

Revisiting orthographic similarity

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

Brandon C. Beltz
Master of Arts
George Mason University, 2004

Director: Christopher T. Kello, Professor
Psychology

Spring Semester 2009
George Mason University
Fairfax, VA

Copyright: 2009 Brandon C. Beltz
All Rights Reserved

ACKNOWLEDGEMENTS

I thank Dr. Chris Kello for all his support during my graduate school career. It has been an interesting and ultimately transforming journey these many years. I also thank Daragh Sibley for many discussions on orthography and his willingness to provide computer programming insights at critical moments during this project.

TABLE OF CONTENTS

	Page
List of Tables.....	v
List of Figures.....	vi
Abstract.....	vii
Chapter 1: Introduction to orthographic similarity and orthographic neighborhoods	1
Chapter 2: Exploring lexical structure through orthographic networks.....	18
Chapter 3: Validating the revised orthographic similarity metrics using the form-priming lexical decision task	42
Chapter 4 : General discussion	64
Appendix A: Words used in the orthographic networks	70
Appendix B: Stimuli pairs used in experiments 1 and 2.....	95
References.....	113

LIST OF TABLES

Table	Page
Table 1: Relative rankings of potential orthographic neighbors for the word ‘STOP’	14
Table 2: Correlations between the orthographic similarity metrics on their ratings of 2000 pairs of letter string.....	17
Table 3. Average percent of similarity among words (k in and k out links)	28
Table 4. Mean response times, standard deviations, standard error of the means, and % trial errors.....	48
Table 5: Correlations between word target reaction times and the orthographic similarity metrics	50
Table 6: Correlations between word target reaction times and the orthographic similarity metrics when factoring out the effects of word frequency	50
Table 7: Correlations between word target reaction times and the orthographic similarity metrics when factoring out the effects of target length	51
Table 8: Correlations between nonword target reaction times and the orthographic similarity metrics	51
Table 9: Correlations between nonword target reaction times and the orthographic similarity metrics when factoring out the effects of target length	52
Table 10: Correlations between word target reaction times, the original Ngram, the two revised Ngram, and the SOLAR metrics	54
Table 11: Correlations between nonword target reaction times, the original Ngram, the two revised Ngram, and the SOLAR metrics	54
Table 12. Mean response times, standard deviations, standard error of the means, and % trial errors.....	58
Table 13: Correlations between word target reaction times and the orthographic similarity metrics	58
Table 14: Correlations between word target reaction times and the orthographic similarity metrics when factoring out the effects of target length	59
Table 15: Correlations between nonword target reaction times and the orthographic similarity metrics	59
Table 16: Correlations between nonword target reaction times and the orthographic similarity metrics when factoring out the effects of target length	60

LIST OF FIGURES

Figure	Page
Figure 1. The spatial coding scheme of the SOLAR model for the words STOP, POST, and SOAP	11
Figure 2. Example of a network of semantic associations among a set of words.....	20
Figure 3: Orthographic neighbors for the word ANT using the orthographic substring metric.....	24
Figure 4: Depiction of in and out links in the orthographic networks.....	27
Figure 5: Link distributions of the Levenshtein Distance, Ngram, and SOLAR orthographic networks.....	29
Figure 6: Power law link distributions from orthographic and phonological networks in Kello and Beltz (in press).....	30
Figure 7. Link distributions of the orthographic substring networks.....	32
Figure 8: Link distributions of the dichotomous Levenshtein Distance, Ngram, and SOLAR directed-out networks.	37
Figure 9: Link distributions of the dichotomous Levenshtein Distance, Ngram, and SOLAR directed-in networks.	38
Figure 10: A typical trial in the form-priming lexical decision experiment.....	47

ABSTRACT

REVISITING ORTHOGRAPHIC SIMILARITY

Brandon C. Beltz, Ph.D.

George Mason University, 2009

Dissertation Director: Christopher T. Kello, Ph.D.

The goal of this dissertation was to better understand theoretical and empirical issues regarding how humans process the visual, orthographic forms of words. Central to this goal was a clarification of the concept of *orthographic similarity* which has been used by researchers for several decades for model building and theory testing across a variety of visual word processing contexts. Two general approaches were taken over the period of inquiry in examining orthographic similarity as it applies to visual word processing.

The first approach sought to expand models of wordform interaction in memory through the use of a graph theoretic, network-based approach. This approach sought to expand the conception of orthographic neighborhoods and explore how orthographic neighborhood structure compares to power law, scale free structures found in other language domains such as semantics and syntax. The results of the orthographic network analyses show that power law structure in orthographic word-forms arises in specific

circumstances in which the degree of letter overlap is very high and occurs in contiguous letter positions among words.

The second approach used a form-priming, lexical decision task to test the priming effects of orthographic similarity during word processing in order to validate the orthographic similarity metrics used to define orthographic networks in the first approach. The experimental results found facilitory orthographic priming effects for primes and word targets of different lengths which supports the claim that the concept of orthographic neighborhood should be revised to include words that differ in length. The facilitory priming effects in the experiments were limited to conditions with 67 ms prime durations in which orthographic similarity was measured by the SOLAR metric (Davis, 2006). The facilitory priming effects found in this study supports the “sloppy”, relative-letter position encoding of the visual system (Davis & Bowers, 2006) during the early stages of word recognition.

CHAPTER 1: INTRODUCTION TO ORTHOGRAPHIC SIMILARITY AND ORTHOGRAPHIC NEIGHBORHOODS

Current theories of visual lexical processing often must account for a multitude of orthographic variables found to influence word processing. One variable that has received a lot of attention in the previous three decades is the effects that *orthographic similarity* among words has on lexical processing. Orthographic similarity is typically defined by the number of shared letters and their respective positions within words. For example, the English words FAME, CAME, and SAME have a high degree of orthographic similarity because they share a large portion of letters in the same positions. In contrast, words such as FAME, JUNE, and TILT do not share letters and may be considered more orthographically distinct from one another.

The orthographic similarity (or lack of) among sets of words has been cast in terms of *orthographic neighborhoods*. In any given language some number of words may have large orthographic neighborhoods because they look very similar to other words in the language. The size of a word's orthographic neighborhood and its particular neighbors can occur for a variety of reasons such as the morphology of the language (e.g. RAN – RUN; SWIM – SWAM), homophones occurring in a language (e.g. SEA – SEE; FOWL – FOUL) or for unknown, seemingly arbitrary reasons (LOVE - MOVE). The fact remains that for the lexicon of a language, some words have a relatively large number of

orthographic neighbors compared to other words which have very few or no orthographic neighbors.

The importance of orthographic neighborhoods stems from the dynamics of how neighbors interact with one another when they are accessed in memory. Many people have likely experienced the effects of incorrectly identifying a word while reading, sometimes with minor amusement, as in the case of misreading “a *spark* erupted between the lovers” compared to “a *spank* erupted between the lovers”. Some of these word identification errors can be attributed to similar words being mistakenly accessed in memory instead of the correct one.

The use of orthographic neighborhoods in model development.

Understanding the effects of orthographic neighborhoods is critical to theories of visual lexical processing because they shed light on how lexical representations in memory are stored and accessed. According to two influential models of lexical processing, the serial search/verification model (Forster & Dickinson, 1976; Paap & Johansen, 1994) and the interactive activation model (McClelland & Rumelhart, 1981), large orthographic neighborhood sizes are expected to increase response times in lexical processing tasks. According to the serial search/verification model, this is due to words being represented in orthographically similar sets in memory. When a person encounters a specific orthographic stimulus in the environment, it must be compared to previously learned words in a primarily serial fashion in order to find a possible match. If that orthographic stimulus looks similar to a large number of other words stored in memory, it

should take more processing time to distinguish it from the others because the others must be either correctly matched to the stimulus or rejected from the stimulus one at a time. According to the interactive activation model, orthographically similar sets of words compete with one another for activation when a similar orthographic stimulus is encountered in the environment. In order for one of the words represented in memory to be correctly identified as the actual word being countered in the environment, it must receive additional activation above its noisy neighbors' activation. This requires additional visual processing of the orthographic stimulus at lower feature levels which takes more time. In both cases, the predictions from the serial search/verification model and the interactive activation model are that words with many orthographic neighbors should require slightly more time to process than words with fewer orthographic neighbors because of the inhibitory effect neighbors have on one another.

The straightforward prediction by these two models sparked a series of empirical studies which investigated whether orthographic neighbors had an inhibitory or facilitatory effect on processing. One of the most frequent word processing tasks used in these studies was the lexical decision task (Andrews, 1989; Andrews, 1997; Carreiras, Perea, & Grainger, 1997; Grainger & Jacobs, 1996; Mathey, Robert, & Zagar, 2004; Perea, 1998; Perea & Rosa, 2000). This task requires an individual to make a yes/no decision on whether or not a presented orthographic stimulus is one that they recognize as being a word in their language.

The empirical results from the lexical decision studies both confirmed and contradicted the models' predictions. In a few experiments (Balota, Cortese, Sergent-

Marshall, Spieler, & Yap, 2004; Coltheart et al., 1977; Johnson & Pugh, 1994) inhibitory effects in the lexical decision task were found for large neighborhood sizes which confirmed the models' predictions. However, contradictory findings which showed a facilitatory effect for large neighborhoods in lexical decision were found in many experiments (Andrews, 1989; Andrews, 1992; Carreiras et al., 1997; Grainger, 1990; Grainger & Jacobs, 1996; Grainger, O'Regan, Jacobs, & Segui, 1989, 1992; Johnson & Pugh, 1994; Sears, Hino, & Lupker, 1995). Reviews of the literature by Andrews (1997) and later Perea & Rosa (2000), attempted to clear up the seemingly contradictory findings across the experiments. They concluded that inhibitory effects associated with neighborhoods were due to neighbor frequency independent of neighborhood size. That is, words with higher frequency neighbors were processed slower than words with lower frequency neighbors. This finding is consistent with the proposed inhibitory effects caused by lexical competition, because higher frequency neighbors seem to be competing with lower frequency targets.

The empirical studies conducted using orthographic neighborhoods as an independent variable provide a nice test bed for general theories of lexical processing. Given this important theory-testing function, one would hope that conceptions of orthographic neighbors apply to all the words of a lexicon. However, this has not often been the case. One limitation that has been mentioned in the orthographic neighborhood literature (Andrews, 1997; Davis, 2006), but has not been fully explored, is the implications that the definition of "orthographic neighbor" has on models of lexical processing.

Issues surrounding the definition of 'orthographic neighbors'.

At a basic level, the orthographic similarity between two words can be determined by how many letters they share. However, counting the number of shared letters between words can get complicated when one considers the letters' positions. An early and most often used criterion for defining orthographic neighbors is the N metric. In this criterion, two words are orthographic neighbors if they share all letters in the same positions except for one. This similarity metric was adapted from an orthographic input coding scheme of memory in which letters were coded as specific positions or slots (Coltheart, 1977) within words. For example, the word 'jump' was represented in memory as 'J₁ U₂ M₃ P₄' where each position within a word had a particular 'slot' filled by a letter.

The N criterion of orthographic similarity has been criticized as being relatively strict because it does not account for robust facilitory processing effects found with words whose shared letters do not align in a slot-based, absolute position sense (Andrews, 1996). In fact, several studies have found that various "neighbor types" containing shared letters in different positions have measurable effects on each other's processing. These neighbor types include; single substitution neighbors (STOP – SHOP), double substitution neighbors (STOP – SHIP), transposition neighbors (SLAT – SALT), nonadjacent transposition neighbors (STOP – SPOT), neighbors once removed (STOP – SOAP), and reverse anagrams (STOP – POTS) (see Davis 2004, for a discussion).

In addition to the lack of sensitivity to letters which did not share slots, studies of orthographic neighborhoods which utilized the N metric focused their investigations on

comparisons between words of the same length (Andrews, 1989; Andrews, 1992; Grainger & Segui, 1990; Grainger & Jacobs, 1996; Sears, Hino, & Lupker, 1995). In many cases these studies used monosyllabic English words between 4-5 letters in length. The primary reason for this is that when words of different lengths are compared to one another, the issue of letter position alignment becomes paramount in determining which letters are shared between the words. For example, if the words 'JUMP' and 'JUMPS' are aligned along their initial letters they would be coded as $J_1 U_2 M_3 P_4$ and $J_1 U_2 M_3 P_4 S_5$ and would share four letters in common. However, if they were aligned from their last, rightmost letters, they would be coded as $J_4 U_3 M_2 P_1$ and $J_5 U_4 M_3 P_2 S_1$ and would share no letters.

The practice of restricting orthographic neighbors to words of the same length may have served a practical purpose to minimize the issues of alignment, however by restricting orthographic similarity to words of the same length, models based upon those empirical results are limited in the scope of the words they can address because they only consider certain subsets of equal length words at a time. Furthermore, lexical processing models which used slot based coding schemes may be disadvantaged in terms of orthographic neighborhoods because length-restricted similarity metrics tend to restrict neighborhood size as word length increases. Words that are greater in length than the average length of English words (i.e. > 7 letters) typically have no neighbors even though there are a large proportion of words in the lexicon within that length range. To illustrate, the word BASEBALL has no neighbors using the N metric. Yet, one can imagine that the words BASE, BALL, or BASKETBALL could be conceived as neighbors of

BASEBALL using a similarity metric free of length restrictions. Length-restricted similarity metrics also rule out a large class of potentially interesting linguistic structure tied to word morphology beyond compound words. For example, affixes such as the addition of the past tense 'ED' or possession ''S' to words change the length of words, but may not seriously change the orthographic similarity between two words. Orthographic similarity is not solely tied to compound words and affixes, but occurs in morphologically distinct words as well such as CENT, SCENT, ASCENT or BEE, BEET, BEETLE.

Several studies of lexical processing have found evidence that neighbors of differing lengths have measurable effects on one another's processing. De Moor & Brysbaert (2000) found that primes and targets which differ in length by one letter have similar effects as primes and targets of equal length in a lexical decision task for Dutch words. Schoonbaert & Grainger (2004) found that nonword letter strings formed by subtracting single letters from words (e.g. BALNCE, BALACE) serve to prime their word counterparts (e.g. BALANCE). Davis & Taft (2005) found that correct 'no' responses for nonwords in the lexical decision task were slower when the nonwords (e.g. SCOME) had real words that were deletion neighbors (e.g. COME) compared to nonwords that had no real word deletion neighbors. Davis also found that low frequency words with high frequency deletion neighbors were slower than their control word counterparts. Drews (2005) found that Dutch prime words (e.g. KELLER) which were addition neighbors to a target word (e.g. KELLE) inhibited responses in an orthographic form-priming, lexical decision task. The results of these studies provide strong support

that orthographic similarity goes beyond words of the same length and should be considered a factor in determining orthographic neighbors.

Therefore, a primary goal of the dissertation is to explore how several recent orthographic similarity metrics extend the notion of shared letter positions beyond a slot-based, position specific sense. Each of these similarity metrics will be discussed in some detail in the next section. In chapter 2, I will explain how these metrics were instrumental to understanding large scale neighborhood structures when they were used to define neighbor relations among a large set of words. In chapter 3, the empirical validity of these metrics were put to the test in a lexical processing task designed to distinguish what respective features, if any, could empirically ground these metrics as viable measures of orthographic similarity.

Introducing revised methods of measuring ‘orthographic neighbors’

The first metric that will be examined is the Levenshtein distance metric. This metric is based on an algorithm taken from domains of speech recognition, spell checking, and DNA analysis (Yap, 2007; Yarkoni, Balota, & Yap, in prep). The Levenshtein distance computes the similarity of words depending upon the number of basic operations it takes to transform one word into the other. These basic operations are a) letter deletions b) letter insertions or c) letter substitutions. For example, the distance between CASTE and CAT is 2 because 2 letters need to be deleted (‘S’ and ‘E’). By counting the operations between words, the Levenshtein distance introduces orthographic similarity values that can vary from 0 (the same word and thus no change operation) up to

the maximum length of a word in a given lexicon. The upper value of orthographic similarity for English words would be approximately 19 or 20 depending on the lexicon used as a reference. To illustrate, it would require 19 insertion operations to transform the word 'I' to 'institutionalization'.

The second metric is based on an alternative to slot based coding and called context-based coding in which letter positions are coded in a relative context. An early example is Wickelcoding in which the positions of letters are coded in the context of their surrounding letters and not an absolute position specific context (Wickelgren, 1969). For example, the units coded in the word JUMP are _JU, JUM, UMP, MP_ (the '_' representing the word beginning/ending). The important feature of context based coding is that it opens the door for neighbor comparisons of different length words. For example, the orthographic similarity between the words JUMP and JUMPS could be determined based upon the ratio of shared units (_JU, JUM, UMP) to the units that are not shared (MP_, MPS, PS_). The second metric of determining orthographic similarity that will be examined, called the Ngram metric, is adapted from Wickelcoding. The term "n-gram" refers to the contiguous letter strings that form words. For example, the word CAT is formed of the single letters C,A,T, the *bigrams* CA, AT, and the *trigram* CAT. Kello & Beltz (in press) used this metric as a basis to count two words of different lengths as orthographic neighbors if one word was contained within (i.e. was a substring) the other word. The major difference between the Ngram metric and Wickelcoding is that the Ngram metric has no preferred unit size within words. In other words, two words can share single letters, bigrams, trigrams, and any number of consecutive letters up to the

size of the words themselves. Wickelcoding, on the other hand, considers trigrams as the basic unit of words.

The third metric which will be examined is based on the SOLAR (self organizing lexical acquisition and recognition) model. SOLAR uses a sloppy spatial coding scheme to compare letter positions (Davis & Bowers, 2006; Davis & Lupker, 2006). The descriptor “sloppy spatial” coding refers to how the metric is sensitive to the relative positions of letters in strings, but not in the sense of an absolute, slot based method. Letters receive an activation based upon which order they appear in a word. Thus, the overall coding of the word is determined by the pattern of activation of letters across positions. Any similarities and differences in activation patterns are what determine the orthographic similarity of two words (for a detailed description, see Davis 2004). The three important aspects of the orthographic similarity calculation are that a) single letters are the basic unit of comparison b) shared relative positions boost similarity (e.g. X before Y before Z) and c) shared letters in the initial position receive more weight than other positions. Figure 1 below demonstrates the activation patterns of three words based upon their relative letter positions.

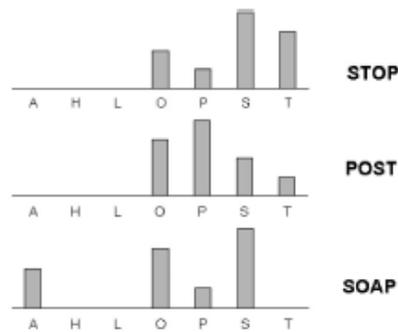


Figure 1: The spatial coding scheme of the SOLAR model for the words STOP, POST, and SOAP from Davis & Bowers (2006)

When comparing the words STOP, POST, and SOAP the words STOP and SOAP are more similar to each other than STOP and POST are. The reason is that although STOP and SOAP share only three letters, those shared letters occur in the same relative positions (i.e. S before O before P). In comparison, STOP and POST share all four letters, but they do not occur in the same relative positions and are thus calculated as slightly less similar than STOP and SOAP. The overlap of the S and P in the first and last positions respectively, plus the proximity of the O in STOP and SOAP factor heavily in how similar the words are.

The fourth metric which will be examined is based on a sequence encoder model of lexical processing (Sibley & Kello, 2004, 2006). The sequence encoder is a serial recurrent network model which uses fixed-width distributed representations to code variable length orthographic words. The model's coding scheme transforms an orthographic wordform into a pattern of conjunctions represented in a size restricted vector. Each pattern corresponds to the conjunction of one letter with its position within a wordform of a given length. For example, the word JUMP is coded as JUMP# and is

represented as five conjunctions: J in the first position, U in the second position, and so on until the # in the fifth position representing the end of the wordform. Over thousands of learning trials, the model attempts to maximize the statistical regularities of English wordforms (e.g. 'u' almost always follows 'q' in English words) in order to efficiently represent thousands of wordforms with minimal memory requirements.

The utility of the sequence encoder is the way it represents words as fixed-width vectors. Word similarity can be computed by comparing the Euclidean Distance between any two word representation vectors regardless of any particularities associated with the word the vectors represent. This number allows a relative comparison of orthographic similarity between all words the model has learned. Obviously, factors such as word length and letter position do play a role in the model's representations, but they do not provide hard constraints on which words can be neighbors because the similarity values are continuous and not dichotomous like slot-based metrics.

The four orthographic similarity metrics each have their unique combinations of orthographic factors which may weight more or less heavily in determining orthographic similarity. To make this comparison more clear, the metrics were unleashed on a set of calibration pairs to determine how each metric would rank some potential neighbors of the word STOP. The calibration pairs were designed to include all types of potential neighbors resulting from letter transpositions, letter additions, letter deletions, and shared letters across the beginning, middle, and end. The general calculation method of similarity among the pairs for the Levenshtein Distance, Ngram, and SOLAR metrics is shown in formula x, where OS is the orthographic similarity between two strings i and j ;

L represents the similarity calculation which varied among the metrics (i.e. Levenshtein distance operations, number of shared letters, etc); and l is the length or number of letters of a given word.

$$OS_{ij} = \frac{L_{ij}}{l_j} \quad (1)$$

The orthographic similarity calculation is symmetric for word pairs that are the same length. However, it is not symmetric for word pairs when the words differ in length. For example, the pair ‘CAT : CATS’ does not result in the same value as ‘CATS: CAT’. In the respective pairs, the length of ‘CAT’ (3) and ‘CATS’ (4) provide different denominators for the similarity calculations. The asymmetry in the orthographic similarity calculations reflect the asymmetry in determining which of the word entries serves as the basis of comparison. One potential relationship between the entries in the pair could represent a visual stimulus in the environment as the first entry and a lexical entry in memory as the second entry. The purpose of the pair would be to determine whether the visual stimulus matches the lexical entry in memory. Another possible relationship which is discussed in chapter 3 is the priming paradigm in which the first entry represents a visual prime stimulus and the second entry represents a visual target stimulus. The difference between the prime and target stimuli (first and second entry) in this relationship is how far apart in time the entries are presented to the visual system and whether or not the second entry is impacted in any way by the first. Regardless of how the relationship between the two entries of the pair are defined, for most theoretical

frameworks one entry will always be processed distinctly whether in temporal, spatial, or conceptual dimensions.

Table 1 below depicts how the orthographic metrics rank the set of calibration pairs from the most similar to the least similar when the word STOP was used as the comparison entry of respective pairs.

Table 1: Relative rankings of potential orthographic neighbors for the word ‘STOP’.

Levenshtein Distance		Ngram		SOLAR		Sequence Encoder	
1	stop	1	stop	1	stop	1	stop
2	stopx	1	st	2	sotp	2	sotp
2	stxop	1	sto	2	stxop	3	ptos
2	xstop	4	stopx	4	stopx	4	stox
5	stox	4	xstop	4	stopxx	5	sxop
5	sxop	4	sp	4	sxop	6	xtop
5	xtop	7	stox	7	stxxop	7	stopx
8	sto	7	xtop	8	sto	8	stxop
8	stopxx	9	ptos	8	stox	9	xstop
8	stxxop	10	stopxx	8	sxxp	10	sto
8	xxstop	10	xxstop	8	xstop	11	stxx
12	ptos	12	sotp	8	xxstop	12	sxxp
12	sotp	12	stxop	13	sp	13	xtox
12	stxx	12	sxop	14	st	14	stopxx
12	sxxp	15	stxx	15	stxx	15	xxstop
12	xtox	15	xtox	15	xtop	16	stxxop
17	sxxx	15	xxst	17	sxxx	17	xxst
18	sp	18	stxxop	18	ptos	18	sxxx
18	st	19	sxxp	19	xtox	19	xxxxs
18	xxst	20	sxxx	19	xxst	20	st
18	xxxxs	20	xxxxs	21	xxxxs	21	sp

Note: The numbers in the left of each column are the relative rankings of the strings. In some cases the same number appears multiple times indicating the metric does not distinguish among the strings of that number.

In each of the four orthographic similarity metrics, the identity string is the most similar (e.g. STOP-STOP). Starting from the second most similar string to the least, distinctions can be made among the metrics in terms of what letter features they are most sensitive to. The Levenshtein distance metric is relatively coarse grained in that the number of letter addition, deletion, or substitution operations does not distinguish where in the strings those operations take place. Therefore, it does not distinguish between addition, deletion, or transposition neighbors- they are all treated roughly the same as long as their number of operations are the same. The ngram metric is revealing in that the second and third most similar stimuli are shorter than the comparison word STOP and are substrings of it. Because the two strings have 100% letter overlap with the comparison word even though they are shorter in length, they are rated as more similar than strings which contain more shared letters, but have at least one or more non-shared letters. Additionally, because the ngram metric has no preferred unit size and calculates shared letters, shared bigrams, shared trigrams, etc. it is sensitive to contiguous clusters of letters that are shared. To highlight, the input stimuli XSTOP, STOPX, and STXOP all share four letters with the comparison word STOP. The visual input STXOP is rated less similar to STOP than the XSTOP and STOPX because it contains fewer shared bigrams and trigrams. In this example, the unshared X letter serves to break up the larger letter clusters in STXOP.

The SOLAR metric tended to rank the input stimuli that began with ‘S’ as more similar to STOP than those that began with ‘X’. This was not the case with the other three metrics. Furthermore, the SOLAR metric was sensitive to relative letter positions (e.g. S

before T and T before O) but did not require the letters be contiguous as was the case for the ngram metric.

The exact similarity ratings for any given pair calculated from the sequence encoder metric can vary from case to case because the sequence encoder organizes word representations in the context of thousands of other words in its learned lexicon. As was mentioned previously, the similarity between two words is calculated by measuring their Euclidean distance in an n-dimensional representation space. As a result of this calculation, it is rarely the case that two words will ever have the same distance value and thus very fine-grained distinctions can be made between words. This contrasts to the other similarity metrics which tend to cluster words into classes or types. In previous learned lexicons, certain trends have stood out in how the word representations are organized by the sequence encoder. Words are most similar when they share letters in the same position and are of similar lengths (Sibley & Kello, 2006). This is consistent with previous versions of slot based metrics. However, the key difference is that the sequence encoder is not tied to words of the same length and distinctions can be made among them. In comparison, the sequence encoder rated the various instances of four letter input strings which contained three shared letters with the comparison word (e.g. STOP-SXOP) as more similar than the ngram metric which was more sensitive to the contiguity of the shared letters.

Although the metrics are more or less sensitive to various features of orthographic similarity, they do overlap in their rankings to some degree. To further determine how consistently they rank potential neighbors, a correlational analysis was done on how

similar the metrics ranked over 2000 pairs of letter strings that were used in the empirical studies that will be described in detail later. For the current purposes, the letter strings were both real English words and nonwords that were semi-randomly paired with one another. The pairs could differ in length (or not) and could share from zero letters in the same positions up to all of their letters except for one in the same position. The results of these analyses are noted in table 2 below.

Table 2: Correlations between the orthographic similarity metrics on their ratings of 2000 pairs of letter string.

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Levenshtein Distance	1.000**	0.668**	0.812**	0.647**
Ngram		1.000**	0.652**	0.467**
SOLAR			1.000**	0.544**
Sequence Encoder				1.000**

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

Overall, the metrics are highly correlated meaning they tend to rate pairs relatively consistently. However, they are nowhere near perfectly correlated to one another, which leaves open the possibility that whatever orthographic similarity features they are more attuned to could be validated in how well they account for word processing performance in areas that orthographic similarity has been found to have a role.

In the next chapters, I will describe an analytical and empirical approach taken to flesh out the differences among these metrics as they relate to defining orthographic neighborhood structures and accounting for human variability in a visual word processing task.

CHAPTER 2: EXPLORING LEXICAL STRUCTURE THROUGH ORTHOGRAPHIC NETWORKS

This chapter will explore the role that the differences among the orthographic similarity metrics described in the previous chapter have on orthographic neighborhood structure. This will be done by conceptualizing orthographic neighborhoods on a larger-scale than has been done previously. On one hand, this approach is based on the length-restricted orthographic neighborhood studies which have represented neighborhoods as simple tallies of similar sets of words defined by an orthographic similarity criterion. This conception of orthographic neighborhoods as small lists of sets can be expanded by appealing to a growing body of literature that has used graph theoretic and network analysis methods to represent abstract, large-scale, neighborhood-like structures in various language contexts. The graph theoretic approach has used a combination of visual representations and statistical descriptions to describe how large sets of words relate to one another. In this paper, I will explore the statistical aspect of the graph theoretical approach and leave visual exploration to future work. To better understand the graph theoretic, network approach, a brief overview will be given of studies that have turned to using graph theoretic and network analyses techniques to represent relatively complex language phenomena.

Language networks

Several studies have examined the structural relationship of word meanings through the creation of word networks. These networks were formed by linking words together based upon their antonym and synonym relationships in a thesaurus (Holanda, Torres Pisa, Kinouchi, Souto Martinez, & Seron Ruiz, 2004; Kinouchi, Martinez, Lima, Lourenço, & Risau-Gusman, 2002; Motter, de Moura, Lai, & Dasgupta, 2002; Steyvers & Tenenbaum, 2005). For example, neighbors of the word COLOR were HUE, TINT, TINGE, DYE, COMPLEXION, SHADE, TINCTURE, CAST, LIVERY, COLORATION, GLOW, and FLUSH. These words were connected through abstract links in a network which represented their semantic relationships. Other approaches to word meanings have created networks by linking words based upon a cued association task when people were asked to name the first word that came to their mind after being given a cue (Steyvers & Tenenbaum, 2005). For example, when given the word DOCTOR, many people would say NURSE. Figure 2 gives an example of how one might eventually link VOLCANO to ACHE through word associations.



Figure 2: Example of a network of semantic associations among a set of words
From Steyvers & Tenenbaum (2005)

Networks have also been used to examine the syntactic relationships among words. Links in syntactic networks have been created when words occurred within some maximal distance of one another, such as linking words up to two positions away from one another in a sentence (Ferrer i Cancho, Sole, & Kohler, 2004; Sole, 2005; Sole, Murtra, Valverde, & Steels, 2006). For example, in the sentence

THE CAT SAT ON THE MAT.

The words CAT and SAT would be linked because they occurred next to each other. Also the words CAT and ON would be linked because they are within two word distances of each other. However, the words CAT and MAT would not be linked because they occur outside of the two word window. In both the semantic and syntactic networks, words were linked because they occurred together in some type of shared context (in the case of synonyms) or temporal window (in the cases of semantic association and syntactical word order).

One benefit of the network studies is that they used a large sample of words ranging on the order of thousands to study the phenomenon of interest. This is on a scale larger than has been done with orthographic neighborhood studies in the past. Additionally, these studies depict neighbors in the network by creating links between them. This is a very explicit step and provides a means to examine the entire network structure by looking at its links. The orthographic neighborhood literature has done this in a sense by listing the neighbors of the particular word which acts a way of identifying links. However, the neighbors were only considered in isolation within the context of single words and not as a collective network structure in its entirety. An extension of large-scale network techniques to orthographic structure seems fruitful when considering how orthographic similarity has been cast in terms of neighborhoods over the past several decades and neighborhoods can be seen as subsets of larger networks.

A major benefit of graph theory for discussing large-scale structure comes from the ability to describe and compare networks along certain measures. Simple measures include the size of the network such as the number of nodes and links. Individual nodes can vary in how well-connected they are to other nodes in the network. Word nodes with many links in an orthographic network would indicate a word that is highly similar to many other words in the lexicon. Word nodes with few or no links would indicate words that are orthographically unique in the lexicon. The number of links for a given node represents its “neighborhood size” as traditionally conceived in the orthographic neighborhood literature. As a means of comparison, the sizes of orthographic neighborhoods in earlier small scale studies using the N metric of similarity (i.e. words

differ only by one letter) ranged from zero to 24. The largest neighborhood size occurred for the words PAT and COT whose single letters could be substituted to form 24 other words (Davis, 2006). When lifting the restriction of neighbors to words of the same length, neighborhood sizes should increase drastically because word nodes can connect to a much wider range of words in the lexicon.

Network structure

Besides considering the sizes of networks in terms of their nodes and links, network approaches have sought to understand how network structures affect the phenomena they are meant to represent. For example, the information that travels across computer networks is heavily influenced by how the individual computers are linked to one another through the network structure. The application of abstract network structure to a lexicon's orthography makes sense when considering how the common theoretical mechanism of spreading activation (McClelland & Rumelhart, 1981; Forster, 2006) impacts word processing. Spreading activation refers to how mental representations that are associated (linked) in the mind can share or compete for mental processing resources if any of those representations are activated in memory. In reference to orthography, spreading activation covers the range of processes from low-level perception such as feature and letter decoding which ultimately (and very quickly) leads to word recognition. As discussed previously, this process can be affected by how the word representations in memory are structured and associated with one another. More specifically, the structure of orthographic neighborhoods could explicitly define and

allow tests for how similar sets of words may be similarly represented in memory. The appeal of understanding orthographic network structure is that it allows for a more explicit demonstration of how activation could spread for given words in a lexicon and could help verify the predictions of lexical processing theories which rely on the concept of spreading activation (McClelland & Rumelhart, 1981; Forster & Dickinson, 1976; Paap & Johansen, 1994).

Although not designed to examine how spreading activation may occur in orthographic representations, an exploratory study by Kello & Beltz (in press) sought to understand how orthographic networks could be created in a manner free of word length restrictions. In this study, a relatively simple dichotomous linking criterion was used to connect words together based on their orthography. Orthographic networks formed by tens of thousands of words and links were created for several world language lexicons using an orthographic substring metric. For every link in a network, there was a shorter length wordform connected to a longer length wordform. Two wordforms were connected if all the letters within the shorter word were contained in the longer word in their exact order. For example the word ANT in English is its own word but is also contained in words such as ANTE, ANTEATER, and ANTECEDENT. The word ANT would not be connected to ANATOMY because the letters A, N, and T are noncontiguous in the wordform. Figure 3 below gives an example of the neighbors of ANT.

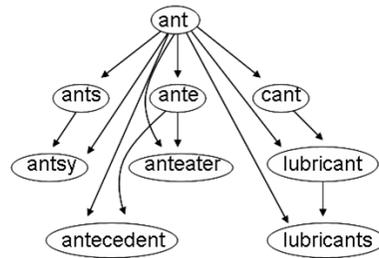


Figure 3: Orthographic neighbors for the word ANT using the orthographic substring metric.

The approach used in the Kello and Beltz (in press) study was inspired by the semantic and syntactic networks and was done as a means to compare the macrostructures of the networks to other language networks. One of the most interesting discoveries of the semantic, syntactic, and substring orthographic networks was that they all displayed a similar large-scale structure. This structure was also found in earlier studies of diverse phenomena such as the structure of the world wide web (Albert, 1999), the spread of disease in social populations (Pastor-Satorras & Vespignani, 2001), ecological food webs (Solé, 2001), and the structure of scientific collaboration networks (Newman, 2001).

In each of the these networks, the structures were all scale free ones, which was apparent in a *power law* distribution of their links. The power law distribution of links meant that there was a relatively small proportion of nodes which had a very large number of links relative to a large proportion of nodes which had a very small number of links. In other words, there were a few very dense neighborhoods relative to the majority of the network which had very small neighborhoods with medium-sized neighborhoods falling proportionately between the two extremes. The term “scale free” in describing

network structure means that the distribution of links proportionately “scales” as the number of nodes and links increase across a network. The result is that there is no typical mean or standard deviation value (i.e. measurement scale) which describes the number of links for all nodes.

It has been argued that the structural benefits of scale free networks in complex physically-based systems allows for an the improved speed of transfer of their basic units (e.g. electronic impulses, disease pathogens, food, or knowledge) (Albert, 1999; Newman, 2001; Pastor-Satorras & Vespignani, 2001; Solé, 2001). Additionally scale free networks have been examined for their life cycle properties such as their robustness against decay (Albert, 2000) and the relatively narrow range of processes which generate their structures (Barabasi, 1999). All of these benefits of scale free structure seem specific to physically instantiated phenomenon, but their relationship to oftentimes abstract, symbolic, and complex phenomenon like human language are less transparent. Before making any unwarranted generalizations of scale-free structural effects on language, it is worthwhile to explore the contexts in which scale-free structure is or is not found in language. Determining the extent and limitations of power law structure served as a primary impetus for a significant portion of the dissertation project and those endeavors will be described next.

Analysis 1: Expanding orthographic network.

The general orthographic network creation method described in Kello & Beltz (in press) was extended to accommodate networks based on the Levenshtein Distance,

Ngram, and SOLAR metrics. The sequence encoder metric was not included in this analysis due to computational limitations in creating networks.

English words were treated as nodes in a network with potential links among them determined by how orthographically similar they were. The words that were analyzed were taken from the corpus used in the semantic network study by Styvers and Tenenbaum (2005) which were adapted from Nelson, McEvoy, & Schreiber (1999) and are included in appendix A. There were 5019 English words ranging in length from 1 to 14 letters. Each of the words was paired with every other word which resulted in over 12.5 million pairs. The orthographic similarity of each pair was calculated using the Levenshtein distance, Ngram, and SOLAR metrics. For all three metrics, a standardized orthographic similarity (OS) value ranging between 0 (no shared letters) and 1 (perfect match) was possible for any given pair. The 12.5 million pairings of the 5019 words were analyzed with each word in the pair serving as both the comparison word and the base word (or denominator). In essence, this resulted in nearly 25 million unique pairs representing unidirectional links in the orthographic network. Each word had two orthographic similarity values for every other word in the network. One value was the ‘in’ directional link containing the value when the word was treated as the base, meaning its length served as the denominator for the orthographic similarity calculation described in the preceding chapter. The other value was the ‘out’ directional link containing the value when the word served as the comparison word and which the other word’s length served as the base. This relationship is shown in figure 4 below

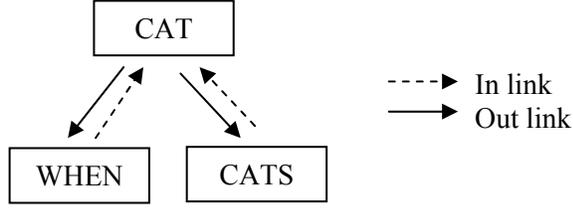


Figure 4: Depiction of in and out links in the orthographic networks
 Note: The end of the arrow indicates which word serves as the base for the similarity calculation of the pair.

Analysis 1: Results and discussion.

Of primary interest for the network structure analysis was an understanding of the distribution of orthographic similarity across the network. The first step was establishing a composite similarity score for each word based on its unidirectional in and out links. This was done by calculating the strength of a word's links k to other words in the network. This is shown in formula 2

$$k_i = \sum_1^n OS_{ij} \quad (2)$$

Where i is a given word; n is the number of words in the network; and OS is the orthographic similarity between the pair i and j .

Table 3 below shows the descriptive statistics of the composite link scores across all of the words for each of the orthographic similarity metric types. The values are depicted as an average percentage of similarity with 0 meaning no shared letters and 1 meaning an exact match to the other words in the network. The composite link score represents the sum of a given word's in or out links to all of the other words in the network. For the mean composite link scores, there were significant differences among

all the metrics with SOLAR rating words as the most similar overall, Ngram the second most, and Levenshtein Distance the least (N = 5019, $p \leq .001$ for all means comparisons).

Table 3. Average percent of similarity among words (k in and k out links)

<i>Link Direction</i>	Mean*		Std. Deviation*		Minimum*		Maximum*		Range*	
	In	Out	In	Out	In	Out	In	Out	In	Out
Lev Distance	0.07	0.07	0.04	0.04	0.00	0.00	0.18	0.17	0.18	0.17
Ngram	0.10	0.10	0.03	0.03	0.03	0.02	0.44	0.22	0.41	0.20
SOLAR	0.17	0.17	0.04	0.04	0.02	0.03	0.26	0.30	0.25	0.26

*N = 5019

Note: Values are % similarity with 0 representing no shared letters and 1 representing an exact match.

While the descriptive level statistics show a difference between the metrics at a broad level, further analyses were conducted to see how the links were distributed among the words. More specifically, how were words that had a high degree of similarity distributed relative to other words which had little similarity? The cumulative probability distribution functions (CPDF) of links k for all of the words was plotted to determine whether the link structure followed anything like a power law found in other language networks. These results are depicted in figure 5.

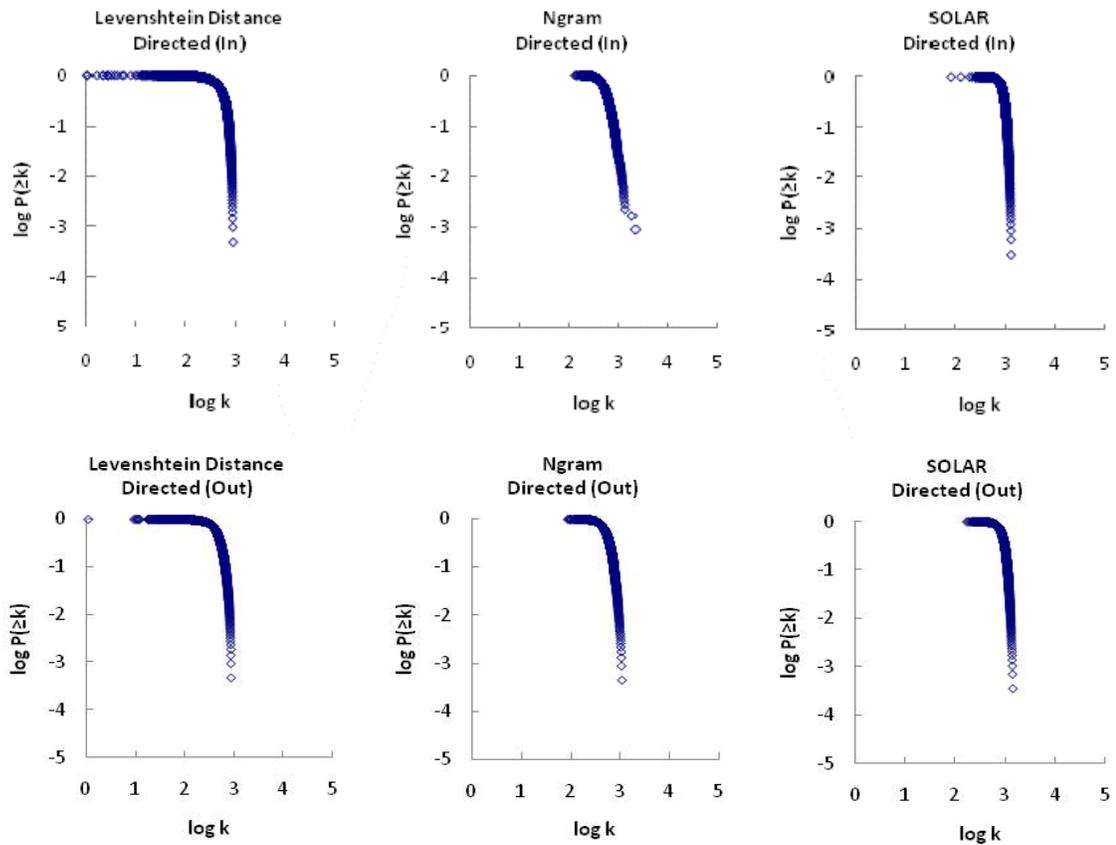


Figure 5: Link distributions of the Levenshtein Distance, Ngram, and SOLAR orthographic networks.

Note: k is the summed link strength a given word. $P(k)$ is the probability of k in the network. All values are expressed in \log_{10}

In Figure 5, the summed link strength of a particular word is represented on the x-axis with values closer to the minimum range towards the left and values closer to the maximum range towards the right. For the y-axis, the probability of link strength k in the network increases from the origin (-5) as one moves up the scale (towards 0). In each of the graphs, as a particularly link strength is reached starting from the minimum (left side) to the maximum (right side), the probability decreases by some amount. The shape and

increments of those decreases is what helps define the structure of the networks. As a point of illustration, the extremely rare, but "highly connected" in-link words such as I, ME, and EAT and the highly connected out-link words such as CLARINET, CREATION, and SENATOR are represented on the extreme right hand sides of their respective graphs. The "less well connected" words such as JUG, FUZZ, AWAY, and BANDANNA are more numerous and appear towards the left hand side of their respective graphs.

The results for all networks indicate that none of the link distributions produced a power law structure. If the links were distributed in a power law fashion there would be a clear linear relationship with slope $\gamma \approx -1$ in the graphs. This clear linear relationship is depicted in figure 6 below from the results from Kello and Beltz using the substring metric on a corpus of over 100,000 English words.

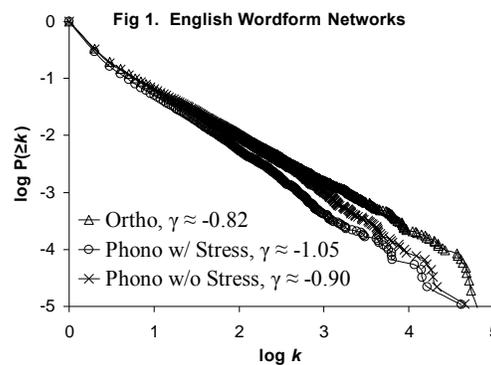


Figure 6: Power law link distributions from orthographic and phonological networks in Kello and Beltz (in press)

Note: k is the summed link strength a given word. $P(k)$ is the probability of k in the network. All values are expressed in \log_{10}

There are two main considerations for why the Levenshtein distance, Ngram, and SOLAR orthographic network link distributions differ from the substring orthographic networks. One is the nodes in the networks and the other is the links. The first possibility is the words that were used as nodes in the respective networks were very different. The substring network of Kello and Beltz used a corpus of over 100,000 words taken from a Carnegie Mellon University pronunciation dictionary (available at <http://www.speech.cs.cmu.edu>) that were verified against a Wall Street Journal corpus (Marcus, Santorini, & Marcinkiewicz, 1993) for correct spelling and usage. The current networks used 5019 words from semantic word association norms established by Nelson, et al (1999) and later adapted by Steyvers and Tennenbaum (2005) for semantic network creation. The second major consideration in the network differences is the linking method determined by the actual orthographic similarity calculations of the respective metrics.

To address these two considerations, a post hoc analysis was done on the 5019 words using the substring metric as the linking method in order to compare all four metrics on the same footing. The results of the substring networks are shown in figure 7 below.

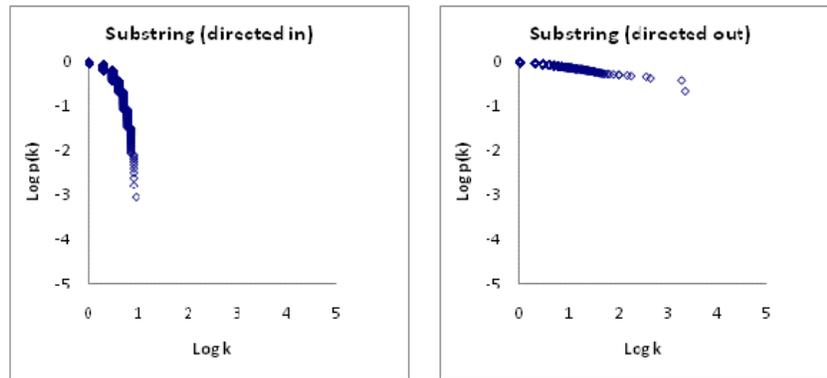


Figure 7: Link distributions of the orthographic substring networks.
 Note: k is the summed link strength a given word. $P(k)$ is the probability of k in the network. All values are expressed in \log_{10}

Neither of the substring networks show a clear power law distribution either. This result indicates that the particular words analyzed did play a significant role in the network structure. However, it does not explain the entire picture given that the substring directed-out network of 5019 words shows a similar, linear shape in the CPDF as the 100,000 word network, but with a much shallower slope, $\gamma = -.17$, than expected of a power law distribution.

One of the more obvious distinctions among all four metric-type networks is that the Ngram and Solar networks have a much more restricted range in k (between $\log 2$ and 3.3) than the range of the Levenshtein distance and substring networks (between $\log 0$ and 3.3). This means that the 5019 nodes in the Ngram and Solar networks are better connected to a larger degree; whereas, the Levenshtein distance and Substring networks have nodes which are more sparsely connected or not connected at all to other nodes. This, in part, can be attributed to a major distinction among the networks in how link values were defined. The Levenshtein distance, Ngram, and SOLAR networks all used

continuous link values ranging between 0 and 1 between nodes to indicate the relative similarity between the words. The substring networks used a dichotomous link value of 0 or 1 to indicate whether or not one word was contained within the other word.

The dichotomous link approach is consistent with how earlier orthographic similarity metrics work such as the slot based, N metric described in the introduction chapter. When the orthographic similarity ratings are defined as continuous as in the Ngram and SOLAR metrics, more "blurring" occurs between words such that the probability of finding words that look nothing like other words is reduced because there is a larger chance (although small) that any randomly selected pair of words will share at least one letter in any position especially as word length increases. For example, the original slot-based N metric of similarity which only considered neighbors if they shared all but one letter seriously restricts the number of neighbors possible for any one word. In comparison to the current results, the N metric's distribution would look similar to the substring directed-in distribution in terms of the range of k (between log 0 and 1). The Ngram and SOLAR continuous metrics have a similar distribution shape, but shift k to the right by a larger degree meaning the "baseline" connection among words is higher.

The larger range of k in the Levenshtein distance networks requires further explanation considering it is a continuous metric in the same vein as the Ngram and SOLAR metrics. The reason the Levenshtein distance networks exhibits a range of k starting from 0 (directed-in network) and from 1 (directed-out network) is due to how orthographic similarity is calculated based on the number of Levenshtein distance operations. In formula 1 for the Levenshtein distance calculation, the numerator consists

of how many Levenshtein operations it takes to change one word into the other, divided by the length of the base word. In all cases of word pairs that differ in length, when a shorter word serves as the base (denominator) and the number of operations required to change it into the longer length word exceeds the length of the shorter word, the orthographic similarity of the pair becomes zero. This is true even if the words share at least one letter. The ultimate result is the Levenshtein Distance metric is not as sensitive to shared single letters among pairs relative to the Ngram and SOLAR metrics; thus, an increase in the range of k on the left-hand side of its link distributions.

The distinct differences in distribution shape of the substring directed-in and substring directed-out networks requires explanation as well. The bottom line in the shape distinction is due to how word length factors into the binary yes/no criterion of whether one word of the pair is a substring of the other word. In the directed-in network, the number of links k of a given word is severely restricted because the length of a word determines how many possible substrings (and links) it may contain. In contrast, the number of links k for the directed-out network is not bounded by word length, but instead is bounded by the number of nodes in the network because any given word could potentially be a substring of any other word in the network. This asymmetry between in versus out links does not occur for the other orthographic metrics and thus their link distributions look similar. To highlight the asymmetry in the substring networks, the words A, I, IN, ON, ME, HE, AD, and BE all have a link strength $k > \log 2$ in the substring directed-out network because they, in fact, exist in a relatively large proportion of the other words. However, in the substring directed-in network, shorter words like

these are greatly restricted in the number of links they can have based upon the number of letters and substrings they can contain (which in some cases is only 1).

Analysis 2: Dichotomous link analysis

The distinctions in distribution structure between the continuous link networks based on the Levenshtein Distance, Ngram, and SOLAR metrics and the dichotomous, directed-in link network based on the substring metric was explored in an additional analysis. The continuous link values of the Levenshtein distance, Ngram, and SOLAR metrics were converted into dichotomous links to reflect the method of the substring networks.

The rationale behind this was to determine whether or not dichotomous link values could push the middle, average orthographic similarity values of the continuous link distributions towards the extreme ends and thus create a larger proportion of highly connected hubs and sparsely connected leaf nodes to more closely reflect a power law distribution in the networks. Additionally, this analysis served to address one confound in the earlier analysis. In the previous analysis, the k value was determined by the summed strength of each node's links to all other nodes. This value combined the number of links (degree) of the nodes as well as the strength of those links; therefore, there was no way to tell the difference between nodes with many low strength links from nodes with a few high-strength links because their k values would be similar.

The continuous orthographic similarity values for the 12.5 million word pairs based on the Levenshtein Distance, Ngram, and SOLAR metrics were converted to

dichotomous values (1 or 0) if those values exceeded a particular ‘neighbor threshold’ value. Three threshold values, low (> 1%), medium (> 33%), and high (> 66%) were used to determine whether a pair of words were connected based on their orthographic similarity ratings. For example if the words ONE and HONEY had the same hypothetical orthographic similarity value of 39% for the Levenshtein Distance, Ngram, and SOLAR metrics, that pair would be connected in the low and medium threshold networks, but not the high threshold network of the respective metrics. By introducing the three threshold levels to the dichotomous linking criterion, the issue of link strength was separated from node degree.

Analysis 2: Results and discussion

The results of the low, medium, and high threshold networks are depicted in figures 8 and 9 below.

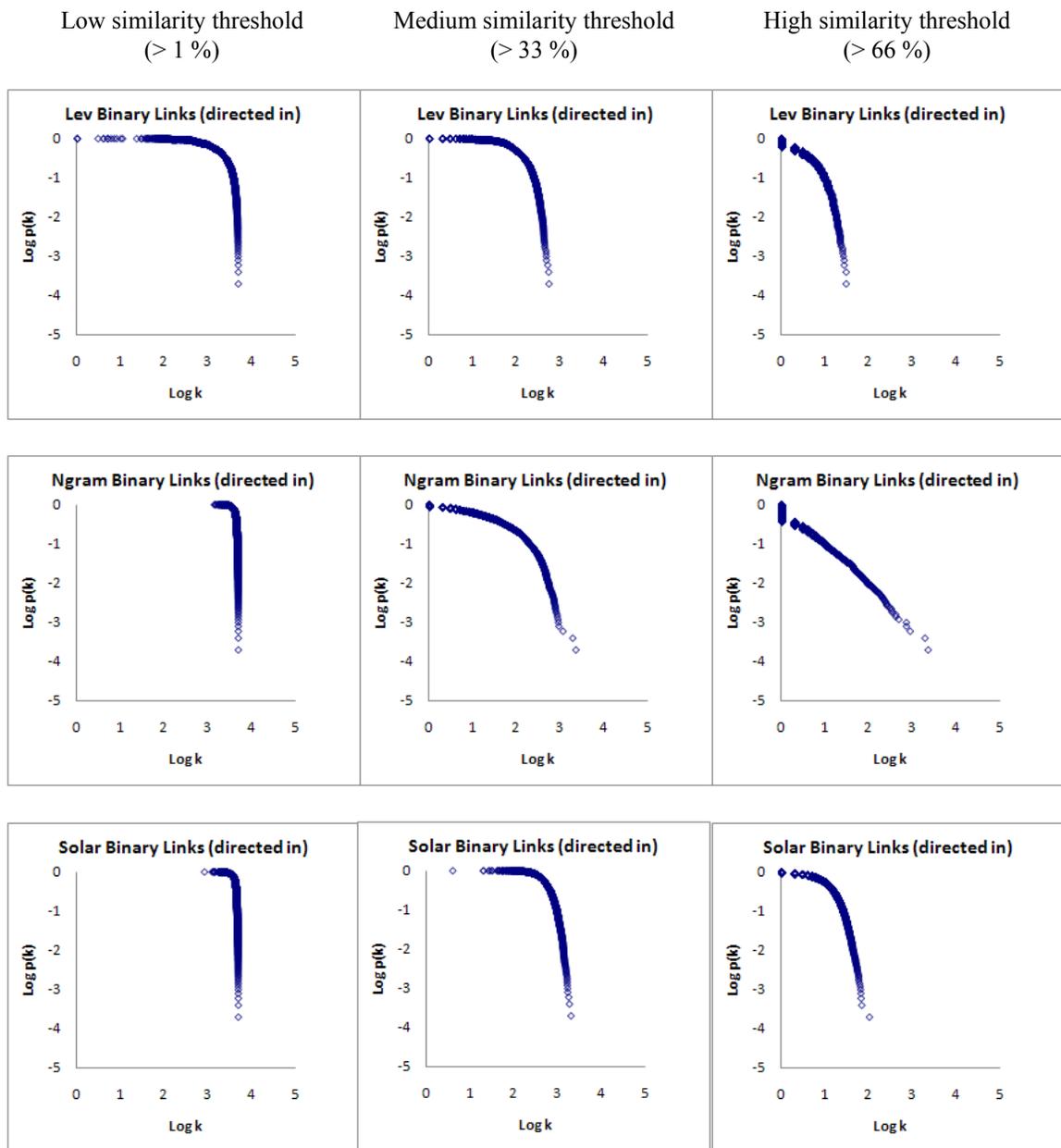


Figure 8. Link distributions of the dichotomous Levenshtein Distance (top row), Ngram (middle row), and SOLAR (bottom row) directed-in networks. The three binary threshold values for each network type are low (left column), middle (middle column), and high (right column).

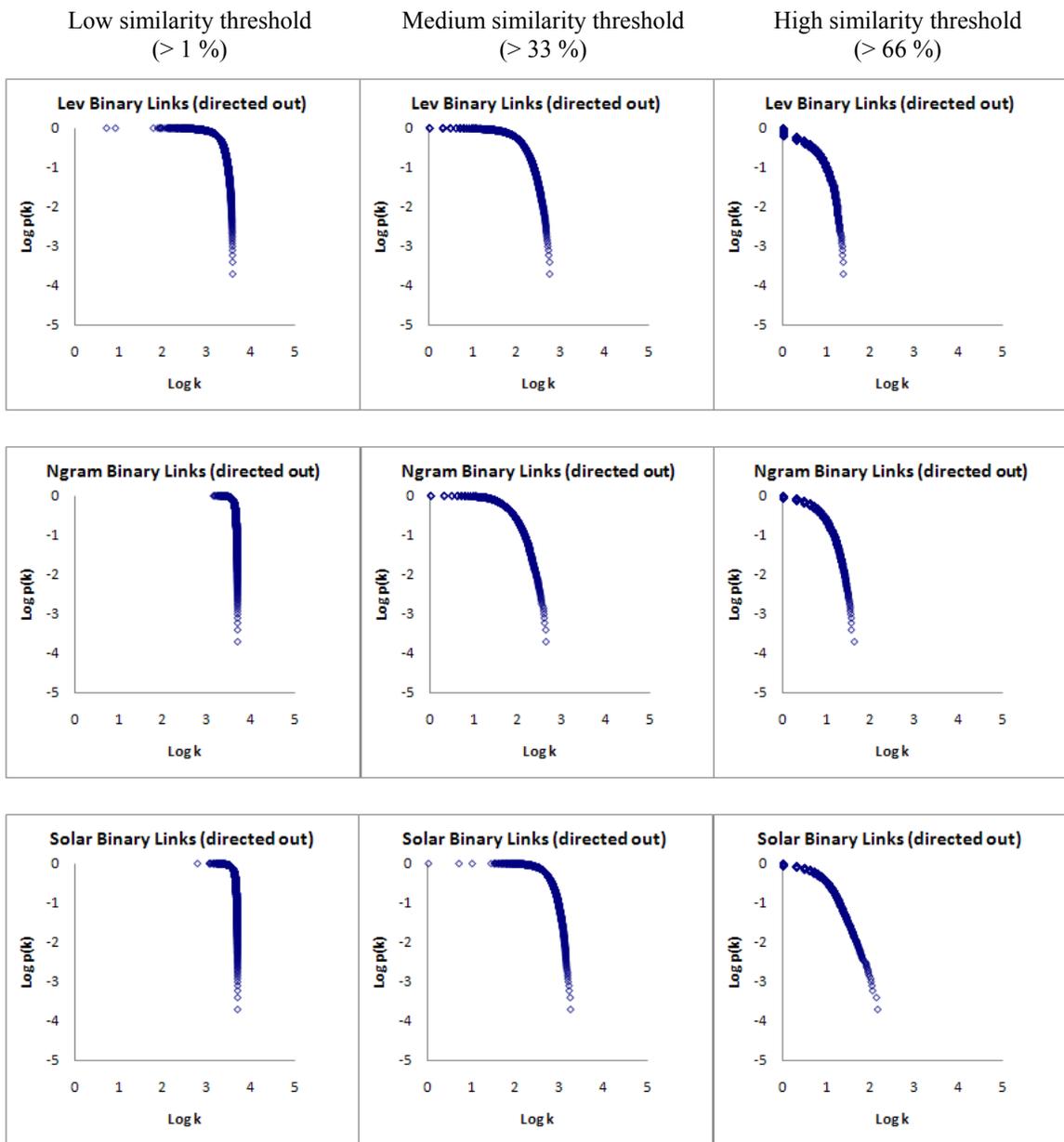


Figure 9. Link distributions of the dichotomous Levenshtein Distance (top row), Ngram (middle row), and SOLAR (bottom row) directed-out networks. The three binary threshold values for each network type are low (left column), middle (middle column), and high (right column).

The most striking finding of the dichotomous link analysis is that the Ngram directed-in networks formed a power law like distribution with increasing threshold values. At the highest threshold value, the Ngram directed-in CPDF showed a linear trend with $\gamma \approx -.88$. All of the other network types showed strong curvature in the CPDFs which is more indicative of log-normal distributions. One, not-so-surprising finding is the power law like distribution of the high threshold Ngram network highlights the similarities between the Ngram metric itself and the substring metric. Namely, that the substring metric is a subset of a more general Ngram approach. When the dichotomous threshold value of the Ngram metric was increased, it acted similarly to the substring metric in that only word pairs which shared a large proportion of letters in the same contiguous positions were considered as neighbors. The only difference between the substring metric and the high threshold Ngram metric is that the high threshold Ngram is slightly less restrictive in this percent of overlap of contiguous letters. The substring metric is essentially 100% contiguous letter overlap; whereas, the high threshold Ngram is $> 66\%$ contiguous letter overlap. These results may be an early indicator that at least 66% overlap of contiguous letters may be a prerequisite for power law structure in dichotomous orthographic networks.

The lack of power law structure in the dichotomous Levenshtein distance and SOLAR networks seems to support the notion that the noncontiguous position matching flexibility of these metrics creates situations in which links among words are still too frequent and too strong around an average similarity value than is required of a power law distribution. Despite the fact that the dichotomous threshold values in these networks

attempted to push the link distributions towards all or nothing type connections, there was still some minimal amount of average similarity due to shared single letters across noncontiguous positions which tended to connect word pairs more often than not and as a consequence, there were not enough less-connected leaf nodes in the networks in proportion to the hubs.

Orthographic networks: general discussion

A theoretical consideration is what exactly the directed nature of links (out or in) means for these networks. The determining factor of the link direction between two words for all orthographic similarity calculations was which of the words served as the base word and which was the comparison word. This distinction is important when considering orthographic recognition processes in the context of orthographic neighbors. The entry-opening model and interactive activation model both predict that words with large orthographic neighborhoods are processed more quickly in lexical processing tasks. The neighborhood size of a particular word would be determined by the sum of its in-link values because that value treats the word as the base word. In a word processing task, such as lexical decision, a word (the base) is physically presented to a person to respond to. Typically, the neighbors of the word are not presented (except in the priming paradigm) and any influence from them is induced from the response to the base word. Any processing influences from neighbors are hypothetical due to factors in how the words are represented and accessed in the brain. The orthographic form-priming experimental paradigm serves as a valuable tool for exploring how neighbors as

comparison words might influence the processing of the base word and that relationship will be better defined and tested in the next chapter.

Overall, the results of the network structure analysis highlights how seemingly straightforward distinctions in how orthographic similarity is measured impacts the larger neighborhood relationships among words in the lexicon. Given the distinctions in large-scale neighborhood structure among the metrics, the question remains which of them actually reflects orthographic similarity as humans perceive it and can better account for empirical data in orthographic processing tasks? The answer to this question will shed light on how and why large-scale neighborhood structure matters for modeling the mental lexicon. This question will be explored in the next chapter.

CHAPTER 3: VALIDATING THE REVISED ORTHOGRAPHIC SIMILARITY METRICS USING THE FORM-PRIMING, LEXICAL DECISION TASK

As mentioned in the introductory chapter, the empirical effects of orthographic similarity and neighborhoods for model development and theory testing have mostly been used in laboratory tasks such as word naming and lexical decision. A similar approach was used in the current study to help ground the revised orthographic similarity metrics in an already well-established empirical context. A masked form-priming, lexical decision methodology was used because it avoids some of the difficulties associated with the standard speeded lexical decision task (LDT). A major criticism of the LDT is that the results often depend upon the item sampling of word and nonword stimuli (Balota et al., 2004; Grainger, 1990). The fierce debates that surround studies of lexical decision often hinge on how well experimenters controlled for factors across the nonword and word stimuli such as word frequency, bigram frequency, and phonological statistics and less about the processes involved with the task.

In recent years, some experimenters have turned fine-tuned masked priming techniques with lexical decision to minimize the issue of controlling for often times complex confounds. The advantage of masked priming for exploring orthographic neighborhoods is that it allows experimenters to hold nonword and word targets constant and instead vary the type of primes which may or may not affect the processing of the

target (Forster, 1998). The priming paradigm helps illuminate the very early processes that occur during visual word processing in terms of memory access and recognition. This is especially true when using primes of very short duration (30-200ms) which serve to focus the measurement on early visual recognition processes by minimizing processes such as semantic based decisions that may occur after a prime is visible and consciously processed (Forster, 1998; Forster, et al. 2003). The intention of using orthographic form-priming is to determine the extent to which a target stimulus's orthographic neighbors have an inhibitory or facilitatory effect on processing in an immediate temporal window when the target is accessed in memory. The nature of these priming effects can validate what factors define orthographic neighbors and their implications for how lexical representations are processed in memory.

Many of the studies that have examined the effects of orthographic neighborhoods on word processing have used an extreme groups design which selected relatively small sets of stimuli based on a particular neighbor type (e.g. single letter substitution vs no shared letter control strings). The purpose of these studies was to methodically sort out which neighbor types have a greater or lesser effect on processing in a relatively constrained context. This issue of control becomes a double edged sword because blocking stimuli into sets based on carefully controlled orthographic features (including sets of the same length) can lose a degree of generalizability when the small sets of stimuli are considered in the larger picture of an adult speaker's lexicon which contains a vast array of words of different types which co-occur in constantly changing contexts.

The experiments that will shortly be discussed were designed to address some of these issues to a small degree. Stimuli were selected to encompass all types of orthographic neighbors and were included in one generic block that did not distinguish among neighbor types. The heavy lifting of distinguishing among fine-tuned letter position matches was left to the Levenshtein distance, Ngram, SOLAR, and sequence encoder orthographic similarity metrics, which were attuned to various position matching schemes and which was ultimately summarized in their orthographic similarity rating of a given stimuli pair. Furthermore, the number of stimuli given to each participant was larger than in previous studies. This ensured that a large number of stimuli in the generic block were seen by a wide variety of people. An overall goal of the study was to determine whether similar effects of orthographic form-priming could be obtained as in previous studies and in a manner that attempted to more closely reflect word stimuli people normally encounter.

Experiment 1

Method

Participants. Twenty five participants were recruited from the George Mason University campus. Participants were either undergraduate psychology majors who participated in the experiment for course credit or persons who responded to advertisements posted around the campus. Those who responded to the advertisements were paid \$15 for their participation. The average age of participants was 22.4 years. Fifteen participants were female and 10 were male. All participants reported that they were native English speakers.

Stimuli and design. 2400 total stimuli pairs containing a prime string and a target string were created for an orthographic form-priming lexical decision task. All of the 2400 primes were nonwords. Nonword primes were used to be consistent with the studies of Davis & Bowers (2006) and Forster, et al. (2003) and reduce complications of prime-target processing due to the prime lexicality effect in which word primes compete with word targets (Forster, 1998). 1200 of the targets were legal English words and 1200 targets were nonword strings. The English word targets were selected from a larger pool of 2500 English words. These words ranged in length from four to eight letters. They had a mean printed frequency of 99 words per million with a range of 33 to 428 words per million (Marcus, Santorini, & Marcinkiewicz, 1993). They consisted of nouns, verbs, adverbs, and adjectives. Both the nonword primes and targets were selected from a pool of 5000 nonwords. The nonwords were created by a computational algorithm which took the entire list of 2500 English words as input and varied the letters to produce nonwords which matched the word input list on letter frequencies and letter length distribution across items. This resulted in nonwords varying from four to eight letters like the words. The 5000 nonwords were selected so they were pronounceable and did not repeat in the nonword pool. For the nonword prime/ word target pairs, the 5000 nonwords were fully paired with the 2500 words which created 12.5 million pairs. The orthographic similarity of each pair using the four metrics discussed previously was calculated. The same procedure was used to create the nonword prime/ nonword target pairs. Each of the 5000 nonwords was paired with all the other nonwords to create a list of 25 million nonwords prime/nonwords target pairs. The orthographic similarity was then computed for each of

those pairs. 1200 pairs were then selected from each of the 12.5 million nonwords prime/word target pairs and 25 million nonwords prime/nonword target pairs respectively. Pairs were selected in a stratified manner to sample the entire range of orthographic similarity values as computed by the orthographic similarity metrics. An additional constraint was that nonwords in each pair were unique and never repeated in any other pair either as a prime or target. Legal pronounceable nonwords were used in order to correspond more to the decision processes participants go through when encountering real English words during the yes/ no lexical decision portion of the task. Illegal nonwords or those which contain illegal letter combinations are responded to much quicker solely due to their orthographic properties less on their lexical properties (Forster, 2003).

Procedure. Each participant was given 1000 trials of a form-priming lexical decision task. The participant's task was to determine whether an indicated target stimulus was an English word or not. Half of the trials consisted of English word targets and half consisted of nonword targets. Each trial consisted of three stimuli. A forward mask of 8 hash marks (#####) appeared in 16 point size Courier New font which remained on the screen for 500 ms. This was followed by a nonword prime in all lower case letters using 12 point size Courier New font. The prime remained on the screen for 67 ms. Finally, the target stimuli appeared on the screen in all upper case letters in 16 point size Courier New font. The target remained on the screen until the participant pressed either the (/) or (Z) key on the keyboard to indicate whether the target was "yes" an English word or "no" it was not a word, respectively. The difference in font size and

case between the prime and target was to ensure that the two stimuli were physically distinct. In cases where the font and case are the same, stimuli pairs such as GOAL and FOAL may be treated the same by the perceptual system (Forster, 2003). The intention of the current study was to ensure the prime and target had the best chance of being processed distinctly by the visual system in order to detect priming in the mental representations of orthographic word-forms and not purely on visual feature overlap alone.

On average, participants completed the set of 1000 trials in 60 minutes. They were allowed a short break in the middle of the experiment. The sequence of an experimental trial is depicted in figure 10 below.

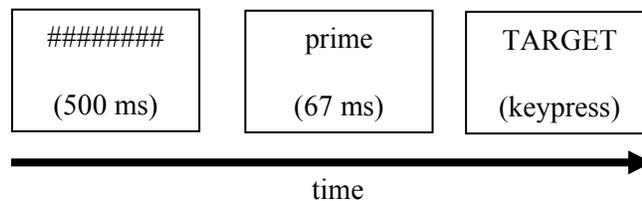


Figure 10: A typical trial in the form-priming lexical decision experiment.
Note: The time duration (in milliseconds) or the subject's response is listed in parentheses below each event.

At the beginning of the experiment, each participant was assigned to one of two testing rooms which contained a personal computer running the experimental software. Each computer used the Microsoft Windows XP operating system. The experimental task was designed and implemented using the DirectRT v2006 software published by Empirisoft (<http://www.empirisoft.com>). The computer keyboards were DirectIN

keyboards, manufactured by Empirisoft. The keyboards were designed to reduce timing errors from button presses to less than 1 millisecond. Each testing computer was attached to a 27 inch cathode ray tube (CRT) monitor with a refresh rate of 60 hertz. The monitors' resolutions were set to 1020 x 768 pixels and were a viewing distance of 30 inches from the participants' chair.

Results

Data trimming. Error trials and those in which reaction times exceeded the reaction time criteria described next were excluded from analysis. The reaction time criteria first removed trials which were faster than 300 milliseconds or exceeded + 4 standard deviations of the mean for all participants. For word targets, the trim value across participants was 4452 ms and for nonword targets the trim value was 5700 ms. A second pass trim was then conducted within each participant's trials that removed the remaining trials which exceeded + 4 standard deviations of the within subject's mean. On average, 2.2% of correct word target trials and 2.4% of correct nonword target trials were removed due to reaction time trimming. This resulted in 11811 word target and 10836 nonword target trials which were analyzed across all 25 participants. The mean response times and percentage of errors of these trials are presented in table 4 below.

Table 4. Mean response times, standard deviations, standard error of the means, and % trial errors.

	Word Targets	Nonword Targets
Mean RT (ms) [‡]	837	1219

Std Dev [‡]	174	335
Std Error [‡]	5.0	9.6
% Errors	3.3	11.1

[‡] Calculated from correct and RT trimmed trials

The effects of orthographic similarity on word target reaction times. For each of the 1200 prime word-target pairs, the mean reaction time was calculated across the participants who responded to a given pair. The mean response time per pair was then correlated with the orthographic similarity metrics. These results are listed in table 5 below.

Table 5: Correlations between word target reaction times and the orthographic similarity metrics

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time	-0.020	-0.135**	-0.082**	0.026

* p < .05, ** p < .01

The ngram and SOLAR orthographic metrics were the only ones which were significantly correlated with reaction times. These results are initially promising because they indicate the metrics could be in line with accounting for word processing performance. However, it is not clear the extent to which they indicate exactly how orthographic similarity impacts word recognition processes because there are other variables which may have a confounding role on reaction times.

The effects of word frequency on word target reaction times. The effects of word frequency are well established in the word recognition literature (Kinsbourne, 1974; Glanzer, 1976). Although the effects of frequency are not a primary focus of the current research questions, their presence (or absence) in the current data would help validate the current experiment. In other words, if similar frequency effects were found in the current data as found in many previous studies, it would lend confidence that the stimuli, measures, and procedures used in the current experiment were viable. The correlation between word frequency ($M = 99.24$, $SD = 65.88$, $N = 1200$) and reaction time ($M = 837.75$, $SD = 174.34$, $N = 1200$) was significantly correlated at $r(1198) = -.14$, $p < .001$. This is consistent with many word recognition studies which find that higher frequency words are responded to more quickly than lower frequency words (Glanzer & Bowles, 1976; Scarborough, Cortese, & Scarborough, 1977; Guttentag & Carroll, 1994). Given the significant correlation between word frequency and reaction time, the question remains whether or not word frequency effects were much stronger than any affects orthographic similarity may have had on reaction times. The variance associated with word frequency was regressed out of the reaction time data and correlated with the orthographic similarity metrics. The results are shown in table 6 below.

Table 6: Correlations between word target reaction times and the orthographic similarity metrics when factoring out the effects of word frequency

Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
-------------------------	-------	-------	---------------------

Reaction Time†	-0.020	-0.131**	-0.080**	0.027
----------------	--------	----------	----------	-------

* p < .05, ** p < .01
† Reaction times with the variance associated with word frequency removed

After regressing out the variance associated with word frequency from reaction times, the ngram and SOLAR metrics were still significantly correlated with reaction times almost to the same degree they were previously. This indicates that any effects of orthographic similarity were not greatly impacted by word frequency.

The effects of length on word target reaction times. The correlation between target length (M = 5.98, SD = 1.4, N = 1200) and reaction time (M = 837.75, SD = 174.34, N = 1200) was significantly correlated at $r(1198) = .32, p < .001$. This is not too surprising considering that one should expect that longer length word targets should take longer for participants to process and respond to. The central question is whether or not the effects of target length override any orthographic priming effects that the prime would have on the target. To examine this question, the variance associated with target length was regressed out of reaction time and then correlated with the orthographic similarity metric ratings. The results are reported in table 7 below.

Table 7: Correlations between word target reaction times and the orthographic similarity metrics when factoring out the effects of target length

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time‡	-0.031	-0.031	-0.062*	-0.011

* p < .05, ** p < .01

‡Reaction times with the variance associated with target length removed

After removing the variance accounted for by target length, the only orthographic similarity metric which was significantly correlated with reaction times was the SOLAR metric. The ngram metric became nonsignificant after the effects of target length were removed.

The effects of orthographic similarity on nonword target reaction times.

Table 8: Correlations between nonword target reaction times and the orthographic similarity metrics

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time	.054	-0.150**	-0.017	0.050

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

The ngram metric was the only one which was statistically significantly correlated with reaction times for nonword targets.

The effects of target length on nonword target reaction times. The correlation between target length ($M = 5.97$, $SD = 1.4$, $N = 1200$) and reaction time ($M = 1219.85$, $SD = 335.18$, $N = 1200$) was significantly correlated at $r(1198) = .40$, $p < .001$. This is the same trend as was found with word targets. The difference being that reaction times to nonword targets were more correlated with length than word targets. The variance

associated with target length was regressed from the reaction times and correlated with the orthographic similarity metric ratings. The results are reported in table 9.

Table 9: Correlations between nonword target reaction times and the orthographic similarity metrics when factoring out the effects of target length

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time‡	-0.004	-0.025	-0.008	0.027

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

‡Reaction times with the variance associated with target length removed

When the effects of target length are factored out of the reaction time variance, the correlations between the metrics and reaction times became nonsignificant. This is similar to the pattern seen with the word targets with the exception that the SOLAR metric was not significantly correlated with nonword reaction times before and after considering nonword target length.

Comparing the factors of the SOLAR metric to the ngram metric. The pattern of results so far regarding the distinction in the statistical significance of the SOLAR metric but not the ngram metric after factoring out the variance associated with word target length begs the question, which of the differences between the metrics helps account for the small correlation in reaction time data? As was described earlier, the main distinctions are that SOLAR a) uses single letters as the unit of analysis b) accounts for relative, noncontiguous positions and c) is sensitive to shared initial letters. To further explore these distinctions, the ngram metric was revised to address two of the main distinctions it

has from SOLAR (a and c). One revision of the ngram metric was created to introduce a weighting scheme that considered pairs of strings as more similar if they shared initial letters relative to a pair that did not (e.g. CATS was more similar to COLD than PARK). A second revision was created which included the first revision plus it considered only single letters as units of analysis and ignored shared ngrams larger than single letters. These two revisions were designed to blend some of the features of the ngram metric while incorporating additional ones from SOLAR. The revised ngram metrics were then correlated with reaction times and the results are reported in table 10 below.

Table 10: Correlations between word target reaction times, the original Ngram, the two revised Ngram, and the SOLAR metrics

	Ngram	Ngram ILW [#]	Ngram ILW-SL ^{##}	SOLAR
Reaction Time	-0.135**	-0.144**	-0.084**	-0.082**
Reaction Time [‡]	-0.031	-0.033	-0.037	-0.062*

* p < .05, ** p < .01

[‡] Reaction times with the variance associated with target length removed

[#] Ngram - initial letter weighting

^{##} Ngram - initial letter weighting, single letter unit of analysis

Table 11: Correlations between nonword target reaction times, the original Ngram, the two revised Ngram, and the SOLAR metrics

	Ngram	Ngram ILW [#]	Ngram ILW-SL ^{##}	SOLAR
Reaction Time	-0.150**	-0.155**	-0.091**	-0.017
Reaction Time [‡]	-0.025	-0.019	-0.027	-0.008

* p < .05, ** p < .01

- ‡ Reaction times with the variance associated with target length removed
- # Ngram - initial letter weighting
- ## Ngram - initial letter weighting, single letter unit of analysis

The revised ngram metric which incorporated an initial letter weighting scheme (Ngram-ILW) had a slightly higher correlation for both word and nonword targets than its original counterpart. On the other hand, the revised ngram metric which incorporated an initial letter weighting scheme and considered only single letter units had a significant drop in correlation for both word and nonword target reaction times similar to the levels of the SOLAR metric. When target length was factored out of reaction times, the correlations for both the revised ngram metrics became nonsignificant which is consistent with the original version. Given the characteristics of the revised ngram metrics and the SOLAR metrics, the only distinction between them is the fact that SOLAR allows for relative position coding that is noncontiguous. For example, SOLAR is sensitive to the relative, noncontiguous letter positions in the pair ABCD and AXBXCXXD; whereas the Ngram metric requires that the matched letters ABCD and any other pair be contiguous as in the case of ABCD and XXABCDXX. These results provide strong evidence that this feature is important in orthographic priming effects for word targets.

Discussion

The ratings of orthographic similarity from the Levenstein Distance and sequence encoder metrics were not significantly correlated with reaction times for word targets. The ngram metric was correlated with reaction times for word targets, but this became

nonsignificant after factoring out the effects of target length. Overall, these results indicate that these metrics may be over and under sensitive to the orthographic features which appear to be relevant for detecting orthographic priming. The results provide support for the notion that the sensitivity of the SOLAR metric to noncontiguous shared letters does account for a portion of the reaction time variance for word targets which the other metrics are not able to do.

None of the orthographic similarity metrics were significantly correlated with nonword target reaction times once the effects of target length were removed. The implication is that orthographic priming does not play a significant role in nonword recognition at least in the measurable context of this experiment. This would be consistent with the entry-opening model of word recognition. When presented with an unfamiliar string of letters that are not represented in a person's mental lexicon, no previous entries exist in which to impact the processing of the nonword target. Any basic effects of shared visual letter overlap between the novel nonword prime and nonword target are not enough to significantly speed up the search time that it takes for a person to scan the lexical entries in their memory in order to determine that an entry does not in fact exist for the presented nonword target. Furthermore, given the slower mean reaction times to nonwords, orthographic priming effects that are typically on the order of 10 to 30 ms would be harder to detect because of an increased reaction time range relative to the more restricted range of word targets.

These results raise some interesting issues that were used as a basis for a second experiment. One issue is whether or not orthographic similarity priming, as detected by

the SOLAR metric, continues to play a role in later stages of visual processing. The prime duration (67 ms) used in experiment 1 was very fast, occurring in the early stages of visual processing of target stimuli. Does orthographic priming continue to play a role when the prime occurs for a longer duration? Another issue is the lack of any measurable orthographic priming for nonword targets. Would an increase in prime duration allow for orthographic nonword primes to speed up the processing of similar nonword targets?

Experiment 2

Method

Participants. 26 participants were recruited from the George Mason University campus. Participants were either undergraduate psychology majors who participated in the experiment for course credit or persons who responded to advertisements posted around the campus. Those who responded to the advertisements were paid \$15 for their participation. The average age of participants was 28.5 years. Fifteen participants were female and 11 were male. All participants reported that they were native English speakers.

Stimuli and design. The stimuli and design were the same as Experiment 1.

Procedure. The procedure of the experiment was the same as experiment 1, with one exception. The duration of the prime was increased from 67 ms to 200ms. This value was selected in order for the prime to be near, but just outside the range in which participants could consciously process and identify the prime. This value was based on several priming studies that found this to be the end of early visual processing (Forster,

2003). Additionally, an informal pilot study was conducted with four different participants with prime durations ranging from 67 to 1000 ms. All participants were able to identify a significant number of the prime's letters at durations > 200 ms, but could not identify them at durations \leq 200 ms.

Results

Data trimming. The same data trimming procedure was used in Experiment 2 as was used in Experiment 1. For word targets, the trim value across participants was 4700 ms and for nonword targets the trim value was 5388 ms. A second pass trim within each participant's trials removed the remaining trials which exceeded + 4 standard deviations of the within subject's mean. On average, 1.9% of correct word target trials and 2.7% of correct nonword target trials were removed due to reaction time trimming. A total of 12,459 word target and 11,971 nonword target trials were analyzed across all 26 participants. The mean response times and percentage of errors of these trials are presented in Table 12 below.

Table 12. Mean response times, standard deviations, standard error of the means, and % trial errors

	Word Targets	Nonword Targets
Mean RT (ms) [‡]	919	1207
Std Dev [‡]	185	303
Std Error [‡]	5.3	8.8
% Errors	2.3	5.35

[‡] Calculated from correct and RT trimmed trials

The effects of orthographic similarity on word target reaction times. The mean response time per prime word-target pair was correlated with the orthographic similarity ratings for the pair. These results are listed in Table 13 below.

Table 13: Correlations between word target reaction times and the orthographic similarity metric ratings

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time	0.016	-0.132**	-0.041	-0.010

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

The ngram was the only metric rating which was significantly correlated with reaction times. The SOLAR metric was not significantly correlated, which differs from the results of Experiment 1.

The effects of word target length on word target reaction times. Similar to the results of Experiment 1, the correlation between target length ($M = 5.98$, $SD = 1.4$, $N = 1200$) and reaction time ($M = 919$, $SD = 185$, $N = 1200$) was significantly correlated at $r(1198) = .37$, $p < .001$. The variance associated with target length was regressed out of reaction time and correlated with the orthographic similarity metric ratings. The results are reported in table 14 below.

Table 14: Correlations between word target reaction times and the orthographic similarity metric ratings when factoring out the effects of target length

Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
----------------------	-------	-------	------------------

Reaction Time‡	0.005	-0.013	-0.015	-0.055
----------------	-------	--------	--------	--------

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

‡Reaction times with the variance associated with target length removed

Again, like experiment 1, the significant correlation between the ngram metric and reaction time became nonsignificant after the effects of target length were removed.

The effects of orthographic similarity on nonword target reaction times.

Table 15 Correlations between nonword target reaction times and the orthographic similarity metrics

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time	.048	-0.153**	-0.013	0.058

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

The ngram metric was the only one which was significantly correlated with reaction times for nonword targets, which mirrors the results of experiment 1.

The effects of length on nonword target reaction times. The correlation between target length ($M = 5.97$, $SD = 1.4$, $N = 1200$) and reaction time ($M = 1207$, $SD = 303$, $N = 1200$) was significantly correlated at $r(1198) = .41$, $p < .001$. The variance associated with target length was regressed from the reaction times and correlated with the orthographic similarity metric ratings. The results are reported in table 16.

Table 16: Correlations between nonword target reaction times and the orthographic similarity metrics when factoring out the effects of target length

	Levenshtein Distance	Ngram	SOLAR	Sequence Encoder
Reaction Time‡	-0.013	-0.025	-0.003	0.035

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

‡Reaction times with the variance associated with target length removed

Similar to the results of the word targets in experiment 2 as well as the nonword target results of experiment 1, the correlations between the metrics and reaction times became nonsignificant.

Discussion

The results from Experiment 2 showed that the facilitory orthographic priming effect found for word targets was limited to a prime duration of 67 ms. The longer prime exposure in Experiment 2 may have placed more work on the visual recognition system to process the prime and thus any minor facilitory priming effects for the target were negated by the extra processing for the prime because it was visible for a longer time. Similar to the results of Experiment 1, nonword targets did not receive any facilitory or inhibitory priming for orthographically similar primes. Thus, any facilitory priming effects due to shared letters for nonword primes and nonword targets is nonexistent for both 67 ms and 200 ms prime durations.

Experiment general discussion

A major goal of Experiments 1 and 2 was to determine the degree to which string length is a significant factor of orthographic similarity and by extension how big a factor

it should be when considering how the mind represents and organizes orthographic features of the lexicon. To this end, the experiments were a success because it appears length is a major factor and should be central to any successful theory of word recognition. This may come as no surprise, but up until this point it was not clear why the assumption that orthographic similarity be limited to words of roughly the same length be supported without extensive investigation.

Overall, the results from Experiments 1 and 2 show a relatively insignificant role in priming that simply sharing a large portion of letters between a prime and target has during visual word recognition (at least in the manner that the Levenshtein distance, Ngram, SOLAR, and sequence encoder measure). The only situation where shared letters between the primes and targets had a statistically significant correlation with reaction time was using a 67 ms prime duration with nonword primes and word targets. In this case, the SOLAR metric seemed most attuned to measuring orthographic priming effects through its sensitivity to relative position coding and shared initial position weighting.

There were several methodological issues in the two experiments which may help explain the paucity of significant findings along this line for the other conditions in the experiments. The extension of length differences between primes and targets up to four letters was not done to my knowledge in previous studies. At most, differences ranged up to two letters. It appears that the differences in the time required for the visual system to process four letter stimuli and eight letter stimuli was large enough that any potential 10-30 ms facilitatory or inhibitory effects due to any shared letters were masked or non-

existent. Research studies which used length differences of two letters or less between primes and targets were in a better position to detect any priming effects due to shared letters (De Moor & Brysbaert, 2000; Schoonbaert & Grainger, 2004; Davis & Taft, 2005; Van Assche & Grainger, 2006) .

Additionally, the procedure of mixing prime-target length differences across trials may have contributed a large amount of unaccountable variance into the data which helped mask any non-length dependent priming effects. The original justification for mixing length differences across trials was to more closely mimic situations when people encounter visual strings outside of a tightly controlled laboratory setting and by extension more closely reflect how words may be accessed in memory in "real" life. When we normally process visual strings, they are not typically paired into similar length strings all the time which is the prevalent method in most orthographic priming studies. The two experiments described were an attempt to move away from this restriction, but it appears it was a step too far and revealed just how important string length is for word recognition reaction times.

CHAPTER 4: GENERAL DISCUSSION

An overarching goal of the dissertation was to expand the conception of orthographic similarity to address word length restriction issues and shared letter position matching issues prevalent in early versions of orthographic similarity metrics, particularly the slot-based N metric. To this end, several recent metrics of orthographic similarity were reviewed in relation to their ability to address length differences and their sensitivities to letter position overlap such as shared initial letter positions and relative versus absolute letter position mappings.

A major question that needed to be answered was if and how the revised orthographic similarity metrics impacted orthographic neighborhood structure relative to previous similarity metrics. This question was explored through the creation of orthographic networks based on the revised metrics. The results of the orthographic network analyses found that the Levenshtein distance, Ngram, and SOLAR metrics produced considerable differences in neighborhood structure than early metrics of orthographic similarity.

A notable, but not surprising finding was that the Levenshtein distance, Ngram, and SOLAR orthographic networks had higher potential neighborhood sizes because the pool of neighbors was not restricted by word length. Thus, it is possible that the words

GLOW and GLOWS could be orthographic neighbors in the revised conception of similarity. Another notable difference was the range of orthographic similarity values the various metrics provided. On average, words in the Ngram and SOLAR orthographic networks had a higher baseline of connectivity to other words because these metrics were sensitive to shared single letters among words; whereas, the Levenshtein distance and N metric of similarity were not and thus words in these networks had a lower degree of similarity. Overall, due to its flexibility in measuring relative position coding, the SOLAR metric showed the highest average degree of similarity among the words in the network.

In addition to measures of neighborhood size and average similarity, the link distribution structure of the current orthographic networks were examined in light of power law, scale free distributions found in other language networks such as semantic ac networks (Holanda, Torres Pisa, Kinouchi, Souto Martinez, & Seron Ruiz, 2004; Kinouchi, Martinez, Lima, Lourenço, & Risau-Gusman, 2002; Motter, de Moura, Lai, & Dasgupta, 2002; Steyvers & Tenenbaum, 2005), syntactic networks (Cancho, Sole, & Kohler, 2004; Sole, 2005; Sole, Murtra, Valverde, & Steels, 2006), and substring orthographic networks (Kello & Beltz, in press). It was found that none of the orthographic networks exhibited a power law distribution of links. This result shows that power law distributions are limited in where they can be found in language phenomena.

The lack of a power law, scale free distribution in the current results does not support the explanation by Kello & Beltz (in press) that the power law, scale free structure of substring orthographic networks could result from a critical trade-off of

memorability and discriminability of wordforms in memory (Cancho & Sole, 2003). This explanation argued that two major competing constraints helped structure the formation of substrings within words across a lexicon. On one hand, evolutionary pressures existed for wordforms to be perceptually distinctive from one another in order to reduce errors during communication (e.g. to avoid confusing phonologically based words such as BIT and PIT or DOES or BUZZ). The pressure for distinctiveness comes at a price in that if every possible entity and concept in the world receives a unique label or wordform in a particular language, the number of wordforms for that language's lexicon would be immense. Learners of the lexicon would then be forced to learn each unique wordform for every entity and concept they encounter which would come at a large cost. This cost introduces an evolutionary pressure for the wordforms of a lexicon to be as "memorable" as possible. In other words, more easily learned and represented in the brain. One way for lexicons to do this is to reuse perceptual units within or even entire wordforms for different entities and concepts. This is where notions of homophones, polysemy, and inflectional morphology play a role in language. Kello and Beltz (in press) argued that the existence of a power law distribution in substring orthographic networks in English and several other world languages was evidence of a critical interplay between these two factors which optimally balanced information entropy in a lexicon.

The current orthographic networks provide some puzzling data for the critical balance of memorability and discriminability explanation. The current study only used 5000 English words which is significantly smaller than the samples ranging from 27,000 up to 150,000 used in Kello and Beltz (in press). However, one hallmark of scale free

distributions is that their structure is not tied to their sample size and the structure should be consistent regardless of the scale of their components. This would indicate that although the current sample was much smaller, characteristics of a power law, scale free distributions should still be apparent- which was not the case. This conflicting data does not rule out the importance of memorability and discriminability in the formation of words of a lexicon, only that their interplay does not necessarily create power law, scale free orthographic structures in the lexicon. The implications of the current orthographic network link distributions for the balance of memorability and discriminability could be tested in future work.

In the third chapter, the revised orthographic similarity metrics were tested in two form priming, lexical decision experiments to examine whether they provided comparable results to previous studies in orthographic priming which found facilitory effects for orthographic neighbors (Andrews, 1989; Andrews, 1992; Carreiras et al., 1997; Grainger, 1990; Grainger & Jacobs, 1996; Grainger, O'Regan, Jacobs, & Segui, 1989, 1992; Johnson & Pugh, 1994; Sears, Hino, & Lupker, 1995). Overall, the empirical results were mixed. The Levenshtein distance, Ngram, and sequence encoder metrics did not significantly correlate with reaction times for either word or nonword targets either in the 67 ms or 200 ms priming conditions once the effect of target length was accounted for. In the 67 ms priming condition, the SOLAR ratings of orthographic similarity were significantly correlated with reaction times such that nonword primes and word targets which had a higher degree of orthographic similarity were responded to faster by subjects. This correlation was quite small relative to other variables such as target length

and word frequency which had a higher degree of correlation and factored more strongly into reaction times. Still, the analyses showed that the SOLAR measure of similarity was independent of length and word frequency. Even more promising for the SOLAR metric, was the nature of the experimental method which loosened some of the traditional controls found in previous priming studies. The current study changed the empirical paradigm to make it more reflective of real-world word processing by mixing primes and targets of considerably different lengths which avoids artificial list effects due to the blocking of similar length stimuli sets. Although the design of this study wasn't perfect in that regard, it was an extension of previous studies (Forster, 2003; Davis, 2006) and provides converging evidence of conditions in which orthographic priming can be measured; namely for word targets under a relatively fast prime duration reflective of early visual processing.

The small facilitory priming effect measured in experiment 1 highlights which feature is most telling for distinguishing orthographic neighbors during early stages of visual word recognition. The most critical feature was the SOLAR metric's measurement of relative positions among shared letters of the primes and targets, which the other metrics were not able to account for. This "sloppy" ability to match letters within some position span (Davis, 2006) is useful during word recognition especially considering how signal degradation of particular letters or features of words may hinder the visual system's ability to correctly identify parts of words. The theme of signal degradation also fits within the scope of memorability, discriminability, and information entropy issues of lexical memory discussed previously.

The conceptual role of orthographic neighbors has helped refine how memory processes play out when attempting to identify a particular word in memory in relation to all words stored in the lexicon. Previous research has shown how words with particular neighborhood structures, such as large neighborhoods under the N definition of orthographic neighbor, could speed up the processing of the target word. This was due to mechanisms such as spreading activation which boosted the lexical representations of the neighborhood set and speeded up access to them as postulated in the entry-opening model (Forster & Dickinson, 1976; Paap & Johansen, 1994) and interactive activation models (McClelland & Rumelhart, 1981). The current findings from the form priming experiments do not substantially change these models, but instead expands the conception of orthographic neighbors and how similar sets of words may be represented in memory. In particular, previous orthographic neighborhood studies treated words of the same length as potential neighbor cohorts in memory. The findings of facilitory priming in the current study with primes and targets of various lengths indicates that the length constraint should be loosened for future conceptions of orthographic neighborhoods.

APPENDIX A: WORDS USED IN THE ORTHOGRAPHIC NETWORKS.
 From Nelson, McEvoy, & Schreiber (1999) and Steyvers and Tennebaum (2005)

a	adjective	air	angle
aardvark	adjourn	air_force	angry
abdomen	adjustment	aircraft	animal
abduct	admire	airplane	animals
ability	admission	airport	anisette
able	admit	alarm	ankle
abnormal	adolescent	album	annihilate
above	adorable	alcohol	announce
absence	adore	ale	annoy
absent	adorn	alert	annoying
abstract	adult	algae	annual
absurd	adultery	algebra	another
abundance	adults	alike	answer
abuse	advance	alive	ant
accelerate	adventure	all	anteater
accept	adventurous	allegiance	antelope
accident	adverb	allergy	antenna
accomplish	adversary	alley	antidote
accomplish	advertisement	alligator	antique
ed	advice	allow	antler
account	advil	almanac	antlers
accumulate	advise	almighty	anxiety
accurate	aerobics	almond	anxious
accuse	affair	alone	anything
ace	affect	alphabet	apart
ache	affection	alter	apartment
achieve	afraid	alternative	apathetic
achievement	africa	alto	apathy
t	after	aluminum	ape
acid	afternoon	always	apex
acknowledg	again	amateur	apology
e	against	amaze	appeal
acorn	age	ambition	appear
acquire	agency	ambulance	appearance
acre	agenda	america	appendage
acrobat	agent	american	apple
act	aggravate	amount	appliance
action	aggressive	amp	application
active	agony	amuse	appointment
activity	agree	analysis	t
actor	agreement	analyze	appraise
actress	agriculture	anarchy	approval
ad	ahead	anatomy	apron
add	ahoy	ancestor	architect
addict	aid	anchor	architecture
addiction	aids	ancient	arctic
addition	aim	angel	area
address		anger	arena

argue
argument
aright
arithmetic
arm
armor
arms
army
aroma
around
arrest
arrive
arrogant
arrow
art
artery
article
artist
arts
ascend
ascent
ash
ashamed
ashes
ashtray
asia
ask
asleep
asphalt
aspirin
ass
assault
assignment
assist
assistance
assistant
associate
assume
asteroid
astrology
astronaut
astronomy
athlete
athletic
atlas
atmosphere
atom
atomic
attach
attack
attempt

attend
attendance
attention
attic
attire
attitude
attorney
attract
attractive
audience
aunt
aura
author
authority
auto
automobile
autumn
avenue
average
avoid
awake
award
aware
awareness
away
awe
awesome
awful
awkward
ax
axon
baby
back
backbone
background
backpack
backward
bacon
bacteria
bad
badge
bag
bagel
baggage
bait
bake
baked
baker
balance
balcony
bald

ball
ballerina
ballet
balloon
ballot
ban
banana
band
bandage
bandaid
bandanna
bandit
bang
banjo
bank
banker
banner
banquet
baptist
bar
barbecue
bargain
bark
barley
barn
barracuda
barrel
barrier
bartender
barter
base
baseball
basement
bashful
basic
basket
basketball
bass
bassinet
bat
bath
bathroom
bathtub
battery
battle
bay
be
beach
bead
beaker
beam

bean
beans
bear
beard
beast
beat
beautiful
beauty
beaver
because
become
bed
bedroom
bee
beef
beer
beet
beetle
before
beg
beggar
begin
beginner
beginning
behavior
behind
being
belief
believe
bell
belly
belong
below
belt
bench
bend
beneath
benefit
bent
berry
best
bet
betray
better
between
beverage
beware
beyond
bias
bible
bicycle

big
bike
bill
billboard
binder
biologist
biology
bird
birds
birth
birthday
biscuit
bishop
bison
bit
bite
bitter
bizarre
black
blackberry
blackboard
blackmail
blade
blame
blank
blanket
blast
blaze
bleach
blend
blender
bless
blessing
blind
blinds
block
blockade
blocks
blonde
blood
bloom
blossom
blot
blouse
blow
blubber
blue
blueberry
bluejay
blueprint
blues

bluff
blunt
blur
blush
board
boat
body
bog
boil
boiled
bold
bologna
bolt
bomb
bond
bondage
bone
bonus
book
bookbag
boom
boot
booth
boots
booze
border
bored
boring
borrow
boss
bother
bottle
bottom
bouillon
boulder
boulevard
bounce
bound
boundary
bounty
bouquet
bourbon
bow
bowl
bowling
box
boxer
boxing
boy
boyfriend
boys

bra
bracelet
braces
brain
brake
bran
branch
brand
brandy
brass
brat
bravado
brave
bravery
brawl
bread
break
breakable
breakfast
breast
breath
breathe
breed
breeze
breezeway
bribe
brick
bride
bridge
brief
briefcase
bright
brilliant
bring
brisk
bristle
britannica
brittle
broad
broccoli
broil
broke
broken
bronze
brook
broom
broth
brother
brown
bruise
brunette

brush
bubble
buck
bucket
buckle
bud
buddy
budget
buffalo
bug
buggy
bugle
build
builder
building
bulb
bull
bullet
bulletin
bullock
bully
bum
bumble
bump
bumper
bumps
bumpy
bun
bunch
bunny
burden
bureau
burglar
burglary
burial
burn
burnt
burst
bury
bus
bush
business
busy
butcher
butler
butt
butter
butterfly
button
buy
buyer

buzz	capture	cell	cheerleader
cab	car	cellar	cheese
cabbage	caravan	cello	chef
cabin	carbon	celsius	chemical
cabinet	card	cement	chemist
cable	cardboard	cemetery	chemistry
caboose	cardinal	cent	cherry
cactus	cards	center	chess
cafe	care	central	chest
cafeteria	career	cents	chew
caffeine	careful	century	chick
cage	careless	ceramic	chicken
cake	caress	ceramics	chief
calcium	caring	cereal	child
calculate	carnival	ceremony	childish
calculator	carpenter	certain	children
calculus	carpet	chain	chili
calendar	carriage	chair	chill
calf	carrot	chairperson	chime
california	carrots	chalk	chimney
call	carry	chalkboard	chimpanzee
calm	cart	challenge	china
calorie	carton	chamber	chinese
calories	cartoon	champagne	chip
camel	case	champion	chipmunk
camera	cash	chance	chips
camp	cashew	change	chirp
campaign	cashier	channel	chisel
camping	casket	chaos	chlorine
campus	cast	chapel	chocolate
can	castle	chapter	choice
canal	casual	character	choir
canary	cat	characteristi	choke
cancer	catastrophe	c	choose
candidate	catch	charcoal	chop
candle	catcher	charge	chore
candy	caterpillar	charity	chorus
cane	cathedral	charm	chowder
cannon	catholic	chart	christ
canoe	cattle	chase	christian
canvas	cauliflower	chauffeur	christmas
canvass	cause	cheap	chromosom
canyon	caution	cheat	es
cap	cautious	check	chuck
capability	cave	checkbook	chunk
capacity	cavern	checkers	church
cape	cavity	checking	cigar
capital	ceiling	cheddar	cigarette
captain	celebrate	cheek	cinema
caption	celebration	cheer	cinnamon
captive	celery	cheerful	circle

circumstance
e
circus
citizen
citrus
city
claim
claims
clam
clamp
clarify
clarinet
class
claw
clay
clean
cleaner
cleaning
clear
clench
clerk
clever
click
cliff
climate
climax
climb
climber
clinic
clippers
cloak
clock
clockwise
clog
clorox
close
closed
closet
closing
cloth
clothes
cloud
cloudy
clove
clown
club
clue
clumsy
coach
coal
coast

coat
cob
cobbler
cobra
cobweb
coca-cola
cocaine
cocktail
cocky
cocoon
cod
code
coffee
coffin
coin
coke
cold
coleslaw
collar
collect
college
cologne
colonel
color
colt
column
coma
comb
combination
n
combine
come
comedian
comedy
comet
comfort
comfortable
comma
command
commander
comment
commercial
commitment
t
committee
common
communication
te
communication
tion
communist

community
compact
companion
company
compass
compassion
compete
competence
competition
complain
complete
complex
complicated
compliment
component
components
compose
compound
comprehend
d
compulsion
compute
computer
con
conceit
conceited
concentrate
concept
concern
concert
conclude
conclusion
concrete
condemn
condition
conditioner
condom
conductor
cone
conference
confess
confession
confidence
confident
conflict
confuse
confusion
congress
connect
conquer
conquest

conscience
conscious
consequence
e
consider
console
constant
constitution
construct
construction
n
constructive
consumer
contact
contacts
container
contemporary
ry
content
contest
context
continent
contract
contractor
control
controls
convent
conversation
n
convict
convince
cook
cookbook
cooked
cookie
cooking
cookout
cool
cooler
cop
copier
co-pilot
copper
copy
coral
cord
core
cork
corn
cornbeef
corner

corporal
corporation
corpse
correct
corridor
corrupt
cosmetics
cost
costume
cottage
cotton
couch
cougar
cough
counsel
count
counter
country
county
couple
courage
courageous
course
court
cousin
cover
covered
cow
coward
cowboy
cowgirl
coyote
cozy
crab
crack
cracker
crackers
crackle
cradle
craft
crafts
cramp
crane
crash
crater
crave
crawl
crayola
crayon
crazy
cream

crease
create
creation
creative
creativity
creator
creature
credit
creek
crescent
crest
crevice
crew
crib
cricket
crime
criminal
crisis
crisp
criterion
critic
critical
criticism
criticize
croak
crochet
crocodile
crook
crooked
crops
cross
crossing
crow
crowd
crown
crude
cruel
cruise
crumb
crunch
crunchy
crush
crust
crutch
cry
crystal
cub
cube
cucumber
cuddle
cue

cuff
culture
cunning
cup
cupboard
cups
curb
cure
curiosity
curious
current
curse
cursive
curtain
curtains
curve
curved
cushion
cuss
custard
custom
customer
cut
cute
cycle
cyclone
cylinder
dad
dagger
daily
dairy
daisy
dam
damage
damn
damp
dance
dancer
dandelion
dandruff
danger
dangerous
dare
daring
dark
darkness
dart
dash
dashboard
data
date

daughter
dawn
day
days
dead
deadly
deaf
deal
death
debate
debt
decade
decay
deceit
deceive
decency
deception
decide
decision
deck
decline
decompose
decorate
decoration
decrease
deep
deer
defeat
defend
defense
define
definition
defrost
degrade
degree
delay
delicate
delicious
delight
delinquent
deliver
demand
demi
democracy
democrat
demolish
demon
denial
denim
dense
dent

dentist
deny
deodorant
depart
department
depend
dependent
depletion
deposit
depressed
depression
depth
deputy
descend
descent
describe
desert
design
designer
desire
desk
despair
despise
dessert
destiny
destroy
destruction
detach
detail
detective
detergent
deteriorate
determine
develop
development
t
device
devil
dew
diagram
dial
diameter
diamond
diaper
diary
dice
dictator
dictionary
die
diet
differ

difference
different
difficult
difficulty
dig
digest
digestion
digger
digit
dignity
dilemma
diligence
dill
dim
dime
diminish
dine
diner
dinner
dinosaur
dip
diploma
direct
direction
directions
director
dirt
dirty
disagree
disagree
ment
disallow
disappear
disappoint
disaster
disbelief
disbelieve
disc
discipline
discomfort
discourage
discover
discovery
discreet
discuss
discussion
disease
disgrace
disguise
disgust
disgusting

dish
dishes
dishonest
dishwasher
disintegrate
dislike
dismay
dismiss
disorder
disown
disperse
display
disprove
disregard
dissimilar
dissolve
distance
distant
distinct
disturb
ditch
dive
diver
divide
diving
division
divorce
dizzy
do
dock
doctor
document
doe
dog
doll
dollar
dollars
dolphin
domain
dominant
dominate
donation
done
donkey
donor
door
doorbell
doorway
dope
dorm
dot

double
doubt
dough
doughnut
dove
down
downstairs
downstream
downtown
doze
dozen
dracula
draft
drag
dragon
drain
drama
drapes
draw
drawer
drawing
dream
drench
dress
dresser
dressing
drift
drill
drink
drip
drive
driver
driveway
drop
drove
drown
drowsy
drug
drugs
drum
drunk
dry
dryer
duck
ducks
due
dull
dumb
dump
dune
dungeon

dunk
duplicate
dusk
dust
dustpan
duty
dwarf
dwarves
dynamite
dynasty
eager
eagle
ear
early
earn
earnestness
earring
ears
earth
earthquake
ease
east
easter
easy
eat
eating
economic
economy
ecstasy
edge
editor
editorial
educate
education
effect
effort
egg
eggs
ego
egypt
eight
einstein
elaborate
elbow
elderly
elders
elect
election
electric
electrician
electricity

electron
elegant
element
elements
elephant
elevator
elf
elimination
elk
elm
embarrass
embrace
emerald
emergency
emotion
emotional
empathy
emperor
empire
employee
employer
employments
empty
encyclopedia
end
ending
endless
enemy
energetic
energy
engage
engagement
engine
engineer
england
english
enjoy
enough
enrage
enter
entertain
entertainment
entrance
entry
envelope
environment
envy

equal
equation
era
eraser
erect
erode
error
erupt
escalator
escape
escargot
essay
essence
esteem
estimate
eternal
eternity
ethics
etiquette
europe
evacuate
evaluate
even
evening
event
ever
everyday
everything
evict
evidence
evil
ewe
exact
exam
examine
example
excedrin
excel
excellent
exceptional
exchange
excise
excited
excitement
exciting
exclamation
_point
excuse
executive
exercise
exhale

exhausted
exist
existence
exit
expense
expensive
experience
experiment
expert
explain
explanation
explode
explorer
explosion
express
expression
extension
extinct
extra
extravagant
extreme
extremity
eye
eyeball
eyebrows
eyelash
eyes
fable
fabric
face
fact
factor
factory
facts
faculty
fad
fade
fail
failure
faint
fair
fairy
fairytale
faith
faithful
fake
fall
false
fame
family
famine

famished
famous
fan
fancy
fangs
fantastic
fantasy
far
farm
farmer
fascism
fashion
fast
fasten
fat
fate
father
fatigue
fattening
faucet
fault
favor
favorite
fawn
fbi
fear
feast
feather
feathers
feature
federal
fee
feeble
feed
feel
feeling
feelings
feet
fell
fellow
felt
female
fence
festival
feud
fever
few
fib
fiber
fiction
fiddle

field
fierce
fig
fight
fighter
figure
file
fill
film
filth
filthy
fin
final
financial
find
fine
finger
fingernail
fingers
finish
finished
fire
fireman
fireplace
firm
first
fish
fishing
fist
fit
fitness
five
fix
flag
flake
flakes
flame
flannel
flap
flash
flashlight
flask
flat
flavor
flea
flee
fleet
flesh
flex
flexible
flick

flight
fling
flip
flipper
float
flood
floor
florida
flour
flow
flower
flowers
flu
fluffy
fluid
flunk
flush
flute
flutter
fly
flying
foam
focus
foe
fog
foggy
foil
fold
folder
follow
follower
folly
fond
food
fool
foolish
foot
football
for
forbid
forbidden
force
forceful
forehead
foreign
foresight
forest
forever
forget
forgive
fork

form
formal
formula
fort
fortune
forward
fossil
foul
found
foundation
fountain
four
fox
fracture
fragile
fragrance
frail
frame
france
frankenstei
n
frantic
fraternity
fraud
fray
freak
freckle
free
freedom
freeway
freeze
freezer
french
frequency
frequent
fresh
friar
friday
fried
friend
friendliness
friendly
friends
friendship
fries
fright
frightened
frigid
frill
frisk
frog

front
frontier
frost
frosted
frosting
frown
frozen
fruit
frustrate
frustration
fry
fudge
fuel
fugitive
full
fumes
fun
function
fund
funeral
fungi
fungus
funny
fur
furniture
furry
further
fury
fuse
fuss
future
fuzz
fuzzy
gag
gain
gal
galaxy
gallon
gallop
galoshes
gamble
game
games
gander
gang
gangster
garage
garbage
garden
garlic
gas

gate
gather
gathering
gauge
gauze
gavel
gazelle
gear
geese
gem
gender
gene
general
generous
genius
gentle
gentleman
genuine
geology
geometry
george
germ
get
geyser
ghetto
ghost
ghoul
giant
gift
gifted
giggle
gin
giraffe
girl
girlfriend
girls
give
give_up
giving
glacier
glad
glance
glands
glare
glass
glasses
glide
globe
gloomy
glory
glove

gloves
glow
glue
go
goal
goat
goblin
god
goddess
godliness
going
gold
golden
golf
gone
goo
good
goodbye
goodness
goods
goose
gorgeous
gorilla
gossip
gourmet
government
governor
gown
grab
grace
graceful
grade
graduate
graduation
graft
grain
grammar
grand
grandma
grandpa
grandparent
s
granite
grant
grape
graph
grasp
grass
grasshopper
grate
grave

gravel
gravity
gravy
gray
graze
grease
great
greatest
greed
greek
green
greet
grief
grill
grime
grind
grip
grit
grizzly
groceries
grocery
groom
gross
ground
grounds
group
grove
grow
grown
grown-ups
growth
guard
guardian
guess
guest
guide
guilt
guilty
guitar
gull
gullible
gully
gulp
gum
gums
gun
guy
guys
gym
gymnast
gymnastics

habit	haunt	hesitate	hopeless
hail	have	hidden	horizon
hair	haven	hide	hormones
haircut	havoc	high	horn
hairspray	hawaii	high_school	hornet
hairy	hawk	highlight	horoscope
half	hay	highway	horrible
hall	haystack	hike	horror
halloween	hazard	hiker	horse
hallucinatio	haze	hiking	hose
n	he	hilarious	hospital
hallway	head	hill	host
halo	head_&_sh	him	hostage
halt	oulders	hindsight	hostess
ham	headache	hinge	hostility
hamburger	headband	hint	hot
hammer	heading	hip	hot_dogs
hammock	heal	hispanic	hotel
hamper	health	history	hotter
hand	healthy	hit	hound
handbag	hear	hitch	hour
handcuffs	hearing	hitchhike	hours
handicap	heart	hive	house
handkerchie	heart_attack	hobby	how
f	heat	hobo	hue
handle	heater	hockey	hug
hands	heaven	hoe	huge
handsome	heavy	hog	hula
hang	hectic	hold	hull
hanger	hedge	holder	hum
happen	heel	hole	human
happening	height	holiday	humble
happiness	heights	holiness	humid
happy	helicopter	holler	humiliate
harbor	helium	holy	humor
hard	hell	home	hundred
hardy	hello	homeless	hung
hare	helmet	homework	hunger
harley	help	honest	hungry
harm	helper	honesty	hunt
harmful	helpful	honey	hunting
harmony	hem	honeymoon	hurricane
harp	hen	honk	hurry
harsh	her	honor	hurt
harvest	herb	hood	husband
hat	herd	hook	husk
hatch	here	hooked	husky
hatchet	heritage	hoop	hustle
hate	hero	hoot	hut
haughty	heroin	hop	hydrogen
haul	heroine	hope	hygiene

hymn	individual	interest	jobless
hyper	indoors	interesting	jock
hypnotize	industry	interior	jog
hypothesis	inebriated	interrupt	join
i	inertia	interstate	joint
ice	inexpensive	interview	joke
ice_cream	infant	intestine	joker
iceberg	infection	intimate	journal
icing	inferior	intoxicate	journey
idea	inferno	intoxicated	joy
ideal	infinite	introduce	joyous
identical	infinity	introvert	judge
identification	inflation	intuition	judgment
identity	influence	invalid	jug
idiot	inform	invent	juggler
idol	information	inventor	juice
igloo	inhale	iron	jump
ignorant	initial	irregular	jumpy
ignore	injection	irritate	jungle
ill	injury	irritating	junior
illegal	ink	island	junk
illness	inlet	isolated	jupiter
illusion	inmate	issue	jury
image	inn	italian	just
imagination	innocence	itch	justice
imagine	innocent	item	justify
imitate	inquire	ivory	juvenile
immature	insane	ivy	kangaroo
immediate	insect	jack	keep
immoral	insecure	jacket	keeper
impact	inside	jail	keg
impatience	insight	jam	ketchup
impatient	inspect	japan	kettle
impediment	instance	jar	key
implode	instant	jaw	keyboard
important	instinct	jaws	keys
impossible	instruct	jazz	kick
impression	instruction	jealousy	kid
impulse	instructions	jeans	kidnap
in	instructor	jeep	kidney
inability	instrument	jello	kids
incense	insult	jelly	kill
inch	insurance	jeopardy	killer
incline	intake	jerk	kilometer
income	integrate	jesus	kin
incorrect	integrity	jet	kind
increase	intellect	jewel	kindness
independent	intelligence	jewelry	kinetic
indian	intelligent	jewish	king
indirect	intensity	jigsaw	kingdom
	intent	job	kiss

kit	lava	liar	loan
kitchen	law	libel	lobby
kite	lawful	liberal	lobe
kitten	lawn	liberty	lobster
kiwi	laws	library	local
kleenex	lawsuit	lice	location
knapsack	lawyer	license	lock
knee	lay	lick	lodge
kneel	layer	lid	loft
knees	lazy	lie	log
knife	lead	lieutenant	logic
knight	leader	life	lollipop
knit	leadership	lifestyle	london
knitting	leaf	lift	lonely
knob	league	light	long
knock	leak	lighter	look
knot	lean	lightheaded	looks
know	leap	lightning	loop
knowledge	learn	like	loose
known	learner	likelihood	loosen
knuckle	learning	lily	lord
lab	lease	limb	los_angeles
label	least	lime	lose
labor	leather	limit	loser
labyrinth	leave	limousine	loss
lace	leaves	limp	lost
lack	lecture	line	lot
ladder	ledge	linen	lots
lady	left	lingerie	lottery
lagoon	leg	link	loud
lake	legal	lint	lounge
lamb	legend	lion	love
lamp	legion	lip	lover
land	legislature	lips	lovers
landscape	legs	lipstick	loving
lane	lemon	liquid	low
language	lemonade	liquor	lower
lap	lend	list	loyal
lapel	length	listen	loyalty
large	lenient	lit	lubricate
laser	lens	literature	luck
lash	leo	litter	luggage
last	leopard	little	lumber
lasting	less	live	lump
latch	lesson	liver	lunch
late	let	living	lung
lather	letter	living_roo	lust
latin	letters	m	luxury
laugh	lettuce	lizard	macaroni
launch	level	load	machine
laundry	lever	loaf	mad

made
mafia
magazine
maggot
magic
magician
magnet
magnify
magnitude
maid
maiden
mail
mailbox
mailman
main
major
majority
make
make_up
maker
malaria
male
mall
malt
mammal
man
managemen
t
manager
manner
manners
mansion
manufactur
e
manufactur
er
many
map
maple
marble
march
mare
margarine
margin
marijuana
marine
marines
mark
marker
market
maroon

marriage
marrow
marry
mars
marsh
mary
mascara
mash
mask
mason
mass
massage
master
mastery
mat
match
mate
material
math
matter
mattress
mature
maximum
maybe
mayonnaise
mayor
maze
mcdonald's
me
meadow
meal
mean
meaning
meaningful
means
measure
measureme
nt
meat
meatballs
mechanic
mechanical
medal
media
medical
medication
medicine
medieval
mediocre
medium
meek

meet
meeting
mellow
melody
melt
member
memorial
memories
memory
men
mend
mental
menthol
meow
merit
mermaid
mess
message
messy
metal
meteor
meteorite
meter
method
methodist
meticulous
metric
mexico
mice
microphone
microscope
microwave
middle
midget
midnight
might
migraine
mild
mildew
mile
military
milk
mill
million
mimic
mind
minded
mine
miner
mineral
minimum

minister
mink
minor
minority
mint
mints
minus
minute
minutes
mirage
mirror
mischief
misery
miss
misses
missile
missing
mission
mist
mistake
mister
mistreat
mitten
mix
mixed
mixture
moat
mob
model
modern
modest
moist
molasses
mold
mole
molecule
molten
mom
moment
monarch
monarchy
monastery
monday
money
mongoose
monk
monkey
monopoly
monotonou
s
monster

month	mussel	neuron	numerous
monthly	must	neutral	nun
monument	mustache	neutron	nurse
moo	mustard	never	nursery
mood	musty	new	nut
moon	mute	new_year	nuts
moonlight	mutton	new_york	nylon
moose	mystery	news	nylons
mop	myth	newspaper	oak
moral	mythology	newsstand	oar
morals	nag	newton	oasis
more	nail	nice	oath
morning	nails	nickel	oatmeal
moron	naive	nicotine	oats
morse	name	niece	obey
mortgage	nap	night	object
mosquito	napkin	nightgown	objective
moss	narrow	nightmare	obligation
most	nasty	nine	obnoxious
motel	nation	nitrogen	oboe
moth	national	no	obscure
mother	native	noise	observe
motion	natural	noisy	obsession
motionless	nature	nomad	obstacle
motor	nausea	none	obtain
motorcycle	nauseous	nonsense	obvious
mountain	navigator	noodles	occupation
mouse	navy	noon	occupied
mousse	near	noose	ocean
mouth	neat	norm	octopus
move	necessary	normal	odd
movement	neck	normal_0	odds
movie	necklace	north	odor
mow	need	nose	off
mower	needle	nosy	offend
much	needy	not	offense
muck	negative	note	office
mud	neglect	notebook	officer
muffin	negotiation	notes	official
mug	neighbor	nothing	often
mule	neighborho	notice	oil
multiply	od	notify	ointment
mummy	nephew	notion	okay
munch	neptune	noun	old
murder	nerd	novel	oligarchy
murderer	nerve	novice	olive
muscle	nerves	now	olympics
museum	nervous	nuclear	omelet
mushroom	nest	nucleus	on
music	net	number	once
musk	network	numbers	one

onion
oodles
open
opener
opening
opera
operate
operation
operator
opinion
opponent
opportunity
opposite
optimistic
option
orange
orange_juice
orchard
orchestra
orchid
order
orderly
ordinary
ore
oregano
organ
organization
orient
original
originate
ornament
orthodontist
other
otter
ouch
ounce
out
out_fox
outcome
outdoors
outfit
outgoing
outlaw
outlet
outline
outrageous
outside
outstanding

oval
oven
over
overcome
overflow
overpower
overview
overwhelm
owe
owl
own
owner
ox
oxygen
oyster
ozone
pack
package
pad
padding
paddle
paddy
page
pail
pain
paint
painter
painting
pair
pajamas
pal
palace
pale
palm
pan
pancakes
pane
panel
panic
panther
pants
pantyhose
paper
paper_towel
paperclip
parade
paradise
paragraph
parcel
parched

pardon
parent
parents
paris
park
parking
parkway
parrot
parsley
part
participate
partner
party
pass
pass_out
passage
passenger
passion
passive
past
pasta
paste
pastry
pasture
pat
patch
path
patience
patient
patio
patriot
patriotic
patrol
pattern
pause
pavement
paw
pawn
pay
payment
peace
peaceful
peach
peak
peanut
peanut_butter
peanuts
pear
pearl
peas

pebble
pecan
pedal
pedestrian
peel
peer
pelican
pelt
pen
pencil
pendulum
penguin
penny
people
pepper
pepperoni
pepsi
perceive
percent
percentage
perception
perch
perfect
perform
performance
perfume
period
periscope
perish
perjury
permission
permit
peroxide
person
personal
personality
personnel
perspire
persuade
pest
pet
petals
petroleum
pets
phantom
phase
philosopher
philosophy
phone
phony

photo
phrase
physical
physician
physics
physiology
piano
pick
pick_up
picket
pickle
pickles
picnic
picture
pie
piece
pieces
pier
pig
pigeon
pigment
pike
pile
pill
pillow
pilot
pimple
pin
pinch
pine
pink
pint
pious
pipe
pirate
pissed
pistol
piston
pit
pitch
pitcher
pitchfork
pity
pizza
place
plaid
plain
plan
plane
planet
planets

planning
plant
plants
plaque
plasma
plaster
plastic
plate
plates
platform
play
play_dough
player
playground
playing
plaything
plaza
plead
pleasant
please
pleasure
pledge
plenty
pliers
plot
plow
pluck
plug
plum
plumber
plus
plush
pluto
poach
pocket
pocketbook
poem
poet
poetry
point
poise
poison
poker
pole
police
policeman
policy
polish
polite
politician
politics

pollution
polyester
pond
ponder
pony
pool
poor
pop
popcorn
pope
popeye
popular
population
porcelain
porch
porcupine
pork
porpoise
port
portion
portrait
portray
position
positive
possess
possession
possibility
possible
possum
post
postage
pot
potato
potatoes
potential
pottery
pouch
pounce
pound
pour
poverty
powder
power
powerful
practice
prairie
praise
prank
pray
prayer
preacher

precious
precipice
precise
predator
predict
prefer
preference
prejudice
premises
prepare
prescription
presence
present
presentation
president
press
pressure
prestige
presume
pretend
pretty
preview
prey
price
priceless
prick
prickly
pride
priest
prim
primate
prime
prince
princess
principal
principle
print
printer
printing
prism
prison
prisoner
private
prize
pro
probability
problem
procedure
process
produce
producer

product	punishment	rainbow	record
production	pupil	rainy	recreation
profession	puppy	raise	recruit
professional	purchase	raisin	rectangle
professor	pure	rake	recycle
profit	purple	ram	red
program	purpose	ranch	reduce
prohibit	purse	range	reef
project	push	rank	reel
prom	pushy	rap	referee
promise	put	rape	reflect
pronoun	put_down	rapid	reflection
proof	puzzle	rare	refrain
proper	pyramid	rash	refreshing
property	q-tips	raspberry	refrigerator
proposal	quack	rat	refuse
proprietor	quail	rate	region
prose	quality	rationalize	register
prosecute	quantity	rattle	regret
prosper	quarrel	raw	regular
protect	quart	ray	regulation
protection	quarter	rays	regulations
protein	queen	razor	reindeer
protestant	quench	reach	reject
proton	quest	reaction	rejoice
protractor	question	read	relation
proud	quick	reader	relationship
prove	quiet	reading	relative
proverb	quill	ready	relativity
providence	quilt	real	relax
provision	quit	real_estate	relaxed
prune	quiver	reality	release
psalm	quiz	realize	reliable
psychology	quotient	reap	relief
pub	rabbi	reappear	religion
puberty	rabbit	rear	religious
public	raccoon	rearrange	relish
publication	race	reason	rely
publisher	rack	rebel	remain
puck	racket	recall	remark
pudding	racquet	receipt	remedy
puddle	racquetball	receive	remember
puff	radiator	recent	remind
puke	radio	receptionist	reminiscence
pull	radish	recipe	e
pump	raft	reckless	remorse
pumpkin	rag	recliner	remove
pun	rage	recognition	remover
punch	rail	recommend	renounce
punctuation	railroad	recommend	rent
punish	rain	ation	repair

repeat	ridicule	rough	same
repentance	ridiculous	round	sample
replace	rifle	routine	sand
reply	right	row	sandals
report	righteous	royal	sandpaper
repress	righteousne	royalty	sandwich
reprimand	ss	rub	santa
reproduce	rights	rubber	sap
reptile	rigid	ruby	sapphire
republic	rim	rude	sardine
republican	ring	rug	sash
repulsive	rinse	ruin	satan
request	riot	rule	satin
rescue	rip	ruler	saturn
research	rise	rules	sauce
reservation	risk	rum	saucer
reserve	ritual	rummy	sausage
resistance	river	rumor	save
resort	roach	run	savior
respect	road	rung	saw
respiration	roam	runner	saxophone
respond	roar	running	say
response	roast	rural	saying
responsibili	rob	rush	scale
ty	robber	russia	scales
responsible	robbery	rust	scallop
rest	robe	rye	scalp
restaurant	robin	saber	scalpel
restore	robot	sack	scapegoat
restriction	rock	sacred	scar
restroom	rocket	sad	scarce
result	rocks	saddle	scarcity
retain	rod	safari	scare
retarded	rodent	safe	scarecrow
retreat	rodeo	safety	scared
return	roll	sage	scarf
revenge	roman	sail	scary
review	romance	sailing	scene
revival	romantic	sailor	scenery
revolution	roof	saint	scent
reward	room	salad	schedule
rhyme	roommate	salary	scheme
rhythm	roost	sale	scholarship
rib	rooster	sales	school
ribbon	root	salesman	science
ribs	rope	saliva	scientific
rice	rose	salmon	scientist
rich	roses	saloon	scissors
riddle	rot	salt	scold
ride	rotten	saltine	scoop
rider	rouge	salute	score

scotch
scout
scramble
scrambled
scrap
scratch
scream
screen
screw
screwdriver
scribble
scroll
scrub
scuba
sculpture
scum
sea
seafood
seagull
seal
seam
search
seashore
season
seasoning
seat
seatbelt
secluded
second
secret
secretary
secretive
section
secure
security
sedative
see
seed
seek
seem
seize
seldom
select
selection
self
selfish
sell
seller
semester
senate
senator

send
sense
sensitive
sentence
sentry
separate
serene
sergeant
serial
series
serious
sermon
serpent
servant
server
service
session
set
seven
several
severe
sew
sex
sexy
shack
shade
shadow
shag
shake
shallow
shame
shampoo
shape
share
shark
sharp
shatter
shave
she
shears
shed
sheep
sheet
sheets
shelf
shell
shelter
shepherd
sheriff
shield
shift

shine
shingle
shiny
ship
shirt
shiver
shock
shoe
shoelace
shoes
shoot
shooting
shop
shopping
shore
short
shortcake
shorts
shot
shoulder
shout
shove
shovel
show
shower
shrimp
shrine
shrink
shrub
shun
shut
shutter
shuttle
shy
sibling
sick
sickle
sickness
side
sidewalk
siding
sight
sign
signal
signature
silence
silent
silk
sill
silly
silver

silverware
similar
simple
sin
since
sincere
sincerity
sing
singer
single
sink
sinker
sinus
sip
siren
sirloin
sister
sit
site
sitting
situation
six
size
skate
skeleton
sketch
ski
skill
skillet
skin
skinny
skip
skirt
skull
skunk
sky
skyscraper
slab
slacks
slander
slap
slave
slay
sleep
sleepy
sleeve
sleigh
slender
slice
slick
slide

slim
slime
slimy
sling
slip
slippers
slippery
slither
sliver
slob
slogan
slope
sloppy
slow
slug
slum
slumber
slurp
sly
small
smart
smash
smear
smell
smelt
smile
smog
smoke
smoking
smoky
smooth
smother
smudge
snack
snail
snake
snap
snatch
sneak
sneaker
sneakers
sneaky
sneeze
sniff
snob
snooze
snore
snorkel
snot
snotty
snow

snuggle
soak
soap
soar
sob
sober
soccer
social
society
sock
socket
socks
soda
sofa
soft
soft_drink
softball
soil
soldier
sole
solid
solo
solution
solve
some
something
sometimes
son
song
sonic
sonnet
soon
soot
soothe
sophisticate
d
sore
sorority
sorrow
sorry
soul
sound
soup
sour
south
southern
sow
space
spade
spades
spaghetti

span
spanish
spank
spare
sparrow
spatula
spawn
speak
speaker
special
specific
speech
speed
spell
spend
sphere
sphinx
spice
spicy
spider
spike
spill
spin
spinach
spine
spirit
spit
splash
splinter
split
spoil
spoiled
sponge
spontaneou
s
spook
spool
spoon
sport
sports
spot
spotless
spots
spouse
sprain
spray
spread
spree
spring
sprinkle
spy

squad
square
squash
squeak
squeal
squeeze
squid
squint
squirrel
stab
stable
stack
stadium
staff
stag
stage
stain
stair
stairs
stairway
stake
stale
stalk
stall
stallion
stamp
stand
standard
standards
standing
staple
stapler
star
starboard
starch
stare
stars
start
starve
starving
state
statement
station
statistics
statue
status
stay
steak
steal
steam
steel

steep
steeple
steer
stem
stench
step
steps
stereo
stereotype
stern
stethoscope
stew
stewardess
stick
sticker
sticky
stiff
still
stimulus
sting
stingy
stink
stir
stitch
stock
stocking
stolen
stomach
stone
stoned
stool
stop
stoplight
stopper
storage
store
storm
story
stove
straight
strain
strand
strange
stranger
strap
strategy
straw
strawberry
stray
stream
street

strength
stress
stretch
strict
strike
string
strip
stripe
stripes
strive
stroke
strong
structure
struggle
stub
stubborn
stuck
stuck_up
stud
student
study
stuff
stumble
stump
stupid
sty
style
sub
subject
subjective
submarine
substance
subtle
subtlety
subtract
subtraction
suburb
subway
succeed
success
succor
suck
suds
sue
suede
suffocate
sugar
suggest
suggestion
suicide
suit

suitcase
suite
sum
summary
summer
summit
sun
sunday
sunny
sunrise
sunset
sunshine
super
superior
superman
supermarke
t
superstition
supervisor
supper
supply
support
suppose
sure
surf
surface
surgeon
surgery
surprise
surrounding
surrounding
s
survey
survive
suspect
suspense
swabs
swallow
swamp
swan
swap
swatter
sway
swear
sweat
sweater
sweep
sweet
sweetheart
swift
swim

swimmer
swing
swiss
switch
swoon
sword
symbol
sympathetic
sympathy
symphony
symptom
synagogue
synopsis
syringe
syrup
system
tab
table
tablet
taboo
tack
tackle
tact
tactful
tag
tail
take
tale
talent
talk
tall
tame
tampa
tan
tangent
tangerine
tangy
tank
tantrum
tap
tape
tapioca
tar
tardy
target
tarnish
tart
tarzan
task
taste
tasty

tavern
tax
taxes
taxi
tea
teach
teacher
team
teapot
tear
tears
tease
technician
technique
technology
teenager
teeth
telephone
telescope
television
tell
teller
temper
temperature
temple
ten
tend
tendency
tender
tennis
tenor
tense
tension
tent
tepee
term
terminal
termite
terms
terrible
terrific
territory
terror
test
test_tube
testify
text
than
thanks
thanksgiving
g

thaw
theater
theft
theme
theorem
theory
therapy
there
thermomete
r
thesaurus
thesis
thick
thicket
thief
thigh
thimble
thin
thing
things
think
thinking
third
thirst
thirsty
thorn
thought
thread
threat
three
threshold
thrift
thriller
throat
throne
through
throw
throw_up
thumb
thunder
tick
ticket
tickle
tide
tidy
tie
tiger
tight
tile
till
timber

time
times
timid
tin
tiny
tip
tire
tired
tissue
title
toad
toast
toaster
toasty
tobacco
today
toe
toenail
toes
together
toil
toilet
toll
tomato
tomb
tombstone
tomorrow
ton
tone
tongue
tonic
tonight
tool
tools
tooth
toothbrush
toothpaste
top
topic
topping
torch
torn
tornado
tortoise
torture
toss
total
tote
touch
touchdown
touchy

tough
tour
tourist
tow
towel
tower
town
toy
toys
trace
track
tractor
trade
tradition
traffic
tragedy
trail
trailer
train
trait
traitor
tramp
trance
tranquil
transparent
transplant
transportati
on
trap
trash
trauma
travel
tray
tread
treason
treasure
treat
treatment
treble
tree
trees
trench
trend
trial
triangle
tribe
tribute
trick
tricycle
trigger

trigonometr	two	unsolved	vegetables
y	tylenol	unsure	vehicle
trim	type	untrue	vein
trip	typewriter	untruthful	velocity
triumph	typhoon	unused	velvet
trombone	typist	unusual	venom
trophy	ugly	up	vent
tropical	ulcer	update	venus
trot	umbrella	upset	verb
trouble	umpire	upstairs	verdict
trousers	unaware	upstream	verse
trout	uncle	uptight	vessel
truce	unclear	uranus	vest
truck	uncomforta	urge	vet
true	ble	urgent	veteran
trumpet	uncommon	usa	vice
trunk	unconsciou	use	victim
trust	s	used	victor
trustworthy	uncooked	useless	victory
truth	undecided	usual	video
truthful	under	usurp	view
try	undergroun	utensil	viking
tub	d	vacancy	village
tuba	understand	vacate	villain
tube	understandi	vacation	vine
tuesday	ng	vacuum	vinegar
tug	underwater	vagrant	viola
tulip	underwear	vague	violation
tumble	uneasy	vain	violence
tumor	unemploym	valentine	violent
tuna	ent	valley	violet
tune	unequal	valor	violin
tunnel	uneven	valuable	virgin
tupperware	unfair	value	virtue
turkey	unhappy	values	virus
turn	unicorn	valve	vision
turnip	uniform	vampire	visit
turtle	union	van	visitor
tusk	unique	vanilla	vitamin
tutu	unit	vanish	vivid
tuxedo	unite	vanity	vocabulary
tweezers	united	vapor	vodka
twelve	universe	variety	voice
twice	university	vary	void
twig	unjust	vase	volcano
twilight	unknown	vaseline	volleyball
twin	unlawful	vast	volt
twine	unload	vault	volume
twinkle	unnatural	veal	volunteer
twist	unpleasant	veer	vomit
twister	unseen	vegetable	voodoo

vote	wax	widow	worship
vow	way	width	worst
voyage	weak	wife	worth
vulgar	weakness	wiggle	worthless
vulnerable	wealth	wigwam	wound
waffles	weapon	wild	wrap
wag	wear	wilderness	wreck
wage	weary	will	wrench
wager	weather	willing	wrestling
wagon	weave	wimp	wrinkle
waist	web	win	wrist
wait	wed	wind	write
waiter	wedding	winding	writer
waitress	wednesday	window	writing
wake	weed	windshield	wrong
walk	week	wine	xerox
walker	weekend	wings	yacht
walkway	weekly	winner	yard
wall	weep	winter	yarn
wallet	weigh	wipe	yawn
walnut	weight	wire	year
walrus	weights	wisdom	yearly
wand	weird	wise	yearn
wander	welcome	wish	years
want	welfare	wit	yell
wanted	well	witch	yellow
war	well-being	with	yes
wardrobe	west	withdrawn	yesterday
warehouse	wet	without	yield
warm	whale	witness	yogurt
warmer	what	wolf	yoke
warmth	wheat	woman	yolk
warn	wheel	women	you
warning	when	wonder	young
warrior	where	wonderful	yours
wart	whip	wood	youth
wary	whiskers	woods	yo-yo
wash	whiskey	woodwind	yuck
washcloth	whisper	wool	yummy
washer	whistle	word	zebra
wasp	white	words	zenith
waste	who	work	zero
wasted	whole	worker	zest
watch	whom	workout	zipper
water	why	world	zit
waterfall	wick	worm	zone
watt	wicked	worn	zoo
wave	wicker	worry	zucchini
waves	wide	worse	

APPENDIX B: STIMULI PAIRS USED IN EXPERIMENTS 1 AND 2

Note: Primes are in lowercase and targets are in uppercase

WORD TARGETS

helsen ABSENCE
 demigns
 ACADEMIC
 anilety ACADEMY
 accesed ACCEPT
 effest ACCOUNTS
 dreems ACCURACY
 achioered ACCUSED
 reptiod ACHIEVED
 acit ACID
 quited ACQUIRED
 saitor ACTOR
 onts ACTS
 bams ADAM
 damane ADAMS
 mopends
 ADEQUATE
 perfect ADMIT
 snit ADMITTED
 voited ADOPTED
 ouction ADOPTION
 adelts ADULTS
 advine ADVANCED
 favilite ADVISE
 arfented AFFECTED
 arenda AGENDA
 apers AGES
 agraid AGREED
 hade AHEAD
 aiss AIDS
 permet AIRPORT
 aparm ALARM
 blaims ALBUMS
 laten ALIEN
 avents ALIENS
 healthly ALIVE
 balley ALLEGED
 orat ALLOW
 alcohol ALLOWED
 loling ALLOWING
 fallows ALLOWS
 amoung AMOUNT
 boaded ANALYSIS
 calarate ANALYST

amening ANIMALS
 slors ANSWERS
 antenne ANTENNA
 anymoty
 ANYBODY
 anymare
 ANYMORE
 anymaty
 ANYWAYS
 amparint APART
 apare APPEAR
 empecred
 APPEARED
 dears APPEARS
 victery APPLYING
 authed APPROVED
 anchive ARCHIVES
 gree ARGUMENT
 wartiong ARMOR
 hars ARMS
 roive ARRIVE
 artinled ARTICLES
 kart ARTS
 wearts ASPECT
 abtaived ASSIGNED
 shume ASSUME
 prining ASSUMING
 ethnis ATHEIST
 catuar ATTACK
 antanled
 ATTACKED
 aptacts ATTACKS
 antimpts
 ATTEMPTS
 patends ATTEND
 buttere ATTENDED
 cuncrion ATTITUDE
 cortion AUCTION
 praduage
 AUCTIONS
 abjnce AUDIENCE
 utes AUDIO
 auctiong AUTHOR
 autes AUTO
 sedual AWFUL
 whead AWHILE

batues BABIES
 bably BABY
 weirt BACKED
 plack BACKUP
 baly BADLY
 objests BAGS
 elliance
 BALANCED
 kale BALL
 tons BANDS
 cang BANG
 diceates BANK
 gans BANKS
 basitell BASEBALL
 babic BASIC
 cass BASS
 bave BATCH
 bayer BATHROOM
 detters BATTERY
 arlept BATTLE
 lelis BEAM
 reast BEAST
 ameme BECAME
 comededy BECOME
 beying BECOMING
 beyong BEDROOM
 derree BEER
 bems BEGINS
 allohel BEHAVIOR
 bevefit BELIEF
 beal BELL
 belonds BELONGS
 frest BELT
 bepht BEND
 impatial BEYOND
 birs BIAS
 horrible BIBLE
 prives BIDDING
 egge BIGGER
 bilps BILLS
 bary BINARY
 bloos BIRD
 marthday
 BIRTHDAY
 burts BITS
 leoples BITTER

clame BLADE
ppeased BLESSED
aclack BLOCKS
clow BLOW
boit BOLT
pone BONE
roparted BOOT
bont BOOTS
borded BORDER
prodened BORING
sorn BORN
chought BOUGHT
lound BOUND
joal BOWL
bods BOXES
onti BREAKING
hought BRIGHT
istome BROTHERS
goutry BROUGHT
nowsere BROWSER
diarts BUCKS
buty BUDDY
falm BUFFALO
bube BUGS
poarding BUILDING
milker BUILT
bult BULK
crield BURIED
buro BURN
cotteder BURNED
fornet BURNS
tuly BUSY
utet BUTT
bute BUTTON
roverty BUTTONS
bumer BUYER
cate CAGE
gake CAKE
buflalo CALLS
caber CAMERA
mamp CAMP
campaign
CAMPAIGN
opetacol CANCEL
hance CANCER
cande CANDY
kalk CAPABLE
butron CARBON
kare CAREFUL
healily CARRY
cloor CATEGORY

gattened CAUSED
mells CELLS
stuss CENSUS
guent CENT
crerinal CENTRAL
chaims CHAIRMAN
dedde CHAMPION
tange CHANGES
chanter CHANNEL
chansed
CHANNELS
warter CHARGE
shecked CHECKED
precks CHECKS
bess CHESS
macked CHICKEN
crief CHIEF
richers CHOICES
hosa CHOOSE
chome CHOSEN
cintle CIRCLE
epitides CITIES
citian CITIZEN
filizess CITIZENS
clanses CLAIMED
brool CLAIMING
celping CLAIMS
fusicad CLASSIC
anvoples CLEAN
plernang
CLEANING
claiteng CLIENT
fatients CLIENTS
hecrer CLINICAL
crecks CLOCK
humang CLOSED
rolir CLOSER
clet CLUE
toack COACH
boch COCK
spors CODES
bink COIN
cond COLD
allerer COLLEGE
recoll COLUMN
clained COMBINED
dycle COMEDY
coritics COMICS
compage COMPARE
explain COMPLAIN

comprain
COMPLETE
producy COMPLEX
conterts CONCEPTS
consered
CONCERNS
anticate CONDUCT
condumir
CONFLICT
contrens CONSIDER
conserer CONSOLE
centents
CONTACTS
conthant
CONTAINS
fonteng CONTENT
goarts CONTENTS
contraty CONTEXT
contion CONTINUE
contrels
CONTRARY
conclare
CONVINCE
papler COPPER
corcers CORNER
cerrupt CORRUPT
counte COUNCIL
sont COUNT
quyet COUNTER
hounts COURTS
choser COVER
bullshec
COVERAGE
traffec COVERING
trat CRAFT
frash CRASH
preatle CREATE
pleating CREATING
creme CREATURE
cret CREDIT
cree CREW
cresir CRISIS
bross CROSS
approvon CROWN
cardning CRYING
stital CRYSTAL
cestral CULTURAL
salk CULTURE
reck CURE
courtais CUTS
divint DAILY

dalt DALE
daunched DANCE
recart DARE
paicted DATED
dinerles DEALERS
dibegals DEALS
adeal DEALT
mexas DEBATE
gebt DEBT
decady DECADE
deachens DECENT
ancluded DECIDED
devime DECISION
realured
DECLARED
beaped DEFEAT
affense DEFENSE
buats DEFINED
clade DELAY
reserte DELETE
reliver DELIVER
twils DEMAND
semo DEMO
rerand DEMON
clatined DENIED
boms DENY
pleply DEPENDS
renchibe DESCRIBE
besere DESERVE
servee DESERVES
resiasas DESIGNED
anymody
DESIGNER
dedure DESIRE
desimed DESIRED
besk DESK
falt DESKTOP
initiol DESPITE
destite DETAILS
vilys DEVIL
mapical DIAL
criol DIED
brient DIET
difrer DIFFER
preatire DIRECT
rececter DIRECTED
soundict DISCOUNT
foverty DISCOVER
caised DISEASE
namer DISEASES
dist DISH

occesior DISORDER
distray DISPLAY
diviny DIVIDED
plave DIVINE
disiving DIVISION
pomerial
DOMESTIC
fors DOORS
anymary DOUBLE
fowarts
DOWNTOWN
dronen DOZEN
drat DRAG
bevidit DRAW
dalning DRAWING
discurs DRAWN
droas DREAMS
minerted DRESSED
hamner DRIVEN
critring DRIVING
grosing DROPPING
shounded DROPS
shacts DRUG
drut DRUM
shunk DRUNK
pums DUMB
dums DUMP
huty DUTY
blying DYING
feaching EARN
carning EARNINGS
seak EASE
sead EAST
mattern EASTERN
bealm ECHO
mamicad
ECONOMIC
manch EDITING
emitiom EDITION
edater EDITOR
renectly EFFECT
offerer EFFECTS
effonts EFFORT
affors EFFORTS
eltert EIGHT
electred ELECTRIC
motrods EMOTIONS
apper EMPLOYER
respend ENDS
retien ENEMIES
senger ENERGY

sensines ENGINES
ruty ENJOY
envolved ENJOYED
anjoyint ENJOYING
preeding
ENTERING
enjudies ENTITIES
envalled ENTITLED
interent ENTRY
scake ESCAPE
poterial ETERNAL
sporen EVENING
extrele EVIDENCE
ecisede EXERCISE
stemion EXISTED
exid EXIT
anted EXPAND
patiens EXPENSE
esposes EXPENSES
elpert EXPERT
extarnas EXPLAINS
exprared
EXPLORER
filps EXTRA
rouns EXTREME
claces FACES
seaght FACILITY
smalner FACTS
failury FAILURE
fars FAIR
ricair FAIRLY
porries FAITH
crosint FAKE
holls FALL
asive FASHION
mespend FASTER
chiss FATE
lomont FEATURE
amater FEATURED
eatlies FEATURES
blyint FEED
beed FEEDBACK
lenks FEELINGS
fies FEES
horor FEET
dife FELL
figited FIGHTERS
cooming FIGHTING
filt FILL
birt FINAL
witcting FINDING

stanks FINISHED
fibrary FIRM
formus FIRMS
festacal FISCAL
firds FIXED
slame FLAME
lorer FLAMES
innoment FLAMING
foms FLUID
sirus FOCUSED
wilk FOIL
foly FOLLOW
forlow FOLLOWED
faltows FOLLOWS
gont FONT
knoch FOOL
relisted FOOLS
forcure FORCED
fiter FOREVER
forgen FORGOT
dinstly FORMER
noter FORMERLY
foram FORMULA
steres FORT
foir FORUMS
porward FORWARD
chounded
FOUNDED
roors FOUNDER
crag FRAUD
greak FREAK
freadom FREEDOM
strit FRUIT
freele FULLY
foction FUNCTION
flot FUND
cound FUNDS
gainen GAINED
gley GALLERY
gant GANG
sheatire
GATHERED
geat GEAR
gass GIANTS
logis GIRLS
gaicen GIVES
givint GIVING
gled GLEN
glary GLORY
boal GOAL
damaced GOLD

gork GORE
tosle GOSPEL
borm GOTTEN
vover GOVERNOR
contint GRAB
rader GRADE
larguate
GRADUATE
chaphict GRAPHICS
pratiful GRATEFUL
ritty GRAVITY
archades GREATLY
caysins GRIN
gron GROUND
graus GROUNDS
morting GROWING
bisk GROWN
goards GUARD
acadede GUIDE
guils GUILT
culf GULF
juns GUNS
halit HABIT
ralt HALL
andind HANDED
gnowing
HANDLING
harrer HARDER
rarm HARM
shoted HATED
heague HEADED
apheads HEADS
moling HEALING
omployes
HEALTHY
heants HEARTS
weat HEAT
healile HEAVILY
helay HEAVY
micked HECK
bendle HELPING
bero HERO
pecers HEROES
herserf HERSELF
reperts HIGHER
wigher HIGHEST
pighway HIGHWAY
chint HINT
chir HIRE
hites HITS
droting HOLDING

stoad HOLDS
lole HOLE
shots HOMES
hingest HONEST
hane HONEY
anpor HONOR
hoon HOOD
toporsow HOPING
horing HORN
spiority HOSPITAL
shoughts HOUR
souse HOUSES
huse HUGE
humarg HUMOR
emare HUNTER
dointing HUNTING
airchart HURT
burth HURTS
plade IDEAL
idear IDEAS
imazing IDENTIFY
adiet IDIOT
ignaring
IGNORANT
illners ILLNESS
ttavel IMAGE
animas IMAGES
imple IMPLY
recome INCOME
creames INCREASE
anening INDEED
decate INDICATE
intinted INFINITE
anferved
INFORMED
protonal INITIAL
purpries INJURIES
purfint INPUT
insilt INSPIRED
enstonce INSTANCE
instatt INSTANT
jalm INTENT
ontering INTERNAL
entinted INVENTED
istued ISSUED
essues ISSUES
bava JAVA
bazz JAZZ
fets JETS
joat JOBS
joid JOIN

protens JUDGES
dest JUMPED
jumpeng JUMPING
tunk JUNK
accept KEPT
lomins KEYBOARD
kney KEYS
sidding KIDDING
asleer KILLER
flonk KILLING
dength KITCHEN
rives KNIGHT
antiming KNOWING
labic LABOR
vady LADY
jaid LAID
slake LAKE
plame LAME
hintreds LANDING
lange LANE
ledies LARGEST
cought LAUGH
marters LAWYERS
seaming LEADING
leadens LEADS
wealts LEAF
balele LEAVES
imagide LEAVING
gace LEGACY
councus LEGENDS
glorn LEGS
raceng LENGTH
suggent LESSONS
watter LETTER
deaters LETTERS
burtint LETTING
loar LIAR
enitals LIBERAL
lemels LIBERALS
recensed LICENSED
slet LIES
diek LIKES
binks LINKS
seather LISTED
sisten LISTEN
drating LISTING
irns LOAD
poas LOCATION
slock LOCK
losimal LOGICAL
lool LOGO

looth LOOP
lolo LOOSE
lorind LORD
gosper LOSER
rall LOSS
rosses LOSSES
poud LOUD
funce LOVELY
palking LOVING
loil LOWER
leck LUCKY
dably MAINLY
daictain MAINTAIN
sarlds MAKER
editeng MANAGE
meniged
MANAGED
mealing
MANAGING
ander MANNER
paps MAPS
mixas MARKETS
arks MARKS
meseerts MASTERS
doment MATE
minimam
MAXIMUM
meak MEAL
preating MEANING
sloime MEASURED
preasure
MEASURES
meni MEDIA
redium MEDIUM
senting MEETING
preeing MEETINGS
teets MEETS
memorior MEMORY
mens MENU
cheamer MERCURY
fequises MERCY
burely MERELY
mewes MESS
feasudes
MESSAGES
minate METER
breet METHOD
celds METHODS
acomomic MICRO
bindle MIDDLE
bire MILE

minumam
MINIMUM
mintakir MINISTER
minaty MINORITY
easies MINUTE
bise MISS
meighber MISSED
mesk MISSING
refeng MOBILE
momede MODE
moder MODEM
odel MODERN
freatem MODIFIED
mentied MOMENT
moner MONKEY
coods MOOD
damon MOON
morar MORAL
mouthern MOTHER
crir MOUNTAIN
mouce MOUSE
momer MOUTH
roys MOVES
bulti MULTI
pliple MULTIPLE
mystere MYSTERY
dush MYTH
aded NAMED
cenation NATION
pations NATIONS
pateral NATURAL
penally NEARLY
teld NEAT
senk NECK
fervous NERVOUS
mewer NEWER
eass NINE
orme NORM
bure NORMALLY
porth NORTH
resuse NOSE
jotes NOTES
doven NOVEL
wens NUDE
sunkers NUMBERS
nits NUTS
oblect OBJECT
shevious OBVIOUS
sesestly OCEAN
doid ODDS
giews OFFENDED

cospared OFFERED
ondering OFFERING
officar OFFICER
fecused OFFICERS
itit OFFICES
solde OLDER
doir OPERA
speater OPERATOR
apent OPPONENT
purposel OPPOSED
omposate
OPPOSITE
relition OPTION
argaring
ORDERING
rinner ORDINARY
morint ORIGIN
sorth ORTHODOX
oetput OUTPUT
drant OWNS
demagns
PACKAGES
larint PAINT
radly PAINTED
sanary PARADISE
heaten PARENTS
vask PARK
deecer PARKING
baties PARTIES
oass PASS
denation PASSION
asite PASTE
imates PATCH
pachion PATH
cathoric PATHETIC
datience PATIENT
patterts PATTERN
pategore
PATTERNS
admented
PAYMENT
meen PAYMENTS
dorain PAYS
reace PEACE
peavo PEAK
nearle PEARL
jort PENALTY
pelish PENIS
sespovec PEOPLES
mopper PEPPER
paiment PERCENT

phicter PERFECT
croar PERFORM
perget PERMIT
sosenay PERSONAL
nipe PHASE
denile PHOENIX
papture PICTURES
keecer PIECE
dile PILE
acedecis PLACED
praces PLACES
viract PLASTIC
prale PLATE
placts PLATFORM
beas PLEASANT
collent PLENTY
pelisies POLICIES
caurts POLITICS
scrir POPE
congict POSSIBLY
fosiety POSTER
sounsict POUND
toas POUR
canks POVERTY
prerinte PRACTICE
bary PRAY
shreats PRAYERS
miract PREFER
phemier PREMIER
daymints PREMIUM
propente PRESENT
peasure PRESSURE
sprit PRIDE
privent PRIEST
pristes PRINCESS
caid PRINTED
panter PRINTER
privary PRIVACY
vireos PRIZE
bammy PRODUCER
provin PROFIT
culture PROJECTS
promise PROMISE
promory PROMO
profen PROOF
promest PROPERTY
printess PROTEIN
skypon PROTEST
promonal
PROTOCOL
emproved PROVE

rerveng PROVEN
revided PROVIDES
cestount PUBLICLY
shull PULL
forn PUMP
pistion PUNCH
fonk PUNK
pustace PURCHASE
brear PURELY
burple PURPLE
curpomer
PURPOSES
dule PUSH
wintting PUTTING
fuackly QUICKLY
duit QUIT
scoted QUOTED
rane RACE
drace RACES
macial RACIAL
daving RACING
nacise RACISM
dention RAGE
bies RAIN
odet RAISE
answens RAISED
reasing RAISING
strange RANGERS
mank RANK
raren RARELY
rapeates RATES
praiting RATING
sark RATIO
swipring
REACHING
cread READS
peady READY
borked REALIZED
dreal REALM
bersons REASONS
recial RECALL
receite RECEIVE
prerects RECENT
recest RECIPE
feasuded
RECORDED
recourle RECORDS
predurer REDUCE
jefer REFER
rellert REFLECT
ramorm REFORM

renuse REFUSE
duss REFUSED
phurthed REGIME
relision REGION
retional REGIONAL
cround REGISTER
serter REJECTED
renate RELATE
rewain RELATION
receate RELATIVE
melele RELEASE
releored RELEASED
antirele RELEVANT
rention RELIGION
cortaing REMAIN
endind REMIND
repends REMINDS
retaiued REMOVED
recoring
REMOVING
renete RENT
rair REPAIR
retealed REPLACE
beplaged
REPLACED
reptied REPLIED
drepring REPLYING
recomed
REPORTED
rasublis REPUBLIC
suprents REQUEST
becuired
REQUIRED
riscue RESCUE
meseercs
RESEARCH
prerered
RESERVED
resiveng RESIDENT
delle RESULT
resund RESUME
perele REVEAL
extreue REVEALED
relearned REVERSE
peard REWARD
bice RICE
rish RICH
prizy RICK
rade RIDE
treder RIDER
ciled RIFLE

inet RING
bringe RINGS
sise RISE
tolss RISKS
kiver RIVER
mobire ROBIN
fogue ROGUE
totalty ROLL
roir ROOF
coors ROOMS
joot ROOT
mortly ROUGHLY
affent ROUND
brinch RUDE
bule RULE
recing RULING
rune RUNS
jank SADLY
seyint SAINT
tarner SALARY
salls SALES
acalyses SAMPLES
dave SAVE
scane SCAN
acquined SCARED
periods SCENARIO
stees SCENES
scroons SCHOOLS
shience SCIENCE
stalnel SCREW
sealer SEALED
seath SEAT
spant SEATS
onds SECONDS
specier SECRET
rellect SECTION
meed SEED
sesk SEEK
staned SENATE
repataty SEPARATE
tervisal SERIAL
rascs SERIOUS
sherever SERVE
skes SERVES
meserve SERVICES
rinagers SERVING
housins SESSIONS
jeel SETS
prowing SETTINGS
murded SETTLE
jerte SETUP

reavins SEVERE
aseal SEXUAL
shais SHADOW
shaor SHAPE
plared SHARED
shars SHARES
stope SHEET
suttints SHIP
shase SHIPPED
soop SHOP
mottes SHOULDERS
loas SHUT
sigs SIGH
rits SIGHT
quactity SIGNED
shirs SILLY
bilter SILVER
bing SING
mins SINS
linter SISTER
sirte SITES
killy SKILL
skir SKIN
oppated SKYBOX
slare SLAVE
liveral SLAVERY
seeint SLEEPING
clide SLIDE
slothe SLOT
stolk SLOW
lold SLOWLY
stalner SMALLER
slell SMELL
smest SMOKE
stardens SNAKE
snin SNIP
finthly SNOW
soan SOAP
sortival SOCIAL
soin SOIL
centein SOLELY
solle SOLVE
arte SOMEBODY
priting SOMEDAY
somehor
SOMEHOW
metional
SOMETIME
midding SORTS
stedint SOUNDING
lare SPARE

bleakers SPEAKER
seater SPEAKERS
reating SPEAKING
rencrabe SPECIES
sheech SPEECH
rinaters SPELL
aluminus SPELLING
cints SPEND
metrods SPENDING
stedints SPLIT
spole SPOKE
selutien SPOKEN
smots SPOTS
cormarts SPRAY
arlowes SPRINGS
squady SQUARE
sungery STANDING
specifes STANDS
tars STARS
sterio STARTING
burs STARTS
tosted STATED
scrates STATUS
rosteres STEP
storer STONE
chorite STONES
ventuce STORED
lories STORIES
daighted STRAIGHT
streed STREET
streem STREETS
scror STRENGTH
rett STRETCH
tricts STRICTLY
loring STRING
stortred STROKE
stror STRONGER
cuimanle
STRUGGLE
pesty STUCK
antedent STUDENT
fart STUDIED
stedicts STUDIES
studint STUDIO
stuly STUDY
reesend STYLE
arterts SUBJECTS
spedier SUDDEN
raugh SUGAR
intedeng SUGGEST
sult SUIT

botworts SUPPORTS
puprise SUPPOSE
durger SURGERY
kurs SURPRISE
surpory SURVEY
shint SURVIVE
foan SWAP
stear SWEAR
tweet SWEET
erable TABLE
rales TABLES
actack TACTICS
sigures TAIL
taly TALE
recerted TALENTED
talt TALKS
thonks TANKS
augunt TAPE
haught TAUGHT
dersus TAXES
toach TEACH
reachess
TEACHERS
bueen TEACHING
peacs TECH
cightens TEND
ingiring TERMINAL
parms TERMS
lertinle TERRIBLE
siver TERRY
stanting TESTING
tade THEE
mench THEME
patecore THEORY
cherapy THERAPY
hiti THINKS
gought THOUGHTS
thousars
THOUSAND
threass THREADS
corth THROW
loys THUS
liticed TICKETS
idet TIED
giarts TIES
pright TIGHT
tild TILL
moming TIMING
tirm TIRE
tetes TITLES
coll TOLL

carron
TOMORROW
tane TONE
vemerang TONIGHT
tolded TOOL
stull TOUR
fowards TOWARDS
werve TOWER
merls TOYS
tactins TRACKING
delping TRADED
corticts TRADING
tain TRAIL
droving TRAINING
helied TRAP
talel TRAVEL
attimpt TREAT
preated TREATED
tret TREK
strir TRIES
slin TRIP
pripole TRIPLE
tolk TROLL
ruke TRUCK
tude TUNE
brotter TURNED
wringing TURNING
turry TURNS
egly UGLY
unarle UNABLE
unerle UNCLE
uninue UNIQUE
concent UNIT
indectly UNLIKELY
junt UNTO
updates UPDATED
dier UPDATES
roparded UPGRADE
duan UPPER
aneless USELESS
disher USER
aliels VALLEY
olave VALVE
mary VARY
creaning VENTURE
aistort VICTOR
suty VICTORY
veedbecs VIDEOS
ratienas VIEWS
miremess VILLAGE
dentury VINYL

bident VIOLENT
difres VIRUS
boise VISA
sille VISIBLE
shon VISION
votel VOTE
caited WAITED
warker WAKE
dealured WALK
comfint WALKER
daltin WALKING
humenite WALL
werle WALTER
silleng WANTING
waro WARD
wark WARM
eledents
WARRANTY
warrion WARRIORS
jash WASH
witched WATCHED
valed WAVE
weaght WEAR
meating WEARING
weirs WEIRD
sisteng WEST
shateved
WHATEVER
enshered
WHENEVER
sheteved
WHEREVER
disiring WHINING
quyer WHITES
moartese WIDE
fenk WIFE
inazing WILLING
sind WIND
dossing WINGS
gins WINS
unirens WIRELESS
distaded WISDOM
wister WISE
wrates WITCH
sint WONT
condrome WOODS
seits WORKER
dorkee WORKERS
purprine WORRIED
dury WORRY
warger WORSE

wonship WORSHIP
torst WORST
morth WORTHY
snun WOUND
preeting WRITING
ards YARDS
yellor YELLOW
cazz YOUNGER
sero ZERO
fone ZONE

**NONWORD
TARGETS**

reartion ABERTION
racice ABUTE
accommate ACCOINT
stast ACCURDED
atheast ACHEAST
ecirions ACHIONS
kask ACHIORER
clou ACLO
commin
ACOMEMIC
antering ACTENG
factory ACTER
acisede ACTIRETE
dars ADARS
metes ADEAN
acean ADEAR
dade ADET
audierce ADIER
admew ADMER
almot ADMOT
envation
ADNATION
alention ADNITION
antanted
ADVANTED
affinted ADVINTED
adien AFFELS
affelt AFFEST
arficess AFFICESS
censent AFFINT
figsted AFFISTED
affitian AFFITIAL
adentity AGENTION
anents AGENTY
collele AGLE
twiss AIRCHANT
apleated ALATE
horkey ALITUATE
olmer ALMER
dolumint
ALUMINUL
elywhere
ALYWHERE
amplaint
AMPURINT
panch ANCH

ancluder
ANCHADES
answent
ANCTANGE
spocks ANCUSACY
anged ANELYST
oriedted ANILETE
antels ANITALS
tende ANJOY
adcorate
ANSODATE
antusacy
ANSOMBLY
ansodute ANSURE
antack ANTANT
anleated
ANTEPTED
engines ANTINEER
antued ANTINLED
actirety ANTIRE
furation ANVATION
assense APPED
fashion APTIONS
arders ARDENS
rets ARELTS
harger ARLER
jarm ARME
senks ARNS
arloged ARPOGED
artica ARRICA
arenue ARTENNE
crots ARTISPT
anthive ARTIVE
arante ASANT
asage ASARE
usefus ASEFES
anemy ASEFUS
reasers ASERS
reekend ASINE
asat ASIT
patherts ASPERTS
advence ASSENCE
easteng ASTEND
anterney
ASTERNE
asan ASTUAN
astuer ASTUED
footbact ATAT
astent ATET
datioved ATNORED

antacted
ATTATHED
attinder ATTINDED
amoult AUGULT
kuss AUSS
nuather AUTHER
mapamer AWFOR
fent BAFELY
singler BAGE
tenight BAITH
bolic BALIC
bant BALT
beel BASIBELL
brast BAST
bover BAVER
belay BEACK
blaw BEACS
decoring BECONG
beants BEINTS
matir BEMILER
lemils BEMILES
cuffedced BEPTS
bert BERK
becroom BERROOM
bere BESURE
besite BETE
bodier BEVEDIT
dess BEYS
buat BIAT
biberty BIBENTY
suackly BIBLY
kird BICRY
biek BILF
milly BILLY
kiler BILMS
biet BILT
arents BINTS
duse BIRTS
cistount BISSENG
cernedy BLAIR
beap BLARS
doar BLOAR
loase BLOAS
woach BOACH
boster BOADER
boan BOAK
boad BOAM
boip BOID
elloance BOIVE
clayoff BOLF
movent BOMENT

brols BOOF
 outlook BOOP
 boke BOPE
 bors BORD
 orte BORT
 bosite BOTE
 bosle BOUBLE
 gour BOUR
 bugar BRAH
 breat BREAL
 breek BREELE
 reratere BRERTER
 briant BRIAL
 everince BRINCE
 ticcets BRINCT
 bint BRINTE
 bret BRIT
 eludents BUBEN
 sudgrent BUDE
 bumb BUGAL
 sultom BULER
 bueer BUMNER
 abume BURRE
 buse BUSELY
 lade CADE
 caiser CAIPES
 suoty CAJOVITY
 clarle CALE
 legind CALENTED
 colk CALK
 calenter CALLENT
 soir CALOVITY
 cact CALT
 comeres CAMERES
 amplain CAMPAINS
 fans CANTS
 cerefal CAREFAL
 calleng CARLENG
 capalle CATHERIC
 cosathan
 CATHINAS
 acan CAWN
 cardling CEADING
 ceats CEADS
 burply CEALILE
 edampred CEARD
 cactured
 CECTURED
 conters CENTENGS
 sepply CEPPLY
 rosiced CERTLE

losis CESTS
 crake CHAKE
 chale CHAME
 voites CHANLINT
 shair CHAPHICS
 chartes CHARE
 crade CHARP
 fless CHARS
 entarnas CHAS
 chey CHAY
 chield CHEAD
 aspeel CHEERLY
 redier CHOEDER
 chosed CHOITED
 prosical CHOSICAL
 cackaged CHOSS
 crot CHOT
 hout CHOU
 uncly CHULY
 chue CHUS
 clut CHUT
 pightant CIGHTINS
 connacte
 CINNACTE
 dirtince CINTINUE
 both CITHOD
 mitizen CITIEN
 clir CLAIR
 crained CLAISED
 nolody CLARSIT
 clabs CLATS
 claple CLAZE
 dycre CLIDY
 caiming CLIMING
 cair CLINT
 crip CLIT
 protuan CLOAD
 costings CLOATING
 loos CLOOSE
 colars COADS
 treme COCTRIRE
 cegime COGIC
 tole COLE
 cuimable COLLILE
 colding COLTINTE
 cormence COMEDE
 cherinal COMERIAL
 mogir COMORING
 pompheX
 COMPHEX

compase
 COMPLATE
 collite CONCLATE
 concenn
 CONCRING
 conttict CONDUCT
 mantas CONTANLY
 coghting CONTENG
 concers CONTERED
 coucher
 CONTHERN
 contett CONTRINT
 corked COREM
 cormince
 CORMANTS
 corminly
 CORMINCY
 goverver
 CORPOMER
 cortents CORRICTS
 cortrire CORTAINT
 cosper CORTER
 tren CORTERT
 coartese CORTEST
 carkniss CORTNESS
 wats CORTS
 eserage COSERAGE
 bosk COSK
 hothom COTHOD
 mothod COTHOM
 pafte COTTEN
 bolc COUNCED
 councer COUNCUL
 pountins
 COUNTINS
 critedio CRAITED
 crols CRALS
 fortise CRAME
 rars CRANT
 crayer CRARE
 crard CRARS
 increate CREADE
 rearter CREATER
 rere CREERE
 crenk CREN
 cerpomer CRIOR
 octesior CRITEGIO
 nelsen CRIVENT
 criof CROOF
 shoor CROOR
 crouble CROPARLE

fruis CROUGHT
 prucs CRUCS
 tosh CTOLS
 cooss CTOSS
 cottom CULTOM
 melicery
 CUNDWARE
 suprise CURPRINE
 lawyed DABLE
 ahean DACKAGED
 dapacite DAGAZIRE
 sturio DAILE
 relieved DAIVE
 dulk DALK
 famities DAMATIES
 domned DAMNER
 catury DAMOR
 dans DANT
 essees DAPEATES
 jass DASS
 treant DEANT
 demp DEAP
 keas DEAS
 diteates DEATENS
 recuraty DECATY
 seecer DECECTER
 difetire DECIRE
 desaived
 DECUIVER
 rede DEDE
 defeel DEFENG
 devited DELIGRED
 celumn DELUMINT
 mecame DEME
 degs DEMS
 dinerves
 DENERRED
 anywaus DENSUS
 replate DEPLATE
 epply DEPPLY
 lerated DERITED
 oppoted DERSERF
 harple DERVOUS
 resease DESEASE
 decensed DESENSE
 devire DESILE
 dere DESKTOR
 shome DESTILE
 shant DESUASE
 aletuate DESUATE
 slen DEVIEN

ceapon DIAMENG
 duas DIAS
 falarate DILATATE
 bananted
 DINANTED
 doncing DINCING
 davins DINS
 dinteng DINTING
 diat DIRACT
 sispers DISCERS
 siss DISS
 postade DISTADES
 souce DISTTINT
 dets DITS
 divence DIVERCE
 cilizess DLEEMS
 dand DOMANE
 abention DONTION
 toop DOOP
 doat DOOT
 dorems DORAIS
 dettors DORNTORS
 doren DORON
 dorned DORREE
 porms DORUMS
 coster DOSTER
 derand DRAN
 doap DRAP
 protind DREGOD
 fliming DREPTING
 droven DRIGGER
 arloss DROT
 duar DUAT
 tirk DURK
 duts DURS
 lear EARD
 bartons EARTINGS
 easine EASICY
 deat EATE
 station EATINT
 factor ECENOMY
 eccesed ECUCATED
 edistes EDITER
 seddinly EDVINY
 buicting EGICTING
 dilence ELATE
 letion ELECIONS
 drecod ELECTHED
 eltra ELTRE
 eprty ELTRY
 semale EMANE

berivion EMITION
 amounds
 EMOUNDS
 smaced EMPACTED
 expensed
 EMPECTED
 exack ENACK
 bunchase ENCH
 seached ENCHES
 enerle ENERLY
 legar ENGARE
 engace ENGLE
 enitiol ENITIOM
 ansault ENSAULT
 easied ENSPITED
 mentival ENTI
 antive ENTIVE
 enterad ENTRA
 equarly EQUARLE
 vireas EQUAS
 orarly ERARLE
 rerle ERERLE
 demense ERINA
 tround ESTAPE
 edernal ETHNIR
 coveming EVERING
 coltince EVETINCE
 emecions
 EVINIONS
 exin EXIL
 exilts EXISTES
 expent EXPANG
 expont EXPANT
 fisited FAIGHTED
 afreed FALE
 finoraty FALORATE
 matis FAMATIES
 acadecis FAMES
 fanch FANCER
 cank FANK
 fanter FANT
 suflalo FASS
 govenver FAVED
 fand FAYMINTS
 mends FEAS
 fequised FECUIVER
 feaming FEERING
 fenteng FENTING
 foid FEREIGS
 tererted FEREVED
 fosk FESK

diamend FESPEND
cided FIDED
falk FIEK
siet FIET
sighte FIGHTED
volling FILLION
binaty FINIRATY
dinasted FINSTER
fing FINT
origis FIPS
mirner FIRER
figned FIRIGHED
tirt FIRT
pinked FIRTE
fick FISK
sitht FITH
folled FLAN
fows FOAT
foon FOILURY
whinecer FOINCER
folt FOIT
forsow FOLSOW
formur FOMOR
bons FONS
unge FONVERT
silps FOPS
limiced FOREVED
afferer FORIM
oratis FORITICS
formelle FORMIL
finner FORNER
fown FOROR
toper FOSPER
foor FOTOR
footbacs FOUT
selk FREADEM
creap FREAP
renes FREMES
frief FRIOF
holled FULIE
belds FULLS
ingue FUNGE
finting FUNTING
funs FUNTS
hesting FUSTING
pubsy FUTS
baip GAIP
gaiter GAITET
tare GARE
ange GEAN
edil GENUIVE

lagends GIVIND
glet GLADY
parguate GLAT
glastes GLATS
egle GLES
onch GLOWN
gack GOACK
combived GOLDED
fock GONK
joor GOOR
onvious GOUNTE
prahan GRAHAN
gask GRALE
aneme GRAME
gras GRAN
graud GRARD
ralm GRASH
dreats GREATENS
prefen GREN
aret GRET
grile GRIDE
vialeng GRIKE
foreigs GRIOL
grivity GRIVILY
broory GROOR
quiter GUITER
wondaged HACED
mags HAGS
binch HAINCH
loir HALIR
phattire HANT
palde HARDCISE
varde HARDLE
balkey HARKEY
sheching HEAMING
fecter HECTER
centing HELTING
merrer HERRY
hespical
HESTACAL
hictory HICTURE
hins HINE
heapes HIPES
haplens HIPLENS
poir HIRS
hustin HISTTINT
hictung HITTUNG
lealtly HOAT
corars HORWARS
loter HOTER
howded HOWDER

cule HULE
foulting HULTING
funch HUNE
hucred HURE
rirus HURMS
buet HUSTEN
odat IDAT
kickes IDIET
ignarang IGNIRANG
astiming IMPHASIC
emproach
IMPLOYEE
waiming INCIMING
fites INDECTRY
sungere INDER
inspect INDOCITE
spiorite INDORITE
irds INDS
onterved INFERNED
ondinate INFINITE
vintor INPOR
iccomate INPOT
ansare INSARE
intenter INTEED
hixtere INTERENG
binter INTINTER
antirior INTIRING
intideng INTURENT
soll INUSUAL
racame INUSUAN
extress IRMS
typong IRPROVE
cirts IRTS
impect ISPENT
cesumes ISSUME
intant ITAT
lanch ITIL
evencies IVENCIES
genuite IVENTITE
jand JANE
shead JARE
jark JART
searby JARY
agenly JENE
harging JERTY
sender JETER
jons JOFT
johnses JOHNSER
boir JOIR
johnte JONT
cordar JORDAR

poak JUAK
 illegal JUAL
 suat JUAT
 rumped JUMB
 numbens JUND
 rergest JURJ
 vilyl KALL
 tark KARK
 sken KENG
 fens KENS
 chisped KIECE
 kilsen KILLEN
 cancal KINSEN
 kink KINT
 learter KISH
 misk KISK
 ceaght KNAGHT
 kolovo KOLOTO
 sturies KTOSING
 cuns KUNS
 leaper LABER
 setutien LABES
 barkur LAKEN
 offains LALLS
 lintined LANTINED
 rorar LARTIES
 bundy LATENDS
 ictertas LATES
 hatted LATTED
 reanched LEACHER
 jums LEAGHT
 reat LEAT
 reyboard LEMAR
 allogal LEMAS
 prentint LENTING
 sleaking LENVING
 lede LERE
 burk LERGEST
 lenk LERK
 felt LERT
 leavins LEVING
 luet LIET
 sighty LIGHLY
 laiming LIMIN
 itel LINNER
 diberals LISERAL
 misten LISTENG
 paster LISTER
 boap LOAP
 allaring LOARDING
 morded LODED

nomer LOMER
 lomics LOMIN
 gond LOND
 korrer LOROR
 bloody LORRY
 lested LOSTER
 lout LOUGHT
 invoples LOXES
 quink LUEN
 encluded LUER
 bick MACK
 minected MAINTA
 dagarite MAIT
 mintal MANCAL
 manile MANISTED
 anto MANT
 cramo MAPAMED
 apant MAPENT
 mardle MASTLE
 maclate
 MATABATE
 datacite
 MATAACATE
 mainle MATTLE
 feathing MEAMING
 cleating MEARING
 meart MEARS
 mecked MECK
 sespoved
 MECOVED
 mecrets MECTETS
 medey MEDEA
 mentince
 MENATING
 renar MENCY
 minety MENECLY
 dellared
 MENENDED
 prinical MENICAL
 menaming
 MERACINE
 redical MERECLY
 dacks MERKS
 mertary MERTURY
 mestere MESERTE
 reselve MEVAL
 micker MICK
 hemenite MIFETIRE
 minner MIGGER
 fikel MIKEL
 migned MINEL

stroctly MINESTLY
 mang MING
 moster MINSTER
 dont MINTHLY
 aninue MINTINUE
 requency MINU
 sheap MINUMUM
 merform MIRMS
 prailang MIRRAR
 vigure MISE
 sisplay MISPLAY
 pister MISTAR
 vistor MISTOVER
 ritched MITCHED
 entre MITCHEN
 riverge MIVERCE
 lomes MOMAL
 cullshec MOMBIT
 romes MOMES
 ownor MOMOR
 totanty
 MONTANLY
 lools MOOLS
 founcer MOPEM
 mofies MORIES
 cold MORMULA
 mosper MORNER
 mistage
 MORTEAGE
 losted MOTED
 mouts MOUNS
 mintakin
 MOURTAIS
 progremms MOVEN
 montiply
 MUNTIPLY
 mits MYTS
 natiovel NACIVE
 weacter NEACTED
 raby NEARBE
 jeet NEET
 nearre NELARIRE
 norrilly NERRILLY
 helt NESTY
 acler NETER
 ninance NIMANCE
 coind NOMAR
 tentince
 NONTENCE
 mogic NOFIL
 soat NOWSURE

rute NUNE
surious NURECOUS
aute NUTE
dilk ODIL
offiges OFFISES
chetined OLDED
ortested ONTINTED
ontiors ONTIONS
aslier OPLIER
oellor OPLIOR
oppear OPPEAN
sebt OPPOSENT
tilk OPTIOR
mealings ORITIS
orms ORNS
opetatol
ORTHOTON
arti ORTO
aspeal OSPEAN
istain OSTAIN
oudio OUDIE
trols OUTTOON
zere OVERELL
oventeas OVERELS
comfont OWARA
fachers PACHERS
gank PAINTUL
possing PAISING
ankint PALKINT
fialodue PALLE
relling PALLING
anteptad PALT
prive PAME
sunkess PANKIND
panoc PANOR
pract PARACTEL
rask PASK
bissing PASSTAND
suan PASSWAND
gathened PATTNED
pleeing PELLING
penelly PEREDICE
rewais PERISIES
enspired PERSOREC
peradice PERTH
bevien PEVEN
pedern PEVER
dowarful PHATIFUL
quactite PHATTITE
speamint
PHEAMINT

pheased PHEORIED
phicken PHICTEN
physens PHYSECS
wicart PIAVO
porpinly PIBLICLE
leapon PIBLICTE
oarn PIBRARY
tinight PIDNEGHT
bist PIST
sarus PLANDED
paplened PLANED
prannels PLANSES
coughly PLASSIC
plandes PLASTES
acatyse PLATRORS
sheasing PLEAKING
friendly PLENLY
funging PLINGING
phoedix PLOCKS
prowsed PLOSED
doan POAN
mevil POLIME
approvol POLOS
commert POMMERS
pappened
POMPINLY
ponk PONT
dactery PONTETT
comoving
POOMING
pardeng POPENDS
cordumir POPULER
purpore PORPOSEL
pombived PORRIED
potior PORROR
derious PORS
micial POTCH
pradit PRAIN
assuer PRASSFER
pastage PRASUAGE
patiods PRAVIOUS
prame PREAK
preslere PREATER
profite PREATITE
prefit PRECIT
wede PREDUSER
printint PRENKINT
creede PREREDED
prequent
PRERENTS
pramo PRERIUS

primitel PRICTER
prike PRILE
masisted PRISTED
procs PROCT
prored PRODEDED
propodes
PROMOSEL
rafely PROMOTY
profer PROPED
plophes PROPOVES
porceng PRORRENS
ratter PROTTERS
retuale PROUBLE
provate PROVATY
progrint PROVIT
engly PRYSTAL
puick PUINK
pude PULE
luak PULIAN
numpens PUMPENG
buals PURPLIES
cryong PURPORY
curning PURTING
pustage PUSTICE
quine QUANE
naning QUATIFY
cest RACEST
acho RACH
carguate
RACOURLE
regind RAIND
maighint RAINT
racipe RAIVE
rilk RALK
cremes RAMES
cran RANS
rary RARK
bart RART
rease RASE
angined RATAINED
raut RAUR
tazy RAZZ
cealily REALILE
realily REALING
pernang REANING
reasong REASEND
beave REAVE
adente REBT
remate RECAME
remards RECARD
recotere RECATE

perceng RECENG
 mechonce
 RECHONCE
 impent RECOMERE
 seppoved
 RECOVERED
 prerites REDIES
 bish REDURED
 retiests REETS
 degreas REGREAS
 objents REIRS
 relete RELEARED
 rellies RELEIVES
 berivior RELICIOR
 rellien RELIER
 presleme RELIME
 belk RELIUM
 resure RELUME
 rereral REMARTS
 remend REMENG
 reterant
 REMERATY
 remoll REMILL
 remete REMOME
 renerred
 RENERATY
 afraid RENIAD
 reny RENS
 redure RENURED
 resense REQUENCE
 serrar REQUENT
 stols REQUINE
 reratire RERAIR
 reseate REREAATE
 reredue RERENUE
 apening RERISTED
 realings RERLING
 eged RESAGREE
 reakers RESEERCS
 rero RESHONCE
 covemeng
 RESIDENG
 resiced RESIGNER
 resiased RESIGREE
 reserds RESSERS
 nener RESTRAL
 sequine RETAIR
 reseils RETEILS
 revimes REVICES
 serian RIAN
 diar RIAR

drils RIAT
 wicret RICKET
 dirtincy RISC
 rirs RISS
 ripair RIST
 resaived RIVED
 roas ROAL
 grop ROAT
 lobic ROBIC
 gava ROCOSM
 horder RODER
 rotion ROFION
 hollar ROLAR
 adectity ROLLE
 prorremms ROLMS
 partned ROPARNED
 corara RORAN
 oret RORE
 roram ROTHT
 mouls ROULS
 rout ROUNT
 smick RULIND
 cruly RULY
 dure RURE
 aurs RURS
 rady SADY
 shappeng SAIN
 saldS SALD
 hanter SALTER
 singent SANGESTS
 shechint SANT
 alare SARARE
 barele SARDLE
 tarler SARGER
 staries SARGES
 dask SASK
 saxinum SAXIMUM
 caicer SCACE
 epedyday
 SCEDYING
 schete SCHEMELE
 screted SCHETELE
 scrint SCRANTE
 scrics SCRING
 strongts SCRINGS
 encher SEACHER
 slan SEAN
 seant SEANTS
 albud SEANTUM
 feasure SEASURED
 restrel SEASUREL

exints SECINTS
 ords SECONTS
 sutecier SECTER
 social SEDCINLY
 bede SEDE
 seduas SEDIAL
 sailt SEILT
 beleives SEITH
 fungs SELT
 sapture SELTUREL
 semerime SEMELE
 seague SEMEWKET
 segens SENS
 essies SENSIGES
 shenting SENTION
 servey SERGED
 senies SERTIENS
 fereins SERTIONS
 ventival SERTIVAL
 prerent SERVENG
 lanner SESTY
 setes SETEP
 resently SETTLY
 entudies SETUTIES
 relind SEYIND
 segins SEYING
 enjoyint SEYINS
 shapring SHAIN
 shait SHAIRTEN
 shamow SHAOS
 saidy SHAPPER
 arange SHAPRENG
 sharad SHARAN
 shat SHARTS
 asars SHARY
 criters SHATS
 shass SHAUS
 smels SHEALS
 specifed SHEDIED
 wheelt SHEELS
 shroring SHEORIED
 fangests SHESS
 sheative SHETIVED
 celps SHETOS
 hink SHINK
 usert SHIRIT
 shicked SHISPED
 geddess SHISS
 choence SHOENCE
 moring SHOPRING
 showes SHORS

shisit SHOST
 shereo SHREW
 suickly SHRICTLY
 sathroon SHROR
 shuly SHULS
 bessons SHUS
 silely SIAN
 victirs SICTIRS
 hasa SIDDAN
 siller SILTER
 statring SINTING
 crat SINU
 scimadio SISA
 sish SIST
 fitch SITCH
 islant SLANS
 slacts SLAS
 adlow SLEAD
 sleap SLEAK
 expocure SLEAR
 aseless SLESS
 loots SLOCTS
 loak SLOLK
 slocks SLORES
 spouses SLOVES
 sied SMICED
 mips SMITS
 snon SMOSS
 unet SNET
 stacts SOATS
 doit SOIT
 silling SOLLING
 toly SOLLY
 gools SONCORN
 monds SONDS
 siggent SONE
 senjerts SONTERS
 sinures SORDIENS
 jerde SORDS
 corera SORER
 sorms SORNS
 posiety SOSIETY
 comecody SOST
 fources SOUNCES
 somer SOWER
 affort SPAME
 suedsing SPECLING
 stannel SPIDE
 poicted SPITCTED
 smoits SPOCTS
 stacen STANDENS

suardion STARION
 tass STAS
 blass STASS
 mamical STATICAL
 stareng STATIONG
 nelarive STEDIED
 stevious STEDIES
 instart STELL
 bosten STEN
 sigered STERED
 sterter STERIES
 stintint STITRING
 stinting STITRINS
 stosted STORTED
 shosing STOSING
 stritere STREERS
 stigring STRINT
 stome STROSE
 slest STUCS
 spactrum
 SUANTUM
 suitide SUATIDE
 sactions SUCLIONS
 silder SUDDER
 luyer SUDLER
 nontance
 SUETANCE
 saisor SUIDY
 nule SULE
 bualt SULL
 sicture SULTUREL
 slaner SUMNER
 bune SUNE
 unte SUNT
 sute SUOTE
 poisod SUPPISE
 sebjerts SUPRERTS
 supes SUPRESE
 animary SURFACY
 firty SURFARY
 sury SURS
 sert SURT
 roplaing SURVIDE
 imigide SURVILE
 feedbecs SUSE
 sisk SUSK
 soas SUSS
 sutep SUTES
 emprains
 SYMPRINS
 tableer TABLY

phactire TACTIRS
 rait TAIT
 pontade TANCE
 arrer TARGER
 masser TASER
 tats TATE
 melented
 TECERTED
 ecedyday TEDEY
 stedects TELDS
 coch TELIELED
 prerium TENKS
 righers TERLS
 hespacial TERTISAL
 choites THIPS
 tonks THONTS
 ranner TINE
 firik TIRS
 bitht TITHT
 classit TOBLY
 domechac TOCH
 detimed TOMAL
 flons TONK
 cont TONT
 pruptet TOPES
 topil TOPOL
 slavery TORK
 ktoseng TORSS
 ctosery TOSER
 mathinas TOUNS
 tean TRAN
 hodse TRARING
 greateng TRATING
 teack TRAW
 tleers TREERS
 treoph TREOPS
 arkest TRIEST
 thror TRIN
 arvise TRUISE
 ttayer TTANEL
 treet TTAREL
 stule TULE
 woul TULK
 timeng TYPENG
 sude UDET
 buald ULBUD
 blient ULIEN
 intikely ULTICATY
 dicense UMIRENSE
 puseph UNETS
 Nile UNIL

dund UNTI
updity UPDATY
speated UPLATED
upsir UPSIT
useas USEAL
realices VEALILE
receal VEMERANS
netion VERTIONG
visuam VESUAM
legly VIGHWAY
vigere VIOMERTE
fistam VIRTIN
unktown VISTAGE
camol VOCATIOL
pottered VOLTARE
valde WAID
watles WALE
porwars WALY
pensorec WANS
acceint WARKEN
corrents
WARRENTY
velume WARVEL
baught WAUGHT
waby WAZY
ctadius WEDDINT
agenle WEENLE
werkers WERKENS
argurint WERLY
stame WHASE
wheap WHERAPY
porsess WISHOS
witters WITLES
joil WOIL
wollover
WOLLOVED
opdated
WONDATED
ankent WONSELE
tolds WOOTS
borten WORKEN
cennedy
WORRENTY
sote WOTE
worter WRATER
writed WRIVES
honter YOUNTER
moncenn ZORE

REFERENCES

REFERENCES

- Albert, R., Hawoong, J., Barabasi, A.L. (1999). Diameter of the World Wide Web. *Nature*, 401, 130.
- Albert, R., Hawoong, J., Barabasi, A.L. (2000). Error and attack tolerance of complex networks. *Nature*, 406, 378-381.
- Andrews, S. (1989). Frequency and Neighborhood Effects on Lexical Access: Activation or Search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 802-814.
- Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2), 234-254.
- Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, 35(6), 775-800.
- Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, 4(4), 439-461.
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology-General*, 133(2), 283-316.
- Barabasi, A. L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286, 509-512.
- Cancho, R. F. I., & Sole, R. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Science*, 100, 788-791.
- Cancho, R. F. I., Sole, R. V., & Kohler, R. (2004). Patterns in syntactic dependency networks. *Physical Review E*, 69(051915), 1-8.
- Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of the orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 857-871.
- Coltheart, M., Davelaar, E., Jonasson, J.F., & Besner, D. (1977). Access to the Internal Lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535-555).

Hillsdale, NJ: Erlbaum.

- Davis, C. J. (2006). Orthographic input coding: a review of behavioral evidence and current models. In S. Andrews (Ed.), *From Inkmarks to Ideas*.
- Davis, C. J., & Bowers, J. S. (2006). Contrasting Five Different Theories of Letter Position Coding: Evidence From Orthographic Similarity Effects. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 535-557.
- Davis, C. J., & Lupker, S. J. (2006). Masked Inhibitory Priming in English: Evidence for Lexical Inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 668-687.
- Davis, C. J., & Taft, M. (2005). More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review*, 12(5), 904-910.
- De Moor, W., & Brysbaert, M. (2000). Neighborhood-frequency effects when primes and targets are of different lengths. *Psychological Research*, 63(2), 159-162.
- Drews, E., & Zwitserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 1098-1116.
- Forster, K. (1998). The Pros and Cons of Masked Priming. *Journal of Psycholinguistic Research*, 27(2), 203-233.
- Forster, K. I. (1987). Form-priming with masked primes: The best match hypothesis. . In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 27-146). Hillsdale, NJ: Erlbaum.
- Forster, K. I., & Dickinson, R. G. (1976). Accessing the Mental Lexicon. In E. Wlaker (Ed.), *New Approaches to Language Mechanisms* (pp. 257-287). Amsterdam: North-Holland.
- Forster, K. I., Mohan, K., Hector, J., Kinoshita, S., & Lupker, S. J. (2003). The mechanics of masked priming. In *Masked priming: The state of the art*. (pp. 3-37). New York, NY US: Psychology Press.
- Forster, K. I., & Taft, M. (1994). Bodies, antibodies, and neighborhood-density effects in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 844-863.
- Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, 2(1), 21-31.

- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, 29(2), 228-244.
- Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103(3), 518-565.
- Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics*, 45(3), 189-195.
- Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception & Psychophysics*, 51(1), 49-56.
- Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, 47(2), 191-198.
- Guttentag, R. E., & Carroll, D. (1994). Identifying the basis for the word frequency effect in recognition memory. *Memory*, 2(3), 255-273.
- Holanda, A., Torres Pisa, I., Kinouchi, O., Souto Martinez, A., & Seron Ruiz, E. E. (2004). Thesaurus as a complex network. *Physica A: Statistical Mechanics and its Applications*, 344(3-4).
- Johnson, N. F., & Pugh, K. R. (1994). A cohort model of visual word recognition. *Cognitive Psychology*, 26(3), 240-346.
- Kello, C. T., & Beltz, B. C. (in press). Scale-free networks in phonological and orthographic wordform lexicons. In i. C. Chitoran, C., Marsico, E. & Pellegrino, F. (Ed.), *Approaches to Phonological Complexity*: Mouton de Gruyter.
- Kinouchi, O., Martinez, A., Lima, G., Lourenço, G., & Risau-Gusman, S. (2002). Deterministic walks in random networks: an application to thesaurus graphs. *Physica A: Statistical Mechanics and its Applications*, 315(3-4), 665-676.
- Kinsbourne, M. (1974). The mechanism of the word frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior*. 13(1), 63-69.
- Marcus, M., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. . *Computational Linguistics*, 19, 313-330.
- Mathey, S., Robert, C., & Zagar, D. (2004). Neighbourhood distribution interacts with orthographic priming in the lexical decision task. *Language and Cognitive*

Processes, 19(4), 533-559.

- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375-407.
- Motter, A., de Moura, A., Lai, Y., & Dasgupta, P. (2002). Topology of the conceptual network of language. *Physical Review E*, 65(6), 65102.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida word association norms. from <http://w3.usf.edu/FreeAssociation>.
- Newman, M. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2), 404-409.
- Paap, K. R., & Johansen, L. S. (1994). The case of the vanishing frequency effect: A retest of the verification model. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1129-1157.
- Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, 86(14), 3200-3203.
- Perea, M. (1998). Orthographic neighbours are not all equal: Evidence using an identification technique. *Language and Cognitive Processes*, 13(1), 77-90.
- Perea, M., & Rosa, E. (2000). The effects of orthographic neighborhood in reading and laboratory word identification tasks: A review. *Psicológica International Journal of Methodology and Experimental Psychology*, 21(3), 327-340.
- Perea, M., & Rosa, E. (2000). Repetition and form priming interact with neighborhood density at a brief stimulus onset asynchrony. *Psychonomic Bulletin & Review*, 7(4), 668-677.
- Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3(1), 1-17.
- Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, 19(3), 333-367.
- Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21(4), 876-900.
- Sibley, D. E., & Kello, C. T. (2004). A computational exploration of double

dissociations: modes of processing instead of components of processing.
Cognitive Systems Research, 6(1), 61-69.

Sibley, D. E., & Kello, C. T. (2006). *Learning representations of orthographic word forms*. Paper presented at the 28th Annual Conference of the Cognitive Science Society.

Sole, R. V. (2005). Syntax for free? *Nature*, 434(7031), 289-289.

Solë, R. V. (2001). Complexity and fragility in ecological networks. *Proceedings of the Royal Society B: Biological Sciences*, 268(1480), 2039-2045.

Sole, R. V., Murtra, B. C., Valverde, S., & Steels, L. (2006). Language Networks: their structure, function and evolution. *Sante Fe Working Paper*.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1), 41-78.

Wickelgren, W. A. (1969). Auditory or articulatory coding in verbal short-term memory. *Psychological Review*, 76, 232-235.

Yap, M. (2007). *Visual word recognition: Explorations of megastudies, multisyllabic words, and individual differences*. . Unpublished PhD thesis, Washington University, St. Louis, MO.

Yarkoni, T., Balota, D. A., & Yap, M. J. (in prep). *Levenshtein distance measures of orthographic and phonological similarity in visual word recognition*. Unpublished manuscript.

CURRICULUM VITAE

Brandon C. Beltz earned a Bachelor of Science degree in Philosophy from the University of Utah in Salt Lake City, UT in 2000. He earned a Master of Arts degree in Psychology with an emphasis in Human Factors and Applied Cognition from George Mason University in Fairfax, VA in 2004. This dissertation was submitted in 2009, in partial completion of a Ph.D. in Psychology from George Mason University. Brandon currently works as a Cognitive Scientist for Aptima, Inc. in Washington, DC.