked Priming: The State of the Art*, 57–94.

McClelland, James L., and David E. Rumelhart. 1981. "An Interactive Activation Model of Context Effects in Letter Perception: Part i. An Account of Basic Findings." *Psychological Review* 88 (5): 375–407.

Nievas, Francisco. 2010. "The Frequency Attenuation Effect in Identity and Associative Priming." *The Spanish Journal of Psychology* 13 (1): 30–62. https://doi.org/10.1017/s1138741600003668.

Norris, Dennis, and Sachiko Kinoshita. 2008. "Perception as Evidence Accumulation and Bayesian Inference: Insights from Masked Priming." *Journal of Experimental Psychology: General* 137 (3): 434–55. https://doi.org/10.1037/a0012799.

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 (1): 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, Jon Sprouse, and Diogo Almeida. 2023. "Asymmetries in the Stem and Suffix Masked Priming Response in a Large-Scale Online Study." *Quaderni Di Linguistica e Studi Orientali*, no. 49: 177–94. https://doi.org/10.13128/QUL-SO-2421-7220-15154.

Potvin, Patrick J., and Robert W. Schutz. 2000. "Statistical Power for the Two-Factor Repeated Measures ANOVA." *Behavior Research Methods, Instruments, & Computers* 32 (2): 347–56. https://doi.org/10.3758/bf03207805.

Rajaram, Suparna, and James H Neely. 1992. "Dissociative Masked Repetition Priming and Word Frequency Effects in Lexical Decision and Episodic Recognition Tasks." *Journal of Memory and Language* 31 (2): 152–82. https://doi.org/10.1016/0749-596x(92)90009-m.

Scarborough, Don L., Charles Cortese, and Hollis S. Scarborough. 1977. "Frequency and Repetition Effects in Lexical Memory." *Journal of Experimental Psychology: Human Perception and Performance* 3 (1): 1–17. https://doi.org/10.1037/0096-1523.3.1.1.

Segui, Juan, and Jonathan Grainger. 1990. "Priming Word Recognition with Orthographic Neighbors: Effects of Relative Prime-Target Frequency." *Journal of Experimental Psychology: Human Perception and Performance* 16 (1): 65–76. https://doi.org/10.1037/0096-1523.16.1.65.

Sereno, Joan A. 1991. "Graphemic, Associative, and Syntactic Priming Effects at a Brief Stimulus Onset Asynchrony in Lexical Decision and Naming." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17 (3): 459–77. https://doi.org/10.1037/0278-7393.17.3.459.

Spence, Jeffrey R, and David J Stanley. 2016. "Prediction Interval: What to Expect When You're Expecting… a Replication." *PloS One* 11 (9): 1–22. https://doi.org/10.1371/journal.pone.0162874.

Stockall, Linnaea, and Alec Marantz. 2006. "A Single Route, Full Decomposition Model of Morphological Complexity: MEG Evidence." *The Mental Lexicon* 1 (1): 85–123. https://doi.org/https://doi.org/10.1075/ml.1.1.07sto.

Taft, Marcus, and Kenneth I. Forster. 1975. "Lexical Storage and Retrieval of Prefixed Words." *Journal of Verbal Learning and Verbal Behavior* 14 (6): 638–47.

W3 Counter. 2023. "Browser & Platform Market Share - November 2023." https://www.w3counter.com/globalstats.php?year=2023&month=11.

Yap, Melvin J., and David A. Balota. 2009. "Visual Word Recognition of Multisyllabic Words." *Journal of Memory and Language* 60 (4): 502–29.

Zevin, Jason D, and Mark S Seidenberg. 2002. "Age of Acquisition Effects in Word Reading and Other Tasks." *Journal of Memory and Language* 47 (1): 1–29.

Zimmerman, Robert, and Pablo Gomez. 2012. "Drawing Attention to Primes Increases Inhibitory Word Priming Effects." *The Mental Lexicon* 7 (2): 119–46.

Wordlists

Experiment 1

related	unrelated prime	word	RT (to repetition) mean	RT (to repetition) SD	RT (to unrelated) mean	RT (to unrelated) SD
low frequency condition						
arrow	hunch	arrow	590	130	587	124
pitch	sneak	pitch	576	126	612	122
hatch	widow	hatch	621	151	639	148
shark	brief	shark	573	125	590	138
tooth	sharp	tooth	536	125	565	116
booth	grief	booth	572	136	627	157
pound	sting	pound	551	127	572	127
weigh	thief	weigh	593	167	636	164
blank	avoid	blank	571	139	596	124
crush	award	crush	554	128	592	136
bench	smack	bench	573	132	601	129
fetch	brand	fetch	622	156	658	146
cheek	salad	cheek	561	141	602	142
brush	swamp	brush	564	130	600	128
march	depth	march	559	125	580	123
bleed	flesh	bleed	560	148	577	146
cliff	harsh	cliff	602	130	645	137
fraud	creep	fraud	621	147	628	132
cloud	plead	cloud	536	115	551	101
fluid	thumb	fluid	605	140	678	162
trash	creek	trash	554	127	560	128
flush	blond	flush	576	123	617	140
porch	stink	porch	587	136	620	160
stiff	patch	stiff	626	154	678	156
cough	sweep	cough	564	142	601	141
smash	squad	smash	570	129	587	126
high frequency condition						
blood	chief	blood	541	130	551	104
bunch	child	bunch	585	148	617	145
catch	board	catch	545	116	562	130
stuff	tough	stuff	555	119	585	137
break	stand	break	545	107	561	124
speak	beach	speak	545	131	573	129
stick	hotel	stick	562	128	598	138
sleep	angel	sleep	538	113	559	119
wrong	truth	wrong	563	143	565	132
grand	quick	grand	571	127	582	143
mouth	world	mouth	543	125	556	119
knock	extra	knock	560	134	631	136
guard	think	guard	580	132	590	134
small	thing	small	557	130	577	125

check	round	check	558	135	562	121
watch	proud	watch	541	128	546	110
group	smell	group	559	127	576	142
month	earth	month	555	120	572	123
south	relax	south	575	139	611	133
lunch	truck	lunch	547	119	557	125
clock	throw	clock	548	132	574	124
sound	death	sound	538	127	552	103
drink	north	drink	559	129	556	122
touch	young	touch	541	122	573	121
laugh	weird	laugh	546	119	568	121
black	reach	black	553	131	563	114
non-word						
alkew	grack	alkew	599	153	591	140
agink	furob	agink	626	148	614	141
ruzak	begro	ruzak	577	130	584	142
sondo	labok	sondo	625	142	612	149
guesh	gazzo	guesh	702	184	721	194
fadio	criam	fadio	618	149	604	146
plich	coreb	plich	650	162	640	159
sgrew	docab	sgrew	626	182	638	182
sceak	colob	sceak	675	154	683	171
ghisk	isloo	ghisk	588	139	593	139
deirm	ahuck	deirm	589	142	596	139
villo	flurb	villo	632	182	615	181
tidow	pikto	tidow	648	167	624	160
drick	aliom	drick	684	168	681	172
phick	purso	phick	643	160	637	165
nello	borno	nello	625	156	612	151
feach	pacaw	feach	730	201	720	192
tello	rilth	tello	651	175	644	171
dolio	caveb	dolio	602	148	610	165
gorgo	swysh	gorgo	643	164	619	170
whilo	lanjo	whilo	612	137	604	150
stanf	drief	stanf	611	134	617	133
crulk	ocheb	crulk	671	162	665	169
phumb	tunch	phumb	645	160	633	148
sirth	steaf	sirth	612	141	618	145
slerk	nohew	slerk	640	153	634	163
vitbo	nualm	vitbo	593	151	596	154
sunch	ofium	sunch	665	165	665	161
soeth	croik	soeth	589	141	589	130
eltow	valuo	eltow	628	171	606	158
framo	sorgo	framo	617	146	618	146
lumpo	shavo	lumpo	630	162	635	172
spuff	oceab	spuff	672	169	667	183
gatob	tolio	gatob	599	139	606	155
nosom	theck	nosom	598	155	604	139
gezzo	tooch	gezzo	592	136	586	131

afoub	slonk	afoub	582	133	589	128
wateb	salch	wateb	633	151	619	133
nelch	raceb	nelch	601	144	594	145
dahoo	ahack	dahoo	598	132	595	146
driek	fideo	driek	606	145	606	143
gnask	fluko	gnask	612	171	604	153
brosk	cyrrh	brosk	629	159	647	175
duvez	revuo	duvez	580	152	580	155
fielm	cempo	fielm	609	146	611	151
pumph	exulk	pumph	669	162	685	176
gerif	kleck	gerif	584	137	588	149
racef	bonth	racef	618	151	622	156
pheek	scook	pheek	640	155	644	176
pruaw	slork	pruaw	593	133	592	135
guilm	whilf	guilm	603	142	598	142
lairf	drosh	lairf	587	144	600	150

Experiment 2

			RT (to repetition)		RT (to unrelated)	
related	unrelated prime	word	mean	SD	mean	SD
low frequency condition						
arrow	hunch	arrow	573	130	591	128
pitch	sneak	pitch	565	132	608	146
hatch	widow	hatch	606	163	644	164
shark	brief	shark	560	132	587	124
tooth	sharp	tooth	533	126	560	111
booth	grief	booth	566	147	618	163
pound	sting	pound	552	139	570	134
weigh	thief	weigh	583	158	649	186
blank	avoid	blank	556	132	583	132
crush	award	crush	542	126	583	142
bench	smack	bench	565	143	601	138
fetch	brand	fetch	612	149	663	146
cheek	salad	cheek	556	138	597	148
brush	swamp	brush	559	136	593	125
march	depth	march	560	149	575	122
bleed	flesh	bleed	552	137	568	135
cliff	harsh	cliff	595	149	662	157
fraud	creep	fraud	596	131	634	113
cloud	plead	cloud	539	132	552	108
fluid	thumb	fluid	601	167	690	173
trash	creek	trash	564	157	566	126
flush	blond	flush	570	133	624	155
porch	stink	porch	584	155	617	153
stiff	patch	stiff	617	163	677	161
cough	sweep	cough	552	136	591	145

smash	squad	smash	553	133	590	142
high frequency condition						
blood	chief	blood	534	131	558	112
bunch	child	bunch	584	156	619	146
catch	board	catch	529	111	568	141
stuff	tough	stuff	556	128	576	133
break	stand	break	542	128	564	123
speak	beach	speak	543	131	570	132
stick	hotel	stick	556	136	599	130
sleep	angel	sleep	525	111	566	137
wrong	truth	wrong	555	142	569	137
grand	quick	grand	564	129	588	147
mouth	world	mouth	534	123	552	116
knock	extra	knock	549	131	615	139
guard	think	guard	565	118	590	118
small	thing	small	549	137	567	130
check	round	check	546	125	563	120
watch	proud	watch	531	115	545	133
group	smell	group	553	126	579	142
month	earth	month	549	126	576	128
south	relax	south	563	138	603	144
lunch	truck	lunch	549	141	563	124
clock	throw	clock	547	133	574	130
sound	death	sound	534	127	559	147
drink	north	drink	550	124	563	116
touch	young	touch	545	145	567	120
laugh	weird	laugh	535	113	556	114
black	reach	black	547	129	578	129
non-word						
alkew	grack	alkew	603	150	606	160
agink	furob	agink	632	161	615	155
ruzak	begro	ruzak	580	149	582	155
sondo	labok	sondo	619	168	621	178
guesh	gazzo	guesh	707	189	748	195
fadio	criam	fadio	632	169	602	171
plich	coreb	plich	645	148	646	180
sgrew	docab	sgrew	637	197	648	184
sceak	colob	sceak	692	179	692	165
ghisk	isloo	ghisk	595	149	594	152
deirm	ahuck	deirm	599	147	595	147
villo	flurb	villo	637	210	615	178
tidow	pikto	tidow	651	175	641	177
drick	aliom	drick	696	194	685	186
phick	purso	phick	648	176	639	168
nello	borno	nello	631	178	637	189
feach	pacaw	feach	748	201	745	208
tello	rilth	tello	665	185	662	174
dolio	caveb	dolio	618	175	609	157

gorgo	swysh	gorgo	651	179	631	178
whilo	lanjo	whilo	619	150	615	164
stanf	drief	stanf	613	147	627	134
crulk	ocheb	crulk	678	173	679	183
phumb	tunch	phumb	655	175	648	180
sirth	steaf	sirth	623	146	616	136
slerk	nohew	slerk	644	167	638	172
vitbo	nualm	vitbo	590	143	596	154
sunch	ofium	sunch	664	162	676	172
soeth	croik	soeth	590	131	589	146
eltow	valuo	eltow	638	170	603	153
framo	sorgo	framo	626	165	631	162
lumpo	shavo	lumpo	634	172	635	164
spuff	oceab	spuff	683	182	665	204
gatob	tolio	gatob	607	167	594	154
nosom	theck	nosom	603	156	608	160
gezzo	tooch	gezzo	601	160	600	164
afoub	slonk	afoub	583	136	587	133
wateb	salch	wateb	640	159	641	160
nelch	raceb	nelch	609	163	595	145
dahoo	ahack	dahoo	611	157	597	154
driek	fideo	driek	613	152	610	151
gnask	fluko	gnask	606	154	606	159
brosk	cyrrh	brosk	637	165	649	180
duvez	revuo	duvez	590	168	576	154
fielm	cempo	fielm	621	181	614	160
pumph	exulk	pumph	687	184	697	202
gerif	kleck	gerif	587	154	586	147
racef	bonth	racef	619	164	621	151
pheek	scook	pheek	641	157	635	169
pruaw	slork	pruaw	599	157	598	143
guilm	whilf	guilm	620	168	603	163
lairf	drosh	lairf	590	138	597	146