Are there Task-specific Effects in Morphological Processing?

Examining Semantic Transparency Effects in Semantic Categorization and Lexical Decision

Qian Wen Chee and Melvin J. Yap

National University of Singapore

Author Note

This research was supported by the National University of Singapore Heads and Deanery Research Support Scheme Grants R-581-000-236-101 and R-581-000-250-101 awarded to M.J.Y. We thank Aaron Lim, An Qi Lim, Hui Ning Tnay, and Nicole Fong for assistance in programming and data acquisition. Correspondence concerning this article should be addressed to Qian Wen Chee, Department of Psychology, National University of Singapore, Singapore 117570. Email: qianwen.chee@u.nus.edu. We thank Simona Amenta and an anonymous reviewer for constructive comments on an earlier version of this paper.

Abstract

Current theories of morphological processing include form-then-meaning accounts, form-with-meaning accounts, and connectionist theories. Form-then meaning accounts argue that the morphological decomposition of complex words is based purely on orthographic structure, while form-with meaning accounts argue that decomposition is influenced by the semantic properties of the stem. Connectionist theories, on the other hand, argue that morphemes are encoded as statistical patterns of occurrences between form and meaning. The weight of evidence from the literature thus far suggests that morphological decomposition is best explained by form-then-meaning accounts. That said, conflicting empirical findings exist, and more importantly, semantic transparency effects in morphological processing have been examined almost exclusively with the lexical decision task, in which participants discriminate between words and nonwords. Consequently, the extent to which observed results reflect the specific demands of the lexical decision task remains unclear. The present study extends previous work by testing if the processing dynamics of early morphological processing are moderated by task requirements. Using the masked morphological priming paradigm, this hypothesis was tested by examining semantic transparency effects for a common set of words across semantic categorization and lexical decision. In both tasks, priming was stronger for transparent (e.g., painter-PAINT) than opaque (e.g., corner-CORN) prime-target pairs; these results speak against form-then-meaning accounts. These findings further inform theories of morphological processing and underscore the importance of examining the interplay between task-general and task-specific mechanisms.

Keywords: morphological processing, masked priming, semantic categorization, lexical decision, word recognition, task-specific effects

Are There Task-specific Effects in Morphological Processing?

Examining Semantic Transparency Effects in Semantic Categorization and Lexical Decision

The majority of English words are complex words which comprise more than one morpheme (e.g., “swimmer”, “happiness”). Complex words are typically formed by the concatenation of a base morpheme with at least one affix; the base morpheme, also known as the stem, gives the word its principal meaning (e.g., “swim”, “happy”), while affixes modify the meaning of the base morpheme (e.g., “-er”, “-ness”). The processes and mechanisms underlying how complex words are recognized and processed have been extensively studied by researchers (see Amenta & Crepaldi, 2012, for a review).

Theories of Morphological Processing

Major theoretical accounts of morphological processing can be broadly classified into form-then-meaning accounts, form-with-meaning accounts, and connectionist theories. According to form-then-meaning accounts (e.g., Rastle, Davis, & New, 2004; Taft & Forster, 1975), complex words are initially processed through the pre-lexical parsing of its letter string into morphemes. Variants of form-then-meaning accounts argue that complex words are automatically decomposed into their morphological constituents based purely on orthographic structure, such that even words with a mere appearance of morphological complexity (e.g., “flower”, “irony”, etc.) will be decomposed into a stem and affix (e.g., “flower” parsed into “flow” and “-er”). These words are commonly referred to as semantically opaque words because their meanings cannot be derived from their base morphemes. According to form-then-meaning accounts, semantics do not play a role in the early processing stages of complex words; that is, decomposition is *morpho-orthographic* and takes place indiscriminately for both semantically opaque words and actual complex words, regardless of any semantic relationship between the stem and the whole word.

In contrast to form-then-meaning accounts, form-with-meaning accounts posit an early role of semantics in morphological processing, such that decomposition is *morpho-semantic* and involves a semantically-based search for morphemes (e.g., Feldman, O'Connor, & Moscoso del Prado Martin, 2009). These accounts propose that morphological decomposition takes place only for semantically transparent words in which the meaning of the complex word (e.g., "running") is derived from the meaning of its stem (e.g., "run"). This means that decomposition takes place for "swimmer" but not for "flower", since "flow" does not inform the meaning of the word "flower".

The fundamental distinction between form-then-meaning and form-with-meaning accounts is therefore their assumptions on the role of semantics in the morphological decomposition process. Specifically, form-then-meaning accounts propose that complex words are decomposed based only on orthographic form regardless of semantics, whereas form-with-meaning accounts argue that decomposition only occurs when the stem is semantically related to the word itself.

In contrast with form-then-meaning accounts and form-then-meaning accounts, connectionist theories argue that morphemes are not explicitly represented in our lexicon, and so there is no morpho-orthographic nor morpho-semantic decomposition process for complex words. Instead, morphemes are encoded as patterns of statistical co-occurrences between a word's orthography and its meaning (e.g., Gonnerman, Seidenberg, & Andersen, 2007; Joanisse & Seidenberg, 1999; Rueckl & Raveh, 1999). According to the connectionist perspective, words are networks of units that represent the mappings between a word's orthography, phonology, and semantics. These units capture the consistencies in the mappings between orthography, phonology, and semantics in the lexicon (see Rueckl & Seidenberg, 2009, for a review). Morphologically related words that consistently share both orthography and semantics (e.g., "swim", "swimmer", "swimming") co-activate similar mappings in these networks, compared to words that share only orthography (e.g., “corn”, “corner”, “cornet”) or semantics (e.g., “flower”, “plant”, “grass”). Morphemes thus reflect a superadditive effect of shared form and meaning, and are an emergent property of statistical learning, rather than discrete representations in our lexicon (Jared, Jouravlev, & Joanisse, 2017). This perspective is also in line with the argument that words that are morphologically related tend to have very systematic form-to-meaning mappings (Marelli, Traficante, & Burani, 2020), otherwise referred to as orthographic-semantic consistency (i.e., the same spellings usually refer to the same meanings) (Marelli & Amenta, 2018; Marelli, Amenta, & Crepaldi, 2015).

Are Semantic Transparency Effects Robust?

To adjudicate between these accounts, researchers have relied on findings from masked priming experiments, wherein participants respond to target words which are preceded by masked, briefly presented (≤ 50ms) primes. Specifically, researchers have examined the extent to which morphological masked priming effects are dependent on semantic transparency, which is the extent to which the meaning of a complex word can be predicted from the meaning of its base morpheme (Bell & Schäfer, 2016). In these studies, stem words are presented as target words which are preceded either by semantically transparent primes (e.g., painter-PAINT), quasi-transparent primes (i.e., primes that share a moderate amount of semantic relatedness to the target, e.g., bookish-BOOK), semantically opaque primes (e.g., corner-CORN), or orthographically related words that have the stem words embedded but no affix (e.g., cashew-CASH). The basic premise is that, if morphological decomposition is indiscriminate and based exclusively on orthographic form, then responses to the stem word should be comparably facilitated by transparent, quasi-transparent, and opaque primes. If morphological decomposition depends on semantics, then responses to the stem word should be facilitated by both transparent and quasi-transparent primes, but not opaque primes. Finally, if morphological processing can be accounted for by the consistency in mappings between form and meaning, then there should be graded effects of semantic transparency, such that priming effects should be greatest for transparent primes, followed by quasi-transparent primes, then opaque primes.

Despite extensive work, the empirical evidence remains mixed. Several studies have demonstrated that transparent and opaque primes afford statistically equivalent facilitation (e.g., Beyersmann *et al.*, 2016; Marslen-Wilson, Bozic, & Randall, 2008; Rastle, Davis, & New, 2004). That said, some studies have reported priming effects only for transparent primes (e.g., Feldman, O'Connor, & Moscoso del Prado Martin, 2009; Giraudo & Grainger, 2001), while others have reported largest priming effects for transparent primes, followed by opaque primes, then orthographic primes (e.g., Diependaele, Sandra, & Grainger, 2005; Diependaele, Sandra, & Grainger, 2009; Feldman *et al.*, 2015). A few studies have also reported graded effects of facilitation that were related to the degree of semantic transparency between prime-target pairs, such that priming effects were largest for transparent primes, moderate for quasi-transparent primes, and smallest (or inhibitory) for opaque primes (e.g., Jared, Jouravlev, & Joanisse, 2017). It therefore remains unclear which theoretical account of morphological processing best describes how complex words are recognized during word recognition.

Does the Task Matter?

It is noteworthy that semantic transparency effects have been examined almost exclusively with the lexical decision task, wherein participants classify letter strings as words or nonwords (e.g., is "cat" a word or nonword?). However, Kinoshita and Norris (2012) have argued that masked priming effects can depend dynamically on task-specific demands, and researchers have also cautioned against drawing conclusions about lexical processing from a single task because observed effects may reflect task-specific mechanisms (see Figure 1; Jacobs & Grainger, 1994). To clarify the interplay between task-general and task-specific mechanisms in morphological processing, semantic transparency effects should also be examined in other lexical processing tasks.

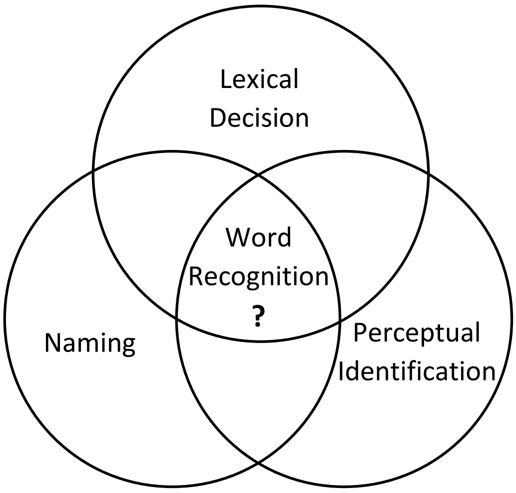

Figure 1. Venn diagram illustrating the concept of functional overlap between tasks. From "Models of Visual Word Recognition: Sampling the State of the Art," by Jacobs, A. M. and Grainger, J., 1994, *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), p. 1329. Copyright 1994 by the American Psychological Association, Inc.

There is some evidence from the literature that semantic transparency effects can be strengthened by experimental conditions which emphasize semantic processing. Consider Marelli *et al.*'s (2013) eye-tracking study, where Italian masked primes were followed by a target word adjacent to a digit. On about 15% of the trials, participants answered a question about the target (e.g., is it a tool?) or the number (e.g., is it odd?). In this semantic decision task, first-fixation and gaze durations on target words were facilitated *only* by transparent primes. Importantly, when the same stimuli were presented in lexical decision, priming was significant for both transparent and opaque targets, with stronger priming in the former.

What happens if we turn to a task which minimizes lexicosemantic access? This was explored in Duñabeitia *et al.*'s (2011) study with the cross-case same-different task, which requires participants to decide if an uppercase target in Spanish (e.g., LUCHA *'fight'*) is the same word as a lowercase reference (e.g., lucha). Results revealed equivalent facilitation by transparent, opaque, and orthographic primes, suggesting that morphological decomposition does not even occur when lexical access is non-obligatory.

That said, the picture is not yet entirely clear. In a more recent event-related potential (ERP) study, Jared, Jouravlev, and Joanisse (2017) observed qualitatively similar graded effects of semantic transparency (i.e., transparent priming > quasi-transparent priming > opaque priming) in both English lexical decision and semantic decision. Two other studies have also demonstrated that the meanings of embedded word stems in Italian (e.g., “pea” in “peace”) are activated when complex words are observed directly (i.e., not as masked primes) in a sentence reading task (Amenta, Marelli, & Crepaldi, 2015) and semantic decision task (Hasenäcker, Solaja, & Crepaldi, 2020) regardless of semantic transparency, thus challenging the assertion that semantic transparency effects are moderated by task requirements.

The Present Study

Despite mixed support for task-specificity in morphological processing, between- and within-study variability in experimental designs and dependent measures complicate cross-task comparisons. In Marelli *et al.* (2013), participants were required to respond to only 15% of trials in semantic decision, but to 100% of trials in lexical decision. The two tasks also collected different dependent measures - eye movements for semantic decision and response times (RTs) for lexical decision. For these reasons, it is difficult to tell if cross-task differences in Marelli *et al.* partly reflect differences in decision-related strategies or dependent measures. Similarly, Jared, Jouravlev, and Joanisse’s (2017) semantic decision findings were based only on ERP data for no-go responses, again making direct comparisons difficult (for more discussion, see Kuperman *et al.,* 2013).

Other than differences in experimental designs and dependent measures, extant empirical discrepancies may also be driven by how semantic transparency has thus far been measured and controlled for in the literature. For instance, while Jared, Jouravlev, and Joanisse (2017) classified their prime-target pairs as transparent or opaque based on subject ratings and latent semantic analysis (LSA; Laudauer , Foltz, & Laham, 1998), it is unclear how the prime-target pairs in Marelli *et al.* (2013) were selected and classified into their respective conditions. Selection of stimuli based on intuition alone is problematic because using opaque prime-target pairs that are semantically related may spuriously yield significant opaque priming that is driven by semantics (Feldman, O'Connor, & Moscoso del Prado Martin, 2009), leading researchers to claim that morphological decomposition is semantically "blind". For instance, when Morris *et al.* (2007) removed 22 of their 108 opaque primes found to be slightly related to their targets (e.g., manage-MAN, secretary-SECRET) based on a post-test ratings check, the priming effect observed for opaque primes was no longer significant.

The aim of the present study was thus to extend earlier work by examining whether semantic transparency effects are indeed moderated by task demands. To this end, we examined semantic transparency effects using the masked morphological priming paradigm in both semantic categorization (e.g., is "cat" abstract or concrete?) (Experiment 1a) and lexical decision (Experiment 1b). To facilitate comparison of the data across tasks, both experiments used identical stimuli, similar experimental procedures, and common dependent measures.

Further, to increase methodological rigor of the work, we collected post-experimental ratings of semantic relatedness between prime-target pairs to verify that our stimuli were appropriately assigned to each semantic transparency condition. As described below, we ensured that the semantic relatedness of transparent prime-target pairs was significantly higher than those of both opaque and orthographic prime-target pairs. Crucially, the semantic relatedness of opaque and orthographic prime-target pairs were also matched to ensure that opaque primes and their stems did not share any meaning, and that the suffixes in opaque primes did not function as real suffixes.

Method

Participants

Participants were 202 undergraduates from the National University of Singapore (NUS) who were reimbursed SGD10 or received course credit for their participation. 125 of these participants took part in the semantic categorization task (Experiment 1a), and 77 participants took part in the lexical decision task (Experiment 1b). All participants reported English as their first language and scored at least 50% on the Ghent English vocabulary test (Brysbaert *et al.*, 2016). All participants also had normal or corrected-to-normal vision.

Materials and Design

A set of 360 related prime-target pairs were selected from Brysbaert, Warriner, and Kuperman’s (2014) concreteness norms. Half of the target words were concrete words with mean ratings of more than 3.0, and the other half were abstract words with mean ratings of less than 3.0. These prime-target pairs were equally divided into each semantic transparency condition (transparent, opaque, orthographic) for each category (concrete vs. abstract). Care was taken to ensure that the prime-target pairs were, as much as possible, phonologically similar (i.e., ponder-POND and not brother-BROTH). These primes and targets were matched across conditions within each category for length, word frequency, orthographic neighborhood density, and orthographic neighborhood frequency. We attempted to match targets on orthographic-semantic transparency (Marelli & Amenta, 2018) but data were not available for 14 of our targets. Another 360 unrelated control primes were selected from the English Lexicon Project (ELP; Balota *et al.*, 2007) and matched with the related primes for the same variables using Match (van Casteren & Davis, 2007) (see Appendix A for the list of stimuli).

In addition, 360 corresponding legal nonwords and their unrelated primes were created using the pseudoword generator Wuggy (Keuleers & Brysbaert, 2010) for the lexical decision task (Experiment 1b). The nonwords and their unrelated primes matched their word counterparts in terms of number of letters and syllables, and at least two out of three sub-syllabic segments. Related primes for the nonwords were then formed by adding suffixes to the nonwords that were the same as their corresponding related word primes (e.g., the related nonword prime-target pair for “runner-RUN” was “renner-REN”).

All prime-target pairs were divided equally into two lists in order to counterbalance prime-target relation (related vs. unrelated) between participants. For all the target words (and nonwords for Experiment 1b) appearing in one list, half of them were preceded by a related prime, while the other half were preceded by an unrelated prime, with an equal number from each category and each semantic transparency condition. These prime-target associations were reversed in the second list, such that targets that were preceded by a related prime in the first list were preceded by an unrelated prime in the second list, and vice versa.

Procedure

All instructions, item presentation, and data collection were computer controlled. Participants were tested individually in a sound-attenuated cubicle, and each experimental session lasted approximately 45 minutes. Both the semantic categorization and lexical decision tasks were programmed using DMDX (Forster & Forster, 2003), and stimuli were presented at a refresh rate of 100 Hz. On each trial, a fixation point was presented in the center of the screen for 400ms, followed by a forward mask (#######) for 500ms. The prime was then presented in lowercase for 40ms[1], immediately followed by the target in uppercase. The target remained on screen for 5000ms, or until a response was made. In Experiment 1a, participants were asked to classify each target word as concrete or abstract by pressing the right or left <SHIFT> key on the keyboard, while in Experiment 1b, participants were asked to classify each target as a word or nonword by pressing the right or left <SHIFT> key on the keyboard. There was no mention of the masked prime, and participants were asked to respond as quickly and as accurately as possible.

For both Experiments 1a and 1b, participants completed 20 practice trials before receiving one of the two experimental lists. In Experiment 1a, approximately half of the participants (*n* = 63) were presented the first list, and the remaining (*n* = 62) were presented

---

[1] This 40ms SOA is in line with other masked priming studies investigating the effects of semantic transparency at the earliest stages of processing (e.g., Beyersmann, Coltheart, & Castles, 2012; Diependaele, Sandra, & Grainger, 2005; Diependaele, Sandra, & Grainger, 2009; Whiting, Cowley, & Bozic, 2017, etc.).

the second list. This was similar in Experiment 1b, where approximately half of the participants (*n* = 38) were presented the first list, and the remaining (*n* = 39) were presented the second list. Each participant therefore saw each target word only once, either preceded by a related or unrelated prime. Stimuli were divided into four blocks of 90 trials in Experiment 1a, and into nine blocks of 80 trials in Experiment 1b. Participants were allowed a short break between blocks, and order of trials within each list was randomized.

Similar to Jared, Jouravlev, and Joannise (2017), participants in Experiment 1a also completed a semantic relatedness ratings task for the prime-target pairs[2]. The purpose of the ratings procedure was to ensure that the semantic similarity between primes and targets were appropriately manipulated for each condition. The semantic relatedness ratings task was programmed using E-Prime 2.0 (Schneider, Eschman, & Zuccollotto, 2002). The 720 related and unrelated prime-target pairs were divided equally into four lists of 180 prime-target pairs each, with an equal number of related and unrelated prime-target pairs from each category and each semantic transparency condition in each list. Each participant received one of the four lists. On each trial, participants were presented a prime-target pair in the center of the screen (e.g., valley-FREE), and were asked to rate how closely related the pair of words were in meaning on a 9-point scale (1 = not at all related, 5 = somewhat related, 9 = very related). Responses were collected via the keyboard, and there was no time limit for each trial. Order of trials within each list was randomized.

Results

Data Cleaning

Transparent prime-target pairs with semantic relatedness ratings of ≤ 7.0 and orthographic, opaque, and unrelated pairs with ratings of ≥ 3.0 were first excluded; this left 39 to 48 words in each semantic transparency condition for both concrete and abstract words.

---

[2] Ratings were collected within the same session, instead of a pretest, to ensure that the estimates of semantic relatedness were obtained from the same group of participants that completed the experiments. This was the same procedure as in Feldman, O’Connor, and Moscoso del Prado Martin (2009).

For this set of items, semantic relatedness was significantly different between the transparent and opaque conditions, but matched between the opaque and orthographic conditions, for both abstract and concrete words (see Table 1).

Table 1.

Means and standard deviations for the control variables and semantic relatedness ratings for related primes and targets across conditions within each category. Standard deviations are in parentheses.

	Semantic Transparency Condition		
	Orthographic	Opaque	Transparent
Concrete Category			
Semantic Relatedness Ratings	1.82 (0.51)	1.88 (0.47)	8.41 (0.41)
Targets			
Length	3.94 (0.64)	4.12 (0.71)	4.19 (0.61)
Word frequency	2.81 (0.74)	2.53 (0.58)	2.73 (0.68)
Orthographic neighborhood density	12.04 (6.34)	11.17 (5.88)	10.19 (5.55)
Orthographic neighborhood frequency	2.55 (0.68)	2.44 (0.45)	2.39 (0.40)
Orthographic-semantic consistency	0.43 (0.30)	0.35 (0.29)	0.73 (0.22)
Primes			
Length	6.34 (1.09)	6.31 (0.78)	6.02 (0.87)
Word frequency	2.03 (0.75)	2.39 (0.81)	2.06 (0.62)
Orthographic neighborhood density	2.36 (2.23)	3.19 (2.92)	3.23 (3.14)
Orthographic neighborhood frequency	1.43 (0.87)	1.68 (0.80)	1.66 (0.92)
Abstract Category			
Semantic Relatedness Ratings	1.84 (0.52)	2.05 (0.47)	8.10 (0.45)
Targets			
Length	4.14 (0.82)	4.49 (1.17)	4.48 (0.74)
Word frequency	2.70 (1.28)	2.96 (1.14)	3.14 (1.07)
Orthographic neighborhood density	11.11 (7.39)	8.56 (7.14)	7.54 (6.88)
Orthographic neighborhood frequency	2.32 (0.70)	2.43 (0.73)	2.30 (0.82)
Orthographic-semantic consistency	0.39 (0.32)	0.48 (0.31)	0.75 (0.31)
Primes			
Length	6.52 (1.09)	6.97 (1.31)	6.79 (0.94)
Word frequency	2.18 (0.87)	2.61 (0.85)	1.93 (0.82)
Orthographic neighborhood density	1.93 (1.96)	2.38 (2.54)	1.90 (2.22)
Orthographic neighborhood frequency	1.38 (1.00)	1.63 (1.01)	1.27 (1.04)

Note. Orthographic-semantic consistency data were not available for 14 of our targets. The values presented here are based on data for the remaining items.

Fifteen targets (annotated in the Appendix) with overall error rates of more than 50% in either Experiment 1a and/or Experiment 1b were also excluded from analyses. For Experiment 1a, data from three participants were excluded because their average RTs were greater than 1500ms and/or their average accuracy was lower than 60% (1Q - (1.5 × IQR)).

Data from one participant were lost due to an experimenter error, yielding a final sample of 121 participants. For Experiment 1b, data from four participants were excluded because their average accuracies were lower than 80% (1Q - (1.5 × IQR)), yielding a final sample of 73 participants. RT data were then cleaned to remove incorrect responses, and responses faster than 200ms or slower than 3000ms. Remaining RTs which were within 2.5 SDs of each participant's overall mean were included in the analyses of RTs.

Data Analyses

Raw, untransformed RTs and accuracy rates were analyzed using linear mixed effects (LME) models using *R* (R Core Team, 2019). Data were fitted using the lme4 package (Bates *et al.*, 2015); *p*-values were obtained using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016). Contrast analyses on the interactions were performed using the car package (Fox *et al.*, 2020) and phia package (De Rosario-Martinez & Fox, 2015). Trial number, prime and target length, word frequency (lgSUBTLWF; Brysbaert & New, 2009), orthographic neighborhood density and orthographic neighborhood frequency (Balota *et al.*, 2007), target morphological family size (Sánchez-Gutiérrez *et al.*, 2018), target concreteness (Brysbaert, Warriner, & Kuperman, 2014), target category (abstract vs. concrete), condition (orthographic, opaque, transparent) and prime type (related vs. unrelated) were included as fixed effects. Random intercepts for participants and targets were included. Since condition has three levels, orthographic targets were treated as the reference group. All continuous variables were centered before analyses. Table 2 presents the correlations between the predictors Experiments 1a and 1b.

Table 2.

Pearson correlation coefficients for the lexical predictors in Experiments 1a and 1b. p-values are indicated by asterisks (*).

	1	2	3	4	5	6	7	8	9
1. Target length	-								
2. Target frequency	-.25***	-							
3. Target neighborhood density	-.62***	.30***	-						
4. Target neighborhood frequency	-.49***	.22***	.44***	-					
5. Prime length	.57***	-.12	-.42***	-.32***	-				
6. Prime frequency	-.05	.23***	.13*	.16**	-.13*	-			
7. Prime neighborhood density	-.22***	.08	.27***	.25***	-.49***	.13*	-		
8. Prime neighborhood frequency	-.13*	.08	.14*	.21***	-.43***	.30***	.45***	-	
9. Concreteness	-.17**	-.13*	.18**	.13*	-.21***	-.04	.20**	.10	-
10. Morphological family size	.01	.15*	.03	.01	.09	.05	-.04	.01	.08

Experiment 1a (Semantic Categorization)

Interactions between condition, prime type, and category were only included if they improved model fit in a forward stepwise model selection procedure. Chi-squared log-likelihood ratio tests revealed that including a condition × prime type interaction significantly improved model fit, $\chi^2(2) = 6.93$, $p = .03$, indicating a significant interaction between condition and prime type. The inclusion of a condition × prime type × category (concrete vs. abstract) interaction did not significantly improve model fit, $\chi^2(2) = 0.59$, $p = .75$, indicating that the priming effects did not differ between categories[3]. Simple effects analyses revealed that the effect of prime type was significant for transparent targets, $\chi^2(1) = 6.22$, $p = .038$; responses to targets were faster when preceded by related, compared to unrelated, primes. Priming was not reliable for opaque targets, $\chi^2(1) = 0.03$, $p = .86$, and orthographic targets, $\chi^2(1) = 1.24$, $p = .53$. Analyses of accuracy rates revealed no significant priming effects.

Experiment 1b (Lexical Decision)

For Experiment 1b, chi-squared log-likelihood ratio tests also revealed that including a condition × prime type interaction significantly improved model fit, $\chi^2(2) = 27.10$, $p < .001$, indicating a significant interaction between condition and prime type. Simple effects analyses

---

[3] Analyses of the concrete and abstract categories separately revealed no significant priming effects for all three conditions. This is potentially due to a lack of power.

revealed that the effect of prime type was significant for transparent targets, $\chi^2(1) = 92.09$, $p < .001$, opaque targets, $\chi^2(1) = 3.87$, $p = .049$, and orthographic targets, $\chi^2(1) = 12.71$, $p < .001$; responses to targets were faster when preceded by related, compared to unrelated, primes. Priming was also significantly larger for transparent targets than opaque targets ($p < .001$), but statistically equivalent between opaque and orthographic targets ($p = .33$). Analyses of accuracy rates revealed significant priming only for transparent targets, $\chi^2(1) = 29.53$, $p < .001$, and not orthographic targets, $\chi^2(1) = 0.16$, $p = .69$, or opaque targets, $\chi^2(1) = 3.16$, $p = .15$.

Comparison of Experiments 1a and 1b

To further test whether priming effects differed between Experiments 1a and 1b, the data from both experiments were combined. Two additional LME models were then run, one including a condition × prime type × task (SCT vs. LDT) three-way interaction, and one without the interaction. Chi-squared log-likelihood ratio tests revealed that the inclusion of the interaction term did not significantly improve model fit, $\chi^2(2) = 1.68$, $p = .43$, indicating that the *pattern* of priming (i.e., transparent > opaque = orthographic) did not differ between the two tasks[4].

Additionally, some researchers have explained the effects of morphological processing in terms of orthographic-semantic consistency, and argue that past studies may not have adequately controlled for the degree of form-to-meaning mapping in their stimuli (Marelli & Amenta, 2018; Marelli, Amenta, & Crepaldi, 2015). We could not match orthographic-semantic consistency across conditions because the data are not available for 14 of our targets. However, when we re-ran our analyses with the remaining items and included

---

[4] One may argue that cross-task comparisons should be interpreted with caution because ‘word’ targets in Experiment 1b were always preceded by ‘word’ primes, while the targets in Experiment 1a were not always preceded by category-congruent primes. However, Fernández-López, Marcet, and Perea (2019) demonstrated that there are no effects of response congruency in LDT when nonwords are orthographically legal. Lexical decision responses are therefore unlikely to have been influenced by the lexical status of the primes in Experiment 1a.

orthographic-semantic consistency (Marelli & Amenta, 2018) as a fixed effect in our LME models, the pattern of results for both tasks remained the same. Also, when we repeated our analyses with z-scored RTs (ZRT), reciprocal RTs (1/RT), and log-transformed RTs, the pattern of results for both tasks again remained the same and are thus not reported here.

Table 3 presents the results for the linear mixed model estimates, and Table 4 presents the mean RTs for target words in each condition, for Experiments 1a and 1b.

Table 3.

Linear mixed model estimates of coefficients, standard errors, and t-statistics for fixed effects in Experiments 1a and 1b. p-values are indicated by asterisks (*).

	Experiment 1a			Experiment 1b		
Fixed effects	Est.	*S.E.*	*t*	Est.	*S.E.*	*t*
Intercept	879	21.35	41.16***	615	13.14	46.80***
Trial number	-0.15	0.02	-9.09***	0.00	0.01	0.26
Target length	-2.36	9.08	-0.26	-21.41	5.75	-3.72***
Target frequency	-34.76	5.62	-6.19***	-45.63	3.56	-12.82***
Target orthographic neighborhood density	1.78	0.97	1.82	-0.51	0.62	-0.83
Target orthographic neighborhood frequency	-6.43	9.27	-0.69	3.92	3.85	1.02
Prime length	10.16	6.51	1.56	4.34	4.14	1.05
Prime frequency	7.85	6.56	1.20	6.82	4.16	1.64
Prime orthographic neighborhood density	0.52	2.28	0.23	0.16	1.45	0.11
Prime orthographic neighborhood frequency	11.26	6.07	1.86	-0.96	5.87	-0.16
Concreteness	-37.27	12.35	-3.02**	-4.41	7.75	-0.57
Morphological family size	0.11	0.47	0.23	-0.83	0.29	-2.82**
Condition (Contrast 1)	5.80	13.03	0.45	-6.53	8.39	-0.78
Condition (Contrast 2)	-17.35	12.53	-1.39	-49.27	8.04	-6.13***
Prime type	-6.22	5.59	-1.11	14.95	4.19	3.57***
Target type	-10.88	28.96	-0.38	-4.75	18.11	-0.26
Condition (Contrast 1) × Prime type	5.09	8.38	0.61	-5.98	6.20	-0.97
Condition (Contrast 2) × Prime type	20.00	7.87	2.54*	22.91	5.76	3.97***

Note. *** $p < .001$, ** $p < .01$, * $p < .05$. Contrast 1 – opaque vs. orthographic; Contrast 2 – transparent vs. orthographic.

Table 4.

Mean response times (in milliseconds) and error rates for target words preceded by related and unrelated primes in each condition in Experiments 1a and 1b. Error rates are in parentheses.

	Experiment 1a			Experiment 1b		
	Orthographic	Opaque	Transparent	Orthographic	Opaque	Transparent
Unrelated	840 (.90)	851 (.89)	843 (.88)	628 (.96)	615 (.97)	602 (.98)
Related	846 (.91)	852 (.88)	829 (.88)	613 (.97)	606 (.98)	564 (.99)
Priming effect	-6 (.01)	-1 (-.01)	14* (.00)	15*** (.01)	9* (.01)	37*** (.01***)

Note. *** $p < .001$, ** $p < .01$, * $p < .05$.

Quantile Plots

To characterize the observed effects in a more fine-grained level, we also generated quantile plots to explore the influence of variables on different portions of the underlying response time distribution (Balota & Yap, 2011). To do this, response times for each participant were used to obtain quantiles (.1, .3, .5, .7, .9) for the different experimental conditions (Ratcliff, Gomez, & McKoon, 2004). Quantiles were then averaged across participants, and priming effects derived by computing the difference between related and unrelated trials at each quantile. Figure 2 plots the priming effects for transparent, opaque, and orthographic trials across the response time distribution; fastest quantiles are found on the left and slowest quantiles on the right.

In semantic categorization (Experiment 1a; top panel), masked priming was evident only for transparent targets and was approximately the same size (~ 20ms) across most of the quantiles. That is, masked transparent priming was reflected by a shift of the entire response time distribution. Turning to lexical decision (Experiment 1b; bottom panel), the different forms of masked priming were associated with distinct distributional signatures. Orthographic priming was relatively constant in size (~ 12ms) across quantiles, indicating distributional shifting. Similar trends were observed for opaque priming, with the key difference being that priming was attenuated in the slowest quantile. Finally, transparent priming showed a modest *increase* across the distribution, i.e., priming was stronger for slower targets. We discuss these findings in the Discussion.

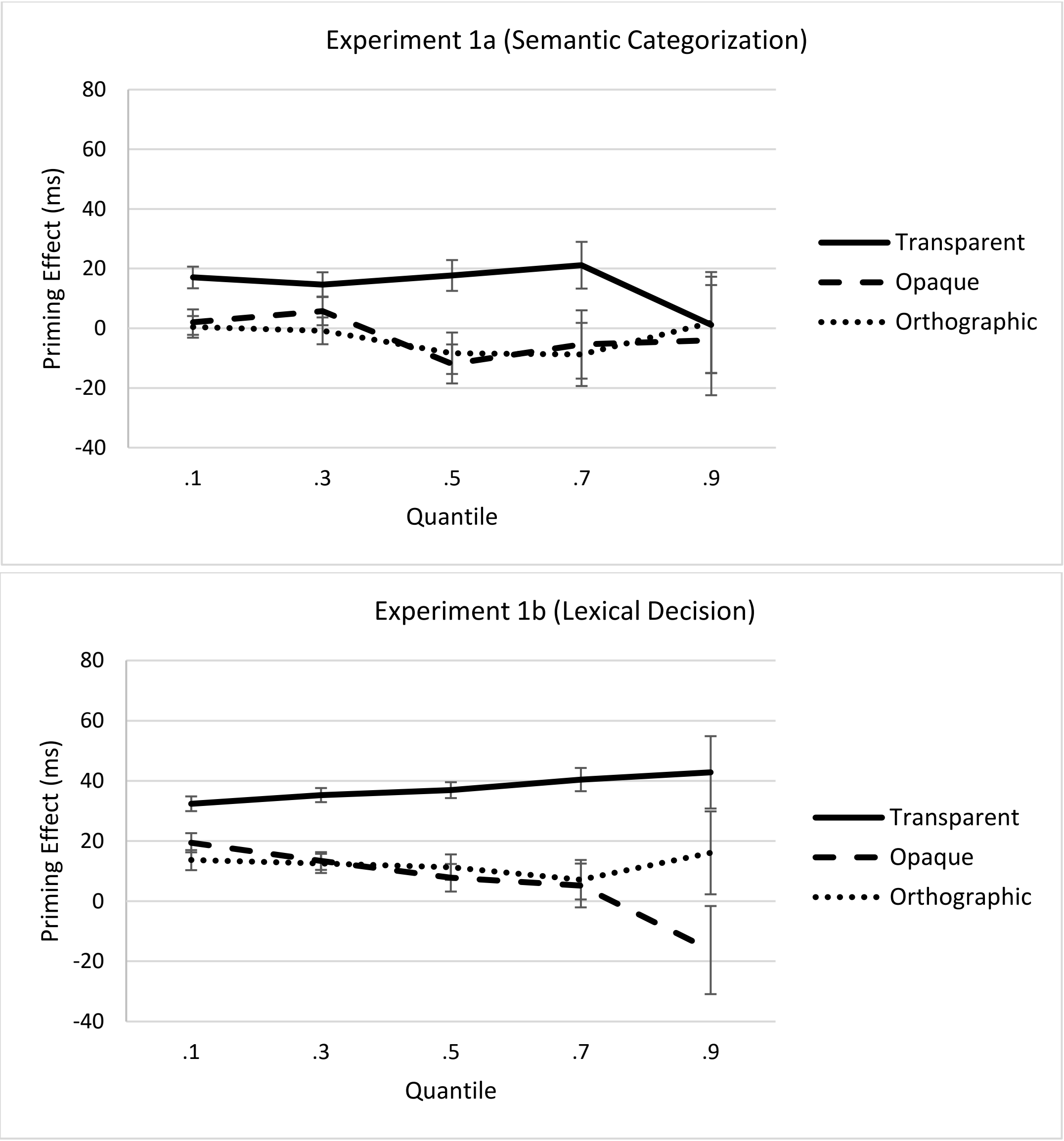

Figure 2. Quantile plots for the priming effects observed in semantic categorization (Experiment 1a; top panel) and lexical decision (Experiment 1b; bottom panel).

Discussion

This study builds on and extends earlier work (e.g., Duñabeitia *et al.*, 2011; Jared, Jouravlev, & Joanisse, 2017; Marelli *et al.*, 2013) by examining semantic transparency effects in morphological processing across semantic categorization (Experiment 1a) and lexical decision (Experiment 1b), using identical stimuli, similar experimental procedures, and common dependent measures. Additionally, the study was based on a large set of validated and carefully controlled prime-target pairs that were presented to a relatively large sample of participants.

Comparing both experiments, it is clear that priming was strongest in the transparent condition, and was significantly weaker and of comparable magnitude in the orthographic and opaque conditions. Specifically, in semantic categorization (Experiment 1a), response times to target stem words were significantly facilitated only by transparent primes, but not by opaque or orthographic primes. In lexical decision (Experiment 1b), transparent priming was substantially larger than opaque and orthographic priming, while the latter two forms of priming were of comparable magnitude. Although opaque priming was significant in the lexical decision task, we argue that this facilitation is better attributed to orthographic overlap, rather than to the processes underlying morphological decomposition. This is because, in the absence of any semantic relationship between the opaque prime and the target, any observed facilitation for opaque primes can only reflect either orthographic overlap and/or morphological decomposition (Longtin, Segui, & Halle, 2003); the inclusion of the orthographic control helps to further constrain the locus of the effect (Pastizzo & Feldman, 2002). Given that orthographic and opaque priming were of statistically significant magnitudes in the lexical decision task (p = .33; see Table 3), this suggests that opaque priming can be more parsimoniously explained by orthographic overlap, rather than morphological decomposition.

The present findings are therefore very difficult to reconcile with form-then-meaning accounts, which assume that complex words are decomposed in a semantically “blind” manner at the earliest stages of processing. Instead, both experiments strongly indicate that morphological processing depends on a semantic relationship between a complex word and its stem, because transparent priming was stronger than any observed opaque priming in both tasks. Our study also raises the possibility that the graded effects observed in earlier studies (e.g., Marelli *et al.*, 2013) may partly reflect the use of opaque prime-target pairs that were more transparent than their orthographic controls. As described earlier, we relied on subject ratings of semantic transparency as a manipulation check and were able to verify that the critical orthographic and opaque conditions were matched.

Next, we observed the same *qualitative* pattern of semantic transparency effects (i.e., transparent > opaque = orthographic) across the lexical decision and semantic categorization tasks, which are remarkably consistent with Jared, Jouravlev, and Joanisse's (2017) study reporting significant facilitation by transparent primes, but no statistical difference between opaque and orthographic primes ($p > .05$) in both go/no-go semantic decision and lexical decision. The present findings indicate that early morphological processing is not driven by idiosyncratic task demands but is instead generalizable to different lexical processing tasks. These results also provide an intriguing contrast against the other effects in our study which were moderated by task. For example, lexical influences (e.g., frequency, length) on target processing were significant in lexical decision, whereas semantic influences (e.g., concreteness) were significant in semantic categorization. The foregoing between-task differences are consistent with the idea that a lexical decision response can be driven by orthographic familiarity (Balota & Chumbley, 1984) or pre-lexical orthographic representations (Grainger & Jacobs, 1996), whereas a semantic categorization response requires a word's meaning to be accessed.

That said, although both tasks yielded qualitatively similar trends with respect to semantic transparency, there were a couple of interesting *between-task* differences. First, RTs were longer and error rates were all higher in semantic categorization than lexical decision (see Table 4), which is unsurprising given that the semantic categorization task is more challenging (e.g., the meaning of the word needs to be retrieved) than the lexical decision task, and participants typically take longer to make a semantic categorization response than a lexical decision response for the same set of stimuli (see Yap *et al.*, 2012).

Second, opaque and orthographic priming were significant in lexical decision but not in semantic categorization (see Table 4). To account for these affects, we turn to the non-morphological process of embedded stem activation (Grainger & Beyersmann, 2017), which is the notion that the orthography of an embedded stem (e.g., “cash” in “cashew”, “pea” in “peace”) is activated when contained in a longer word. Studies have shown that embedded stem activation occurs automatically, and facilitates responses in masked nonword priming paradigms regardless of whether it is combined with an affix (e.g., cheapize-CHEAP) or non-affix (e.g., cheapstry-CHEAP) (e.g., Beyersmann *et al.*, 2016; Hasenäcker, Beyersmann, & Schroeder, 2016; Heathcote *et al.*, 2018). We argue that embedded stem activation is more likely to facilitate lexical decision than semantic categorization, because the activated orthography of the embedded stem can drive a correct lexical decision response based on orthographic familiarity (Balota & Chumbley, 1984; Grainger & Jacobs, 1996). In contrast, in semantic categorization, even if the orthography of an embedded stem has been activated, its meaning must be retrieved before a response can be made. While studies have shown that the automatic orthographic activation of embedded stems also activates its semantics (e.g., Bowers, Davis, & Hanley, 2005; Hasenäcker, Solaja, & Crepaldi, 2020), these findings have *only* been observed when the word is consciously processed (i.e., directly observed), and there is currently no evidence that such semantic activation is reliable under masked priming conditions.

To our knowledge, no study has examined task differences in the strength of embedded stem activation effects using the same set of stimuli. In sum, the stronger transparent, opaque, and orthographic priming observed in lexical decision could potentially reflect the greater utility of embedded stem activation in lexical decision than semantic categorization in masked priming paradigms. That is, while embedded stem activation occurs in both lexical decision and semantic categorization, it is more likely to facilitate responses under masked priming conditions in the former. This is consistent with the previous research that have demonstrated that the strength and direction of lexicosemantic effects can be systematically modulated by task-specific demands (Yap *et al.,* 2011), which is in line with the perspective of a flexible lexical processor which leverages attentional mechanisms to optimize information processing and performance on a given task (Balota & Yap, 2006). Of course, this account is *post-hoc* and speculative and needs to be empirically verified in future work.

Distributional Analyses

Supplementary response time distributional analyses were also conducted to shed more light on the influence of masked morphological primes on different portions of the response time distribution. To our knowledge, Andrews and Lo (2013) and Hasenäcker, Beyersmann, and Schroeder (2016) are the only published studies that have reported such analyses in masked morphological priming. Specifically, in lexical decision, both studies reported that both transparent and opaque priming were associated with distributional shifting. In the present study, transparent priming in semantic categorization was similarly reflected by distributional shifting, consistent with a relatively modular head-start mechanism in which primes pre-activate the lexical representations of target words by some constant amount (Balota *et al.*, 2008).

Turning to lexical decision, the distributional signatures for orthographic and opaque priming were very similar, except that opaque priming was attenuated in the slowest quantile. In contrast, masked transparent priming became larger for slower targets. This dissociation is intriguing and seems compatible with other empirical demonstrations (e.g., Feldman *et al.*, 2015; Rastle, Davis, & New, 2004) in which longer stimulus onset asynchronies are associated with increased transparent priming and decreased opaque priming. That said, we should emphasize that the foregoing analyses are tentative, and should be cautiously interpreted because the present experimental design is *not* optimized for such analyses. For example, the different conditions did not have a fixed number of observations after applying the various exclusion criteria, and the quantile analyses do not control for the covariates included as fixed effects in the LME models. We also need to acknowledge that our lexical decision quantile plots diverge from those reported by Andrews and Lo (2013) and Hasenäcker, Beyersmann, and Schroeder (2016), but the methodological differences (e.g., more observations per cell, stimulus validation through semantic relatedness ratings) between these studies complicate comparisons. In any case, we look forward to future work that verify the foregoing findings with more appropriate designs and more granular distributional analyses.

Limitations and Future Directions

While our results speak against form-then-meaning accounts of morphological decomposition, we note that our results do not distinguish between form-with-meaning accounts and connectionist theories of morphological processing, which both predict larger priming effects for transparent primes than opaque primes. Future research on task-specific effects in morphological processing can consider including a quasi-transparent condition which would allow the observation of any graded effects, which may help differentiate between the two theoretical accounts.

On a related note, semantic transparency was defined and treated as a categorical variable in our study (i.e., transparent vs. opaque). Subject ratings were used to ensure that transparent prime-target pairs were semantically related, while opaque and orthographic prime-target pairs were not. This stringent control was to ensure that our stimuli had been correctly classified, to facilitate interpretation of the data and to adjudicate between theoretical accounts of morphological processing. As a case in point, when we included analyses of all items based on our original classification (i.e., no targets were excluded based on subject ratings), the condition × prime type interaction was no longer significant in both tasks. Importantly, our subject ratings revealed that some transparent prime-target pairs that we had selected were not sufficiently related (e.g., boxer-BOX), while some opaque prime-target pairs were semantically related (e.g., crooked-CROOK), which demonstrates that researchers' intuitions may not be entirely reliable when it comes to stimuli selection.

However, we acknowledge that semantic transparency may be better defined as a graded, continuous variable (e.g., corner-CORN is more opaque than splinter-SPLINT). While the use of subject ratings for stimuli classification is not uncommon (e.g., Jared, Jouravlev, & Joanisse, 2017; Rastle *et al.*, 2000; Rastle, Davis, & New, 2004), one may argue that perceived semantic relatedness may not entirely reflect semantic transparency (e.g., flawless-FLAW had an average rating of 5.9, even though the meaning of "flawless" can be inferred from the meaning of "flaw"). Differences in how studies have defined and measured semantic transparency may have thus contributed to the mixed findings of semantic transparency effects in the current literature (see Amenta, Guenther, & Marelli, 2020, for a discussion). Some studies (e.g., Feldman *et al.*, 2009; 2015; Whiting, Cowley, & Bozic, 2017) have used distributional semantic models such as LSA (Laudauer, Foltz, & Laham, 1998) to compute semantic similarity based on the cosine value of vectors representing the meaning of primes and their targets. Marelli and Baroni (2015) have also developed a new framework for semantic transparency based on compositional distributional semantics, which takes into the ease of morpheme combination and the degree of the transformation brought about by the affix. As such, while we emphasize the need for semantic transparency to be clearly defined and measured, it would be of interest to examine if other measures of semantic transparency can also predict the priming effects observed in our study.

Lastly, it should be noted that the findings reported here may not be generalizable to languages other than English. Different languages have different morphological systems, and the mechanisms underlying morphological processing may differ across languages. For instance, languages can differ in the consistency with affixes can be identified based on their spellings (Jared, Jouravlev, & Joanisse, 2017). English, like French, contains a large number of pseudo-prefixes (e.g., “pre-” in “present”), and one may argue that semantically “blind” morphological decomposition in such a language may lead to more processing difficulty rather than efficiency (Diependaele, Grainger, & Sandra, 2012). On the other hand, languages like Spanish and Italian are considered morphologically productive; affixes are commonly used to form new words with bound stems (e.g., *casa/casita/casucha* – house/small house/hovel in Spanish and *gatto/gatta/gatti* – cat/female cat/cats in Italian). Morphological decomposition may therefore play a more important role in such languages, compared to English (Beyersmann, Coltheart, & Castles, 2012). We therefore do not rule out the possibility that the differences in our results and Marelli *et al.* (2013) may be due to language-specific differences, and look forward to future work examining the influence of language systems on morphological processing.

In conclusion, our study provides strong evidence that early morphological processing of complex words depends on the semantic properties of its stem, which is in line with form-with-meaning accounts and connectionist theories of morphological processing. Further, our study demonstrates that the mechanisms underlying morphological processing appears to be generalizable across tasks, but task demands can still affect the strength of priming effects, which is consistent with the notion of a flexible lexical processor (Balota & Yap, 2006). Our results therefore underscore the importance of examining differences in morphological processing across different paradigms, and future research can consider examining morphological decomposition in other task environments (e.g., context experiments) with comparable experimental procedures and dependent measures. This will further inform the interplay between task demands and morphological processing.

References

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: an analytical review of morphological effects in visual word identification. *Frontiers in Psychology, 3*, 232.

Amenta, S., Günther, F., & Marelli, M. (2020). A (distributional) semantic perspective on the processing of morphologically complex words. *The Mental Lexicon, 15*(1), 62-78.

Amenta, S., Marelli, M., & Crepaldi, D. (2015). The fruitless effort of growing a fruitless tree: Early morpho-orthographic and morpho-semantic effects in sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1587-1596.

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language, 68*(3), 279-296.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 340.

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). New York: Psychology Press.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160-166.

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 and Language, 59*(4), 495-523.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445-459.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package 'lme4'. *Convergence, 12*(1), 2.

Bell, M. J., & Schäfer, M. (2016). Modelling semantic transparency. *Morphology, 26*(2), 157-199.

Beyersmann, E., Castles, A., & Coltheart, M. (2012). Morphological processing during visual word recognition in developing readers: Evidence from masked priming. *Quarterly Journal of Experimental Psychology*, *65*(7), 1306-1326.

Beyersmann, E., Cavalli, E., Casalis, S., & Colé, P. (2016). Embedded stem priming effects in prefixed and suffixed pseudowords. *Scientific Studies of Reading*, *20*(3), 220-230.

Beyersmann, E., Coltheart, M., & Castles, A. (2012). Parallel processing of whole words and morphemes in visual word recognition. *Quarterly Journal of Experimental Psychology, 65*(9), 1798-1819.

Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016). Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review*, *23*(2), 533-539.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a "hat" in "that"?. *Journal of Memory and Language, 52*(1), 131-143.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods, 41*(4), 977-990.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant’s age. *Frontiers in Psychology, 7*:1116.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904-911.

De Rosario-Martinez, H., & Fox, J. (2015). Package ‘phia’. *CRAN repository.* Retrieved 15 Mar 2020.

Diependaele, K., Grainger, J., & Sandra, D. (2012). Derivational morphology and skilled reading. In M. J. Spivey, K. McRae, & M. F. Joanisse (Eds.), *The Cambridge Handbook of Psycholinguistics* (pp. 311-332). Cambridge University Press.

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, *20*(1-2), 75-114.

Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition, 37*(6), 895-908.

Duñabeitia, J. A., Kinoshita, S., Carreiras, M., & Norris, D. (2011). Is morpho-orthographic decomposition purely orthographic? Evidence from masked priming in the same–different task. *Language and Cognitive Processes, 26*(4-6), 509-529.

Feldman, L. B., Milin, P., Cho, K. W., Moscoso del Prado Martín, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience, 9*, 111.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, *16*(4), 684-691.

Fernández-López, M., Marcet, A., & Perea, M. (2019). Can response congruency effects be obtained in masked priming lexical decision?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(9), 1683-1702.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124.

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., ... & Gorjanc, G. (2019). R Core Team (2018). car: Companion to applied regression. *R package version 3.0-2.*

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review, 8*(1), 127-131.

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, *136*(2), 323-345.

Grainger, J., & Beyersmann, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 67, pp. 285-317). Academic Press.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518-565.

Hasenäcker, J., Beyersmann, E., & Schroeder, S. (2016). Masked morphological priming in German-speaking adults and children: Evidence from response time distributions. *Frontiers in Psychology*, *7*, 929.

Hasenäcker, J., Solaja, O., & Crepaldi, D. (2020). Food in the corner and money in the cashews: Semantic activation of embedded stems in the presence or absence of a morphological structure. *Psychonomic Bulletin & Review*, *27*(1), 155-161.

Heathcote, L., Nation, K., Castles, A., & Beyersmann, E. (2018). Do 'blacheap'and 'subcheap'both prime 'cheap'? An investigation of morphemic status and position in early visual word processing. *Quarterly Journal of Experimental Psychology*, *71*(8), 1645-1654.

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition: sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1311-1334.

Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(3), 422.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences, 96*(13), 7592-7597.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627-633.

Kinoshita, S., & Norris, D. (2012). Task-dependent masked priming effects in visual word recognition. *Frontiers in Psychology, 3*, 178.

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology*, *66*(3), 563-580.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). Tests in linear mixed effects models. *R package version*, *2,* 33.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259-284.

Longtin, C. M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*(3), 313-334.

Marelli, M., & Amenta, S. (2018). A database of orthography-semantics consistency (OSC) estimates for 15,017 English words. *Behavior Research Methods*, *50*(4), 1482-1495.

Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review*, *122*(3), 485-515.

Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of Orthography-Semantics Consistency on word recognition. *Quarterly Journal of Experimental Psychology*, *68*(8), 1571-1583.

Marelli, M., Amenta, S., Morone, E. A., & Crepaldi, D. (2013). Meaning is in the beholder's eye: Morpho-semantic effects in masked priming. *Psychonomic Bulletin & Review*, *20*(3), 534-541.

Marelli, M., Traficante, D., & Burani, C. (2020). Reading morphologically complex words: Experimental evidence and learning models. In V. Pirrelli, I. Plag, & W. U. Dressler (Eds.), *Word Knowledge and Word Usage* (pp. 553-592). De Gruyter Mouton.

Marslen-Wilson, W. D., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes, 23*(3), 394-421.

Morris, J., Frank, T., Grainger, J., & Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, *44*(4), 506-521.

Pastizzo, M. J., & Feldman, L. B. (2002). Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 244-249.

R Core Team. (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes, 15*(4-5), 507-537.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother’s brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*(6), 1090-1098.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*(1), 159-182.

Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language, 68*(1-2), 110-117.

Rueckl, J. G., & Seidenberg, M. S. (2009). Computational modeling and the neural bases of reading and reading disorders. In K. Pugh & P. McCardle (Eds.), *How children learn to read: Current issues and new directions in the integration of cognition, neurobiology and genetics of reading and dyslexia research and practice* (pp. 101–133). Psychology Press.

Sánchez-Gutiérrez, C. H., Mailhot, H., Deacon, S. H., & Wilson, M. A. (2018). MorphoLex: A derivational morphological database for 70,000 English words. *Behavior Research Methods, 50*(4), 1568-1580.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). E-Prime 2.0 reference guide manual. *Pittsburgh, PA: Psychology Software Tools.*

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 638-647.

van Casteren, M., & Davis, M. H. (2007). Match: A program to assist in matching the conditions of factorial experiments. *Behavior Research Methods*, *39*(4), 973-978.

Whiting, C. M., Cowley, R. G., & Bozic, M. (2017). The role of semantic context in early morphological processing. *Frontiers in Psychology, 8*, 991.

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6*, 72.

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review, 18*(4), 742-750.

Appendix A

List of concrete and abstract targets, and the related and unrelated primes used in Experiments 1a and 1b, organized by semantic transparency.

Orthographic			Opaque			Transparent		
Target	Related	Unrelated	Target	Related	Unrelated	Target	Related	Unrelated
Concrete Category								
ANT	anthem	shalom	ANGLE	angler	pallet	BAKE	baker	waves
APE	apex	opus	ARC	arcade	fibers	BALM[b]	balmy	fauna
BAND	bandit	aspect	ARCH[b]	archer	buckle	BANK	banker	shaved
BAR	barley	turret	ASS	assist	tarzan	BOIL	boiler	timber
BEE	beech	retch	BADGE	badger	tuning	BOMB	bomber	endure
BELL	bellow	stills	BARB	barber	relate	BONE[b]	boneless	exertion
BIG	bigot	churn	BEAK	beaker	humped	BOOR[b]	boorish	stetson
BILL	billiard	syllable	BLAZE	blazer	cavern	BOX[b]	boxer	asset
BOAR	board	judge	BOOT	booty	voted	BREW	brewer	buzzed
BOW	bowel	haunt	BRAND	brandy	breast	BRIDE	bridal	fusion
BROTH	brothel	markets	BUTT	butter	cared	COLOR	colorful	remotely
BROW	brown	taste	CENT	center	taught	CORD[b]	cordless	speckled
BUCK	bucket	maniac	CLAM	clamor	deaden	CUBE	cubic	ebony
BUFF[b]	buffalo	psychic	CLOVE[b]	clover	nicked	CURL	curly	pause
BULLET	bulletin	brunette	CORN	corner	passed	DIVE	diver	rover
BUSH	bushel	realms	COW	cower	blare	DOSE	dosage	naples
CARD	cardinal	athletic	CRATE[b]	crater	drapes	DRUM	drummer	stunned
CART[b]	carton	rubble	CROOK[b]	crooked	denial	DUST	dusty	ropes
CASH	cashew	captor	CRYPT[b]	cryptic	jeepers	EAR[b]	earless	sultans
CHAP[b]	chaplain	hysteria	DENT	dentist	profile	ERUPT[a]	eruption	headlong
DIAL	dialect	labeled	DRAW[b]	drawer	defeat	FARM	farmer	earned
DISC[b]	disco	debts	FIG	figment	optical	FAX	faxing	flocks
DISH	dishevel	aeronaut	FIN[b]	finish	asking	FELON	felony	resume
DOLL	dollop	pixels	FLEET[b]	fleeting	depended	FILTH	filthy	donald
EXAM	example	fortune	FLOUR	flourish	nautilus	FISH	fisher	rolled
FUSE[b]	fuselage	trillion	FORT	forty	drops	FLY	flying	client
GRAPH	graphite	longboat	FRUIT[b]	fruitless	mezzanine	GLIDE	glider	mailer
HALO	halogen	eschews	GARB[a]	garbage	members	GORE[a]	gory	spry
HARE	harem	pinto	GLOSS[b]	glossary	paroxysm	GREEN	greenish	scarcity
HELM[b]	helmet	horror	GRATE	grateful	audience	HARP	harpist	edified
HORN[b]	hornet	remake	INFANT	infantry	mystical	HAZE	hazy	awry
JACK	jacket	market	IRON	irony	cocky	HERB	herbal	hooves
MUTT[b]	mutton	pianos	LIMB[b]	limber	drifts	HYMN[b]	hymnal	svelte
PARK	parka	aglow	LIST[b]	listless	crannies	ICON	iconic	shelve
PEN	pendant	follies	LITER[a]	literal	saddled	INJURE[a]	injury	rhythm
PILL	pillow	retire	MAN[b]	many	town	LEAK	leakage	genteel
PLANK	plankton	recalled	MARSH	marshal	vehicle	LOCK	locker	skills
PULP	pulpit	elicit	MESS	message	contact	MARK	marker	spells

PUMP	pumpkin	stretch	MILL	million	mistake	MILE[b]	mileage	galahad
SCALP	scalpel	cologne	MIST	mister	stayed	MOLD[b]	molding	ditches
SCAN	scandal	spoiled	NUMB[a]	number	police	NUDE	nudist	alcove
SCAR	scarlet	bouquet	PAN	panic	drama	PAINT	painter	cleaner
SING	singular	knickers	PARCH[b]	parchment	dominance	PAVE	paving	spares
SLUM[b]	slumber	stashed	PIG	pigment	collate	PIPE[b]	piper	crack
SOCK	socket	trader	PLUM	plumage	midriff	POET	poetic	dalton
SPIN	spinach	matilda	POND	ponder	champs	POST	postal	merits
STAR	starch	ballot	PORT	portion	inspire	RAPE	rapist	filter
STEW	steward	workout	PUCK[b]	pucker	drowns	ROCK	rocky	trunk
STUB	stubborn	everyday	QUART[b]	quarter	secrets	ROLL	roller	classy
STY	stylus	bluish	RAMP	rampage	machete	RUN	runner	region
SURF[b]	surface	trusted	RUST[b]	rustic	exodus	SAIL	sailor	blonde
TAPE[b]	tapestry	interpol	SAND	sandal	arisen	SHIP[b]	shipment	platform
TEA	tease	alter	SEVER[a]	several	nowhere	SPINE	spinal	radius
TORN[b]	tornado	dynasty	TAIL	tailor	tavern	STORM	stormy	homing
VILLA	villain	proving	VAN	vanish	partly	TEAR[b]	tearful	departs
WALL	wallet	pushed	WAND	wander	bushes	TRAIL[b]	trailer	chapter
WIND[b]	window	broken	WHISK	whisky	heroic	TRIBE	tribal	sturdy
WREN[a]	wrench	sewing	WICK	wicked	washed	WASH	washable	illusive
YELL[a]	yellow	remain	WOOF[b]	woofer	blotch	WAX	waxing	reside
ZOO	zoom	thug	YANK	yankee	recipe	WOOL	wooly	abode

Abstract Category

ADVENT[b]	adventure	traveling	ACCESS	accessory	precision	ACT	actor	wheel
AGAIN	against	anybody	ACCORD[b]	accordance	discomfort	ADAPT	adaptive	passbook
ALPHA[b]	alphabet	foreplay	ALLEGE[b]	allegory	fiercest	ADOPT	adoptive	inherent
AND	android	vacancy	AMEN	amenable	gradient	AFTER	afterward	suitcases
ANTI	antic	emote	AUDIT[b]	auditory	janitors	ALLOW[b]	allowance	extension
AURA[b]	aural	bevel	AWE	awful	admit	AVOID	avoidance	mystified
BOO	booze	sober	BAN	banner	bumped	BOOST	booster	longing
BRISK	brisket	relapse	CANDID[b]	candidate	relations	CALM	calmer	thorny
BUT	butler	models	CASTE	caster	pails	CARE	careful	finally
CANT	canteen	negroes	CASUAL	casualty	eyeballs	CHEAT	cheater	statues
CHANCE[b]	chancellor	counseling	COMMIT[b]	committee	mountains	CREATE	creator	flooded
CHRONIC	chronicle	rearrange	COUNT[a]	county	loving	CURE	curable	ironies
CON	concur	measly	CUSTOM	customer	diamonds	DUTY	dutiful	osmosis
COST[b]	costume	whiskey	DECADE	decadence	grappling	EASY	easiest	jupiter
COUP[a]	coupon	corral	DEPART	department	expensive	EERIE	eerily	bigots
COY	coyote	mutant	DOLE[b]	doleful	pharynx	EGO	egoist	secede
CRAM	cramp	roles	EARL[a]	early	ought	ELATE	elation	crumpet
CURT	curtail	bureaus	EARN	earnest	rejects	ETHIC	ethical	goliath
DETER	determine	wandering	EMPIRE[a]	empirical	temptress	FIND	finder	vines
DIRE	direct	genius	EVEN	evening	captain	FLAW[b]	flawless	sicilian
DISS	dissolve	ruptured	EVENT[b]	eventual	proximal	FRISK[b]	frisky	swiped

EASE	easel	skulk	EVER[b]	every	being	GLEE	gleeful	dustbin
ELECT	electron	surfaces	FACT	faction	shaping	GLOOM	gloomy	inhale
EXTRA[b]	extradite	prospered	FACTOR	factory	bullets	GREAT	greater	session
FEAT[b]	feature	rescued	FIT	fitful	warmup	GUILT[b]	guilty	spirit
FORE	foreign	spanish	FOR	forward	mention	HAPPY[b]	happiest	wardrobe
FREE	freeze	valley	GALL[b]	gallant	cynical	HEED[b]	heedless	collages
HABIT[b]	habitat	cabaret	IMPORT	importance	legitimate	HOPE	hopeful	berries
HARM	harmony	genuine	INTER	interest	magazine	IDEAL	idealize	encircle
HELL	hello	away	INVENT[b]	inventory	attorneys	KEEP	keeper	carrot
INFER	inferno	strudel	LUST[b]	luster	sander	LAW[b]	lawful	admits
LACK	lackey	feeder	MAXIM[b]	maximal	granule	LOGIC	logical	colored
LATE	latent	nosing	MAY	mayor	issue	LUCK	lucky	clean
LESS	lesson	clever	MET	metal	walks	MANAGE	manager	battle
MAD[b]	madam	study	MISS	mission	keeping	MILD	milder	leaden
MAIN[b]	maintain	railroad	MUST	muster	divers	MIND	mindful	layover
MANIC	manicure	impostor	NAUGHT	naughty	massive	MOURN	mourner	resound
MID	midget	fuller	OFF	offer	magic	MYTH[b]	mythic	aegean
MODE[b]	modem	matey	PASS	passion	collect	NEAR	nearest	muscles
MULL	mullet	vowels	PAST	pastor	melted	NORM	normal	bottle
MUSE[b]	museum	hidden	POSIT[a]	positive	margaret	OPT[b]	option	sherry
PAR	pardon	record	PROPER	property	criminal	QUIET	quieter	tootsie
PLAIN	plaintiff	supported	PUN	punish	polish	REAL	realist	undying
PLEA[b]	pleat	hearer	QUEST[b]	question	problem	RISK	risky	rumor
SCAM	scamp	plaid	REND[b]	render	mailed	RUDE	rudeness	impaired
SEMI	seminal	barrack	SHIFT[b]	shiftless	toadstool	SENSE[b]	sensory	accents
SHALL[b]	shallow	resolve	SKEW[b]	skewer	weakly	SIN	sinful	baboon
SHUN	shunt	covet	SNEAK	sneaker	relates	SOON	sooner	speech
SINCE	sincere	montana	SPAN	spanner	lifters	START	starter	marries
SPEC[a]	spectrum	hijacked	SUM	summer	seeing	THINK	thinker	drought
STERN	sternum	mundane	SURGE	surgery	william	THRILL	thriller	canister
SUBPAR	subpart	salable	TACT[b]	tactic	jailer	USE	usage	malta
TOIL	toilet	maggie	TELL[b]	teller	defect	VALID	validity	reciting
TRUMP[b]	trumpet	risking	TEMPER[b]	temperate	antitrust	VIGIL[b]	vigilance	grounding
TWIT	twitch	sweets	TRAIT	traitor	comment	VOCAL	vocalist	agonized
WANT	wanton	mantel	VIA	viable	prague	WAIT[b]	waiter	wounds
WHIM	whimper	pansies	VIRTUE	virtual	sparrow	WAIVE	waiver	pranks
WILL	willow	posted	WAY[b]	wayward	gradual	WILD	wilder	insert
WON	wonder	heaven	WIT	witness	soldier	WORTH	worthy	scenes
ZEN[b]	zenith	urchin	WITH	wither	dodged	ZEST	zestful	diverge

[a] Targets excluded from analyses on the basis of overall error rates exceeding 50%.

[b] Targets excluded from analyses based on ratings of semantic relatedness.