ARTICLE IN PRESS

Journal of Memory and Language xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity

Melvin J. Yap [a,*], Chi-Shing Tse [b], David A. Balota [b]

[a] Department of Psychology, National University of Singapore, Singapore
[b] Department of Psychology, Washington University in St. Louis, USA

ARTICLE INFO

Article history:
Received 14 June 2009
Available online xxxx

Keywords:
Semantic priming
Visual word recognition
RT distributional analyses
Individual differences
Lexical quality
Lexical integrity
Lexical decision
Ex-Gaussian analyses

ABSTRACT

Word frequency and semantic priming effects are among the most robust effects in visual word recognition, and it has been generally assumed that these two variables produce interactive effects in lexical decision performance, with larger priming effects for low-frequency targets. The results from four lexical decision experiments indicate that the joint effects of semantic priming and word frequency are critically dependent upon differences in the vocabulary knowledge of the participants. Specifically, across two Universities, additive effects of the two variables were observed in means, and in RT distributional analyses, in participants with more vocabulary knowledge, while interactive effects were observed in participants with less vocabulary knowledge. These results are discussed with reference to [Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 813–840] multistage account and [Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review, 107*, 786–823] single-mechanism model. In general, the findings are also consistent with a flexible lexical processing system that optimizes performance based on processing fluency and task demands.

Introduction

Word frequency and *semantic priming* effects are probably the most studied effects in the visual word recognition literature. Frequently encountered words are recognized faster than rarely encountered words. Targets preceded by related primes (e.g., BREAD–BUTTER) are recognized faster than targets preceded by unrelated primes (e.g., DOCTOR–BUTTER). Importantly, on a theoretical level, the two variables interact, with larger semantic priming effects for low-frequency targets than for high-frequency targets (Becker, 1979; Borowsky & Besner, 1993; Plaut & Booth, 2000).

Multiple stages vs. single mechanism

Different mechanisms have been proposed to account for the priming by frequency interaction (see McNamara, 2005; Neely, 1991, for excellent reviews). We will consider two major perspectives on how the interaction could be accommodated. The first assumes that the word recognition process is best conceptualized as separate, serially organized processing stages and the second assumes that word recognition reflects the operation of a single mechanism within a parallel distributed processing (PDP) network.

The serially organized stage framework is predicated on additive factors logic (Sternberg, 1969), which proposes that an interaction between two variables signifies that the two variables influence *at least* one common stage,

* Corresponding author. Fax: +65 6773 1843.
E-mail address: melvin@nus.edu.sg (M.J. Yap).

0749-596X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2009.07.001

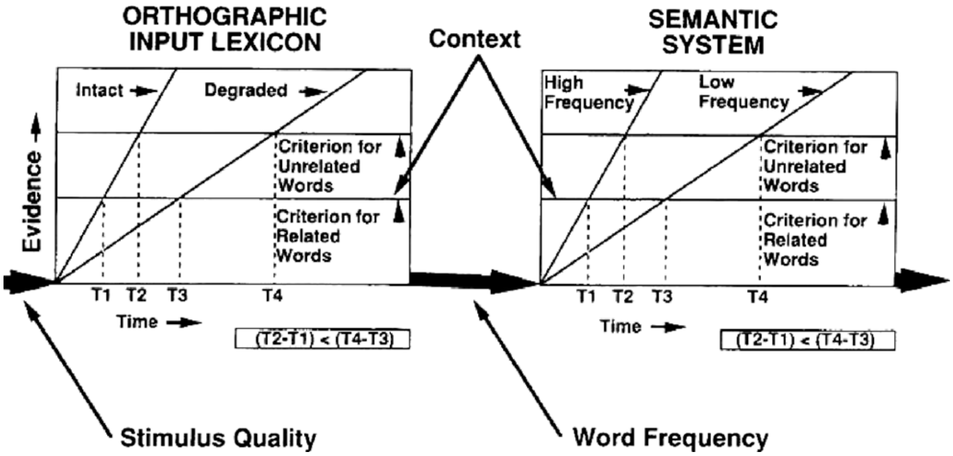

Fig. 1. A multistage activation model of visual word recognition. From "Visual word recognition: A multistage activation model" by R. Borowsky and D. Besner (1993), Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, p. 832.

while additive effects (i.e., two main effects and no interaction) is more likely to indicate that the variables influence different stages. Empirically, priming interacts with stimulus quality (e.g., Stolz & Neely, 1995) and with word frequency (e.g., Stone & Van Orden, 1993), but stimulus quality and word frequency produce robust additive effects in lexical decision (Becker & Killion, 1977; Plourde & Besner, 1997; Yap & Balota, 2007). This complex pattern of data has been interpreted within a multistage model (see Fig. 1 for an example), where stimulus quality affects the first stage (i.e., orthographic input lexicon), word frequency affects the second one (i.e., semantic system), and both stages are sensitive to semantic priming. Specifically, in Borowsky and Besner's (1993) multistage activation model, words are first "cleaned up" before they are processed by a second stage that is sensitive to word frequency. The priming by frequency interaction implies that these two variables jointly influence the second stage.

Importantly, according to the multistage perspective, rather than influencing the word detectors in the orthographic input lexicon, word frequency is postulated to modulate the mappings between the orthographic input lexicon and the semantic system. That is, high-frequency words possess more efficient mappings between the orthographic input lexicon and semantic system, and therefore, evidence for such words accumulates more rapidly than for low-frequency words. A related semantic context lowers the recognition threshold, and for any given change in criterion, larger priming effects will be observed for low-frequency targets (slower activation rate) than for high-frequency targets (faster activation rate) (see Fig. 1's right panel). This predicts the observed priming by frequency interaction. Of course, a central tenet in this model is that activation in the semantic system, not the orthographic input lexicon, drives lexical decisions (see Borowsky & Besner, 1993, p. 833) for further discussion of this assumption). While this and other assumptions in the multistage model may seem *post hoc*, Borowsky and Besner have argued that these assumptions are necessary, given the complex joint effects of priming, frequency, and stimulus quality described earlier.

Plaut and Booth's (2000) PDP account of the combinatorial influence of these variables provides an important alternative. Unlike the stage-based accounts, which incorporate thresholded processing and multiple stages, Plaut and Booth's model accounts for the priming by frequency interaction and other empirical effects in semantic priming with a single mechanism that mediates input and output processes. Specifically, a non-linear sigmoid mapping between input and output allows equal differences in input to be reflected by equal or unequal differences in the output, depending on the portion of the sigmoid function being examined (see Fig. 2). Because high-frequency and related targets possess higher input strengths (i.e., they are located higher on the input continuum), high-frequency targets yield smaller priming effects than low-frequency targets.

There is an ongoing debate about whether the joint effects of priming, word frequency, and stimulus quality are better accommodated by a multistage mechanism or by a single mechanism. Borowsky and Besner (2006) have argued that the single-mechanism model has problems discriminating words from orthographically matched nonwords, that its reliance on semantics for carrying out lexical decision is inconsistent with neuropsychological evidence, and that it is unable to accommodate additive and interactive effects within the same range of *RT*s typically observed (see also Besner & Borowsky, 2006; Besner, Wartak, & Robidoux, 2008). In response to these criticisms, Plaut and Booth (2006), after carrying out additional modeling, argued that none of these issues are truly problematic for their model. For example, they demonstrated that the model could distinguish consonant–vowel–consonant (CVC) words from CVC nonwords very accurately. The full details of this debate are outside the scope of this paper, but it seems reasonable to conclude that there is currently no consensus on whether a multistage mechanism or a single mechanism better accommodates the extant data. We will revisit this debate in 'General discussion', when we evaluate our data against the two classes of models.

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

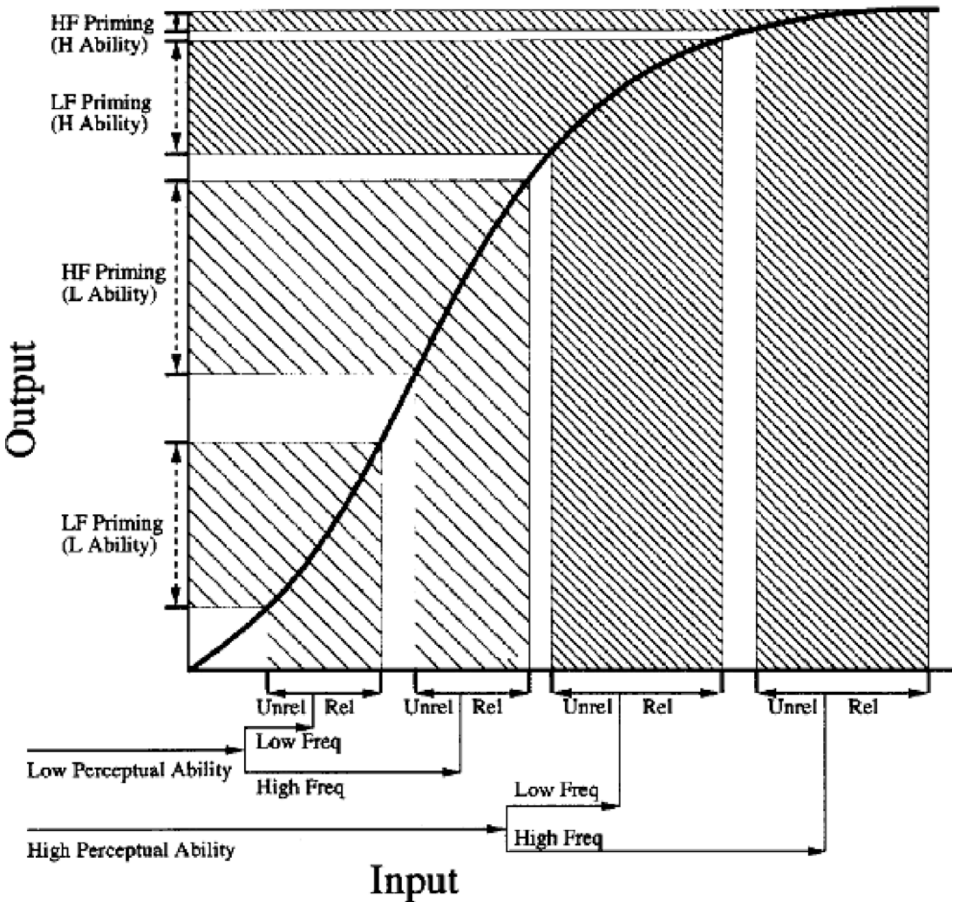

Fig. 2. The sigmoid activation function of the Plaut and Booth (2000) model. From "Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing" by D. C. Plaut and J. R. Booth, 2000, Psychological Review, 107, p. 832.

Do priming and frequency always interact? The role of individual differences

Although it has traditionally been assumed that priming and frequency produce interactive effects, Plaut and Booth (2000) have demonstrated that these two factors do not always interact. Specifically, they observed *interactive* effects of priming and frequency in participants with high-perceptual-ability, and *additive* effects in participants with low-perceptual-ability, as measured by a standard psychometric test of processing speed called the Symbol Search Test of the Wechsler Intelligence Scale for Children (Wechsler, 1991). In this paper-and-pencil matching to sample task, participants are required to indicate, as quickly and accurately as possible, whether either of two meaningless symbols on the left is present in a row of five meaningless symbols on the right. Plaut and Booth decided to focus on perceptual ability because of its links to reading proficiency (Vernon, 1987), abnormal language development (Farmer & Klein, 1995), and early reading acquisition (Detterman & Daniel, 1989).

Plaut and Booth further demonstrated that their single-mechanism model could parsimoniously yield the same three-way interaction between perceptual ability, word frequency, and semantic priming, via the non-linear sigmoid activation function (see Fig. 2). For low-perceptual-ability readers, located on the lower left-hand portion of the graph, the input–output relationship is relatively linear, producing equal-sized priming effects for low- and high-frequency targets. In contrast, for high-perceptual-ability readers, located on the upper right-hand portion of the graph, the input–output function is more logarithmic in shape, and this yields larger priming effects for low- than for high-frequency targets.

Lexical integrity vs. perceptual ability

The present study extends the work by Plaut and Booth (2000) by considering the role of *lexical integrity* on the effects of priming and frequency. By lexical integrity, we are referring to the strength and quality of the underlying lexical representations. Lexical integrity is conceptually very similar to Perfetti and Hart's (2002) lexical quality, where "quality" is defined by fully specified orthographic representations and fully redundant phonological representations. High-integrity representations, due to their coherence and stability, are more likely to be retrieved in a fluent manner. How might lexical integrity modulate the joint effects of priming and frequency? If one assumes that the flow of activation in word recognition is fundamentally an interactive process, low-frequency words, compared to high-frequency words, should have more opportunity to benefit from a related semantic context, since they are further from recognition threshold. For individuals with relatively rich lexical representations (high-lexical integrity), one *a priori* assumes that *for the same word* these individuals would be closer to recognition threshold than individuals with relatively poor lexical representations (low-lexical integrity). Specifically, a medium-frequency word for a high-lexical-integrity individual is likely to be a low-frequency word for a low-lexical-integrity individual. So, one might actually expect individuals with *lower* integrity representations to show a larger influence of semantic context than those with higher integrity representations, all other things being equal. We will henceforth refer to this position as the *lexical integrity hypothesis*.

In our study, lexical integrity is assessed by vocabulary knowledge (i.e., knowledge of word forms and word meanings). There is evidence that the size of an individual's vocabulary is positively related to the precision (Perfetti, 2007; Perfetti & Hart, 2002; Verhoeven & Van Leeuwe, 2008) and stability (Kinoshita, 2006; Kinoshita & Mozer, 2006; Paap, Johansen, Chun, & Vonnahme, 2000) of underlying lexical representations. Lexical integrity contrasts well with Plaut and Booth's (2000) perceptual ability, which primarily has to do with an individual's speed at processing new information (Tulsky, Saklofske, & Zhu, 2003), and implicates lower-level processes that encode not only letters and words, but also digits, pictures, and objects. Indeed, in Plaut and Booth's sample, symbol search performance was uncorrelated with vocabulary knowledge ($r = .09$), as measured by the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), indicating that these two instruments tap distinct abilities.

Before discussing the specific predictions of the lexical integrity hypothesis, some other findings are relevant. When Tainturier, Tremblay, and Lecours (1992) examined the relationship between educational level (and by extension, vocabulary knowledge) and the magnitude of word frequency effects in lexical decision, they found that frequency effects were *smaller* for more educated participants. This indicates that processing differences between low- and high-frequency words are *smaller* for the readers with (presumably) more vocabulary knowledge. Seidenberg (1985) also reported that rapid decoders produced smaller frequency effects than slow decoders in speeded pronunciation, a pattern that has been replicated by Schilling, Rayner, and Chumbley (1998). These results would appear to predict that the effects of word frequency and semantic priming should be additive for high-lexical-integrity readers and interactive for low-lexical-integrity readers. Specifically, for low-integrity readers, low-frequency words, compared to high-frequency words, are less strongly represented and are processed more effortfully; low-frequency words should therefore benefit more from a related prime. For high-lexical-integrity readers, low-frequency words are so well represented that high- and low-frequency words are processed very efficiently, such that both classes of words will benefit to the *same* extent from a related prime.

To summarize, in the present study, we explore the joint effects of priming, frequency, and vocabulary knowledge. As discussed, interactive effects of priming and frequency should be associated with readers with less lexical integrity, i.e., less vocabulary knowledge. The individual differences issue, a major theme in this paper, seems timely given researchers' growing interest in the effects of individual differences on semantic priming. For example, Hutchison (2007) examined the role of attentional control and the relatedness proportion effect in semantic priming. As the proportion of related prime–target pairs in an experiment increases, priming effects become larger. This relatedness proportion effect reflects participants' effortful generation of likely targets when they encounter a prime. Interestingly, Hutchison reported a positive linear relationship between participants' attentional control and the magnitude of their relatedness proportion effects, suggesting that individual differences in attentional control modulate strategic processes in semantic priming.

Priming, frequency, and distributional analyses

In addition to individual differences, the present study explores the characteristics of response time (*RT*) distributions to better understand the nature of the interactive effects of these variables. While the joint effects of frequency and priming place important constraints on models of word recognition and priming, these effects are not well-understood at the level of underlying *RT* distributions. Although mean *RT*s are faster for semantically related targets than for unrelated targets, differences in *mean RT*s can be reflected by distributional shifting, skewing, or a mixture of shifting and skewing (Balota & Spieler, 1999; Heathcote, Popiel, & Mewhort, 1991). A recent study reported that semantic priming effects are reflected by distributional shifting (Balota, Yap, Cortese, & Watson, 2008; see Roelofs, 2008, for a discussion of distributional effects in priming for semantic categorization). However, it is still unclear if this shifting applies only to targets with high-integrity lexical representations (i.e., high-frequency words) or to targets in general (i.e., high- and low-frequency words). In contrast, word frequency effects are consistently reflected by both a shifting and skewing of the *RT* distributions (e.g., Andrews & Heathcote, 2001; Balota & Spieler, 1999; Yap & Balota, 2007). Because understanding the joint effects of priming and frequency at the distributional level will help impose finer constraints on extant models, the second major theme of the current study is to explore whether or not the priming effects for high- and low-frequency words show qualitatively similar *RT* distributional profiles across individuals with higher and lower levels of lexical knowledge.

Currently, Distributional analyses can be carried out by fitting *RT*s to a theoretical distribution like the ex-Gaussian distribution (see Van Zandt, 2000, for a discussion of *RT* distributional analyses), or by averaging *RT* distributions across a number of participants. In this paper, both tech-

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

niques are employed. Fitting individual raw *RT* data to the ex-Gaussian distribution, a three-parameter (μ, σ, τ) function, allows differences in means to be partitioned into distributional shifting (μ) and an estimate of distributional skewing (τ); importantly, the algebraic sum of μ and τ is the mean of the fitted ex-Gaussian distribution. *Vincentizing* is a non-parametric technique which computes a number of vincentiles for each participant, where a vincentile is defined as the mean of observations between neighboring percentiles. For example, to obtain 10 vincentiles, the *RT* data within each condition for a participant is first sorted (from fastest to slowest responses), and the first 10% of the data is then averaged, followed by the second 10%, and so on. Individual vincentiles are then averaged across participants. Vincentizing makes no assumptions about the shape of the underlying *RT* distribution and examines the raw data directly.

Controlling for prime–target associative strength

In a factorial experiment manipulating priming and frequency, it is critical that low- and high-frequency targets are equally related to their related primes. Interestingly, in virtually every published study examining the priming by frequency interaction, the associative strength of high- and low-frequency targets were matched using a rating procedure. For example, Becker (1979) presented participants with word pairs, and asked them to indicate, on a seven-point scale, the likelihood of generating the second word, given the first word. In the current study, we used Nelson, McEvoy, and Schreiber's (2004) free association norms to select primes (see also Tse & Neely, 2007). This approach possesses two major advantages. One, rating two items (e.g., A and B) as highly related does not indicate if there is a strong A to B connection, or a strong B to A connection (Nelson et al.). Free association norms allow both the magnitude and *direction* of associations to be taken into account (see Hutchison, Balota, Cortese, & Watson, 2008, for a discussion of variables that predict priming in a large database). More importantly, the Nelson et al. norms, which are based on the responses of more than 6000 participants, should provide more reliable estimates of associative strength, compared to rating norms based on relatively small samples of participants.

Effect of nonword context on semantic priming

Nonword type can be manipulated in order to modulate word–nonword discrimination difficulty. Using pseudohomophones (e.g., BRANE), compared to legal nonwords (e.g., FLIRP), increases the similarity between words and nonwords, which yields slower lexical decision *RT*s and larger word frequency effects (Stone & Van Orden, 1993; Yap, Balota, Cortese, & Watson, 2006). Essentially, pseudohomophones make it more difficult to discriminate between words and nonwords, which in turn exaggerates the magnitude of effects. In this study, we also manipulated nonword type, with the *a priori* prediction that priming and frequency effects, along with the priming by frequency interaction would increase in the context of pseudohomophones compared to pronounceable nonwords.[1] This would provide additional leverage for our exploration of the influence of these variables across different levels of vocabulary knowledge.

Overview of experiments

Word frequency and semantic priming were factorially manipulated in four lexical decision experiments, and effects were analyzed both at the level of the mean and at the level of *RT* distributional characteristics. Experiments 1 and 3 (E1 and E3) featured legal nonwords (i.e., orthographically and phonologically plausible, e.g., FLIRP), while Experiments 2 and 4 (E2 and E4) featured pseudohomophonic nonwords (i.e., sound like real words, e.g., BRANE). Note that E3 and E4 were literal replications of E1 and E2, respectively, with an independent pool of participants, with varying levels of vocabulary knowledge. Recruiting participant pools from different universities allowed us to test the way in which individual differences modulate the priming by frequency interaction. As a preview, participants in E1 and E2 were associated with faster, more accurate word recognition performance, and higher vocabulary scores than those in E3 and E4. Given that college students in general are already selected for their vocabulary knowledge, this implies that the empirical patterns observed in E3 and E4 are more representative of typical readers, while participants in E1 and E2 are more likely to represent individuals with very high-vocabulary knowledge.

General method

Participants

One hundred and fifty-six undergraduates participated in the four experiments for course credit or $5 (see Table 1 for a summary of participant characteristics). All participants had normal or corrected-to-normal vision and were recruited from participant pools at Washington University (WUSTL, E1 & E2) and the University at Albany, State University of New York (SUNY-A, E3 & E4). Collapsing across experiments, participants from the two universities were significantly different in years of education, $t(144) = 5.91$,

[1] Shulman and Davison (1977) reported larger semantic priming effects in lexical decision when legal, compared to illegal, nonwords were used, but did not examine priming effects in the context of pseudohomophones. The study that comes closest to addressing this issue is one by Milota, Widau, McMickell, Juola, and Simpson (1997), who used the primed lexical decision task to prime real words, legal nonwords, or pseudohomophones. For example, the participant could see *doctor* (real word), *docton* (legal nonword), or *docter* (pseudohomophone) primed by either *nurse* (related) or *win* (unrelated). Interestingly, Milota et al. reported that pseudohomophone distracters, compared to legal nonword distracters, attenuated semantic priming, and attributed this effect to participants strategically suppressing the influence of a prime when it was less helpful for word-nonword discrimination. Specifically, there was a prime–target relationship for both word (*nurse–doctor*) and nonword (*nurse–docter*) trials. Note, however, that Milota et al.'s paradigm is clearly different from ours. Half their pseudohomophones were primed by related words, inducing strategic suppression of prime information, whereas our pseudohomophones are never related to their primes. Hence, whether pseudohomophones increase the priming main effect and the priming by frequency interaction remain open empirical questions.

Table 1
Mean age, year of education, and vocabulary score of participant (standard deviations in parentheses).

Experiment	Testing site	*N*	Age	Years of education	Vocabulary age	Vocabulary score
1	Washington University	40	19.33 (1.17)	13.03 (1.06)	18.47 (1.04)	32.07 (3.45)
2	Washington University	48	19.88 (1.40)	13.45 (1.15)	18.97 (0.73)	33.77 (2.38)
3	University at Albany	40	18.93 (1.10)	11.95 (0.75)	17.09 (1.45)	27.4 (4.19)
4	University at Albany	28	20.71 (3.72)	12.50 (1.37)	17.54 (1.24)	30.68 (4.59)

Note: Due to the missing data in the Shipley test, the mean age, years of education, vocabulary age, and vocabulary score of participants in Washington University are based on 36 and 42 participants in Experiments 1 and 2, respectively.

Table 2
Descriptive statistics for the word and nonword stimuli used in Experiments 1–4.

Word stimuli	High-frequency targets ($N = 150$)		Low-frequency targets ($N = 150$)	
	Mean	*SD*	*Mean*	*SD*
Log HAL frequency	11.51	.81	8.29	1.02
Raw HAL frequency per million	1102.99	1334.08	47.42	42.35
Raw KF frequency per million	280.03	284.84	19.09	21.42
Orthographic neighborhood size	5.42	4.86	5.03	4.63
Length	4.79	1.11	4.98	1.09
Forward associative strength	.56	.21	.54	.19
Backward associative strength	.20	.24	.19	.20
Nonword stimuli	Pronounceable nonwords ($N = 300$)		Pseudohomophonic nonwords ($N = 300$)	
Log HAL *baseword* frequency	–	–	9.13	2.13
Orthographic neighborhood size	4.70	3.96	3.68	3.66
Length	4.89	1.10	4.94	.84

Note: All values were based on Balota et al. (2007) and Nelson et al. (2004). The baseword frequency is the frequency of the word (e.g., *brain*) for the pseudohomophonic nonword (e.g., *brane*). KF Frequency refers to the Kučera and Francis (1967) word frequency counts.

$\eta_p^2 = .20$, in vocabulary scores, $t(144) = 7.68$, $\eta_p^2 = .29$, but not in age, $t < 1$; the WUSTL participants had more years of education and higher vocabulary scores than the SUNY-A participants.

Design

In each experiment, Priming (related or unrelated) and Frequency (high or low) were manipulated within participants. Across experiments, Nonword Type (legal in E1 and E3 or pseudohomophonic in E2 and E4) and University (WUSTL in E1 and E2 or SUNY-A in E3 and E4) were manipulated between participants. The dependent variables were *RT* and error rate.

Stimuli

Descriptive statistics for the word and nonword stimuli are presented in Table 2. High- and low-frequency targets were matched on length and orthographic neighborhood size (Coltheart, Davelaar, Jonasson, & Besner, 1977). Primes were selected using the Nelson et al. (2004) free association norms, and prime–target associative strengths were matched for high- and low-frequency targets in both directions (prime-to-target and target-to-prime). Nonwords were orthographically legal and pronounceable in E1 and E3, and pseudohomophonic in E2 and E4. Word and nonword length were matched across all four experiments, while word ($M = 5.23$, $SD = 4.74$) and nonword ($M = 4.70$, $SD = 3.96$) orthographic neighborhood sizes were matched in E1 and E3. In E2 and E4, the mean orthographic neighborhood size of pseudohomophones ($M = 3.68$, $SD = 3.66$) was significantly lower than that of words, $p < .001$.[2] Of course, orthographic neighborhood size does not reflect the effect of *orthographic* neighbors *per se* but may also reflect the influence of phonological neighbors (i.e., neighbors obtained by substituting a single phoneme). As Mulatti, Reynolds, and Besner (2006) have pointed out, orthographic and phonological neighborhood sizes are highly correlated; for the 300 words in the present study, this correlation was .729. Overall, there were 150 high-frequency words, 150 low-frequency words, and 300 nonwords. Within each frequency range, targets were either primed by a related or unrelated prime per participant, resulting in 75 observations per participant cell. Four counterbalancing lists were created, each of which was randomly and equally assigned to 10, 12, 10, and 7 participants in E1 to E4, respectively. No item was repeated within a participant.

Procedure

PC-compatible computers running E-prime software (Schneider, Eschman, & Zuccolotto, 2001) were used to control stimulus presentation and to collect data. All stimuli were displayed at the center of the computer screen and

[2] In order to secure high-quality pseudohomophones, the 300 pseudohomophones used in E2 and E4 were selected from the appendices of published articles. The limited pool of available pseudohomophones, and the need to match these stimuli to words on length, imposed constraints that made it impossible for us to match the orthographic neighborhood of words and nonwords as closely as one would like.

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

participants' responses were made on a computer keyboard. Participants were tested individually in sound-attenuated cubicles, sitting about 60 cm from the screen. Participants first provided demographic information (chronological age, years of education) and completed the vocabulary subscale (40 item vocabulary test) of the Shipley Institute of Living Scale (Shipley, 1940; Zachary, 1992) on the computer. The full Shipley scale was originally devised to provide a quick measure of intellectual functioning, and contains vocabulary knowledge (reliability coefficient = .87) and abstract thinking (reliability coefficient = .89) subscales. In our study, we administered the vocabulary subscale, and participants' vocabulary knowledge was estimated using raw Shipley scores. Despite its age, the Shipley continues to be widely used by researchers and it has been shown to correlate highly with most standard intelligence tests (see Zachary, Paulson, & Gorsuch, 1985, for a review).

Participants were then instructed, on each trial, to silently read the first word, and to then decide whether the subsequently presented letter string formed a word or nonword by making the appropriate button press. Participants were encouraged to respond quickly, but not at the expense of accuracy. Twenty practice trials were then presented, followed by six experimental blocks of 100 trials, with mandatory breaks between blocks. The order in which stimuli were presented was randomized anew for each participant. Stimuli were presented in uppercase 14 point Courier, and each trial consisted of the following order of events: (a) a fixation point (+) at the center of the monitor for 2000 ms, (b) the prime for 150 ms, (c) a blank screen for 650 ms, and (d) the target. (Thus, the prime–target SOA was 800 ms.) The target remained on the screen for 3000 ms (i.e., response deadline) or until a response was made. Participants made their lexical decisions by pressing the *apostrophe* key for words and the *A* key for nonwords. Each correct response was followed by a 450 ms delay (i.e., intertrial interval). If a response was incorrect, a 170 ms tone was presented simultaneously with the onset of a 450 ms presentation of the word "Incorrect" (displayed slightly below the fixation point).

Results

For all experiments, errors and *RT*s faster than 200 ms or slower than 3000 ms were first excluded, and the overall mean and standard deviation of each participant's word and nonword *RT*s were then computed. The overall error rates were 5.3%, 6.3%, 6.6%, and 9.3% in E1 to E4, respectively. Of the remaining responses, any *RT*s 2.5 *SD*s above or below each participant's respective mean (across all conditions) were removed. The percentage of correct responses that were eliminated due to being designated as an outlier were 2.6%, 2.9%, 2.8%, and 2.8% in E1 to E4, respectively. In general, outlier rates were slightly higher for unrelated (M = 2.8%) than for related (M = 2.1%) targets, but the relative difference between related and unrelated conditions for high- and low-frequency targets was relatively stable across the four experiments.

To perform the distributional analyses, ex-Gaussian parameters (μ, σ, τ) were estimated for each participant across the different experimental conditions, using the quantile maximum likelihood estimation procedure in QMPE 2.18 (Cousineau, Brown, & Heathcote, 2004; Heathcote, Brown, & Mewhort, 2002). This procedure provides unbiased parameter estimates and has been demonstrated to be more effective than continuous maximum likelihood estimation for small samples (Heathcote & Brown, 2004; Speckman & Rouder, 2004). Mean vincentiles for the data were also plotted, providing a graphical complement to the ex-Gaussian fits. As discussed in the Introduction, vincentizing averages *RT* distributions across participants (Andrews & Heathcote, 2001; Ratcliff, 1979; Rouder & Speckman, 2004; Vincent, 1912) to produce the *RT* distribution for a typical participant. Note, for each plot, that empirical vincentiles are represented by data points and standard error bars, while the vincentiles for the respective best-fitting ex-Gaussian distribution are represented by lines. The theoretical vincentiles were computed by line search on the numerical integral of the fitted ex-Gaussian distribution (Heathcote, personal communication, 5th January 2009). The goodness of fit between the empirical and theoretical vincentiles reflects the extent to which the empirical *RT* distributions are being captured by the ex-Gaussian parameters.

The mean *RT*s, error rates, and ex-Gaussian parameters of the *RT* data were submitted to a Priming (related or unrelated) × Frequency (high or low) repeated-measures analysis of variance (ANOVA), with participants treated as random effects. For mean *RT*s and error rates, an ANOVA was also conducted, with items treated as random effects, and minF' treating both items and participants as random effects (see Clark, 1973). We will first consider the results for each experiment. This will be followed by cross-experiment analyses which include Nonword Type (legal or pseudohomophonic) and University (SUNY-A or WUSTL) as between-subject variables.

Experiment 1 (WUSTL, nonword type: legal)

The mean *RT*s, error rates, and ex-Gaussian parameters are displayed in Table 3. The test statistics for the omnibus ANOVA by participants and by items are presented in Table 4. Importantly, there were large main effects of priming and frequency, but these two variables did not interact in the *RT* data. It is also interesting to note that the priming

Table 3
Mean *RT*s, % errors, and ex-Gaussian parameters and their 95% confidence intervals as a function of Frequency, and Priming in Experiment 1 (WUSTL, legal nonwords).

	RT	% Error	μ	σ	τ
Low-frequency targets					
Unrelated	635	8.0	494	61	142
Related	588	4.3	447	59	143
Priming effect	*47 ± 10*[*]	*3.7 ± 1.4*[*]	*47 ± 14*[*]	*2 ± 10*	*−1 ± 11*
High-frequency targets					
Unrelated	607	4.5	475	56	133
Related	567	3.1	434	52	135
Priming effect	*40 ± 7*[*]	*1.4 ± 1.0*[*]	*41 ± 15*[*]	*4 ± 13*	*−2 ± 13*
Interaction	*7 ± 11*	*2.3 ± 1.6*[*]	*6 ± 15*	*−2 ± 12*	*1 ± 14*
Nonwords	678	5.7	537	61	141

[*] $p < .05$.

Table 4
ANOVA table for Experiment 1 (WUSTL, legal nonwords).

Source	*MSE*	*F*	*p*	η_p^2
RT (participant analyses)				
PRIMING	451.36	164.74	<.01	.81
FREQ	381.57	62.11	<.01	.61
PRIMING × FREQ	277.74	1.99	.17	.05
RT (item analyses)				
PRIMING	1930.20	169.64	<.01	.36
FREQ	3805.73	28.73	<.01	.09
PRIMING × FREQ	1930.20	1.76	.19	.01
RT (minF' analyses)	*df*	*F*	*p*	
PRIMING	(1, 135)	83.58	<.01	
FREQ	(1, 242)	19.64	<.01	
PRIMING × FREQ	(1, 152)	.93	.34	
% Error (participant analyses)				
PRIMING	9.02	27.72	<.01	.42
FREQ	6.04	37.10	<.01	.49
PRIMING × FREQ	5.91	8.70	.01	.18
% Error (item analyses)				
PRIMING	27.25	34.41	<.01	.10
FREQ	47.01	17.87	<.01	.06
PRIMING × FREQ	27.25	7.07	.01	.02
% Error (minF' analyses)	*df*	*F*	*p*	
PRIMING	(1, 117)	15.35	<.01	
FREQ	(1, 236)	12.06	<.01	
PRIMING × FREQ	(1, 162)	3.90	.05	
μ (participant analyses)				
PRIMING	1406.47	55.30	<.01	.59
FREQ	337.33	30.07	<.01	.44
PRIMING × FREQ	574.92	.64	.43	.02
σ (participant analyses)				
PRIMING	910.00	.44	.51	.01
FREQ	450.00	3.50	.07	.08
PRIMING × FREQ	376.16	.17	.68	.00
τ (participant analyses)				
PRIMING	982.97	.04	.85	.00
FREQ	516.96	5.83	.02	.13
PRIMING × FREQ	506.34	.04	.83	.00

Note: The *df*s are (1, 39) and (1, 298) for the participant and item analyses, respectively.

effect was fully mediated by μ (distributional shifting). This implies that semantic priming was produced *purely* by distributional shifting (replicating the pattern reported by Balota et al., 2008), and this shift was of similar magnitude for high- and low-frequency targets. It is also noteworthy that there was evidence of a priming by frequency interaction in error rates (in both participant and item analyses), with larger priming effects for low-frequency targets than for high-frequency targets.

Vincentile analysis

The vincentile plots provide converging support for distributional shifting as a function of prime relatedness. As can be seen in Fig. 3, the priming effects for both high- and low-frequency targets were approximately the same size and constant across the vincentiles. To further explore the reliability of this pattern, we conducted an ANOVA with vincentile as a within-subject variable.[3] This revealed

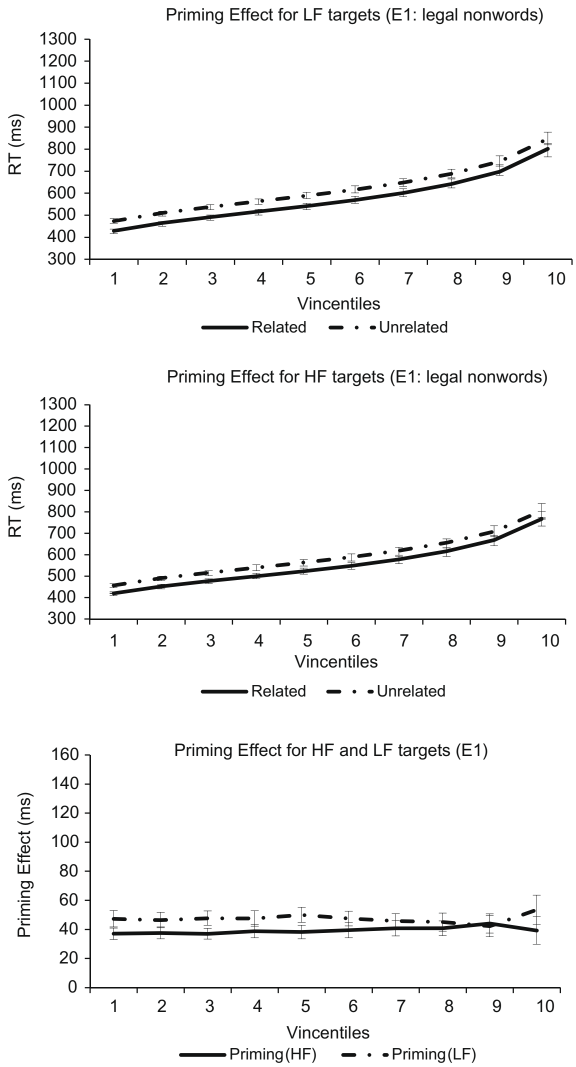

Fig. 3. Lexical decision performance from Experiment 1 (WUSTL, pronounceable nonwords) as a function of Priming and Vincentiles for low-frequency targets (Top Panel) and high-frequency targets (Middle Panel), along with the priming effect as a function of Vincentiles (Bottom Panel). For top two panels, empirical Vincentiles are represented by error bars while fitted ex-Gaussian Vincentiles are represented by lines. Error bars in the bottom panel reflect the standard errors of the difference scores.

that neither the Priming × Vincentile (p = .88) nor the Priming × Frequency × Vincentile interaction (p = .55) approached significance, confirming that priming is relatively invariant across the *RT* distribution (i.e., reflecting a simple distributional shift).

Experiment 2 (WUSTL, nonword type: pseudohomophonic)

The mean *RT*s, error rates, and ex-Gaussian parameters are displayed in Table 5. The test statistics for the omnibus ANOVA by participants and by items are presented in Table 6. Again, there were clear and large additive effects of Priming and Frequency in *RT*s and μ. There was also a

[3] In the present and subsequent analyses involving the Vincentile, we used the Greenhouse–Geisser correction for potential violations of sphericity.

Table 5
Mean *RT*s, % errors, and ex-Gaussian parameters and their 95% confidence intervals as a function of Frequency, and Priming in Experiment 2 (WUSTL, pseudohomophones).

	RT	% Error	μ	σ	τ
Low-frequency targets					
Unrelated	676	8.5	496	58	182
Related	640	4.8	461	62	180
Priming effect	*36 ± 12**	*3.7 ± 1.4**	*35 ± 14**	*−4 ± 9*	*2 ± 13*
High-frequency targets					
Unrelated	633	4.8	472	51	163
Related	603	3.2	446	52	159
Priming effect	*30 ± 12**	*1.6 ± 0.9**	*26 ± 8**	*−1 ± 6*	*4 ± 12*
Interaction	*6 ± 14*	*2.1 ± 1.8**	*9 ± 15*	*−3 ± 13*	*−2 ± 15*
Nonwords	709	7.2	533	55	176

* $p < .05$.

Priming × Frequency interaction in error rates, with larger priming effects for low-frequency targets. In two experiments, additive effects of priming and frequency were observed in *RT*s, but interactive effects of the two variables were observed in accuracy rates. This is an interesting pattern that will be discussed in greater detail in the General Discussion. Finally, like E1, priming effects were fully mediated by μ (i.e., distributional shifting), and were qualitatively similar for low- and high-frequency targets.

Table 6
ANOVA table for Experiment 2 (WUSTL, pseudohomophones).

Source	*MSE*	*F*	*p*	η_p^2
RT (participant analyses)				
PRIMING	1060.86	50.15	<.01	.52
FREQ	711.56	106.53	<.01	.69
PRIMING × FREQ	618.65	.70	.41	.02
RT (Item analyses)				
PRIMING	5017.22	57.58	<.01	.16
FREQ	1971.00	89.81	<.01	.23
PRIMING × FREQ	1971.00	.72	.40	.00
RT (minF' analyses)	*df*	*F*	*p*	
PRIMING	(1, 147)	26.80	<.01	
FREQ	(1, 184)	48.73	<.01	
PRIMING × FREQ	(1, 159)	.35	.56	
% Error (participant analyses)				
PRIMING	8.04	41.59	<.01	.47
FREQ	12.69	26.34	<.01	.36
PRIMING × FREQ	9.04	5.61	.02	.11
% Error (item analyses)				
PRIMING	57.31	18.23	<.01	.06
FREQ	24.39	42.82	<.01	.13
PRIMING × FREQ	24.39	6.50	.01	.02
% Error (minF' analyses)	*df*	*F*	*p*	
PRIMING	(1, 278)	12.67	<.01	
FREQ	(1, 116)	16.31	<.01	
PRIMING × FREQ	(1, 146)	3.01	.09	
μ (participant analyses)				
PRIMING	859.12	52.79	<.01	.53
FREQ	1042.07	18.41	<.01	.28
PRIMING × FREQ	620.83	1.38	.25	.03
σ (participant analyses)				
PRIMING	182.75	1.11	.30	.02
FREQ	755.38	4.26	.04	.08
PRIMING × FREQ	467.20	.12	.73	.00
τ (participant analyses)				
PRIMING	1186.35	.29	.59	.01
FREQ	1106.70	17.60	<.01	.27
PRIMING × FREQ	688.70	.12	.73	.00

Note: The *df*s are (1, 47) and (1, 298) for the participant and item analyses, respectively.

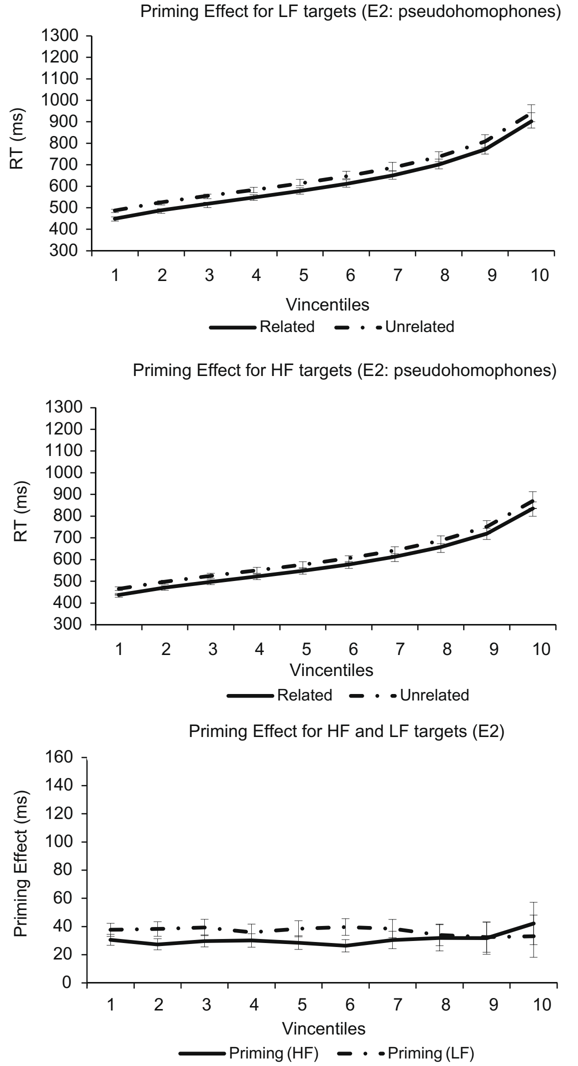

Fig. 4. Lexical decision performance from Experiment 2 (WUSTL, pseudohomophonic nonwords) as a function of Priming and Vincentiles for low-frequency targets (Top Panel) and high-frequency targets (Middle Panel), along with the priming effect as a function of Vincentiles (Bottom Panel). For top two panels, empirical Vincentiles are represented by error bars while fitted ex-Gaussian Vincentiles are represented by lines. Error bars in the bottom panel reflect the standard errors of the difference scores.

Vincentile analysis

The vincentile plots (see Fig. 4) show that the priming effect was relatively constant across the *RT* distribution, and of the same magnitude for high- and low-frequency targets. Neither the Priming × Vincentile ($p = .82$) nor the Priming × Frequency × Vincentile interaction ($p = .46$) was significant.

Experiments 1 and 2 (WUSTL)

To examine the effect of nonword type on the effects of word frequency and semantic priming, we included Nonword Type as a between-participants factor and submitted the mean *RT*s, error rates, and ex-Gaussian parameters to a Priming × Frequency × Nonword Type (legal or pseudohomophonic) mixed-factor ANOVA. Only interaction effects with Nonword Type are reported.

RTs and error rates

For mean *RT*s, the Priming × Nonword Type interaction did not reach statistical significance [$F(1, 86) = 2.68$, $MSE = 784.46$, $p = .11$, $\eta_p^2 = .03$]; the priming effect was 11 ± 13 ms numerically *smaller* when nonwords were pseudohomophones than when they were legal. The Frequency × Nonword Type interaction was significant [$F(1, 86) = 9.21$, $MSE = 561.91$, $p < .01$, $\eta_p^2 = .10$]; the word frequency effect was 16 ± 10 ms larger when nonwords were pseudohomophones. Neither the main effect of Nonword Type nor the three-way interaction was significant, all $Fs < 2.35$, $p > .12$. For error rates, the Nonword Type main effect was not significant and Nonword Type also did not interact with any other variable, all $Fs < 1$.

Ex-Gaussian parameters

For μ, the Nonword Type main effect was not significant but the Priming × Nonword Type interaction was marginally significant [$F(1, 86) = 3.52$, $MSE = 1107.33$, $p = .06$, $\eta_p^2 = .04$]. Interestingly, the priming effect in μ was 14 ± 15 ms *smaller* when nonwords were pseudohomophones. Neither the main effect of Nonword Type nor any other interaction was significant, all $Fs < 1.92$, $p > .17$. Turning to σ, the Nonword Type main effect was not significant and Nonword Type did not interact with any other variable, all $Fs < 1.16$, $p > .28$. Finally, for τ, the main effect of Nonword Type approached significance [$F(1, 86) = 3.41$, $MSE = 27176.18$, $p = .07$, $\eta_p^2 = .04$], but this was qualified by the marginally significant Frequency × Nonword Type interaction [$F(1, 86) = 3.42$, $MSE = 839.26$, $p = .07$, $\eta_p^2 = .04$]. The word frequency effect was 12 ± 13 ms larger in τ when nonwords were pseudohomophones. Nonword Type did not interact with any other variables, all $Fs < 1$. In summary, when pseudohomophones were used as nonwords, word frequency effects became *larger* in the tail of the *RT* distribution, but priming effects became *smaller* in the modal portion of the *RT* distribution.

Experiment 3 (SUNY-A, nonword type: legal)

The mean *RT*s, error rates, and ex-Gaussian parameters are displayed in Table 7. The test statistics for the omnibus ANOVA by participants and by items are presented in Table 8. The main effects of Priming and Frequency were highly reliable in *RT*s, error rates, and μ. More importantly, in the participants analyses, the Priming × Frequency interaction was significant (or approaching significance) in *RT*s, error rates, μ, and σ. (For the item analyses, the interaction was significant in error rates, and approached significance in *RT*s). In general, there was greater priming for low-frequency targets than for high-frequency targets in these measures.

Table 7
Mean *RT*s, % errors, and ex-Gaussian parameters and their 95% confidence intervals as a function of Frequency, and Priming in Experiment 3 (SUNY-A, pronounceable nonword type).

	RT	*% Error*	μ	σ	τ
Low-frequency targets					
Unrelated	730	7.5	545	83	183
Related	665	3.8	482	65	181
Priming effect	*65 ± 15**	*3.7 ± 1.3**	*63 ± 19**	*18 ± 18**	*2 ± 17*
High-frequency targets					
Unrelated	690	4.4	501	55	187
Related	650	2.7	469	60	181
Priming effect	*40 ± 14**	*1.7 ± 1.1**	*32 ± 11**	*−5 ± 12*	*6 ± 12*
Interaction	*25 ± 18**	*2.0 ± 1.7**	*31 ± 23**	*23 ± 24*	*−4 ± 20*
Nonwords	802	8.7	610	75	192

* $p < .05$.

Table 8
ANOVA table for Experiment 3 (SUNY-A, pronounceable nonword type).

Source	*MSE*	*F*	*p*	η_p^2
RT (participant analyses)				
PRIMING	1258.51	86.78	<.01	.69
FREQ	1441.23	20.51	<.01	.35
PRIMING × FREQ	783.99	7.52	.01	.16
RT (item analyses)				
PRIMING	2948.97	106.83	<.01	.26
FREQ	6937.45	21.77	<.01	.07
PRIMING × FREQ	2948.97	3.20	.07	.01
RT (minF′ analyses)	*df*	*F*	*p*	
PRIMING	(1, 118)	47.88	<.01	
FREQ	(1, 132)	10.56	<.01	
PRIMING × FREQ	(1, 254)	2.24	.14	
% Error (participant analyses)				
PRIMING	7.33	39.31	<.01	.50
FREQ	6.85	26.62	<.01	.41
PRIMING × FREQ	6.65	6.32	.02	.14
% Error (item analyses)				
PRIMING	29.87	34.94	<.01	.10
FREQ	49.26	14.19	<.01	.05
PRIMING × FREQ	29.87	5.00	.03	.02
% Error (minF′ analyses)	*df*	*F*	*p*	
PRIMING	(1, 151)	18.50	<.01	
FREQ	(1, 221)	9.26	<.01	
PRIMING × FREQ	(1, 165)	2.79	.10	
μ *(participant analyses)*				
PRIMING	1197.00	75.21	<.01	.66
FREQ	824.91	40.32	<.01	.51
PRIMING × FREQ	1234.67	7.83	.01	.17
σ *(participant analyses)*				
PRIMING	1018.57	1.76	.19	.04
FREQ	1163.34	9.24	<.01	.19
PRIMING × FREQ	1265.99	3.95	.054	.09
τ *(participant analyses)*				
PRIMING	1070.08	.60	.44	.02
FREQ	1354.49	.10	.75	.00
PRIMING × FREQ	988.95	.17	.68	.00

Note: The *df*s are (1, 39) and (1, 298) for the participant and item analyses, respectively.

Moreover, the results from the ex-Gaussian analyses indicated that the priming for high-frequency targets (40 ms) was mediated largely by shifting (32 ms). In contrast, priming for low-frequency targets (65 ms) was mediated by shift-

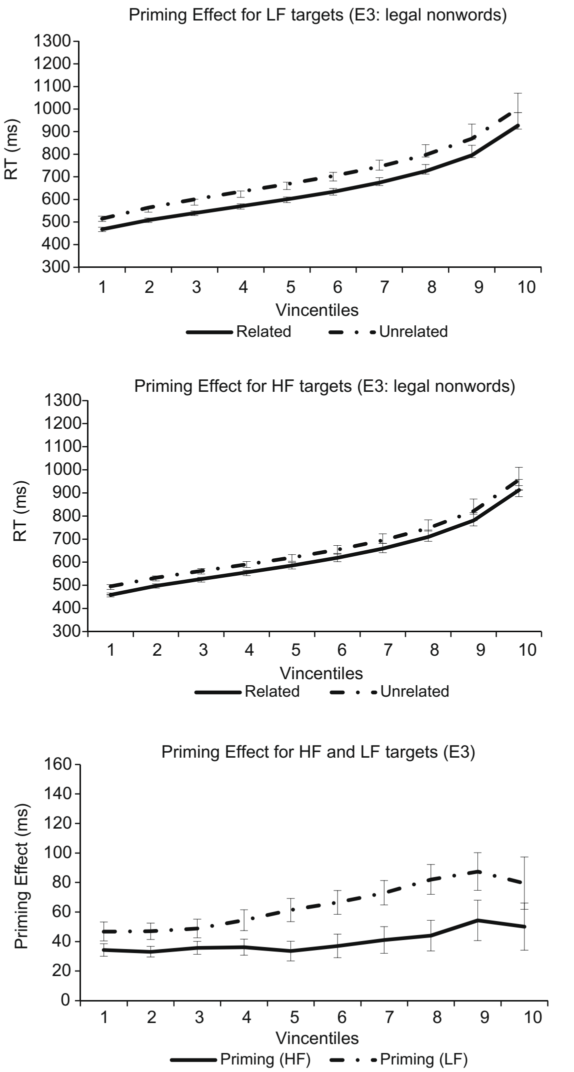

Fig. 5. Lexical decision performance from Experiment 3 (SUNY-A, pronounceable nonwords) as a function of Priming and Vincentiles for low-frequency targets (Top Panel) and high-frequency targets (Middle Panel), along with the priming effect as a function of Vincentiles (Bottom Panel). For top two panels, empirical Vincentiles are represented by error bars while fitted ex-Gaussian Vincentiles are represented by lines. Error bars in the bottom panel reflect the standard errors of the difference scores.

ing (63 ms) and σ (18 ms), but not by stretching the tail of the distribution. In other words, the larger priming effect for low-frequency targets was driven by both μ (distributional shifting) and σ (greater variability in the modal *RT*s).

Vincentile analysis

The results from the ex-Gaussian analyses are broadly consistent with the vincentile plots (see Fig. 5). Although the Priming × Frequency × Vincentile interaction was not significant ($p = .34$), this may have been due to noise in the final two vincentiles. When we restricted our analyses to the first eight vincentiles, the three-way interaction was significant, $p = .036$, indicating that across the *RT* distribution, priming effects were constant in magnitude for high-frequency targets, but increasing (due to greater variability of the modal *RT*s) for low-frequency targets. Despite the *post hoc* nature of this analysis, it is noteworthy that the three-way interaction holds for the first eight vincentiles, which constitute a clear majority of the dataset. In marked contrast, reexamining the Priming × Frequency × Vincentile in the first two experiments, using only the first eight vincentiles, yielded non-significant interactions for both E1 and E2, $Fs < 1$.

Table 9
Mean *RT*s, % errors, and ex-Gaussian parameters and their 95% confidence intervals as a function of Frequency, and Priming in Experiment 4 (SUNY-A, pseudohomophonic nonword type).

	RT	% *Error*	μ	σ	τ
Low-frequency targets					
Unrelated	788	10.4	526	65	265
Related	723	7.2	502	74	223
Priming effect	$65 \pm 19^*$	$3.2 \pm 2.1^*$	$24 \pm 13^*$	-9 ± 12	$42 \pm 17^*$
High-frequency targets					
Unrelated	720	5.6	496	62	227
Related	683	3.9	470	56	215
Priming effect	$37 \pm 17^*$	$1.7 \pm 1.0^*$	$26 \pm 11^*$	6 ± 10	12 ± 18
Interaction	$28 \pm 26^*$	1.5 ± 2.6	-2 ± 13	$-15 \pm 14^*$	$30 \pm 28^*$
Nonwords	848	11.8	632	92	217

* $p < .05$.

Experiment 4 (SUNY-A, nonword type: pseudohomophonic)

The mean *RT*s, error rates, and ex-Gaussian parameters are displayed in Table 9. The test statistics for the omnibus ANOVA by participants and by items are presented in Table 10. Main effects of Priming and Frequency were observed in *RT*s, error rates, μ, and τ. There was also a Priming × Frequency interaction in *RT*s, σ, and τ, in both participant and, where applicable, item analyses. In all of these measures except σ, priming effects were larger for low-frequency targets than for high-frequency targets. In E4, the interaction between Frequency and Priming occurred in σ and τ, but not in μ, in contrast to the findings in E3 (legal nonwords), where the interaction was mediated by μ and σ. Specifically, priming for high-frequency targets (37 ms) reflected mainly distributional shifting (26 ms). Low-frequency targets yielded the opposite pattern, where priming (65 ms) reflected some shifting (24 ms) but mostly skewing (42 ms). As a result, the Priming × Frequency interaction in mean *RT*s observed in the SUNY-A students can be attributed to the slow *RT*s in the tail of the distribution.

Vincentile analysis

These trends are compatible with the vincentile plots (see Fig. 6), which confirm that the Priming × Frequency interaction was indeed largest in the slowest vincentiles. The Priming × Frequency × Vincentile interaction was significant ($p = .006$), and follow-up analyses revealed that the Priming × Vincentile interaction was significant for low-frequency targets ($p < .001$), but not for high-frequency targets ($p = .18$), reinforcing the idea that priming for high-frequency words was mediated mainly by distributional shifting, but priming for low-frequency words was mediated by a mixture of skewing and shifting.

Table 10
ANOVA table for Experiment 4 (SUNY-A, pseudohomophonic nonword type).

Source	*MSE*	*F*	*p*	η_p^2
RT (participant analyses)				
PRIMING	1085.89	67.85	<.01	.72
FREQ	2170.22	37.07	<.01	.58
PRIMING × FREQ	1124.26	4.58	.04	.15
RT (item analyses)				
PRIMING	7483.50	51.87	<.01	.15
FREQ	10711.86	54.22	<.01	.15
PRIMING × FREQ	7483.50	4.62	.03	.02
RT (minF' analyses)	*df*	*F*	*p*	
PRIMING	(1, 125)	29.40	<.01	
FREQ	(1, 73)	22.02	<.01	
PRIMING × FREQ	(1, 98)	2.30	.13	
% Error (participant analyses)				
PRIMING	6.96	23.27	<.01	.46
FREQ	29.61	15.68	<.01	.37
PRIMING × FREQ	11.07	1.38	.25	.05
% Error (item analyses)				
PRIMING	37.16	22.94	<.01	.07
FREQ	77.10	31.55	<.01	.10
PRIMING × FREQ	37.16	2.40	.12	.01
% Error (minF' analyses)	*df*	*F*	*p*	
PRIMING	(1, 100)	11.55	<.01	
FREQ	(1, 59)	10.47	<.01	
PRIMING × FREQ	(1, 65)	.88	.35	
μ (participant analyses)				
PRIMING	629.26	28.27	<.01	.51
FREQ	773.54	34.13	<.01	.56
PRIMING × FREQ	294.62	.05	.82	.00
σ (participant analyses)				
PRIMING	475.95	.26	.61	.01
FREQ	424.05	7.19	.01	.21
PRIMING × FREQ	326.16	4.56	.04	.14
τ (participant analyses)				
PRIMING	738.40	27.31	<.01	.50
FREQ	2240.50	6.85	.01	.20
PRIMING × FREQ	1322.57	4.59	.04	.15

Note: The *df*s are (1, 27) and (1, 298) for the participant and item analyses, respectively.

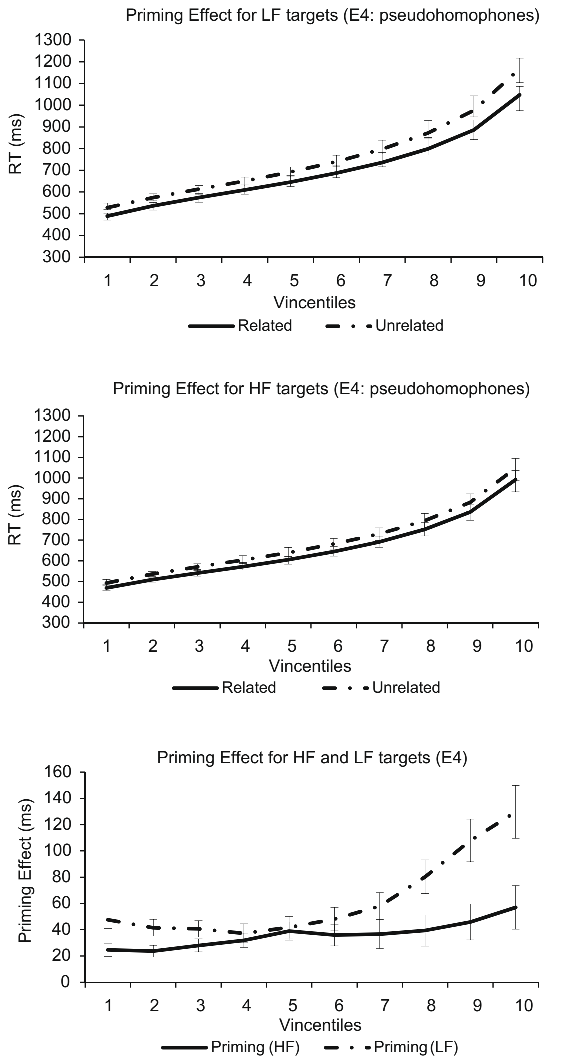

Fig. 6. Lexical decision performance from Experiment 3 (SUNY-A, pseudohomophonic nonwords) as a function of Priming and Vincentiles for low-frequency targets (Top Panel) and high-frequency targets (Middle Panel), along with the priming effect as a function of Vincentiles (Bottom Panel). For top two panels, empirical Vincentiles are represented by error bars while fitted ex-Gaussian Vincentiles are represented by lines. Error bars in the bottom panel reflect the standard errors of the difference scores.

Experiments 3 and 4 (SUNY-A)

In order to directly investigate the influence of nonword type, the data from Experiments 3 and 4 were combined. The mean *RT*s, error rates, and ex-Gaussian parameters were submitted to a Priming × Frequency × Nonword Type mixed-factor ANOVA. Only the effects associated with Nonword Type are reported.

RTs and error rates

For mean *RT*s, the Frequency × Nonword Type interaction was significant [$F(1, 66) = 6.61$, $MSE = 1739.45$, $p < .05$, $\eta_p^2 = .09$]; the word frequency effect was 27 ± 21 ms greater when nonwords were pseudohomophones. Neither the main effect of Nonword Type nor other interactions associated with Nonword Type were significant, all *F*s < 1.96, $p > .16$. For error rates, the main effect of Nonword Type was significant [$F(1, 66) = 4.84$, $MSE = 64.28$, $p < .05$, $\eta_p^2 = .07$]; the error rate was 2.2 ± 2.0% higher when nonwords were pseudohomophones. The marginally significant Frequency × Nonword Type interaction [$F(1, 66) = 3.82$, $MSE = 16.16$, $p = .06$, $\eta_p^2 = .06$] further showed that the word frequency effect was 2.0 ± 2.0% larger when nonwords were pseudohomophones. Neither the main effect nor other interactions associated with Nonword Type approached significance, all *F*s < 1.

Ex-Gaussian parameters

Consistent with the results from the Washington University sample, for μ, the Priming × Nonword Type interaction was significant [$F(1, 66) = 8.44$, $MSE = 964.74$, $p < .01$, $\eta_p^2 = .11$]; priming effects were 23 ± 16 ms *smaller* when nonwords were pseudohomophones. In addition, there was a significant Priming × Nonword Type × Frequency interaction [$F(1, 66) = 5.13$, $MSE = 850.11$, $p < .05$,

$\eta_p^2 = .07$], which was due to Priming and Frequency yielding interactive effects in the legal nonword condition and additive effects in the pseudohomophone condition. The main effect of Nonword Type was not significant nor did it interact with any other variable, all *F*s < 1. Turning to σ, the Priming × Nonword Type × Frequency interaction was significant [$F(1, 66) = 6.37$, $MSE = 881.52$, $p < .05$, $\eta_p^2 = .09$]. In the legal nonword condition, priming effects were 22.4 ± 22.8 ms larger in σ for low-frequency targets than for high-frequency targets. However, in the pseudohomophone condition, the pattern was opposite, with 15 ± 14 ms larger priming effects in σ for high-frequency than low-frequency targets. The main effect of Nonword Type was not significant nor did it interact with any other variable, all *F*s < 1.60, $p > .21$. Finally, for τ, the main effect of Nonword Type [$F(1, 66) = 4.64$, $MSE = 34286.76$, $p < .05$, $\eta_p^2 = .07$] was significant; τ was 50 ± 46 ms larger when nonwords were pseudohomophones. Both the Priming × Nonword Type [$F(1, 66) = 9.20$, $MSE = 934.40$, $p < .01$, $\eta_p^2 = .12$] and the Frequency × Nonword Type interactions [$F(1, 66) = 6.12$, $MSE = 1716.94$, $p < .05$, $\eta_p^2 = .09$] were significant. In the presence of pseudohomophones (compared to legal nonwords), priming effects were 23 ± 15 ms larger and word frequency effects were 25 ± 20 ms larger. The Priming × Nonword Type × Frequency interaction was also significant [$F(1, 66) = 4.12$, $MSE = 1125.43$, $p < .05$, $\eta_p^2 = .06$]. The three-way interaction in τ was due to Priming and Frequency yielding interactive effects (i.e., larger priming effects for low-frequency words) *only* when nonwords were pseudohomophones. To summarize, when pseudohomophones were used as nonwords, priming effects became smaller in the modal portion of the *RT* distribution while word frequency effects became larger in the tail of the *RT* distribution, replicating the trends observed in the first two experiments. In addition, for E3 and E4, one also observes larger priming effects and a Priming × Frequency interaction in the tail of the distribution when pseudohomophones were used.

Experiments 1–4 (WUSTL vs. SUNY-A)

As shown in Table 1, SUNY-A participants and WUSTL participants differed in their years of education and vocabulary knowledge. To verify if University modulates the Frequency × Priming interaction in our experiments, we included the data from all four experiments and submitted the mean *RT*s, error rates, and ex-Gaussian parameters to a Priming × Frequency × Nonword Type × University (WUSTL or SUNY-A) mixed-factor ANOVA.[4] Only the effects associated with University are reported.

RTs and error rates

For mean *RT*s, the main effect of University was significant [$F(1, 152) = 19.22$, $MSE = 60002.91$, $p < .01$, $\eta_p^2 = .11$]; the WUSTL participants were 88 ± 39 ms faster than the SUNY-A participants. The Priming × University interaction was significant [$F(1, 152) = 7.19$, $MSE = 959.73$, $p < .01$, $\eta_p^2 = .05$]; priming effects were 14 ± 10 ms larger in SUNY-A. Importantly, the three-way interaction was also significant [$F(1, 152) = 5.11$, $MSE = 663.36$, $p < .05$, $\eta_p^2 = .03$]; the Priming × Frequency interaction (i.e., greater priming for low-frequency targets) was 20 ± 17 ms larger for SUNY-A participants than for WUSTL participants. None of the other interactions associated with University was significant, all *F*s < 2.45, *p*s > .12. For error rates, neither the main effect of Nonword Type nor any interaction associated with University approached significance, all *F*s < 2.69, *p*s > .10.

Ex-Gaussian parameters

For μ, the main effect of University and the University × Frequency interaction were significant [$F(1, 152) = 11.58$, $MSE = 14462.20$, $p < .01$, $\eta_p^2 = .07$ and $F(1, 152) = 6.91$, $MSE = 758.31$, $p < .01$, $\eta_p^2 = .04$, respectively]. More importantly, for μ, the University × Priming × Nonword Type × Frequency interaction was also significant [$F(1, 152) = 4.05$, $MSE = 708.31$, $p < .05$, $\eta_p^2 = .03$]. Although WUSTL participants produced additive effects of Frequency and Priming for both legal nonwords and pseudohomophones, SUNY-A participants produced additive effects for pseudohomophones but interactive effects for legal nonwords. No other interaction associated with University was significant, all *F*s < 1. Turning to σ, the main effect of University was significant [$F(1, 152) = 6.88$, $MSE = 1587.45$, $p < .01$, $\eta_p^2 = .04$]. Like μ, the University × Priming × Nonword Type × Frequency interaction was also significant [$F(1, 152) = 5.25$, $MSE = 623.86$, $p < .05$, $\eta_p^2 = .03$]. This four-way interaction was driven by two opposing Priming × Nonword Type × Frequency interactive effects in the SUNY-A participants, discussed above in the analyses for Experiments 3 and 4. None of the other interactions associated with University was significant, all *F*s < 2.00, *p*s > .16. Finally, for τ, the main effect of University and the University × Priming interaction were significant [$F(1, 152) = 14.14$, $MSE = 30265.47$, $p < .01$, $\eta_p^2 = .09$ and $F(1, 152) = 7.72$, $MSE = 1025.10$, $p < .01$, $\eta_p^2 = .05$, respectively]. Importantly, the University × Priming × Nonword Type × Frequency interaction was also significant [$F(1, 152) = 4.02$, $MSE = 831.61$, $p < .05$, $\eta_p^2 = .03$]. This four-way interaction was driven by SUNY-A participants producing a Priming × Frequency interaction (i.e., greater priming in τ for low-frequency targets) *only* when nonwords were pseudohomophones.

Composite analyses

Overall, these analyses provide preliminary evidence that vocabulary knowledge predicts the word frequency by semantic priming interaction, as reflected by between University comparisons. However, these results may also reflect other pre-existing differences between the two samples. In order to assess the influence of vocabulary

[4] The WUSTL sample had more years of education and higher vocabulary scores than the SUNY-A sample. In order to address this confound, we included years of education as a covariate and re-ran all the ANOVAs that included University as a between-subject factor. Importantly, these analyses yielded qualitatively similar trends, confirming that vocabulary knowledge, rather than years of education, was driving the group differences. This is consistent with a follow-up analysis where we found a significant difference in vocabulary scores between participants from the two universities, even after chronological age and years of education are controlled for [$F(1, 142) = 40.96$, $MSE = 8.017$].

knowledge more directly, we also conducted analyses of covariance (ANCOVAs) as a function of vocabulary knowledge (lower vs. higher), with nonword type as a covariate. Participants from the two universities were combined, effectively ignoring the University variable. The three-way Priming × Frequency × Vocabulary Knowledge interaction was not significant. However, using a median split to dichotomize Vocabulary Knowledge, a continuous measure, would have diminished the statistical power of our analysis (Cohen, 1983; Humphreys, 1978; Maxwell & Delaney, 1993). When we used the top third and bottom third of the Shipley scores to define high- and low-vocabulary-knowledge participants, respectively, while at the same time ensuring that each counterbalancing list was equally represented across participants in each group, the three-way interaction was indeed reliable, $F(1, 101) = 4.45$, $MSE = 607.43$, $p < .05$, $\eta_p^2 = .04$. The Priming × Frequency interaction was reliable for the low-vocabulary-knowledge group, who showed greater priming for low-frequency targets ($d = 60$ ms) than for high-frequency targets ($d = 33$ ms), $p < .01$. High-vocabulary-knowledge participants showed more similar priming for low-frequency ($d = 44$ ms) and high-frequency ($d = 34$ ms) words, $p = .096$. Importantly, the significant three-way interaction was *totally* mediated by the τ parameter, $F(1, 101) = 6.99$, $MSE = 870.85$, $p < .01$, $\eta_p^2 = .07$, which is consistent with the idea that the tail of the *RT* distribution is especially sensitive to the integrity of lexical representations.

We further pursued this pattern by conducting hierarchical multiple regression analyses to determine whether participants' vocabulary scores predicted the magnitude of their Priming × Frequency interaction after controlling for appropriate variables. We entered participants' age and years of education in the first step, whether they received legal nonwords (i.e., 0) or pseudohomophones (i.e., 1) in the second step, and their vocabulary raw score in the final step. The dependent measures were the "interaction" scores (i.e., the difference in the priming effect for low-frequency targets and for high-frequency targets) in mean *RT*s, error rates, μ, σ, and τ. (We also conducted parallel analyses using *z*-transformed *RT*s to control for overall differences in processing speed, see Faust, Balota, Spieler, & Ferraro, 1999, and found qualitatively identical findings.) Importantly, after partialling out variance accounted for by age, years of education, and nonword type, vocabulary scores predicted the interaction score in *RT*s [$\beta = -.168$, $p = .06$] and τ [$\beta = -.195$, $p = .03$], but not in error rates [$\beta = .012$, $t < 1$], μ [$\beta = .037$, $t < 1$], or σ [$\beta = .001$, $t < 1$]. The negative regression coefficients indicate that participants with more vocabulary knowledge produced *smaller* priming differences between low- and high-frequency targets in both mean *RT*s and in τ.[5] Together with the ANOVAs described earlier, these regression analyses provide converging evidence that higher-vocabulary knowledge readers are more likely to produce *additive* effects of priming and frequency, while lower-vocabulary knowledge readers are more likely to produce *interactive* effects of the two variables, with the interaction primarily occurring in the tail of the *RT* distribution.

Interestingly, we also have access to a new dataset that provides converging support for the findings reported in this study. This dataset is based on an in-progress multi-university primed word recognition megastudy (http://spp.montana.edu) that includes participants from WUSTL and SUNY-A. Importantly, using tests from the Woodcock–Johnson Tests of Achievement (WJ III; Woodcock, McGrew, & Mather, 2001), participants' vocabulary knowledge was assessed by asking them to generate synonyms and antonyms for printed words, and having them complete analogies (e.g., *elephant* – *big*; *mouse* – ?). As before, participants from WUSTL and SUNY-A were combined, and for each participant, a composite measure of vocabulary knowledge based on WJ III scores (Synonyms, Antonyms, and Analogies) was computed. We used the top third ($n = 82$) and bottom third ($n = 82$) of the WJ III scores to define high- and low-vocabulary-knowledge participants, respectively, while ensuring that each counterbalancing list was equally represented across participants in each group. In lexical decision performance, the Priming × Frequency interaction was larger for low-knowledge participants ($d = 27$ ms) than for high-knowledge participants ($d = 15$ ms). More importantly, the critical three-way interaction between priming, frequency, and vocabulary knowledge approached or reached significance for raw *RT*s ($p = .092$) and more importantly, in the *z*-transformed *RT*s ($p = .016$). Hence, the three-way interaction can be replicated on an independent sample of participants using a different measure of vocabulary knowledge.

General discussion

This study yielded the following noteworthy findings. First, in line with Plaut and Booth (2000), semantic priming and word frequency do *not* always interact in lexical decision performance. However, in contrast to Plaut and Booth, whose *low*-perceptual-ability participants yielded additive effects, additive effects were associated with *higher*-vocabulary-knowledge readers while interactive effects were associated with *lower*-vocabulary-knowledge readers. Second, the *RT* distributional analyses revealed interesting new constraints on the semantic priming effect that replicate and extend the distributional effects reported in Balota et al. (2008).

Semantic priming and word frequency do not always interact in RTs

Since Becker (1979) first reported larger semantic priming effects for low-frequency targets compared to high-frequency targets, the semantic priming by frequency interaction has become a benchmark finding in the semantic priming literature (see Neely, 1991; McNamara, 2005, for reviews). The present study, along with Plaut and Booth

[5] After removing one outlier participant (whose vocabulary score was 15, more than three *SD*s below the overall mean, i.e., 32.6), we found qualitatively similar results in the hierarchical regression analyses. That is, after partialling out variance accounted for by age, years of education, and nonword type, vocabulary scores still predicted the interaction score in *RT*s [$\beta = -.165$, $p = .068$] and τ [$\beta = -.229$, $p = .011$], but not in error rates, μ, or σ, $ts < 1$.

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

(2000), indicates that the interaction may not be as robust as researchers have heretofore assumed. Instead, whether the two variables interact or not seems to depend on individual differences, as reflected by perceptual ability and vocabulary knowledge. In Plaut and Booth's study, high-perceptual-ability participants produced interactive effects while low-perceptual-ability participants produced additive effects. In our study, we observed a different pattern; additive effects when participants had *more* vocabulary knowledge, and interactive effects when participants had *less* vocabulary knowledge. If perceptual ability and vocabulary knowledge both broadly reflect the fluency of lexical processing, then one would expect the same three-way interaction in both studies. The puzzling discrepancy will be discussed in greater depth later.

Returning to our results, the collective analyses suggest that the semantic priming by frequency interaction is more likely to emerge for low-vocabulary-knowledge participants.[6] For high-vocabulary-knowledge participants, priming and frequency effects were additive at the level of the mean response latency, and this additive pattern persisted whether legal nonwords or pseudohomophones were used as distracters. More intriguingly, the distributional analyses (see Tables 3 and 5 and Figs. 3 and 4) indicate that semantic priming was primarily reflected by a shift in the *RT* distribution, replicating the findings reported by Balota et al. (2008). Distributional shifting is most consistent with a simple head-start mechanism in lexical processing, whereby the effect of the prime is to pre-activate the target representation, which then speeds up lexical access by some constant amount of time. Interestingly, the *RT* distributions for high- and low-frequency targets were shifted *to the same extent* by semantic priming, likely reflecting the fact that we controlled for associative strength across high- and low-frequency words.

In contrast, for participants with relatively less vocabulary knowledge, priming and frequency clearly interacted, with larger priming effects for low-frequency targets. Obviously, these results are more consistent with the extant literature, where greater priming for low-frequency targets is usually reported. When one considers the legal nonword condition (E3), the distributional analyses (see Table 7 & Fig. 5) revealed that priming for high-frequency targets was reflected predominantly by distributional shifting, while priming for low-frequency targets was reflected by shifting and greater variability in modal *RT*s. Specifically, for high-frequency targets, the magnitude of the priming effect was relatively invariant across vincentiles, while for low-frequency targets, priming effects increased in size as *RT*s became longer. These trends are even clearer when word–nonword discrimination difficulty was increased by using pseudohomophonic nonwords (see Table 9 & Fig. 6). Here, even high-frequency targets showed some evidence of distributional skewing in priming (although these trends were not statistically significant), while for low-frequency targets, priming effects were relatively stable across the first five vincentiles, but increased dramatically (from 40 ms to 120 ms) in the slower vincentiles.

Lexical integrity, priming, and frequency

The present findings can be reconciled with the lexical integrity hypothesis in a straightforward manner. For low-lexical-integrity participants with less vocabulary knowledge, pure distributional shifting, and its attendant head-start mechanism, was observed only when targets were strongly represented (i.e., high-frequency words). When targets (i.e., low-frequency words) had relatively less integrity, target processing was further from threshold, and there was greater reliance on prime information for resolving these difficult targets, with reliance being proportional to the difficulty of the trial. For high-lexical-integrity participants with more vocabulary knowledge, high- and low-frequency words were fluently processed due to their equally strong representations; here, priming reflected a simple head-start mechanism. In fact, these results mesh well with Balota et al.'s (2008) study of the joint effects of target stimulus quality and priming on lexical decision and speeded pronunciation performance. In that study, when target words were presented clearly, priming produced a simple shift in the *RT* distribution, but when words were visually degraded, priming effects became larger as *RT*s became slower. Degrading target words increased processing difficulty, which in turn increased reliance on prime information. According to Balota et al., this is consistent with the idea that when target processing is relatively degraded, the system increases reliance on (or retrieval of) the prime information to resolve the degraded target, consistent with recent arguments by Bodner and Masson (2001). Collectively, these findings can also be seen as compatible with the interactive compensatory framework (Stanovich, 1980), which proposes that priming is more automatic for fluent lexical processing and more strategic for less fluent lexical processing. Of course, we need to acknowledge that vocabulary knowledge, as measured by Shipley raw scores alone, is at best a relatively crude proxy for the integrity of underlying lexical representations. Future work examining lexical integrity should consider using more global measures of vocabulary knowledge (e.g., tests of synonyms, antonyms, and lexical analogies on the WJ III, Woodcock et al., 2001). Interestingly, using the WJ III measures to quantify individual differences in lexical integrity yielded the critical three-way interaction, as shown in the composite analyses. More notably, lexical integrity is multidimensional and reflects the quality of the orthographic, phonological, and semantic constituents of a representation, as well as the mapping between these constituents (Perfetti & Hart, 2002). A constellation of tasks that examine spelling performance, retrieval of

[6] One might contend that the lack of an interaction in the WUSTL participants is simply due to their being less sensitive to the word frequency and semantic priming manipulations, compared to the SUNY-A participants. In other words, is the frequency range used in the present study simply not sufficient for detecting an interaction in the WUSTL sample? We do not think so, for the following reasons. As pointed out, WUSTL participants showed robust main effects of priming and frequency in both E1 and E2. Moreover, if we compare E1 and E3 (across the two samples), the priming effects for high-frequency targets were practically identical in the two experiments (E1: 40 ms, E3: 40 ms), along with the main effects of target frequency (E1: 25 ms, E3: 28 ms).

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

pronunciations, and identification of meanings should therefore yield a more fine-grained measure of individual differences in lexical integrity.

It is important to note that although pure distributional shifting is compatible with a head-start mechanism of priming, distributional shifts can also be produced by changes in the decision criterion, i.e., the amount of evidence required before a decision is made. For example, in evidence accumulation models such as the random-walk model, altering the decision criterion affects the μ component (distributional shifting) but has no effect on σ (scaling) or τ (skewing) (Spieler, Balota, & Faust, 2000; Yap et al., 2006). Can the results in the present study be reconciled with a simple criterion-based mechanism of priming, whereby participants set a higher response threshold for targets preceded by unrelated primes? This account is problematic for two reasons. First, it is difficult to explain, in a principled manner, why priming shifts decision criteria to the same extent for high- and low-frequency words in readers with *more* vocabulary knowledge, but shifts them to different extents in readers with *less* vocabulary knowledge. Second, *even* if we assume that priming purely reflected changes in decision criteria, this clearly does not accommodate the pattern in E3, where priming was mediated by μ and σ for low-frequency words, or in E4, where priming reflected changes in both μ and τ, particularly for the low-frequency targets.

To summarize, these findings suggest that the nature of priming mechanisms may be modulated by the fluency of target processing. The priming data from E1 and E2 *always* reflected a shift, regardless of target frequency or nonword type. Distributional shifting is most easily reconciled with a relatively modular lexical processing system in readers with high-quality underlying lexical representations, where the effect of the prime is to afford the same head-start to all targets. In contrast, when lexical processing becomes more difficult, the system becomes increasingly sensitive to useful contextual information, and flexibly relies more on prime information (see Balota & Yap, 2006, for a discussion of flexible lexical processing).

Interactive effects of priming and frequency in accuracy

So far, our discussion has focused on the effects of priming and frequency on *RT*s. In *RT*s, one observes additive effects for readers with more vocabulary knowledge, and interactive effects for readers with less vocabulary knowledge. The trends are less clear when *accuracy* is the dependent variable. In E1 and E2, despite additivity in *RT*s, there was an overadditive priming by frequency interaction in accuracy. To further explore these results, we calculated the magnitude of the Priming $\times$ Frequency interaction in accuracy and *RT*s for each participant. We then correlated *RT* and accuracy interactions, and found that the correlations were not significant in both E1 ($r = -.046$) and E2 ($r = .041$), confirming that the additive effects are not simply an artifact of a speed–accuracy tradeoff. It is also worth noting that vocabulary knowledge did *not* predict the size of the interaction in error rates.

One might argue that the differences between the higher- and lower-vocabulary-knowledge readers can simply be attributed to a shift in response criteria. That is, higher-knowledge participants may be responding faster but making more errors in cases where lower-knowledge participants are responding slower but more accurately. We are skeptical that the present findings can be fully accommodated by this account. First, mean accuracy rates for higher- and lower-knowledge participants were very similar across the different experimental conditions, and in fact did not differ significantly ($F < 1$). If response criteria indeed varied as a function of vocabulary knowledge, then one would expect accuracy rates to be significantly *lower* for higher-knowledge participants for the most difficult trials (i.e., the low-frequency unrelated targets). Second, the account implies a speed–accuracy tradeoff for difficult trials, with lower-knowledge participants sacrificing speed for accuracy and higher-knowledge participants sacrificing accuracy for speed. However, again, there was no evidence of a speed–accuracy tradeoff in the difficult low-frequency unrelated condition. Specifically, the correlations (all non-significant) between *RT*s and accuracy were $-.198$, .039, $-.068$, and .135, respectively in the four experiments.

More importantly, the present analyses indicate that classifying the joint effects of priming and frequency as *either* additive or interactive is probably too inflexible. It might be more useful to conceptualize additivity and interactivity as poles of a continuum, with many intermediate positions in between. It is likely that the effects produced by the WUSTL and SUNY-A samples represent different points on the continuum, and participants can be "pushed" to show greater additivity or interactivity depending on a constellation of factors, including word frequency, vocabulary knowledge, perceptual degradation, and possibly prime–target associative strength. In this framework, one can consider the SUNY-A participants more "interactive" because they produce a significant interaction in *both RT*s and accuracy, whereas the WUSTL participants are more "additive" because they produce a significant interaction in accuracy and a non-significant trend towards greater priming for low-frequency words in *RT*s. In fact, vocabulary knowledge and word frequency are continuous variables, and we have obviously only selected two levels of both in the present study. In principle, if one had a larger range of word frequencies than in the present study, it is likely that we could have produced an interaction in *RT*s even for our high-knowledge readers. Again, this is consistent with the Balota et al. (2008) study in which stimulus degradation produced a reliable interaction with priming in the tail of the distribution, even for the WUSTL sample. Although this may sound remarkably similar to the Plaut and Booth (2000) sigmoid function, we will discuss in the next section how the specific results in the present study are not that easy to reconcile with that function. Ultimately, to simultaneously accommodate *RT* and accuracy data in a principled manner, one needs an explicit model, such as Ratcliff, Gomez, and McKoon's (2004) diffusion model of lexical decision performance. Such an approach may also provide insights into the differences between higher- and lower-knowledge readers.

Individual differences in word recognition performance

As discussed, the interaction between semantic priming and word frequency has been considered one of the benchmark effects in the word recognition literature. However, our results show that this "benchmark effect" is modulated by individual differences. Participants with (relatively) less vocabulary knowledge produce an interaction, while higher vocabulary knowledge participants produce additive effects. We tested this more rigorously across *all* participants by examining whether vocabulary knowledge, as measured by Shipley raw scores, predicted the magnitude of the Priming × Frequency interaction (i.e., larger priming effects for low-frequency targets), after controlling for chronological age, years of education, and nonword type. These regression analyses confirmed that vocabulary knowledge and the magnitude of the Priming × Frequency interaction were negatively correlated in *RT*s and τ (measure of distributional skewing), providing converging evidence that participants with less vocabulary knowledge were indeed more likely to produce larger priming effects for low-frequency targets. Furthermore, the influence of vocabulary knowledge on the interaction was primarily mediated by participants' slowest *RT*s, which reflect the most difficult trials for a participant.

Interestingly, this pattern is the exact opposite of what Plaut and Booth (2000) found. If we assume that vocabulary knowledge and perceptual ability map onto lexical input strength in the same way in the single-mechanism model (see Fig. 2), how might one account for the discrepancy? Perhaps the inconsistency between the two studies is more apparent than real. One way for the model to accommodate the data is to assume that low-knowledge readers are represented at the leftmost steep portion of the activation function, and high-knowledge readers are represented at the middle, gradual portion of the function. This will allow low-knowledge readers to show interactive effects and for high-knowledge readers to show additive effects. There are two problems with this "solution". If low-knowledge readers are positioned at the leftmost end of the curve, where the input–output function resembles a power function, then this should yield larger priming effects for high-frequency targets, because effects are larger for stronger inputs on this portion of the continuum. In our study, however, the low-knowledge readers produced larger priming effects for low-frequency targets. The second problem reflects the predictions made for readers with very high-vocabulary knowledge (i.e., higher than the current WUSTL sample), who should be represented at the rightmost portion of the function. The function predicts interactive effects of priming and frequency for such readers. This pattern seems most improbable given the present results.

Alternatively, it is possible that the low-knowledge participants in our study actually correspond to the high-perceptual-ability participants in Plaut and Booth's (2000) study. Hence, these participants produced an interaction because they are located within the portion of the curve where there is a Priming × Frequency interaction. In contrast, the high-knowledge readers are located further up in the asymptotic portion of the curve where *RT*s are faster but the interaction is smaller due to the ceiling effect. In fact, Plaut and Booth (2006) used this explanation to account for Borowsky and Besner's (1993) finding that visually degrading words strengthened rather than weakened the priming by frequency interaction. Of course, the foregoing discussion is based on a somewhat simplistic approach to accommodating empirical effects within the sigmoid function. Given the flexibility of the function, it is important to impose appropriate constraints when evaluating it, and this is more challenging than typically assumed (see Besner & Borowsky, 2006; Plaut & Booth, 2006, for more discussion). More specifically, the sigmoid function does *not* literally describe the operations of the single-mechanism model implemented by Plaut and Booth (2000). Rather, it is at best a metaphor for the actual behavior of the model (Plaut & Booth, 2006). In fact, Plaut and Booth (2006) demonstrated that their model could simulate empirical results which were inconsistent with the most straightforward interpretation of the sigmoid function. Hence, our criticisms of the sigmoid function may not apply to the actual implemented model.

The discrepancy between Plaut and Booth's (2000) study and ours may also be due to the fact that perceptual ability reflects amodal decoding speed while vocabulary knowledge reflects the integrity of underlying lexical representations. As we have suggested earlier, the extent to which prime information is retrospectively retrieved depends on how effortful it is to resolve a lexical target. This type of effort may be related to the integrity of lexical representations (tapped by vocabulary knowledge) but not to perceptual decoding speed (tapped by perceptual ability). In fact, Plaut and Booth's (2000) study provides some support for this dissociation. First, as mentioned in the Introduction, perceptual ability and vocabulary knowledge (as measured by the PPVT-R) were uncorrelated in their sample ($r = .09$), indicating that perceptual ability and vocabulary knowledge are measuring distinct constructs. Second, some aspects of their data may actually mirror our general findings. In Experiment 2, they examined the joint effects of priming and frequency in children, with age (3rd grade vs. 6th grade) as the between-participants variable. Presumably, age should be a good proxy for vocabulary knowledge. Interestingly, they reported that sixth graders produced overadditive effects of priming and frequency (with larger priming effects for low-frequency words), while third graders produced a more additive pattern (p. 797). On initial consideration, this seems quite consistent with the three-way interaction they obtained when perceptual ability was the between-participants variable. However, results presented in other portions of Plaut and Booth's paper suggest that contrary to what they reported, their third graders (i.e., the readers with less vocabulary knowledge) were actually showing *more* interactive effects of priming and frequency. Specifically, in both Fig. 5 (p. 797) and Appendix (p. 823), third graders appear to be showing larger priming effects for low-frequency ($d = 75$ ms) than high-frequency ($d = 47$ ms) words, whereas sixth graders showed more similar priming effects for low-frequency ($d = 37$ ms) and high-frequency ($d = 46$ ms) words. Plaut and Booth also conducted another analysis which compared adults to children. Here, they found that both their adults and children produced similarly sized overadditive effects of priming and frequency. If children have less vocabulary knowledge than adults, our account predicts stronger interactive effects for the children, relative to the adults; however, this was not the observed pattern. Plaut and Booth (p. 798) indicated that "this comparison is difficult to interpret" because the children and adults did not receive the same set of words (difficult words were eliminated for the children), and children and adults were also different on perceptual ability and other factors. It is plausible though that the magnitude of the interaction was similar for children and adults because the difficulty of the items was calibrated for their respective vocabulary knowledge. Our study, in contrast, used the *same* words for both the WUSTL and SUNY-A readers, and these words must have been relatively more difficult for the SUNY-A than for the WUSTL sample.

However, we need to acknowledge that our account does not offer an obvious explanation for Plaut and Booth's (2000) findings, i.e., additive effects for low-perceptual-ability readers and interactive effects for high-perceptual-ability readers. We have argued that fluent lexical processors are more likely to yield additive effects, but there is no principled reason why low-perceptual-ability readers should be more fluent lexical processors, especially since they were actually substantially slower than the high-perceptual-ability readers on the lexical decision task (see Plaut & Booth, Fig. 3). Rather than contriving a *post hoc* explanation for Plaut and Booth's results, we suggest that the distinct patterns of results associated with perceptual ability and vocabulary knowledge provide interesting questions for future research. At the very least, the dissociations between the two individual differences measures indicate that it may be misleading to map them onto a single dimension (e.g., input strength on the single-mechanism model), and attempting to accommodate both measures under a unified theoretical framework may not be the best approach.

To recapitulate, the present findings, along with the results from Balota et al. (2008) indicate that as processing fluency decreases, due to low-integrity representations, visual degradation, or increased task demands, priming effects become larger, particularly for the most difficult targets in the tail of the distribution. These results further underscore the importance of considering individual differences in visual word recognition. An effect that is identified in a particular sample may not generalize to other sites, making it important to replicate novel effects across samples that may vary with respect to processing fluency (cf. Yap, Balota, Tse, & Besner, 2008). In addition, it is clear that the lexical processing system is remarkably flexible and adaptive, and can show greater reliance on the semantic context as target processing becomes more difficult. This could be viewed as consistent with Hutchison's (2007) finding that participants, particularly those who are high in attentional control, are sensitive to relatedness proportion in a priming experiment, producing larger priming effects as relatedness proportion becomes higher. However, further work is needed to better understand the extent to which the present effects are under strategic control.

Why are priming effects smaller when pseudohomophones are used?

The nonword type manipulation (legal nonwords vs. pseudohomophones) produced an intriguing counterintuitive finding with respect to priming effects. The presence of pseudohomophones attenuated the magnitude of priming effects. More specifically, the priming effect was smaller in μ for both the WUSTL and SUNY-A samples when a pseudohomophone context was used, compared to a legal nonword context. It is indeed intriguing for an effect to become *smaller* in the lexical decision task as discrimination becomes *more* difficult. In contrast, the effect of nonword context on the word frequency effect was precisely as predicted, i.e., we replicated the well-established finding of a larger frequency effects in the pseudohomophone context compared to the legal nonword context (see Stone & Van Orden, 1993; Yap et al., 2006). Hence, the present results yielded the noteworthy pattern that the presence of pseudohomophones simultaneously increased word frequency effects and decreased semantic priming effects.

Why would priming effects become smaller in a pseudohomophone context? To address this question, it is first necessary to reiterate that the lexical decision task is primarily a binary discrimination task whose difficulty is a function of the overlap between words and nonwords. If we conceptualize the word recognition system as a collection of processing modules and pathways that support the computations mediating orthography, phonology, and meaning (Balota, Paul, & Spieler, 1999), the type of nonwords used in a lexical decision task may engage attentional control systems that appropriately adjust the weights between different modules. Specifically, in the standard lexical decision task where legal nonwords are used, participants are attempting to discriminate between familiar/meaningful words and relatively unfamiliar/meaningless nonwords, and therefore emphasize the connections between orthography and meaning. However, pseudohomophones increase the familiarity/meaningfulness overlap between words and nonwords, and hence this dimension becomes less informative for word–nonword discrimination. In fact, using familiarity/meaningfulness tends to increase the false alarm rate since pseudohomophones (e.g., BRANE), by design, are constructed to activate meaning-based information, via an orthographic–phonological pathway. In such a situation, the system may deemphasize the pathway between orthography and meaning, and rely less on meaning, thereby reducing the semantic priming effect.

This prediction is consistent with the connectionist triangle model perspective (Plaut, 1997; Seidenberg & McClelland, 1989), where task demands modulate the extent to which participants attend to different types of lexical information (i.e., orthographic, phonological, and semantic) in lexical decision. For example, when nonwords are illegal (e.g., BRNTA), orthographic familiarity is sufficient for driving word–nonword discrimination. With legal nonwords (e.g., BRONE), phonological, rather than orthographic, familiarity is recruited. However, when distracters are pseudohomophones (e.g., BRANE), which look and sound like real words, only semantic familiarity is viable for decision-making. Hence, this account suggests that the decreased priming could be due to the attenuation of the phonology → semantics pathway, which would help suppress false alarms in the context of pseudohomophones. The foregoing discussion is necessarily speculative, but it does suggest that our counterintuitive finding can be accommodated either within a flexible lexical processing framework (Balota & Yap, 2006) or some version of the triangle model where pathway control is implemented. Both accounts of this intriguing finding merit further study.

Stages vs. single mechanism

One of the original objectives of the present study was to establish whether the interactive effects of priming and frequency were more consistent with multiple independent stages (Sternberg, 1969) or with a single non-linear mechanism (Plaut & Booth, 2000). As it turns out, the answer to this question was less clear-cut than anticipated. Most critically, the interaction between priming and frequency was stronger for our low-vocabulary knowledge readers than our high-vocabulary knowledge readers. As discussed, it is not easy to reconcile these findings with Plaut and Booth's (2000) sigmoid activation function (see Fig. 2), although as pointed out in a previous section, criticisms directed against the sigmoid function may not necessarily apply to the actual implemented model (see Plaut & Booth, 2006). In principle, the data *can* be accommodated by the multistage perspective, using additive factors logic to revise extant assumptions. It must be emphasized, however, that these modifications are *post hoc* and need to be empirically verified in future studies.

To recapitulate, in E1 and E2, additive effects of priming and frequency were observed, which is consistent with priming and frequency influencing *independent* stages. In E3 and E4, interactive effects were observed, which is consistent with priming and frequency influencing a *common* stage. This suggests that for high-lexical-integrity participants, priming *only* influences an earlier perceptual stage by providing a head-start for subsequently presented targets, while frequency influences a later lexical retrieval stage. For low-lexical-integrity participants, priming exerts effects on both the early stage as well as the later lexical retrieval stage. How does priming influence the subsequent lexical retrieval stage? As we have argued in previous sections, as target processing increases in difficulty, the reliance on the prime information increases, especially for the low-frequency targets.

The precise mechanisms which mediate the effects of target difficulty on semantic priming influence remain unclear. Possibly, when targets are difficult to process, it is more likely that there is episodic retrieval of the prime (Bodner & Masson, 1997), hence increasing its influence. Alternatively, if we adopt the perspective of the multistage activation model (Borowsky & Besner, 1993), related primes lower the response criterion (see Fig. 1). In addition, the more difficult the target is, the more the criterion is lowered. Since the rate of evidence accumulation is steeper for high-frequency targets than for low-frequency targets, a constant change in criterion for both classes of words should yield larger priming effects for low-frequency targets. Let us further assume that for all high-frequency targets, the lowering in criterion due to the related prime is relatively invariant, since none of the high-frequency targets are particularly difficult. On the other hand, low-frequency targets are more variable with respect to difficulty, and one expects the criterion to be lowered more for more difficult items. This will explain why priming effects become larger across the *RT* distribution for low-frequency, but not high-frequency, targets.

There is an alternative stage-based account that does not require an appeal to episodic retrieval of the prime (Bodner & Masson, 1997). The multistage activation model (Borowsky & Besner, 1993) contains feedforward and feedback pathways between the orthographic input lexicon and the semantic system. Importantly, there is evidence that feedback from the semantic system to the orthographic input lexicon is neither mandatory nor automatic (Stolz & Neely, 1995). Instead, the feedback mechanism operates *only* when it is beneficial. For example, Stolz and Neely reported additive effects of priming and stimulus quality when relatedness proportion was low ($RP = .25$) but an overadditive interaction when relatedness proportion was high ($RP = .50$); interactive effects indicate semantic feedback while additive effects indicate no feedback. These results are consistent with the idea that when relatedness proportion is low, feedback from the semantic system is eliminated, because this feedback is not useful on the majority of trials. Similarly, one could argue that there is less semantic feedback for readers with high-integrity lexical representations, because such readers process lexical targets fluently and hence there is relatively less benefit from related primes. Again, the results attest to the flexibility of the lexical processing system in accomplishing task goals.

Conclusions

The present study examined the joint effects of semantic priming and word frequency in lexical decision performance. The intriguing finding was that these two factors do not always interact. In fact, whether priming and frequency interact depends on the vocabulary knowledge of the participant. Readers with less vocabulary knowledge show larger priming effects, particularly for difficult low-frequency targets that fall into the tail of the *RT* distribution, and this is consistent with the idea of a flexible lexical processing system that optimizes task performance by emphasizing task-relevant information. In contrast, the lexical processing system of readers with greater vocabulary knowledge appears to be more modular in nature, whereby the effect of a prime is primarily to provide a head-start that is independent of a target's difficulty, i.e., shifts the *RT* distribution. From a methodological point of view, this study also underscores the need to extend visual word recognition research by considering both individual differences and by analyzing how variables influences the underlying *RT* distribution.

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

Acknowledgments

This research was supported by National Institute of Aging Grant AG03991 and National Science Foundation Grant BCS0001801 to D.A. Balota. We thank Derek Besner, Andrew Heathcote, Debra Jared, Jim Neely, and Dave Plaut for their constructive comments on an earlier version of this paper, Viviana Benitez for her assistance with stimuli development and data collection, Jim Neely and Matt Thomas for their help with data collection in SUNY-Albany, and Keith Hutchison for providing us with access to data from the primed lexical decision megastudy project.

Appendix

Primes	LF targets	Primes	LF targets	Primes	LF targets	Primes	LF targets
STOMACH	ACHE	SNAP	CRACKLE	BRIDE	GROOM	SPOILED	ROTTEN
RAGE	ANGER	SOB	CRY	CAP	HAT	DICTATOR	RULER
SPRAIN	ANKLE	SNUGGLE	CUDDLE	HITCH	HIKE	UNHAPPY	SAD
KNIGHT	ARMOR	MOM	DAD	HULA	HOOP	FRIGHT	SCARE
VEIN	ARTERY	CLOAK	DAGGER	SADDLE	HORSE	YELL	SCREAM
AWAKE	ASLEEP	DUSK	DAWN	SCRATCH	ITCH	QUIVER	SHAKE
TYLENOL	ASPIRIN	DOE	DEER	DENIM	JEANS	JAWS	SHARK
HELIUM	BALLOON	RELY	DEPEND	PUN	JOKE	SLEEVE	SHIRT
GAUZE	BANDAGE	SOIL	DIRT	MUSTARD	KETCHUP	SOCKS	SHOES
KIDNEY	BEAN	TRENCH	DITCH	RUNG	LADDER	ILL	SICK
HIVE	BEE	SCUBA	DIVE	MOLTEN	LAVA	SATIN	SILK
BUCKLE	BELT	FLIPPER	DOLPHIN	MOWER	LAWN	SMUDGE	SMEAR
PEDAL	BIKE	DRAWER	DRESSER	LIME	LEMON	ODOR	SMELL
MOSQUITO	BITE	INTOXICATED	DRUNK	FIB	LIE	ESCARGOT	SNAIL
CLOROX	BLEACH	WASHER	DRYER	ROAR	LION	COBRA	SNAKE
BRUNETTE	BLONDE	STUPID	DUMB	NOISY	LOUD	LATHER	SOAP
SKIRT	BLOUSE	YOLK	EGGS	BUTLER	MAID	OREGANO	SPICE
ROW	BOAT	VOLCANO	ERUPT	SHOPPING	MALL	STIFF	STARCH
ATOM	BOMB	FLUNK	FAIL	DWARF	MIDGET	HUNGRY	STARVE
SONIC	BOOM	BANNER	FLAG	PLUS	MINUS	ROB	STEAL
LEND	BORROW	FROSTED	FLAKES	CHIMPANZEE	MONKEY	SKUNK	STINK
COMB	BRUSH	TICK	FLEA	ELK	MOOSE	MARSH	SWAMP
PAIL	BUCKET	RAFT	FLOAT	HAMMER	NAIL	BROOM	SWEEP
MARGARINE	BUTTER	TULIP	FLOWER	NIECE	NEPHEW	FLAVOR	TASTE
ICING	CAKE	CLARINET	FLUTE	DIME	NICKEL	GUMS	TEETH
WICK	CANDLE	SPOON	FORK	CASHEW	NUT	RACKET	TENNIS
CAUTIOUS	CAREFUL	DELICATE	FRAGILE	SALIVA	SPIT	THICK	THIN
CELERY	CARROT	TOAD	FROG	TROUSERS	PANTS	LOOSE	TIGHT
GUM	CHEW	SMILE	FROWN	COBBLER	PEACH	CAVITY	TOOTH
OYSTER	CLAM	BET	GAMBLE	INK	PEN	AUNT	UNCLE
CIRCUS	CLOWN	GOOSE	GANDER	SALT	PEPPER	NOUN	VERB
MINER	COAL	TRASH	GARBAGE	DILL	PICKLE	BASKET	WEAVE
JACKET	COAT	GHOUL	GHOST	VENOM	POISON	SLIPPERY	WET
MUSK	COLOGNE	TONIC	GIN	OUNCE	POUND	CORK	WINE
REEF	CORAL	LENS	GLASSES	TAPIOCA	PUDDING	LOSER	WINNER
HUSK	CORN	MITTEN	GLOVE	HANDBAG	PURSE	ANNUAL	YEARLY
SOFA	COUCH	PASTE	GLUE	EGYPT	PYRAMID		
SALTINE	CRACKER	VINE	GRAPE	SHINGLE	ROOF		

Primes	HF targets	Primes	HF targets	Primes	HF targets	Primes	HF targets
WILLING	ABLE	SCISSOR	CUT	THRONE	KING	BASIC	SIMPLE
MISTREAT	ABUSE	ALIVE	DEAD	ACRE	LAND	FIVE	SIX
CHECKING	ACCOUNT	VERDICT	DECISION	MEDIUM	LARGE	TINY	SMALL
BEFORE	AFTER	LUNCH	DINNER	TARDY	LATE	ROUGH	SMOOTH
SOLO	ALONE	PUPPY	DOG	TEACH	LEARN	APOLOGY	SORRY
ZOO	ANIMAL	KNOB	DOOR	EVACUATE	LEAVE	TONE	SOUND
QUESTION	ANSWER	UP	DOWN	MORE	LESS	NORTH	SOUTH
LEGS	ARMS	PLANET	EARTH	BREATH	LIFE	ASTRONAUT	SPACE
FRONT	BACK	FINAL	END	BULB	LIGHT	PHASE	STAGE
BOUNCE	BALL	GRAMMAR	ENGLISH	LENGTH	LONG	ASTRONOMY	STAR
ORIGINATE	BEGIN	ODD	EVEN	FOUND	LOST	SERVICE	STATION
HUMAN	BEING	SWIFT	FAST	AFFECTION	LOVE	REMAIN	STAY
ABOVE	BELOW	MOTHER	FATHER	ROBOT	MACHINE	HALT	STOP
WHITE	BLACK	TOUCH	FEEL	MINOR	MAJOR	TALE	STORY
PLASMA	BLOOD	CAMERA	FILM	CASH	MONEY	AVENUE	STREET

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

Appendix (*continued*)

Primes	HF targets	Primes	HF targets	Primes	HF targets	Primes	HF targets
SKY	BLUE	SEEK	FIND	CINEMA	MOVIE	POWERFUL	STRONG
CHALK	BOARD	BLAZE	FIRE	INITIAL	NAME	PUPIL	STUDENT
ANATOMY	BODY	LAST	FIRST	FAR	NEAR	RESEARCH	STUDY
LIBRARY	BOOK	TILE	FLOOR	DAY	NIGHT	DISCUSS	TALK
SELL	BUY	GROCERY	FOOD	DIGIT	NUMBER	QUIZ	TEST
OPENER	CAN	REMEMBER	FORGET	ON	OFF	OBJECT	THING
AUTO	CAR	PAL	FRIEND	NEW	OLD	CLOCK	TIME
CREDIT	CARDS	EMPTY	FULL	CLOSED	OPEN	TOMORROW	TODAY
HAUL	CARRY	FUMES	GAS	INSIDE	OUTSIDE	SUM	TOTAL
EFFECT	CAUSE	GUY	GIRL	CELEBRATE	PARTY	ATTEMPT	TRY
CORE	CENTER	PURPOSE	GOAL	FRAME	PICTURE	OVER	UNDER
RISK	CHANCE	SILVER	GOLD	CRUEL	MEAN	DESIRE	WANT
ALTER	CHANGE	BAD	GOOD	RICH	POOR	OBSERVE	WATCH
ADULT	CHILD	PALM	HAND	GIFT	PRESENT	FLOOD	WATER
OPTION	CHOICE	JOYOUS	HAPPY	UGLY	PRETTY	EAST	WEST
STEEPLE	CHURCH	DIFFICULT	HARD	DILEMMA	PROBLEM	PANE	WINDOW
TOWN	CITY	LISTEN	HEAR	LEFT	RIGHT	DICTIONARY	WORDS
SESSION	CLASS	ASSIST	HELP	JOG	RUN	LABOR	WORK
CLARIFY	CLEAR	LOW	HIGH	BARGAIN	SALE	GLOBE	WORLD
UNIVERSITY	COLLEGE	GRASP	HOLD	PHYSICS	SCIENCE	INCORRECT	WRONG
HUE	COLOR	ADDRESS	HOME	TALL	SHORT	NO	YES
REMARK	COMMENT	PRAIRIE	HOUSE	VISION	SIGHT		
ACCURATE	CORRECT	SLAY	KILL	ALIKE	SIMILAR		
Legal nonwords							
AFTEN	CLANGE	FIRT	HAREFUL	NEPHEE	SCARN	STON	WARROT
AMINAL	CLOM	FLAH	HEETH	NEZ	SCOENCE	STREEB	WEP
ARMOT	CLOOR	FLANCE	HEMON	NOTAL	SCROAM	STRONK	WINT
ARTOPY	CLORN	FLEAR	HERN	OCCAUNT	SEAVE	STUVENP	WOARN
ARUSE	COAB	FLET	HODY	OFE	SELK	SULL	WOME
ASPORIP	COLLICT	FLOAB	HOLK	OLK	SESS	SURSE	WONNER
AYONE	CORRAGE	FLOST	HORAN	ONDLE	SEST	TAWN	WORP
BAME	COVIE	FLOWIP	HORP	ONKLE	SHERT	TEAVE	WORTY
BANDLE	CRICKLE	FODDING	HOUCH	ONSWEN	SHIRP	TER	WRONK
BAPPY	CRO	FOOK	INGER	OREN	SHOIT	THIND	WUMBER
BEEL	CROCKEN	FOON	INT	ORMS	SHORK	THOV	WUST
BEPPER	CROG	FOOTH	JIANS	OUPSITE	SIFE	THURCH	YAN
BEROW	CROUSE	FOWN	JIKE	PACHINE	SIRL	TIGH	YANDER
BEZ	DAU	FRAKE	KIDGET	PAGGER	SIRST	TILK	YARRY
BIE	DEEK	FRAPILE	KIGHT	PATHER	SIUTH	TIMPLE	YATCH
BIMB	DELT	FREEND	KOUND	PIGHT	SKALL	TIND	YATE
BINE	DEVEND	FRESS	KULP	PLETTY	SKOKE	TOISON	YOAL
BLARK	DICKEL	FRINK	LADDEG	POLPHIP	SLAGE	TONNIP	YONEY
BLEARN	DIGHT	FROPE	LAS	PONTS	SLARVE	TOSTE	YONKEY
BLOON	DIMB	FROTE	LETCHUM	POOSE	SLATION	TRE	YOOP
BOAM	DIX	FRUE	LIOP	PORDS	SLONDE	TWAMP	YORN
BOCH	DOCISION	FRUSH	LOMMENT	PRAIL	SLOOTH	UNPER	YORSE
BOCKET	DORT	FUST	LOND	PRAKES	SLUDY	USLEEM	ZALE
BOING	DOTCH	FUTTER	LOOF	PRAPE	SLUS	UTCH	ZEAD
BOLLEEN	DOU	GANDAGE	LOOR	PRELL	SMAR	VATER	ZEAR
BORLD	DRYEM	GICKLE	LOUK	PRUNK	SMEAT	VAUSE	ZEARLY
BORRY	EARCH	GIKE	LOVA	PRYSINT	SMEEP	VICTUNE	ZIME
BOST	EBLE	GIP	MAWN	PYRIMOD	SOARD	VIGHE	ZINDOW
BRESSER	ELUPT	GLOSSEN	MILUS	RAJOR	SOAT	VIGHT	ZINNER
BROBLEN	ENGLITH	GLUB	MIVE	RAR	SORK	VILM	ZOAL
BROVE	EPPS	GOOM	MOID	REGIN	SOVE	VING	ZOAT
CABE	ESEN	GORROW	MOLONGE	RIMILAR	SPEAB	VITY	ZOOR
CARNS	FAIZ	GROOP	MOOM	ROLOR	SPIZ	VODAY	ZOUND
CES	FARGE	GUDDLE	NAID	ROTTEY	SPOCE	VOLD	ZOUSE
CHEY	FEEP	HACE	NALL	RYLER	STACE	VONG	
CHILP	FICK	HAMP	NEACH	SAH	STARRO	VORB	
CHOINE	FIKE	HAO	NEAK	SALL	STAX	VOY	
CINTER	FIRGOT	HARBAGE	NEN	SAMBLE	STIRY	VUN	
Pseudohomophones							
ABAUT	BRAIT	DANSE	GOTE	KOME	OXYDE	SCAIL	TAIM
AFORDE	BRANE	DEELER	GRAE	KOMPLIT	PAGUN	SEAP	TALANT
AJENDA	BREEF	DEFEET	GRANE	KONDUCT	PAIJE	SED	TAYP
AKLAME	BROAK	DEFF	GREAN	KONSEET	PANIK	SEET	TEAZE
AKT	BROOT	DELT	GRONE	KOPY	PANZY	SEEZE	TENCE

(*continued on next page*)

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001

Appendix (*continued*)

Primes	HF targets	Primes	HF targets	Primes	HF targets	Primes	HF targets
ALINE	BURDS	DET	GRUPE	KRONIC	PEAP	SENSER	TERKEY
ALJEE	BURTH	DETALE	HAF	KURSE	PEECE	SERTON	THEEF
ALMUND	BYCEP	DILE	HAYLE	KWICK	PEEPLE	SEZ	THRED
AMBUR	CAIGE	DOAM	HED	LAMM	PERPLE	SHAIR	THRET
AMPER	CAIM	DOLLER	HEERO	LEECE	PHAN	SHEAP	THURD
AMUNG	CANION	DRANE	HEET	LEEDER	PHIRM	SHEAT	TIKEL
ARKAID	CARBUN	DRES	HEEVE	LEEP	PHUNNEL	SHURE	TIPE
ARROE	CAREAR	DUBBLE	HERTS	LEESH	PIRCH	SHUV	TOLE
ATTIK	CEL	DURBY	HOKES	LEEST	PLEED	SIRCH	TOOB
AVALE	CHEAK	EAZY	HOZE	LERCH	PLEEZ	SKALP	TOPIK
AWATE	CHEEF	ELBOE	HUNNY	LOGIK	POIZE	SKARF	TRALE
BABI	CHELLO	ERRUR	IDEEL	LOTTARY	POZE	SKORE	TREET
BAID	CHOAK	FAIM	IKSEPT	LYKE	PROSEED	SLEAK	TROAL
BATTEL	CLAME	FALCE	JAIDE	LYNE	PROZE	SLEAP	TROFF
BAYGE	CLEEN	FALT	JARGEN	MAGIK	PRUFE	SLEAT	TRUPE
BEAF	CLIRK	FAMUS	JEAP	MAYDEN	PRUVE	SLITE	TYDE
BEAP	CLOO	FEER	JENDER	MAYER	PURCH	SNEEK	TYGER
BEED	COFFEY	FEEST	JENTLE	MELUN	RADE	SNEEZ	UNDEW
BEEK	COMFERT	FETHER	JERM	MENT	RANDUM	SNEWS	VACUME
BEEM	COMIDY	FIRN	JIRK	MENY	RATHFOO	SOAL	VALT
BEERD	COMMEN	FLAIM	JOOCE	METUL	REELY	SOKE	VILLEN
BEEST	CONSERN	FLEACE	JOYNT	MITH	REEP	SOOP	VURSE
BELIF	CONSERT	FLURT	KAF	MUTCH	RESORSE	SPANE	WAIGE
BERCH	CORSE	FOME	KANAL	NAWTY	RINKLE	SPEEK	WEET
BERN	COTTEN	FORSE	KANDY	NEAD	RITHUM	SPEER	WEIT
BEWEAR	CRAIN	FOURTY	KANON	NEET	ROAB	SPERT	WHEAL
BIRST	CRAIT	FRALE	KANOO	NEEZE	ROAP	SPRED	WIRTH
BLAIM	CROOD	FREKLE	KEAP	NERSE	ROBBIN	STAIL	WRENT
BLEEK	CRUES	FRUM	KEMIST	NOIZE	ROZE	STOAR	YEELD
BLOTE	CULPRET	GAIM	KLAMP	NOOZE	RUBER	STROLE	
BOR	CUNSENT	GALLEN	KNEAL	NOZE	SAIN	STUK	
BOURNE	DAIT	GARDUN	KNET	NURVE	SALERY	SURVE	
BOYL	DAIZE	GLOE	KNYFE	OTE	SAYLE	SWERL	

References

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 514–544.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128*, 32–55.

Balota, D. A., & Yap, M. J. (2006). Attentional control and flexible lexical processing: Explorations of the magic moment of word recognition. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 229–258). Hove, UK: Psychology Press.

Balota, D. A., Paul, S., & Spieler, D. H. (1999). Attentional control of lexical processing pathways during word recognition and reading. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 15–57). East Sussex, UK: Psychology Press.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory & Language, 59*, 495–523.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445–459.

Becker, C. A. (1979). Semantic context and word-frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 5*, 252–259.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 3*, 389–401.

Besner, D., & Borowsky, R. (2006). Postscript: Plaut and Booth's (2006) new simulations: What have we learned. *Psychological Review, 113*, 194–195.

Besner, D., Wartak, S., & Robidoux, S. (2008). Constraints on computational models of basic processes in reading. *Journal of Experimental Psychology: Human Perception & Performance, 34*, 242–250.

Bodner, G. E., & Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory & Language, 37*, 268–293.

Bodner, G. E., & Masson, M. E. J. (2001). Prime validity affects masked repetition priming: Evidence for an episodic resource account of priming. *Journal of Memory & Language, 45*, 616–647.

Borowsky, R., & Besner, D. (2006). Parallel distributed processing and lexical-semantic effects in visual word recognition: Are a few stages necessary. *Psychological Review, 113*, 181–194.

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 813–840.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior, 12*, 335–359.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*, 249–253.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance, VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Cousineau, D., Brown, S. D., & Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers, 36*, 742–756.

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13*, 349–359.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test-revised*. Bloomington, MN: Pearson Assessments.

Farmer, M., & Klein, R. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review, 2*, 460–493.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information processing rate and amount: Implications

for group differences in response latency. *Psychological Bulletin, 125*, 777–799.
Heathcote, A., & Brown, S. D. (2004). Reply to Speckman and Rouder: A theoretical basis for QML. *Psychonomic Bulletin & Review, 11*, 577–578.
Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9*, 394–401.
Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109*, 340–347.
Humphreys, L. G. (1978). Research on individual differences requires correlational analysis, not ANOVA. *Intelligence, 2*, 1–5.
Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition,, 33*, 645–662.
Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *The Quarterly Journal of Experimental Psychology, 61*, 1036–1066.
Kinoshita, S. (2006). Additive and interactive effects of word frequency and masked repetition in the lexical decision task. *Psychonomic Bulletin & Review, 13*, 668–678.
Kinoshita, S., & Mozer, M. C. (2006). How lexical decision is affected by recent experience: Symmetric versus asymmetric frequency-blocking effects. *Memory & Cognition, 34*, 726–742.
Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin, 113*, 181–190.
McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.
Milota, V. C., Widau, A. A., McMickell, M. R., Juola, J. F., & Simpson, G. B. (1997). Strategic reliance on phonological mediation in lexical access. *Memory & Cognition, 25*, 333–344.
Mulatti, C., Reynolds, M., & Besner, D. (2006). Neighborhood effects in visual word recognition: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception & Performance,, 32*, 799–810.
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264–336). Hillsdale, NJ: Erlbaum.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida word association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*, 402–407.
Paap, K. R., Johansen, L., Chun, E., & Vonnahme, P. (2000). Neighborhood frequency does affect performance in the Reicher task: Encoding or decision? *Journal of Experimental Psychology: Human Perception & Performance, 26*, 1691–1720.
Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357–383.
Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins.
Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review, 107*, 786–823.
Plaut, D. C., & Booth, J. R. (2006). More modeling but still no stages: Reply to Borowsky and Besner. *Psychological Review, 113*, 196–200.
Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language & Cognitive Processes, 12*, 765–805.
Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology, 51*, 181–194.
Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461.
Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111*, 159–182.
Roelofs, A. (2008). Dynamics of the attentional control of word retrieval: Analyses of response time distributions. *Journal of Experimental Psychology: General, 137*, 303–323.
Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the Vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin & Review, 11*, 419–427.
Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*, 1270–1281.
Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.
Seidenberg, M. S. (1985). The time course of information activation and utilization in visual word recognition. In D. Besner, T. Waller, & G. MacKinnon (Eds.). *Reading research: Advances in theory and practice* (Vol. 5). New York: Academic Press.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.
Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology, 9*, 371–377.
Shulman, H. G., & Davison, T. C. (1977). Control properties of semantic coding in a lexical decision task. *Journal of Verbal Learning and Verbal Behavior, 16*, 91–98.
Speckman, P. L., & Rouder, J. N. (2004). A comment on Heathcote, Brown, and Mewhort's QMLE method for response time distributions. *Psychonomic Bulletin & Review, 11*, 574–576.
Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception & Performance, 26*, 506–526.
Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32–71.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II: Proceedings from a symposium on attention and performance*. *Acta Psychologica, 30*, 276–315 [special issue].
Stolz, J. A., & Neely, J. H. (1995). When target degradation does and does not enhance semantic context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 596–611.
Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 19*, 744–774.
Tainturier, M. J., Tremblay, M., & Lecours, A. R. (1992). Education level and the word frequency effect: A lexical decision investigation. *Brain & Language, 43*, 460–474.
Tse, C.-S., & Neely, J. H. (2007). Semantic priming from letter-searched primes occurs for low- but not high-frequency targets: Automatic semantic access may not be a myth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1143–1161.
Tulsky, D. S., Saklofske, D. H., & Zhu, J. (2003). Revising a standard: An evaluation of the origin and development of the WAIS-III. In R. K. Heaton, R. J. Ivnik, D. S. Tulsky, D. H. Saklofske, & G. J. Chelune, et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 43–92). San Diego, CA: Academic Press.
Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7*, 424–465.
Verhoeven, L., & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22*, 407–423.
Vernon, P. A. (1987). *Speed of information processing and intelligence*. Norwood, NJ: Ablex.
Vincent, S. B. (1912). The function of vibrissae in the behavior 2005 of the white rat. *Behavioral Monographs, 1* (Whole No. 5).
Wechsler, D. (1991). *Wechsler intelligence scale for children* (3rd ed.). San Antonio, TX: Harcourt, Brace.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca, IL: Riverside.
Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 274–295.
Yap, M. J., Balota, D. A., Cortese, M. J., & Watson, J. M. (2006). Single versus dual process models of lexical decision performance: Insights from RT distributional analysis. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1324–1344.
Yap, M. J., Balota, D. A., Tse, C.-S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: Evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 495–513.
Zachary, R. A. (1992). *Shipley institute of living scale-revised, manual*. Los Angeles, CA: Western Psychological Services.
Zachary, R. A., Paulson, M. J., & Gorsuch, R. L. (1985). Estimating WAIS IQ from the Shipley Institute of Living Scale using continuously adjusted age norms. *Journal of Clinical Psychology, 41*, 820–831.

Please cite this article in press as: Yap, M. J., et al. Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language* (2009), doi:10.1016/j.jml.2009.07.001