8-17-2016

# Individual Differences in Sensitivity to Morphological Structure in Words and Nonwords

Oliver M. Sawi
*University of Connecticut - Storrs*, oliver.sawi@uconn.edu

# Individual Differences in Sensitivity to Morphological Structure in Words and Nonwords

Oliver M. Sawi

B.A., San Francisco State University, 2013

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

# APPROVAL PAGE

Master of Science Thesis

Individual Differences in Sensitivity to Morphological Structure in Words and Nonwords

Presented by

Oliver Sawi, B.A.

**Major Advisor** ________________________________
Jay Rueckl

**Associate Advisor** ________________________________
Kenneth Pugh

**Associate Advisor** ________________________________
Gerry Altmann

**Associate Advisor** ________________________________
Devin Kearns

**University of Connecticut**

**2016**

## Abstract

Understanding visual word recognition has been a central goal of psycholinguistics from its early beginnings. Examination of the statistical properties of language has uncovered many aspects of words that facilitate recognition. In addition, evidence from both behavior and computational modeling suggests that individual differences in experience and the strength of connections in an individual's reading network affect the sensitivity to these statistical properties in language. Morphology has special properties in this sense as morphologically related items have statistical regularities across both form and meaning. The current study examined whether individual differences in skill modulate sensitivity to morphological structure. Specifically, we looked at the relationship of three established measures of sensitivity to morphological structure (i.e., do they index the same dimension of variability?). We used a visual lexical decision task to simultaneously examine sensitivity to morphological structure in nonwords (nonword complexity effect), and words (two counts of morphological frequency – family size & base frequency). Linear mixed effects modeling was used to assess the main effects of each measure and to extract individual effect slopes to be used in individual differences analysis. Participants also completed an individual skill battery meant to examine exposure to print, vocabulary knowledge, and form (orthographic, phonological) based processing. We found that the nonword complexity effect, base frequency effect, and family size effect show systematic variability. Overall, as skill increased the nonword complexity effect increased and the morphological effects in words decreased. In addition, the nonword complexity effect in reaction time and the family size effect seem to be indexing opposite ends of the same dimension of variability. Base frequency, while closely related to family size, indexes a separate dimension of variability. Implications for the characterization of each of the effects and a possible future direction are discussed.

# Introduction

Understanding the mechanisms driving visual word recognition has been a central goal of modern psycholinguistics from its inception. Countless studies examining thousands of individuals have sought to describe the prototypical skilled reader in hopes of uncovering the elusive underlying processes. While most studies of reading have focused on group-level data using nomothetic analysis techniques to characterize the prototypical reader and generate generalizable theories, a growing body of research suggests that there are individual differences in reading skill.

Examination of the statistical properties driving the connections between orthographic, phonological, and semantic information has uncovered many aspects of words that facilitate recognition. For example, words that occur more often in a language are responded to faster than words that occur less often; commonly known as the word frequency effect (e.g. Forster & Chambers, 1973; Whaley, 1978). Researchers have also examined properties directly linked to the structured relationships between words such as the orthographic and phonological neighborhood density effects, in which words with more neighbors (the number of words that can be produced by changing one letter of the target word) are responded to more quickly than words with fewer (e.g., Andrews, 1992). Further, examinations of nonwords can provide additional insight into the processes underlying word recognition. For example, one can generate the orthographic neighborhood density for nonwords. While the effect is facilitative for words, there is a corresponding inhibitory effect of orthographic neighborhood density for nonwords (Coltheart et al., 1977).

These effects have also been examined in terms of individual differences. Efficient word recognition is driven by the linguistic characteristics of words as learned by particular individuals and the strength of the connections between phonological, orthographic, and semantic information that are developed through experience (Harm & Seidenberg, 2004; also see Perfetti, 2007, Lexical Quality Hypothesis). As individuals gain experience, they strengthen these connections over time and the connections become automatic (e.g., between phonological and semantic information in oral language). This would predict that individuals with weaker connections would have smaller effects of whole-word form frequency and larger effects of lexical and sublexical characteristics in words, as the connections are less automatic and individuals must rely on more granular characteristics (Perfetti, 2007). In nonword processing, however, less interference is generated from word-likeness as individuals with less experience received input from fewer word forms. For example, both good and poor readers show word frequency effects and good readers are faster overall, but the difference in response time between good and poor readers is largest in the lowest frequency words (Ashby et al., 2005; Hawelka et al., 2010; Kuperman & Van Dyke, 2011; Pugh et al., 2008; Shaywitz, 2003). Turning to the orthographic neighborhood density effect in words and nonwords, individuals with smaller vocabularies produce larger neighborhood effects in words (Yap et al., 2012) and smaller neighborhood effects in nonwords (Yap et al., 2015).

Morphology is a special case of how the statistical properties of language input drives the connections between form and meaning, which can in turn affect processing. Like the words related in the orthographic neighborhood sense, morphologically related words share orthographic features. However, morphologically related items have statistical regularities across both form (orthographic and phonological) and meaning. Of interest to the current study is whether individual differences in skill modulate sensitivities to these regularities. Given the previous discussion of how an individual's experience and strength of connections affect the use of and the sensitivity to the lexical characteristics of a word and the special statistical structure of morphologically related items, individuals should vary systematically in sensitivity to morphological structure related to the form and/or meaning overlap of morphologically structured words.

**Effect of Morphological Complexity on Visual Word Recognition**

Morphological aspects of words such as frequency of the stem (e.g., TEACH in TEACHER) (Burani & Caramazza, 1987; Burani, Salmaso, & Caramazza, 1984; Bradley,1979), productivity of the affix, or how often the affix is used to create new words, (e.g., Ford, 2010), and number of morphological family members (e.g., de Jong, et al., 2000; Bertram, et al., 2000) affect word recognition. Words with more information encoded in the constituent morphemes (e.g., higher frequency stem, more productive affixes) are easier to recognize. In other words, words with morphological constituents that occur more often and in a consistent manner are easier to recognize. Further insight into the processes underlying lexical and sublexical mechanisms comes from evidence regarding the processes by which stems and affixes in morphologically complex words are accessed.

For example, accounts of morphological processing posit that morphologically complex items are decomposed into stems and affixes prior to lexical access based in orthographic segmentation (Rastle & Davis, 2008; Rastle et al., 2004) very early in visual word recognition (Larvic, Elchlepp, & Rastle, 2012). Two primary bodies of evidence support the prelexical, orthographically based morphological decomposition. For example, the recognition of base targets is speeded by the prior brief presentation of morphologically related words (masked priming) (e.g., Grainger, Colé, & Segui, 1991; Rastle, Davis, Marslen- Wilson, & Tyler, 2000) such that transparently related pairs (TEACHER-teach) are responded to faster than opaque (CORNER-corn) and opaquely related pairs are responded to faster than form (BROTHEL-broth) but only transparently related pairs (TEACH-teacher) in long-term priming (Rueckl & Aicher, 2008). Further, transposed letter primes also prime related words, but only when the transposed were within morpheme (TAECHER-teacher) and not between morphemes (TEAHCER-teachter), which suggests that morphological decomposition processes also occur with letter position coding (Duñabetia, Perea, and Carreiras (2007). However, while, there are contradictory findings for both the morphological priming (e.g., Rastle, Davis, & New, 2004) and morphological transposed letter priming (e.g., Sánchez-Guitiérrez and Rastle, 2013; McCormik & Rastle, 2013; Rueckl & Rimzhim, 2011), exploring these effects through the lens of individual differences helped adjudicate the inconsistencies because it provided insight into more fine grain differences in morphological processing due to an individual's reading profile. In other words, the findings may be inconsistent because of individual variability in the use of these processes.

For example, Andrews et al. (2013) used linear mixed effects model as the primary analysis to examine transparency effects in morphological priming. Target stimuli included ninety prime-target pairs from Rastle et al. (2004). The ninety pairs were separated into three categories: semantically transparent (teach-TEACHER), opaque (corn-CORNER) and orthographic controls (broth-BROTHEL). Andrews et al., (2013) used vocabulary as a measure of "semantic coherence" (Perfetti, 2007) and spelling as an index of orthographic precision of lexical processing. As spelling and vocabulary were highly correlated, composite scores for each were entered into a PCA to obtain orthogonalized components. The first component was highly related to skill in both Spelling and Vocabulary and reflected overall skill, but did not interact with priming. The second component however, reflected the unique variation differentiating spelling and vocabulary. Andrews et al., (2013) used this component to label individuals as having an "orthographic" or a "semantic" profile. On one end of the component included individuals with superior spelling relative to vocabulary skills ("orthographic") and the other end represented the superior vocabulary relative to spelling ("semantic"). Superior spelling relative to vocabulary in dimension 2 of the PCA was associated with increased priming for opaque pairs and reduced priming for transparent pairs, but higher vocabulary than spelling was associated with stronger priming for transparent than opaque pairs. Individuals with higher a semantic profile, seem to be in line with the graded accounts of morphological priming. However, individuals with an orthographic profile seem to cause problems for this viewpoint in that they have the same level of priming for both opaque and transparent pairs. Again, this seems to suggest that individuals with better Lexical Qualtiy (here individuals with a semantic profile), as referenced in Kuperman & Van Dyke (2011) are more sensitive to the statistical structure of morphology than individuals with worse. This finding provides evidence that there are fine grain differences in the use of morphemes in processing based on various reading and word recognition skills.

Further, Duñabeitia et al., (2014) examined early morphological decomposition of complex words by using a masked priming transposed-letter paradigm. In this paradigm, letters are either switched within a morpheme (e.g., TAECHER) or between morphemes (e.g., TEACEHR) and are used as masked primes for related words (e.g., TEACHER). Duñabeitia et al., (2014), also in line with (Andrews et al., 2013) used individual differences to adjudicate inconsistencies in the literature, in which some groups find greater priming for within than between morpheme transposition (suggesting morphological decomposition) and some groups did not. They found that individual differences in reading speed regulated the difference in the masked transposed letter priming effect between morphemes. Individuals with faster reading times displayed greater priming for within- than between- morpheme transpositions while individuals with slower readings times showed no difference between the two types of transpositions. This suggests that faster readers may be more likely to consistently use decomposition strategies early in processing, while slower readers may not. Individuals with more skill may also be more sensitive to the morphological structure of the primes causing an advantage.

**Effects of Morphological Complexity on Nonword Recognition**

As stated previously, examining nonword processing can provide additional insight into the processes underlying word recognition. Of particular interest to the current study is the effect of morphological structure on nonword recognition. In this case, one can further examine individual effects in early, prelexical morphological decomposition, as nonwords inherently do not have whole-form lexical entries. In addition, to our knowledge, there is only one other study examining individual differences in the effects of morphological complexity on nonwords (Yap et al., 2015).

In particular, the morpheme interference effect, or nonword complexity effect, occurs when nonwords that are created by combining existing morphemes (e.g., GASFUL) are rejected more slowly in a lexical decision task than words that do not have lexical structure (e.g., GASFIL). Taft and Forster (1975) found that nonwords composed of existing prefixes and bound stems (e.g., DEJUVINATE) were rejected more slowly than were nonwords composed of the same prefixes but non-existing stems (e.g., DEPERTOIRE). Similarly, Italian nonwords that were decomposable into morphological constituents produced longer RT latencies than nonwords that were not decomposable (Caramazza et al., 1988). The increased response latency to morphologically decomposable nonwords has been put forth as strong evidence for the early prelexical, obligatory decomposition of morphologically complex words into constituent morphemes. Crepaldi et. al. (2010) extended this finding by examining the effect of the position of the pseudo-affix on lexical decision in English. The critical manipulation was the complexity of the nonword. In the morphologically complex, or decomposable, condition, the nonword includes a baseword with a syntactically legal suffix (e.g. GASFUL). In the morphologically simple, or nondecomposable, condition, one letter in the suffix is changed in order to make it an illegal suffix (e.g. GASFIL). There was a large effect of morphological complexity, but only when the morphemes were in their syntactically legal places (GASFUL/GASFIL v. FULGAS/FILGAS). This suggests that once the morphemes are placed in syntactically legal positions, the word-likeness of the nonword construction produces interference for recognition. Further evidence for the obligatory decomposition of letter strings prelexically, based on orthographic features. Additional sources of interference could be caused by the spreading activation of similar word forms. Due to morphology's special, consistent statistical structure across orthography (e.g., TEACH, TEACHER, TEACHING) phonology, and semantics, greater interference would be caused, particularly when the specific components are placed in syntactically legal positions (i.e., where they would be placed in a real word).

In a related finding, inflectional endings such as -S and -ED had an inhibitory effect on nonword lexical decision RT (Muncer, Knight, & Adams, 2013a). Nonwords that include morphological structure are harder to reject as nonwords (e.g., ZINTED, ZINTS). In a follow up study with the British Lexicon Project database, Muncer, Knight, & Adams (2013b) extended their findings by including additional affixes defined by Fudge (1984) and Pinnell and Fountas (1998) and reported a number of affixes effect. A greater number of affixes had an inhibitory effect on nonword recognition (i.e., slower RT). Further studies using the relatively larger English Lexicon Project database have also shown a number of affixes effect, and a corresponding facilitative effect on word recognition (i.e., faster RT) Yap et al., 2015; 2012). The number of affixes effect has also been taken as further evidence for obligatory decomposition processes. For example, Yap et al., (2015) reasoned that interference was generated due to the initial processes of parsing the nonword letter string based on the affix before rejecting. Therefore, more affixes meant more parsing needed. Taken together, this may also suggest that morphological information produces interference to the recognition of a letter string as a nonword because the nonword is more "word-like" and the morphological information activates competing wordforms (see Harm & Seidenberg, 2004 for a connectionist implementation). While evidence points to morphological information interfering with nonword recognition, research suggests it facilitates word recognition (Yap et al., 2012).

As the processes by which morphological information is accessed early in visual word recognition (obligatory form based decomposition, Rastle & Davis, 2008; Rastle et al., 2004) have been well established for words and nonwords, we will further explore measures that examine sensitivity to morphological structure in terms of indexes of statistical regularity (frequency) in morphological constituents, particularly in terms of the statistical properties of the stem.

**Effect of Statistical Properties of Morphological Constituents on Word Recognition**

Previous research on the identification of morphologically complex words suggests that complex words are recognized using multiple sources of information such as whole lexical forms, their morphological constituents, and morphological families (e.g., Baayen, Dijkstra, & Schreuder, 1997; Burani & Caramazza, 1987; Burani, Salmaso, & Caramazza, 1984; Schreuder & Baayen, 1997; Taft, 1979). For example, TEACHER would be recognized using information related to the whole form TEACHER, its morphological constituents TEACH and –ER , and words to which it is related such as TEACH and TEACHABLE.

Specifically, three sources of information have been well established: Surface Frequency, related to the whole lexical form, and Base Frequency and Family Size, related to the morphological features respectively. Surface Frequency refers to the frequency of the whole-word string (e.g., TEACHER). Family Size refers to the Type count of morphologically related words (e.g., TEACH, TEACHER, TEACHABLE – family size 3)**.** Base Frequency, also referred to as Cumulative Root Frequency and Baseword Frequency, is the cumulative frequency (token count) of morphologically related family members (for a full description, see de Jong, 2000). Note, while theoretically similar (in the decomposition sense), Base Frequency is not to be confused with Stem Frequency, the frequency of the bound stem embedded in a morphologically complex word such as the frequency of TEACH in TEACHER (Burani, Salmaso, & Caramazza, 1984; Burani & Caramazza, 1987).

**Base Frequency.** The base frequency effect, as first described by Taft (1979) is when words with high base frequency are responded to faster and more accurately than words with low base frequency when surface frequency is controlled (e.g., Taft, 1979; Colé, Beauvillain, & Segui, 1989; Schreuder & Baayen, 1997; Baayen, Dijkstra, & Schreuder, 1997; Bertram, Schreuder, & Baayen, 2000). However, the base frequency effect has been shown to vary with specific word types and situations such as affix transparency, productivity, and decomposability. Vannest et al., (2010) examined the effect of decomposability on the base frequency effect. "Decomposable" items or words that are decomposed into stem and affix in recognition (e.g. –able, -less, -ness) were contrasted with "Whole-word" items or words that are not decomposed into stem and affix (e.g. –ity, -ation). Only "decomposable" items showed a base frequency effect. In line with this finding, Xu and Taft (2015) further explored the interaction of semantic transparency and base frequency. In a transparent item, both the baseword and the affix provide information to the meaning of the word (e.g. TEACHER). However, in an opaque item, the baseword and the suffix are both legal, but the combination does not provide additional information (e.g., CORNER). Xu and Taft (2015) found that the base frequency effect became larger with more transparent words (i.e. the transparent words had a larger effect than partially transparent and opaque words). Relatedly, only words with highly productive affixes, affixes that are used in the production of many new words, produce reliable base frequency effects (Ford et al., 2010). This finding was also replicated in Spanish, a language with more transparent print-to-sound mapping than English (Lazaro, 2012). Lastly, the base frequency effect is also sensitive to nonword context. In nonword contexts in which all of the nonwords have complex morphological structure (stem+affix) the base frequency effect actually reverses (Taft, 2004). Taken together, base frequency effects are most robust in decomposable, transparent words, with productive affixes and in mixed nonword contexts. Further as the base frequency effect is so affected by affix productivity, decomposability, and transparency, it has been cast as most related to form based, morphological decomposition processes.

**Family Size.** On the other hand, family size has been cast as a process related to the semantic processes, particularly in Hebrew (see Moscoso del Prado Mart´ın et al. 2005; Baayen, 2014 for review). The family size counts also differ from base frequency in that family size is the type count of morphologically related words and base frequency is the token count (cumulative frequency) of morphologically related words. Visual lexical decision response times to words with larger family sizes (i.e., appearing as a constituent in larger numbers of derived words and compounds) are faster than for words with smaller family sizes. This effect has been shown in monomorphemic, or simplex words in Dutch (Schrueder & Baayen 1997) and in complex words (de Jong, et al., 2000; Bertram, et al., 2000). Bertram et al., further explored the role in inflected and derived words and semantic transparency. In complex words, there is a strong family size effect for a range of inflected and derived words. There was also a strong effect for words that “straddled” the line between inflected and derived words. Interestingly, similar to findings in the base frequency effect, they found that semantically transparent family members drove family size effects and the family size effect was largely absent or attenuated in semantically opaque family members. The family size effect has also been shown to be quite robust in English (Feldman and Pastizzo, 2003; Baayen et al., 2007). However, unlike base frequency, the family size effect is not affected by affix productivity (Ford et al., 2010). Interestingly, several studies in Dutch (Schreuder and Baayen, 1997; Bertram et al., 2000; De Jong et al., 2000) suggest that the relevant predictor for visual lexical decision rection time is the type-count family size measure and not the token-count related frequency measures like base frequency. However, their results could be due to the productivity of the suffixes in their sample (Ford et al., 2010).

**Potential Differences between Base Frequency and Family Size.** While base frequency and family size encapsulate similar information in relation to morphology, there is not consensus in the field regarding their relationship. For example, De Jong et al, (2000), suggests that Family Size, not Family frequency, the token count equivalent similar to base frequency (the cumulative frequency of family members) is the relevant predictor of response latency. However, several studies suggest that, in relation to Surface Frequency and Base Frequency, Family Size is a separate predictor of reaction time. For example, Ford et al., (2010) found that while the base frequency effect only occurred in words with productive affixes, the family size effect occurred regardless of affix productivity. This lead to the conclusion that base frequency is more related to statistical properties of the form of the word and family size is more related to the semantic properties.

Further, several studies in English and Hebrew have put forth family size as semantic in nature (see Baayen, 2014 for review). Xu and Taft (2015) had a similar finding using linear mixed effects modeling to examine the separate effects of surface frequency, base frequency, and family size with the other effects statistically controlled for. Even when including each of the three statistics in the model, the three effects were significant, suggesting that both base frequency and family size facilitate word recognition and are separate predictors. Interestingly, within Kuperman & Van Dyke (2011), both family size and base frequency were included in the LME statistical model. However, family size showed "less predictive power" than base frequency, so it was left out of the analyses completely. This is interesting because 1) there is evidence to suggest Base Frequency and Family Size, while both morphologically related, are related to difference aspects of the statistical structure of morphology (Ford et al., 2010) and 2) both were included in the LME model and Family Size may have shown some predictive power even with the base frequency effect partialled out. This suggests that family size may have been a separate predictor (Xu & Taft, 2015) and could be informative in further analysis.

In summary, according to some reports, both are related to morphology, but base frequency is more closely related to morphological processing related to form (frequency effects, dependence on productivity, strongly occurring in suffixes, less so in prefixes) and family size is more closely related to semantic overlap. As the nonword complexity effect has not been examined for individual differences or compared to other morphological measures, its relationship to word morphological effects is unknown. After we conducted direct correlational analysis, we compared the patterns of correlation with skill measures with the nonword complexity effect and the word effects.

Although group differences in the morphological processing in words and nonwords have been examined thoroughly, relatively few studies have examined the effects of individual differences in reading skill on the sensitivity to morphological structure via the use and integration of various forms of morphological information in words (e.g. surface frequency, base frequency, family size) and interference in nonwords (e.g., nonword complexity effect; morphological decomposition).

**Individual Differences in Sensitivity to Morphological Structure**

Kuperman & Van Dyke (2011) looked at individual differences in use of whole-word and morphemic information and posit a trade-off based on individual skill. They measured eye movements while reading connected text. Target words included 69 English suffixed words (teach + er). Targets were all semantically transparent with productive suffixes (-er,-or,-ist,-ing), the combination according to previous research most likely to generate robust base frequency effects. Linear mixed effects models were used to tease apart effects of word, base, and family size. However, family size was not as predictive as base frequency and was subsequently left out of further analysis. Additionally, direct interactions between base frequency and whole word effects were not significant and were also therefore left out.

Whole-word frequency and a battery of 17 individual differences measures were analyzed with separate models. A similar process was used for the base frequency effect. Word and nonwords segmentation and two comprehension tests were the only tasks that provided significant interactions. Segmentation seemed to be more related to skill based on understanding the form of words (phonological, orthographic) and comprehension was more related to meaning. Overall, fixation time was negatively related to whole-word frequency. The effect was greater for readers with higher segmentation and comprehension scores. This suggested that better readers were faster overall than poorer readers and poorer readers had a more pronounced slope from low frequency to high frequency. Additionally, the effect of baseword frequency was negative for the poorest readers, but positive for the best readers suggesting poorer readers tend to rely more on frequency information related to the morphological constituents (base frequency). Conversely, the positive effect better readers suggested that they have competition or interference from the additional information rather than facilitation. Interestingly, individual strategies and trade-offs were discussed without direct comparisons of the word and morphological variables via an individual interaction term.

While this review has examined several studies exploring individual differences in morphological effects, they are greatly outnumbered by studies examining group-level phenomena. This is most evident in nonword recognition, as there is only one major analysis using data from the English Lexicon Project, a mega-study across six universities, which compiles trial-level data from various lexical decision and naming experiments, (ELP; Balota et al., 2007) and an effect that is not as well established as other measures of nonword complexity, such as the morpheme interference effect, or nonword complexity effect (e.g., Taft, 1975). Yap et al., (2015) examined reaction time and vocabulary scores using data from the English Lexicon Project.

Yap et al., conducted item-level and participant-level analyses on lexical decision data for around 37,000 nonwords in the English Lexicon Project to explore the influence of various psycholinguistic variables on nonword lexical decision performance such as orthographic neighborhood density, length, and morphological characteristics, specifically number of affixes (Muncer et al., 2013a; 2013b). Overall reaction times were slower for nonwords including more affixes. This replicates the findings in the previous explorations of the group-level effects. Yap et al., then extended these findings by exploring the effect of individual differences in vocabulary, taken as a measure of the integrity of lexical forms, on the number of affixes effect. The findings indicate that individuals with more skill in lexical processing, as indexed by vocabulary score and nonwords drift rate (measured by examining how an individual's responses change over time), are more sensitive to number of affixes. The morphological complexity effect increased with vocabulary size where individuals with higher vocabulary sizes were slowed down more by increased numbers of affixes. This seems to be in line with findings in words (particularly, Kuperman et al., 2011), in which individuals with larger vocabularies activate more like-words when presented with morphological structure overall which in turn generates interference for nonword recognition and facilitation for word recognition.

However, this examination, uses data from the ELP, a megastudy without the tight experimental controls that could be afforded with an in-laboratory experiment. Further, this study also lacks the direct comparison with other forms of established morphological information (e.g., base frequency, family size effects). In addition, given that the data are from the ELP, the investigators did not have direct control over the stimuli. Consequently, the ELP nonwords were not selected to investigate morphological effects in particular. Lastly, the structure of the nonwords, with affixes but no stems, while theoretically similar to well-established nonword complexity effects in terms of decomposition processes, are quite different from the structure of the nonwords in the classic nonword complexity effect which had both stems and affixes.

## The Current Study

The current study seeks to extend the literature by simultaneously characterizing systematic individual differences in sensitivity to morphological structure in words and nonwords (potentially morphological decomposition processes and sensitivity to statistical information of the constituent morphemes respectively). A unique contribution of this study is the characterization of individual differences in the family size, base frequency, and nonword complexity effects and the comparison of these effects to each other and to a battery of skill measures to see how individual reading profiles affect sensitivity to morphological structure. For example, we extended the findings regarding the effects of morphological structure on nonwords via a conceptual replication of the Yap et al. (2015) finding using a stronger manipulation and more well established measure of nonword interference.

The morpheme interference effect, or nonword complexity effect (e.g., Taft, 1975; Carramazza et al., 1998; Crepaldi et al., 2010) is a well-characterized measure in terms of nomothetic analysis. Nonwords that are easily decomposable (Vannest et al.,, 2010 sense) and have affixes in syntactically correct positions (Crepaldi et al., 2010) generate the largest and most stable effects. Tests that generate large, stable, and variable effects are ideal for individual difference study. Therefore, we will use the stimuli from Crepaldi et al. (2010) with affixes in the correct positions to see if there are indeed individual differences in this effect. This effect should pattern similarly to the Yap et al., (2015) finding (higher skill, more interference) as both the number of affixes and nonword morphological complexity seem to be related to interference from morphological information. Further, although individual differences in sensitivity to base frequency has been examined (Kuperman & Van Dyke, 2011) in which better readers had an inhibitory effect of base frequency and poorer readers had a strong facilitative effect, family size has not been well-characterized in terms of systematic individual differences. Additionally, while the nonword complexity effect has been posited to provide insight into sensitivities to morphological structure that occur early in visual word recognition, its relation to other well established measures of morphological processing (family size, base frequency) and whole-word processing is not understood (i.e., whether different measures of sensitivity to morphological structure--nonword complexity effect, family size, base frequency effect--are indices of the same underlying dimension of variation). Further, there is disagreement in the literature as to whether the family size effect and base frequency effect are indices of sensitivity to the same or different aspects of morphology. This study looked closely at the relationships between these three effects and examined differences in the literature through the lens of individual differences examination.

To explore the effects of morphological statistics and their relationships to each other, as with the nonwords, we used stimuli with the greatest chance to generate robust and variable effects. As discussed previously, complex words that are transparent (Xu & Taft, 2015), decomposable (Vannest et al., 2010), and have productive suffixes (Ford et al., 2010) produce the most reliable effect for both family size and base frequency. As a result, we used a subset of the words from Ford et al., (2010) with exclusively productive suffixes. In addition, as Ford et al., (2010) and Xu & Taft (2015) first established, family size and base frequency are potentially separate constructs.

Visual lexical decision provides the opportunity to explore both word and nonword effects simultaneously within the same individual unlike reading of connected text. Therefore, direct comparisons of effects (through correlations) and indirect analysis of effects (via patterns of correlation with skill measures) are possible. We used linear mixed effects models to help facilitate this process. LME afforded us the opportunity to simultaneously examine participant and item effects and also extract individual effect slopes for several effects (surface, base, family size, nonword complexity) separately to compare with various skill measures via interaction terms (Kuperman & Van Dyke, 2011) or via correlation of individual effect slopes to skill measures.

## Experiment 1

Before the individual differences in the nonword complexity effect may properly be explored, the robustness and stability of the effect must be established. A primary concern here is the list context in visual lexical decision tasks, particularly regarding morphological effects. For example, the base frequency effect, while widely cited as evidence for the robustness of obligatory decomposition accounts may be reversed depending on the list context. In Taft (2004), the list context was manipulated by using contrasting nonword distractors. All words were matched on surface frequency (low) and varied on base frequency (medium vs. high). Words were also tightly controlled for the ratio between the frequency of the base and whole-word. Interestingly, the words with low surface frequency and high base frequency were atypical in that the affix was unusual relative to the information associated with the stem (e.g., seeming, moons) (Taft, 2004). All nonword distractors were generated with either nonsense stems (e.g., GLEENIFY) or real-word stems (e.g., GREENIFY) similar to the Crepaldi et al, (2010) stimuli. Nonword distractors with nonsense stems generated the classic base frequency effect. However, nonword distractors with real word stems resulted in a reverse base frequency effect (i.e., longer RT latencies for words with high base frequency). Of note, the "reverse" base frequency effect is an effect on words, generated by nonwords. However, examining list context here is important as both effects are presumed to be morphological in nature. Further, the nonword complexity effect has not been examined to see if word context effected the direction of the effect.

Various research groups have looked at the effect of semantic transparency on both long-term and masked (e.g., Grainger, Colé, & Segui, 1991; Rastle, Davis, Marslen- Wilson, & Tyler, 2000) priming. For example, (Rueckl & Aicher, 2008), explored the effect of semantic transparency on long-term priming (e.g. TEACHER, CORNER, BROTHEL words). The critical manipulation in this set of stimuli is the relationship between the baseword and the suffix. In the transparent context, all words are semantically transparent (e.g., TEACHER) and as stated before, adopting a strategy based on morphological decomposition would be beneficial for word recognition, but detrimental to nonword recognition. The mixed condition, however, contains words that are semantically transparent, opaque (e.g., CORNER), and orthographic controls (e.g., BROTHEL). Adopting a strategy based solely on decomposition would not be beneficial to word recognition. Given the Taft (2004) finding that nonword context influenced (actually reversed) the base frequency effect in words (a well-established measure of sensitivity to morphological structure) with surface frequency controlled, confirming that the morpheme interference effect is robust to context i.e., list context, is an important first step into confirming it is a suitable task for further individual difference exploration.

## Method

### Participants

The participants were 51 undergraduate students enrolled in an introductory psychology course at the University of Connecticut who participated for course credit. All were native speakers of American English.

**Design and Materials**

The experimental stimulus set contained two sets of 32 nonwords adapted from Crepaldi et al. (2010). In the baseword-plus-suffix (complex) condition, existing basewords were combined with existing suffixes (e.g, gasful). These combinations were syntactically legal. Nonwords in this condition were constructed by using 16 different suffixes, each of which was attached to four different stems. In the baseword-plus-control (simple) condition, the same basewords were combined with similar suffixes used in the decomposable condition. Nonmorphological endings were created by changing the central letter of each of the suffixes used in the decomposable condition (e.g., gasfil). Since the same morphemes were used across conditions, the experimental nonwords were distributed over two different sets of words, with 32 items per condition so the participants did not see the same stem or same suffix in the same position twice.

In addition, the between-subjects, list-context manipulation included two contexts. The mixed condition included 90 words (3 sets of 30) adapted from Rueckl & Aicher (2008). With transparent (e.g., TEACHER), opaque (e.g., CORNER), and form words (e.g., BROTHEL) as described previously. The transparent condition included the same transparent set from the mixed condition and a set of 60 filler transparent words (following the same rules as the Rueckl & Aicher (2008)). Filler multisyllabic nonwords (36) were selected from the English Lexicon Project. Nonwords varied in length from 6-8 letters to match the average length of the Crepaldi nonwords. Ninety filler, monosyllabic, monomorphemic words and nonwords were also included. Words varied in terms of number of letters (4-6) and frequency. The simple nonwords did not include morphological structure and varied in terms of number of letters, parallel with the filler words

**Table 1. General Characteristics of Critical Stimuli Used in Experiment 1**

| | Critical Nonwords | | | | Critical Words | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Complex | | Simple | | Transparent | | Opaque | | Form | | Transparent Filler | | | |
| **Example** | **GASFUL** | | **GASFIL** | | **TEACHER** | | **CORNER** | | **BROTHEL** | | **LEARNER** | | | |
| Count | 32 | | 32 | | 30 | | 30 | | 30 | | 60 | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | | |
| Frequency | - | - | - | - | 1.02 | 0.90 | 0.86 | 0.86 | 0.74 | 0.83 | 0.72 | 0.66 | | |
| Length | 7.08 | 1.19 | 7.08 | 1.19 | 7.30 | 0.86 | 7.50 | 1.20 | 4.67 | 0.79 | 7.00 | 1.14 | | |
| Syll | 2.36 | 0.48 | 2.34 | 0.48 | 2.33 | 0.54 | 2.37 | 0.66 | 2.30 | 0.78 | 2.26 | 0.58 | | |
| Orth N | 0.19 | 0.62 | 0.03 | 0.18 | 1.43 | 2.33 | 0.97 | 1.73 | 4.13 | 3.94 | 1.26 | 1.94 | | |

*Syll, number of syllables; 'N', neighborhood density; data from CELEX. Frequency natural log transformation.*

**Table 2. General Characteristics of Fillers Used in Experiment 1**

| | Filler Nonwords | | | | Filler Words | |
|---|---|---|---|---|---|---|
| | Multisyllabic/Multimorphemic | | Monomorphemic | | Monosyllabic/Monomorphemic | |
| **Example** | **ARMIGHTY** | | **PLOSIOB** | | **CATCH** | |
| Count | 36 | | 90 | | 90 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Frequency | - | - | - | - | 1.70 | 1.96 |
| Length | 7 | 0.83 | 5.28 | 0.73 | 5.00 | 0.82 |
| Syll | 2.29 | 0.78 | 1.59 | 0.59 | 1 | - |

*Syll, number of syllables; 'N', neighborhood density; data from CELEX. Frequency natural log transformation.*

## Procedure

Participants were tested individually in a quiet room. After giving informed consent, they were told that they would see a series of letter strings presented one at a time and that they would be required to decide as quickly and accurately as possible whether or not each string was a word. Following the instructions, the participants completed a practice session of 20 trials, were given a chance to ask questions, and then completed the rest of the trials. On each trial a fixation point (a cross) was presented for 250 ms, followed by a letter string that remained on the screen for until a response was made. Participants responded by pressing designated computer keys with the index finger of either hand, with the 'yes' response assigned to the dominant hand. The inter-trial interval was 250 ms. The trials during the main session of the experiment were arranged in a random order Participants were offered the opportunity to take a short break after every 94 trials. Stimulus presentation and data collection was controlled using the E-prime software package running on a Pentium 4 personal computer.

## Results

Correct response times (RTs) and error rates (ERs) were analyzed using linear-mixed effects (LME) modeling in R (Baayen, Davidson, & Bates, (2008). Subjects and items were entered as crossed-random factors. Reaction time (RT) data were log transformed. Analysis of reaction time generated t-values. An absolute t-value near two is considered an appropriate indicator of significance (see Baayen et al., 2008 for review). Additionally, following the procedure outlined in Kearns (2016), significance may be determined by examining change in chi-squared. Examining the delta chi-squared enabled us to examine whether a variable explained a significant amount of variance in the model. The analysis of error rates was conducted using the binomial function, which generates z scores from which p values could be directly calculated. The LME coefficient, b, is reported for the effects of interest to provide insight into the relationship between the fixed effect factor and dependent variable (e.g., a negative coefficient signifies a negative), along with the standard error. Fixed factors that were continuous variables were standardized to avoid spurious correlations and to facilitate LME analysis**.** Reaction time was only reported for correct responses. RTs faster than 250ms were removed. For slow reaction times, individual cutoffs were generated by calculating 3 standard deviations from an individual subjects' mean reaction time across target words. Reaction times slower than these cutoffs were replaced with the cutoff value. The individual random effects structure was established using the log-likelihood ratio model comparison test and included participant and item as intercepts. Reaction times were converted via natural log to approximate a normal distribution.

Nonword complexity (Complex, Simple) and context (mixed, transparent) were entered as fixed effects for both reaction time and error rate. A nonword complexity by context interaction term was also entered as a fixed effect. Log previous reaction time, trial order, and previous trial type (i.e. word, nonwords) were entered into the fixed effects to control for any potential influences of these variables.

The primary analysis involved examining the effects of morphological complexity in nonwords for reaction time and error in transparent and mixed contexts. There is a strong effect of nonword complexity in reaction time (b = -0.053, SE = 0.022, |t| = 2.42, $\Delta x^2$ = 8.5). The effect of context (b = 0.0013, SE = 0.066, |t| = 0.02, $\Delta x^2$ = .003) and the interaction between nonword complexity and context (b = -0.009, SE = 0.020, |t| = .48, $\Delta x^2$ = .229) were not significant. A separate analysis was conducted for Error Rate following the same methods for specifying the optimal RT model. Simple nonwords produced fewer errors than complex nonwords (b = -1.85, SE = 0.24, |z| = -6.45, $p < .001$) and the transparent context produced more errors than the mixed context (b = .85, SE = 0.31, |z| = 2.76, $p < .001$). However, the interaction between nonword complexity and context was not significant (b = -0.078, SE = 0.28, |z| = .276, p = .78). Overall, complex nonwords take longer to identify and have a higher error rate regardless of the context, but the transparent context produces more errors overall than the mixed context.

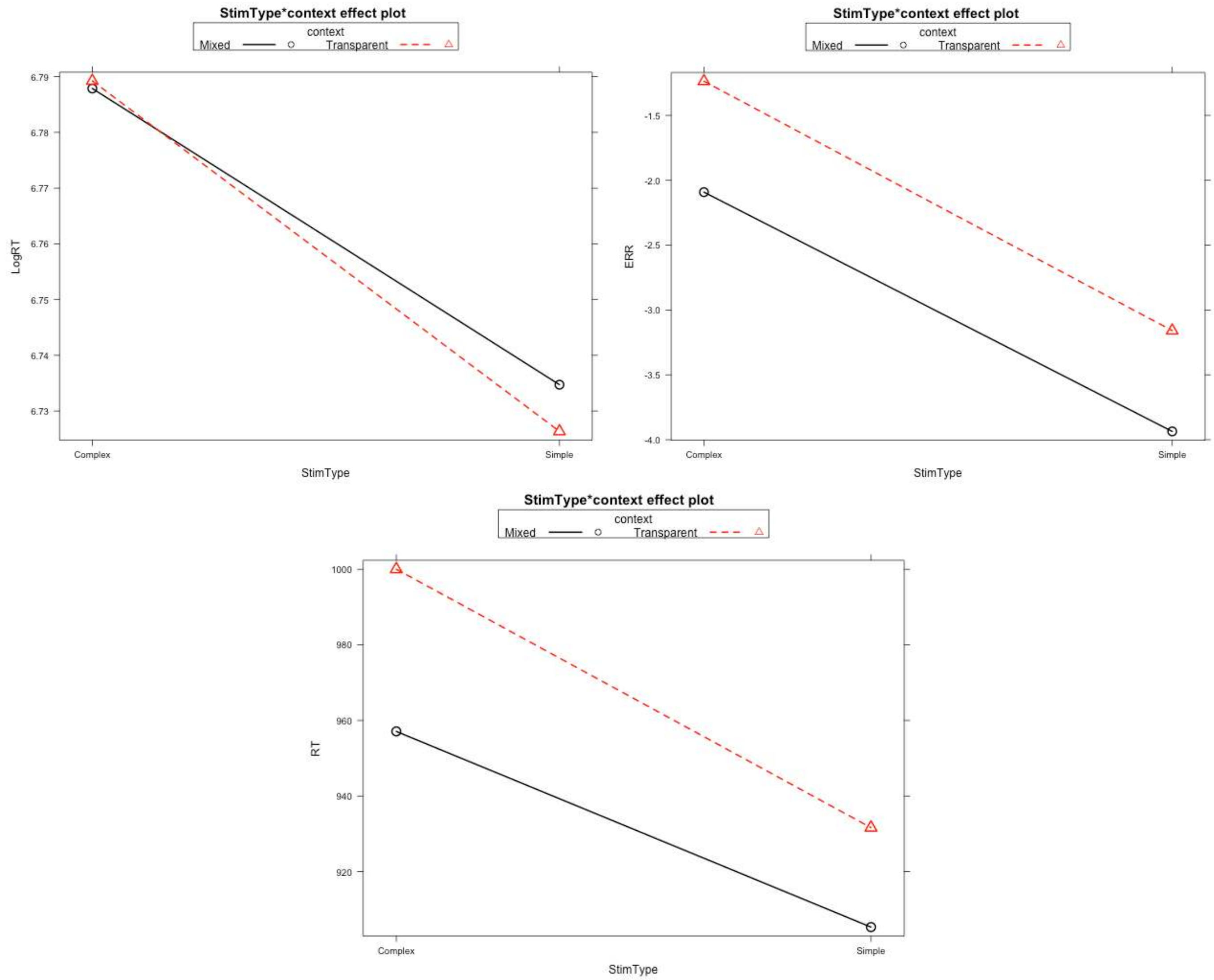



**Figure 1.** Nonword complexity effect in reaction time for the mixed and transparent contexts. Reaction times transformed using the natural log as reaction time data are skewed. Panel 3 shows the raw reaction time data.

## Discussion

The main effect of morphological structure on RT and ER replicates the Crepaldi (2010) finding of the morphological complexity effect using the same stimuli. Additionally, experiment 1 confirmed that the nonword complexity effect generated by the Crepaldi et al. (2010) stimuli was not context sensitive, i.e., the effect did not disappear in either or reverse condition like the base frequency effect (Taft, 2004). Lastly, experiment 1 provided insight into which context generated the largest and potentially variable effect. The main effect of context (higher ER for transparent than mixed) suggests that individuals have more errors overall in the transparent condition. This allowed us to compare the nonword complexity effect to the base and family size effects as the word morphological effect are strongest and most robust when the relationship between the stem and the affix is transparent.

## Experiment 2

Experiment 1 established that the nonword complexity effect was not context sensitive. We therefore chose the context which generated more errors overall on nonwords. The increased errors in nonwords overall helped us avoid ceiling effects in relation to error rates and allowed for the possibility of more variability in the effect. The focus on transparent words additionally allowed us to extend the Kuperman & Van Dyke., (2011) findings regarding individual differences in sensitivity to base frequency effects by including transparent words with productive suffixes from Ford et al., (2010) which varied freely in terms of surface frequency, base frequency and family size. Experiment 2 further extended past literature on sensitivity to morphological structure in nonwords and words by characterizing systematic individual differences in the nonword complexity, base frequency and family size effects and their relationships to each other and to a battery of skill measures.

Experiment 2 examined individual differences in sensitivity to morphological structure in both nonword and word effects, in visual lexical decision, and the relationships between them. More specifically, experiment 2 explored whether three measures of sensitivity to morphological structure in nonwords and words (nonword complexity, family size, base frequency) index the same underlying dimension of variability by comparing individual effects both through correlational analysis and patterns of correlation with individual difference measures (e.g., vocabulary, spelling) meant to examine the quality of the connections in an individual's reading network (orthography, phonology, semantics).

To explore individual differences in the nonword and word effects, individual effect slopes for nonword complexity, family size, and base frequency were extracted from the separate LME models for words and nonwords. We directly compared the nonword complexity effect and the word morphological effects using correlational analysis. Additionally, we explored the relationship of two well-established morphological effects in words, the base frequency effect and family size effect. While these two measures encapsulate information related to the morphological constituents of a word, recent evidence suggests that these measures are separate predictors (Ford et al., 2010; Xu & Taft, 2015).

Individual difference measures were selected from measures established in the literature to vary with individual sensitivity to morphological structure. Each measure also taps into various aspects of an individuals' reading network. For example, skilled reading relies on the complex relationships and connections of information related to orthography (writing), semantics (meaning), and phonology (sound). Given the special statistical properties of morphology, i.e., related words both overlap in terms of form both orthographic and phonological and meaning (e.g., a TEACHER, TEACHES) and the structure is consistent across words (e.g., JUMPING, RUNNING), various differences in the structure of the reading network could affect morphological processing. For example, Yap et al. (2015) established that an individual's vocabulary size positively correlated with the effect of number of affixes in nonwords. Correlational analysis was used to relate the individual nonword complexity effect slopes with the individual difference battery (see Yap et al., 2015). Our correlation analysis functioned as both a conceptual replication and an extension of their design with a stronger manipulation and a controlled experimental design. These comparisons allowed us to explore whether the nonword complexity, family size, and base frequency effect are indexing the same underlying statistical properties and whether well-established measures of sensitivity to morphological structure in words (family size, base frequency) are separate as Ford et al., (2010) and Xu & Taft (2015) suggest.

## Method

### Participants

The participants were 87 undergraduate students enrolled in an introductory psychology course at the University of Connecticut who participated for course credit. All were native speakers of American English.

### Design and Materials

The experimental stimulus set contained the same set of Crepaldi et al. (2010) nonwords, filler morphologically complex nonwords, and filler words and nonwords as Experiment 1. However, to incorporate the base/surface frequency manipulation, the Rueckl & Aicher (2008) words were replaced with 108 semantically transparent words with productive suffixes from Ford et al. (2010), which independently vary in base and surface frequency and morphological family size. Frequency data were obtained from the CELEX database (Baayen et al., 1995). For each word, the base morpheme frequency, derived word-form frequency and family size were obtained. Lemma frequency was obtained to calculate base frequency (cumulative root frequency).

Base frequency numbers and Family Size were obtained using the procedure described by de Jong et al. (2000) for cumulative root frequency and family size respectively. Family numbers were calculated by identifying morphologically related words to the target word (e.g., target – CALCULATE, members – calculate, calculable, calculation, calculator). Compounds (e.g., WATCHTOWER) and hyphenated compounds (e.g., CHECK-IN) were also included per de Jong et al. (2003). Base Frequency numbers were then calculated by adding up the lemma frequency of each confirmed family member. For example, the base frequency for *calculator* would be the summed lemma frequency for *calculate, calculable,* and *calculation*.

**Table 3. General Characteristics of Critical Stimuli Used in Experiment 2**

| | Critical Nonwords | | | | Critical Words | |
|---|---|---|---|---|---|---|
| | Complex | | Simple | | New Transparent | |
| **Example** | **GASFUL** | | **GASFIL** | | **TREATMENT** | |
| Count | 32 | | 32 | | 104 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Frequency | - | - | - | - | 4.28 | 1.25 |
| Length | 7.08 | 1.19 | 7.08 | 1.19 | 7.30 | 0.86 |
| Syll | 2.36 | 0.48 | 2.34 | 0.48 | 2.7 | 0.8 |
| Orth N | 0.19 | 0.62 | 0.03 | 0.18 | 1.1 | 1.8 |

*Syll, number of syllables; 'N', neighborhood density; data from CELEX. Frequency natural log transformation.*

**Table 4. Specific Characteristics of Transparent Words**

| | Surface Frequency | Family Size | Base Frequency |
|---|---|---|---|
| **Ex. High** | **TREATMENT** | **SICKNESS** | **READINESS** |
| **Ex. Low** | **DAFTNESS** | **SCAVENGER** | **DEPORTATION** |
| *Mean* | 4.28 | 7.26 | 6.95 |
| *STDEV* | 1.25 | 4.92 | 1.19 |
| *MAX* | 7.04 | 31.00 | 9.11 |
| *MIN* | 0.00 | 2.00 | 3.76 |

***Note:*** *Frequency data were obtained from the CELEX database (Baayen et al., 1995). For each word, the base morpheme frequency, derived word-form frequency and family size were obtained. Lemma frequency was obtained to calculate base frequency (cumulative root frequency).Base frequency numbers and Family Size were obtained using the procedure described by de Jong et al. (2000)*

**Table 5. General Characteristics of Fillers Used in Experiment 2**

| | Filler Nonwords | | | | Filler Words | |
|---|---|---|---|---|---|---|
| | Multisyllabic/Multimorphemic | | Monomorphemic | | Monosyllabic/Monomorphemic | |
| **Example** | **ARMIGHTY** | | **PLOSIOB** | | **CATCH** | |
| Count | 36 | | 90 | | 90 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Frequency | - | - | - | - | 1.70 | 1.96 |
| Length | 7 | 0.83 | 5.28 | 0.73 | 5.00 | 0.82 |
| Syll | 2.29 | 0.78 | 1.59 | 0.59 | 1 | - |

*Syll, number of syllables; 'N', neighborhood density; data from CELEX. Frequency natural log transformation.*

## Individual Differences Battery

**Test of Word Reading Efficiency (TOWRE).** Participants were presented with an initial list of words, which get increasingly more difficult to read (length, complexity). Participants read the words out loud while the researcher marked incorrect responses. The students were under a 45-second time limit. Next the participants were presented with a list of pseudo words with the same set of instructions and time limit. Participants were also recorded and a separate investigator scored responses. (Torgensen, J. K., Wagner, R. K., & Rashotte, C. A., 1999)

**Nelson-Denny Vocabulary Test.** Participants were given unlimited time to complete a 50-question vocabulary test. Each question was in the form of a sentence with a missing word (e.g. to be intelligent is to be _____) and given four options to complete the sentence. Questions were of increasing difficulty (Nelson, M. J., Brown, J. I., & Denny, M. J., 1960)

**Author Recognition Task.** Participants were presented with a list of 66 authors and non-authors. The task was to indicate which names were authors. There was a penalty for guessing, as each non-author selected incurred a 1-point deduction. (Cunningham, A. E., & Stanovich, K. E., 1990)

**Comprehensive Test of Phonological Processing (CTOPP).** Participants completed two subtests: Blending Words and Nonwords. Participants were given phonemes. Participants were then asked to blend the sounds they were presented. (Wagner, R. K., Torgesen, J. K., & Rashotte, C. A.,1999b).

**Spelling Task.** Participants were presented with a list of words. Each word is written in two ways—one way is correct, and one is a misspelling. Participants were asked to click the correct spelling. Participants were given unlimited time to complete the task.

## Procedure

The lexical decision procedure was the same as experiment 1. Participants then completed the individual differences battery: Test of Word Reading Efficiency, Nelson-Denny Vocabulary Task, Author Recognition Task, Comprehensive Test of Phonological Processing.

## Results

As described in the introduction, we first conducted group-level analysis of the nonword and word effects to determine whether the effects were in line with the literature, then individual effects were extracted. We examined whether there was systematic variability in the nonword complexity, base frequency, and family size effects. Then we compared individual nonword (nonword complexity) and morphological word effects (base frequency, family size). Specifically, we examined individual differences in morphological effects and the relationship between the effects through correlations with each other and a battery of skill measures.

### Group-level Analysis

Correct response times (RTs) and error rates (ERs) were analyzed using linear-mixed effects (LME) modeling in R (Baayen, Davidson, & Bates, (2008). Subjects and items were entered as crossed-random factors. Reaction time (RT) data were log transformed. Analysis of reaction time generated t-values. An absolute t-value near two is considered an appropriate indicator of significance (see Baayen, 2008 for review). Additionally, following the procedure outlined in Kearns (2016), significance may be determined by examining change in chi-squared. Examining the delta chi-squared enabled us to examine whether a variable explained a significant amount of variance in the model. The analysis of error rates was conducted using the binomial function, which generates z scores from which p values could be directly calculated. The LME coefficient, b, is reported for the effects of interest to provide insight into the relationship between the fixed effect factor and dependent variable (e.g., a negative coefficient signifies a negative), along with the standard error. Fixed factors that were continuous variables were standardized to avoid spurious correlations and to facilitate LME analysis**.** Reaction time was only reported for correct responses. RTs faster than 250ms were removed. For slow reaction times, individual cutoffs were generated by calculating 3 standard deviations from an individual subjects' mean reaction time across target words. Reaction times slower than these cutoffs were replaced with the cutoff value. The random effects structure was established using the log-likelihood ratio model comparison test and included participant and item as intercepts. Reaction times were converted via natural log to approximate a normal distribution.

Group-level analyses were conducted for both nonwords and words to first confirm that the overall pattern of results was in line with the previous findings. Separate models were conducted for words/nonwords and reaction time/error rate (4 models in total). In these analyses (and all that follow), continuous predictor variables were scaled prior to entry in the model and reaction times were log transformed**.**

**Table 6. Fit Indices and Model Comparison Test Results for Nonword Complexity Reaction Time Models**

| | | | | | Model Comparison | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Fit Index | | | | Unconditional | | | Control | |
| *Model* | *AIC* | *BIC* | *logLik* | *deviance* | *Chisq* | *Df* | *p* | *Chisq* | *Df* | *p* |
| Unconditional | 2980.2 | 3006 | -1486.1 | 2972.2 | | | | | | |
| Control | 2894.4 | 2926.5 | -1442.2 | 2884.4 | 87.85 | 1 | 2.20E-16 | | | |
| Nonword Complexity (NWC) | 2888.5 | 2927.1 | -1438.2 | 2876.5 | 95.767 | 2 | 2.20E-16 | 7.9167 | 1 | 0.004898 |
| Random Effect of NWC | 2886.1 | 2937.5 | -1435 | 2870.1 | 102.19 | 4 | 2.20E-16 | 14.345 | 3 | 0.002471 |

*Note. AIC Akaike information criterion; BIC Bayesian information criterion; logLik negative log likelihood. a Deviance is equal to 2 logLik. Used for 2 model comparison tests. b Unconditional model includes person and item random effects. c Control model includes random effects plus trial order predictors. d This model cannot be compared to word main effects models because models are not nested.*

**Table 7. Fit Indices and Model Comparison Test Results for Nonword Complexity Error Rate Models**

| | | | | | Model Comparision | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Fit Index | | | | Unconditional | | | Control | | |
| *Model* | *AIC* | *BIC* | *logLik* | *deviance* | *Chisq* | *Df* | *p* | *Chisq* | *Df* | *p* |
| Unconditional | 3864.2 | 3884 | -1929.1 | 3858.2 | | | | | | |
| Control | 3840.5 | 3866.9 | -1916.2 | 3832.5 | 25.701 | 1 | 3.99E-07 | | | |
| Nonword Complexity (NWC) | 3797.5 | 3830.5 | -1893.7 | 3787.5 | 70.686 | 2 | 4.48E-16 | 44.985 | 1 | 1.99E-11 |
| RandomEff NWC | 3785.7 | 3832 | -1885.8 | 3771.7 | 86.475 | 4 | 2.20E-16 | 60.775 | 3 | 4.02E-13 |

*Note. AIC. Akaike information criterion; BIC Bayesian information criterion; logLik negative log likelihood. a Deviance is equal to 2 logLik. Used for 2 model comparison tests. b Unconditional model includes person and item random effects. c Control model includes random effects plus trial order predictors. d This model cannot be compared to word main effects models because models are not nested.*

First, nonword complexity (Complex, Simple) was entered as a fixed effect for reaction time**.** Log previous reaction time, trial order, and previous trial type (i.e. word, nonwords) were entered into the fixed effects to control for any potential influences of these variables. The effect of nonword complexity was also included as a random factor for individual participants**.** The primary analysis involved examining the effects of morphological complexity in nonwords. In line with the LME models conducted in Experiment 1, there was a strong effect of nonword complexity in reaction time (b = -0.067, SE = 0.023, |t| = 2.81, $\Delta x^2$ = 7.92). A separate analysis was conducted for Error Rate using mixed effects logistic regression. Similar to reaction time, nonword complexity (Complex, Simple) was entered as a fixed effect**.** Also in line with Experiment 1, there was a strong effect of nonword complexity in error rate (b = -1.85, SE = 0.24, |z| = 7.531, p < .001). Overall, complex nonwords take longer to identify and have a higher error rate. Figure 2 illustrates the effects of nonword complexity in reaction time and error rate.

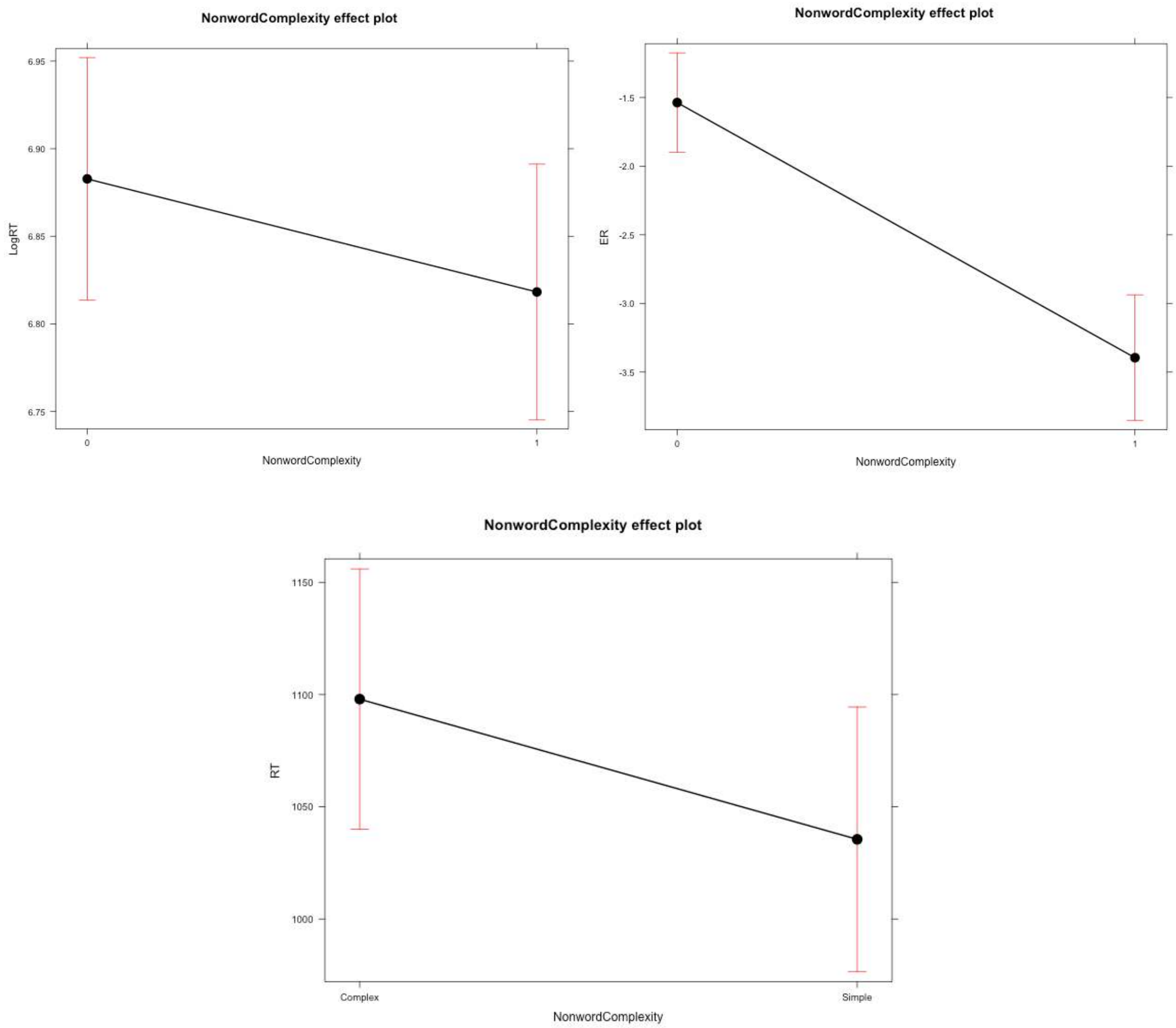
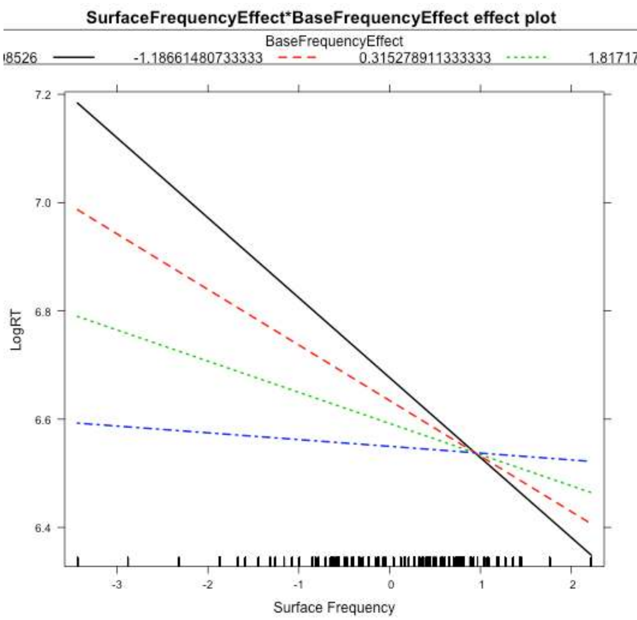


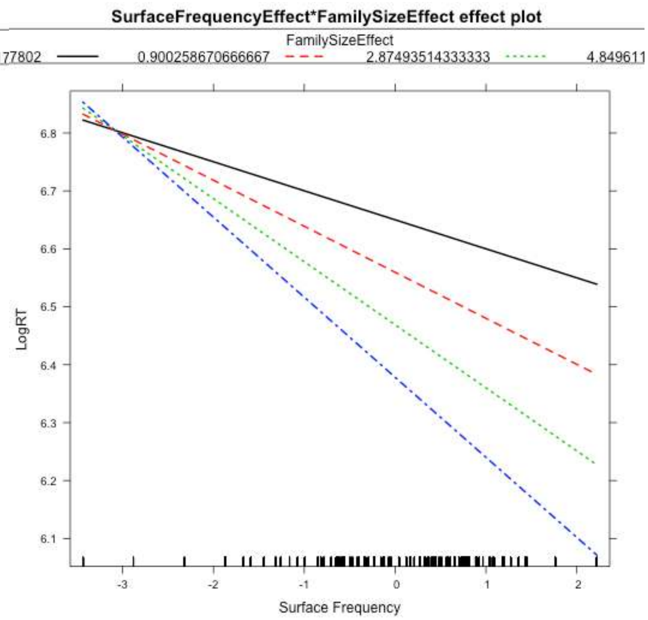

**Figure 2.** Panel 1 and 2 show the nonword complexity effect in reaction time and error rate extracted from the nonword linear mixed effects models. Reaction time in natural log units. Error rate in log odd units. Panel 3 shows raw reaction time data.

Turning to the morphological effects in words, we conducted group-level analysis for word reaction time and replicated the surface frequency, base frequency, and family size effects established in the literature. Surface frequency, base frequency, and family size were entered as fixed effects that interacted. Log previous reaction time and previous trial type (i.e. word, nonwords) were entered into the fixed effects as controls. Participants showed strong morphological effects for both base frequency and family size. Response time decreased as base frequency increased (b = -0.067, SE = 0.014, |t| = 1.96 , $\Delta x^2$ = 10.24) and as family size increased, reaction time decreased (b = -0.028, SE = 0.023, |t| = 2.81, $\Delta x^2$ = 11.89). There was also a strong surface frequency effect (b = -0.067, SE = 0.015, |t| = 4.52, $\Delta x^2$ = 27.18).

It is worth noting that both family size and base frequency were significant in the model and are thus separate predictors of reaction time (Xu & Taft, 2015). Moreover, while both family size and base frequency had a negative relationship with overall reaction time, they entered into a significant three way interaction (b = 0.05, SE = 0.018, |t| = 2.87). Additionally, while surface frequency, base frequency, and family size entered into a three way interaction, it is interesting to note the differential two way interactions with surface frequency. Further, in order to more easily interpret the three way interaction, it is useful to explore the separate two way interactions. For example, base frequency entered into a significant interaction with surface frequency (b = 0.030, SE = 0.015, |t| = 1.96, $\Delta x^2$ = 10.27) and there was no interaction between family size and surface frequency. Therefore at average family size and low base frequency there was a strong negative surface frequency effect (facilitative), but at high base frequency there was no surface frequency effect (Figure 3). In contrast, the effect of surface frequency is the same across family size at average base frequency. The pattern of facilitative to no effect and no interaction at low and high levels of base frequency and family size respectively, accompanied with overall faster reaction times at high levels of base frequency and equivalent reaction times for low surface frequency words for low and high family size, indicate that high levels of base frequency allow for low surface frequency words to be responded to as quickly as high surface frequency words, while family size has no effect on the surface frequency effect.

While family size does not affect the surface frequency effect on its own, family size modulates the interaction between base frequency and the surface frequency effect as indicated in the significant three-way interaction (b = 0.05, SE = 0.018, |t| = 2.87). At low family size, there is a small surface frequency effect at high base frequency and no surface frequency effect at low base frequency. At high family size, the surface frequency effect is actually strongly inhibitory at high base frequency, but strongly facilitative at low base frequency. High levels of both forms of morphological (family size, base frequency) seem to cause interference for the sensitivity to surface frequency information. However, words with high family size, but low base frequency information seem to facilitate the surface frequency effect.



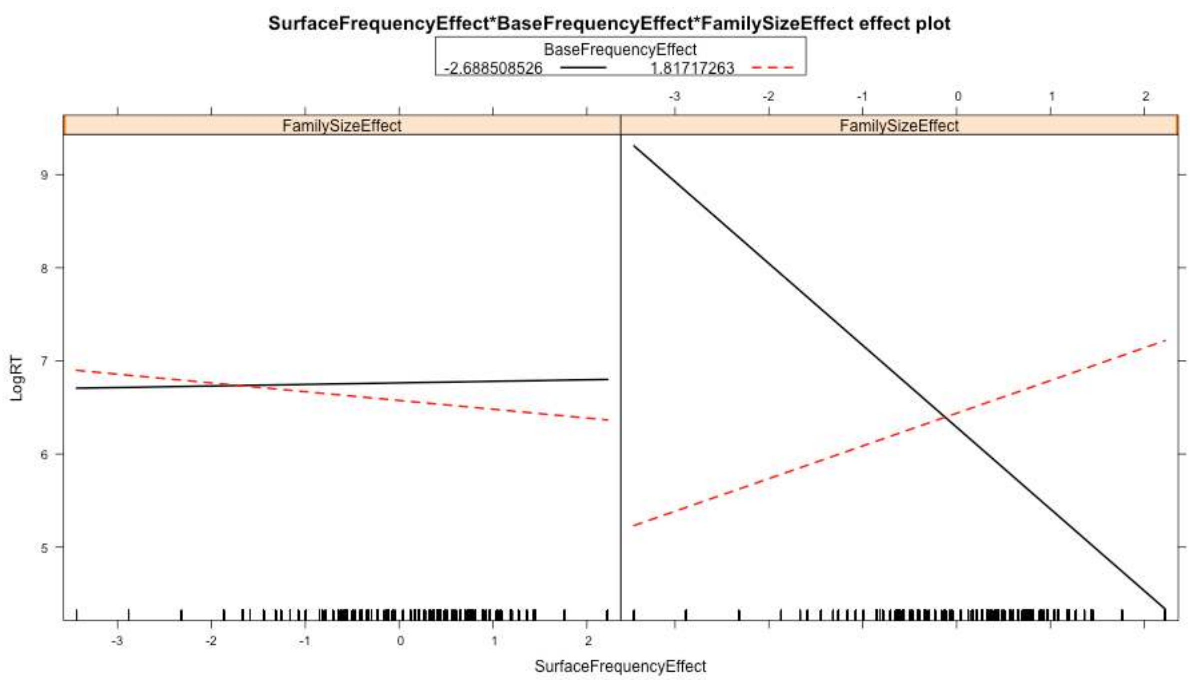



**Figure 3.** Panel 1 and Panel 2 show the interaction of the surface frequency effect and the base frequency and family size effects respectively. Panel 1 shows the surface frequency Effect at low base frequency and at high base frequency on the left and right respectively. Panel 2 also shows the surface frequency effect at low family size and high family size on the left and right respectively.

A separate analysis was conducted for Error Rate using mixed effects logistic regression. The surface frequency by base frequency interaction was not included in the model as including the interaction caused the model to not converge. Only the surface frequency effect (b = -0.79, SE = 0.177, |t| = 4.50, p < .001) was significant. Overall, participant error rates on words were very low and did not have enough systematic variation to conduct proper individual difference analysis and therefore were not examined further.

**Table 8. Fit Indices and Model Comparison Test Results for Word Reaction Time Models- Main Effects**

| | | | | | | Model Comparison | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Fit Index | | | Unconditional | | | Control | | |
| | *Model* | *AIC* | *BIC* | *logLik* | *deviance* | *Chisq* | *Df* | *p* | *Chisq* | *Df* | *p* |
| Main Effects | ModelU | 3087.2 | 3115.3 | -1539.6 | 3079.2 | | | | | | |
| | ModelB | 2872.1 | 2914.3 | -1430 | 2860.1 | 219.1 | 2 | 2.20E-16 | | | |
| | SF | 2846.9 | 2896.2 | -1416.4 | 2832.9 | 246.28 | 3 | 2.20E-16 | 27.182 | 1 | 1.85E-07 |
| | Base | 2845.2 | 2894.5 | -1415.6 | 2831.2 | 247.98 | 3 | 2.20E-16 | 28.879 | 1 | 7.71E-08 |
| | FS | 2856.2 | 2905.5 | -1421.1 | 2842.2 | 237.02 | 3 | 2.20E-16 | 17.913 | 1 | 2.31E-05 |
| | SF + BF | 2837 | 2893.4 | -1410.5 | 2821 | 258.17 | 4 | 2.20E-16 | 39.068 | 2 | 3.28E-09 |
| | SF + FS | 2831.4 | 2887.8 | -1407.7 | 2815.4 | 263.77 | 4 | 2.20E-16 | 44.663 | 2 | 2.00E-10 |
| | SF + BF + FS | 2828.8 | 2892.2 | -1405.4 | 2810.8 | 268.41 | 5 | 2.20E-16 | 49.31 | 3 | 1.12E-10 |
| | SF*BF | 2830.7 | 2901.2 | -1405.4 | 2810.7 | 268.44 | 6 | 2.20E-16 | 49.338 | 4 | 4.96E-10 |
| | SF*FS | 2830.6 | 2901.1 | -1405.3 | 2810.6 | 268.53 | 6 | 2.20E-16 | 49.43 | 4 | 4.75E-10 |
| | BF*FS | 2830.6 | 2901.1 | -1405.3 | 2810.6 | 268.54 | 6 | 2.20E-16 | 49.439 | 4 | 4.73E-10 |
| | SF*BF*FS | 2828.3 | 2919.8 | -1401.1 | 2802.3 | 276.9 | 9 | 2.20E-16 | 57.802 | 7 | 4.14E-10 |

*Note. AIC Akaike information criterion; BIC Bayesian information criterion; logLik negative log likelihood. a Deviance is equal to 2 logLik. Used for 2 model comparison tests. b Unconditional model includes person and item random effects. c Control model includes random effects plus trial order predictors. d This model cannot be compared to word main effects models because models are not nested. SF - Surface Frequency, BF - Base Frequency, FS - Family Size. + indicates additive, * indicates interaction*

## Individual Differences in Sensitivity to Morphological Structure

After the group-level analysis for words and nonwords were conducted, we explored individual differences in the nonword and word effects and then compared them. The random effects structure for each model was established using the log-likelihood ratio model comparison test and included both participant and item as intercepts. In order to test whether adding the random participant effects of our morphological variables (i.e., individual differences in the effects) accounted for unique variance, the person random effect of nonword complexity was added to the nonword models in reaction time and error rate and the person random effects of surface frequency, base frequency, and family size were added to the word model in reaction time. Consistent with the group level analysis, adding random effects for both family size and base frequency explained unique variance in the model suggesting that individuals vary on each effect separately.

**Table 9. Fit Indices and Model Comparison Test Results for Word Reaction Time Models- Random Effects**

| | | | | | | Model Comparison | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Fit Index | | | | | Main Effects | | | Previous | | |
| | *Model* | *AIC* | *BIC* | *logLik* | *deviance* | *Chisq* | *Df* | *p* | *Chisq* | *Df* | *p* |
| Random Effects | Base Main Effects Model | 2828.3 | 2919.8 | -1401.1 | 2802.3 | | | | | | |
| | SF1 | 2802.9 | 2908.6 | -1386.5 | 2772.9 | 29.355 | 2 | 4.22E-07 | | | |
| | Base2 | 2795 | 2900.7 | -1382.5 | 2765 | 37.237 | 2 | 8.21E-09 | | | |
| | FS3 | 2788.3 | 2894 | -1379.2 | 2758.3 | 43.929 | 2 | 2.89E-10 | | | |
| | SF + BF | 2789.6 | 2916.4 | -1376.8 | 2753.6 | 48.692 | 5 | 2.57E-09 | 19.337 | 3 | 2.33E-04 |
| | SF + FS | 2769.9 | 2896.7 | -1367 | 2733.9 | 68.371 | 5 | 2.24E-13 | 39.016 | 3 | 1.72E-08 |
| | SF + BF + FS | 2766.8 | 2921.8 | -1361.4 | 2722.8 | 79.445 | 9 | 2.08E-13 | 30.753 | 4 | 3.44E-06 |
| | SF + BF + FS | 2766.8 | 2921.8 | -1361.4 | 2722.8 | 79.445 | 9 | 2.08E-13 | 11.075 | 4 | 2.57E-02 |

*Note. AIC   Akaike information criterion; BIC   Bayesian information criterion; logLik   negative log likelihood. a Deviance is equal to  2   logLik. Used for 2 model comparison tests. b Unconditional model includes person and item random effects. c Control model includes random effects plus trial order predictors. d This model cannot be compared to word main effects models because models are not nested. SF - Surface Frequency, BF - Base Frequency, FS - Family Size. + indicates additive, * indicates interaction*

Individual differences in skill were determined with the individual differences battery. For each of the individual difference measures, three variables were calculated: speed, accuracy, and efficiency. In order to facilitate ease of interpretation, time was inverted such that larger numbers indicate faster, not slower. A logit transformation was then performed on the raw scores to generate accuracy scores (Mirman, 2014). Lastly, to combine both metrics, time was divided by number of answers correct and the sign was inverted to produce a measure of efficiency. In addition, following the procedure outlined in Andrews et al., (2011), spelling and vocabulary were entered into a principal component analysis (PCA) to examine more fine grain differences.

Before we compared individual differences in sensitivity to morphological structure in nonwords and words, we characterized both nonwords and words in terms of overall reaction time and error rate and the distributions of the associated morphological effects (nonword complexity, base frequency, family size). Table 10 shows the mean and standard deviation of reaction time and error rate to filler words and nonwords. Error rates for the filler words and nonwords were transformed into logits using the empirical logit transformation (Mirman, 2014). Overall, nonwords had longer reaction times and had more variance than words. In addition, individuals that had fast reaction times for words also had fast reaction times for nonwords ($r = .77$, $p < .001$). Within nonwords, there was no speed/accuracy ($r = .133$).

**Table 10. Mean and Standard Deviation of Nonword and Word Reaction Time and Error Rate**

| | | Transformed | | Raw | |
|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* |
| RT | Nonword | 6.7 | 0.25 | 888 | 486 |
| | Word | 6.45 | 0.12 | 655 | 159 |
| ER | Nonword | -1.97 | 0.74 | 0.14 | 0.35 |
| | Word | -2.51 | 0.55 | 0.08 | 0.27 |

*Note: Reaction Time transformed with the natural log. Error Rate transformed into Logits using the Empirical Logit transformation (Mirman, 2014).*

**Table 10a. Mean and Standard Deviation of Critical Nonwords and Words**

| | | | **Complex** | | **Simple** | |
|---|---|---|---|---|---|---|
| | | | *M* | *SD* | *M* | *SD* |
| Nonword Effects | Nonword Complexity | RT | 1079 | 435 | 1031 | 449 |
| | | ER | 0.25 | 0.08 | 0.15 | 0.12 |
| | | | **Quartile 4** | | **Quartile 1** | |
| | | | *M* | *SD* | *M* | *SD* |
| Word Effects | Surface Frequency | RT | 863 | 514 | 880 | 601 |
| | Base Frequency | | 863 | 504 | 874 | 567 |
| | Family Size | | 860 | 562 | 861 | 489 |

*Mean and Standard Deviation raw reaction time and error rate data for Complex and Simple Nonwords and RT for Quartile 4 and Quartile 1for Surface Frequency, Base Frequency, and Family Size information.*

Individual effects of nonword complexity, family size, base frequency, and surface frequency were extracted from the LME models. The individual participant coefficients describe the person-specific sensitivity to morphological structure in words and nonwords and are the primary outcome measures that we used to examine individual differences as a function of various skill measures. As each of the word morphological effects produced negative overall effects, individuals with more negative slopes actually had larger effects of morphological structure. Therefore negative slopes (e.g., family size) were sign reversed (negative to positive) to reflect the magnitude of each effect for ease of interpretation. In other words, larger (more positive) individual effects terms then indicated the magnitude of the effect. Figure 3 shows the density plot of each of the effects. The nonword complexity effect in reaction time varied from almost no effect to a relatively large effect. However, most individuals have a large effect while relatively few individuals have a small or no effect. In error rates, participants also varied from almost no effect to a relatively large effect. There seemed to be a more even distribution for error rate than for reaction time. Most individuals have small base frequency effects relative to the family size and surface frequency effects. However, whereas all individuals have relatively strong surface frequency effects, some individuals have no base frequency or family size effect. Lastly, the family size effect had the most variability in the word effects.

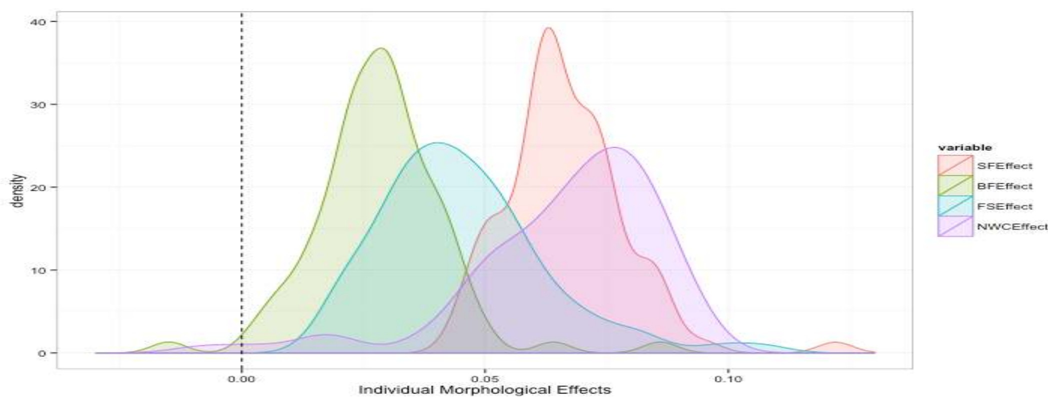



**Figure 3.** Density plot for individual morphological effects extracted from the nonword and word LME analyses. Red – Surface Frequency, Green – Base Frequency, Blue – Family Size, Purple – Nonword Complexity

Turning to the relationship between the nonword and word effects, Table 11 shows the correlations between the nonword and word morphological effects and overall reaction time and error rates. First, the overall nonword reaction time and the nonword complexity effect were highly negatively correlated meaning that individuals that were faster overall also had larger nonword complexity effects. Similarly, as overall nonword error rate increased, the nonword complexity effect in error rate also increased. Conversely, overall word reaction time was highly positively correlated with the surface frequency effect and family size effect, but not base frequency effect. Next, we directly compared the nonword complexity effect and established morphological effects and whole-word effects in words. Interestingly, while both family size and base frequency were highly positively correlated with surface frequency, they were not strongly related to each other (Figure 4). Similarly, the nonword complexity effect in reaction time was highly negatively correlated with family size and surface frequency effects but not correlated with base frequency effect (Figure 5). This relationship mirrors (opposite of) the relationship between the word morphological measures (family size and base frequency effects) and the surface frequency effect.

This further supports the evidence from the group-level model that 1) each word effect is a separate predictor that accounts for unique variance and 2) each effect has a different pattern of interaction with surface frequency. The nonword complexity effect in error rate did not reliably correlate with any of the other effects, however there was a trending correlation with base frequency. In summary, as the surface frequency effect increases, both the family size and base frequency effects increase and the nonword complexity effect in reaction time decreases, which seems to suggest that both family size and base frequency are closely related. However, the family size and base frequency effects are not strongly correlated. Further, while the family size effect is strongly related to the overall word reaction time and the nonword complexity effect, the base frequency effect is not related to either. Altogether, these data suggests that, while the nonword complexity and family size effects have been put forth as measures of form based and semantic processing in morphology respectively, nonword complexity and family size pattern together. However, the nonword complexity effect and the base frequency effect – both purported as measures of form based processing – do not pattern similarly or correlate strongly. This provides evidence that 1) the nonword complexity effect may be semantic in nature, similar to family size and 2) the family size and base frequency effects are separate predictors of reaction time.

**Table 11. Bivariate Correlations Between Overall Nonword/Word RT and ER and Nonword and Word Effects**

| | | | Overall | | | | Effect | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Reaction Time | | Error Rate | | Reaction Time | | | |
| | | | Nonword | Word | Nonword | Word | Nonword Complexity | Family Size | Base Frequency | Surface Frequency |
| **Overall** | RT | Nonword | -- | | | | | | | |
| | | Word | **0.770**** | -- | | | | | | |
| | ER | Nonword | 0.133 | **-0.253*** | -- | | | | | |
| | | Word | **-0.323*** | -0.191 | -0.056 | -- | | | | |
| **Effect** | RT | Nonword Complexity | **-0.821**** | **-0.562**** | -0.205 | 0.154 | -- | | | |
| | | Family Size | **0.803**** | **0.902**** | -0.106 | -0.104 | **-.649**** | -- | | |
| | | Base Frequency | 0.158 | 0.089 | 0.008 | -0.048 | -0.028 | **.239*** | -- | |
| | | Surface Frequency | **0.630**** | **0.636**** | -0.130 | -0.169 | **-.419**** | **.647**** | **.667**** | -- |
| | ER | Nonword Complexity | -0.039 | 0.118 | **-0.542**** | -0.142 | 0.085 | 0.008 | 0.158 | 0.150 |

*Note: Reaction Time transformed with the natural log. Error Rate transformed into Logits using the Empirical Logit transformation (Mirman, 2014). Effects are derived from individual random slopes extracted from separate LME Models for Nonwords (RT, ER) and Words (RT)*

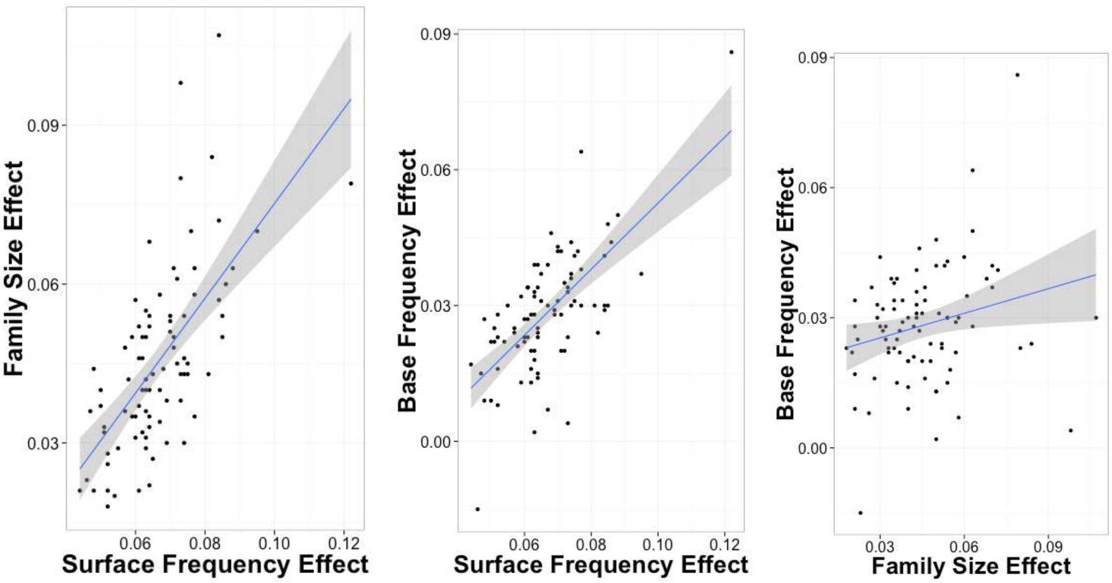



**Figure 4.** Scatterplots of relationships between individual whole-word (Surface Frequency) and morphological effects (Base Frequency, Family Size) effects extracted from the word LME model.

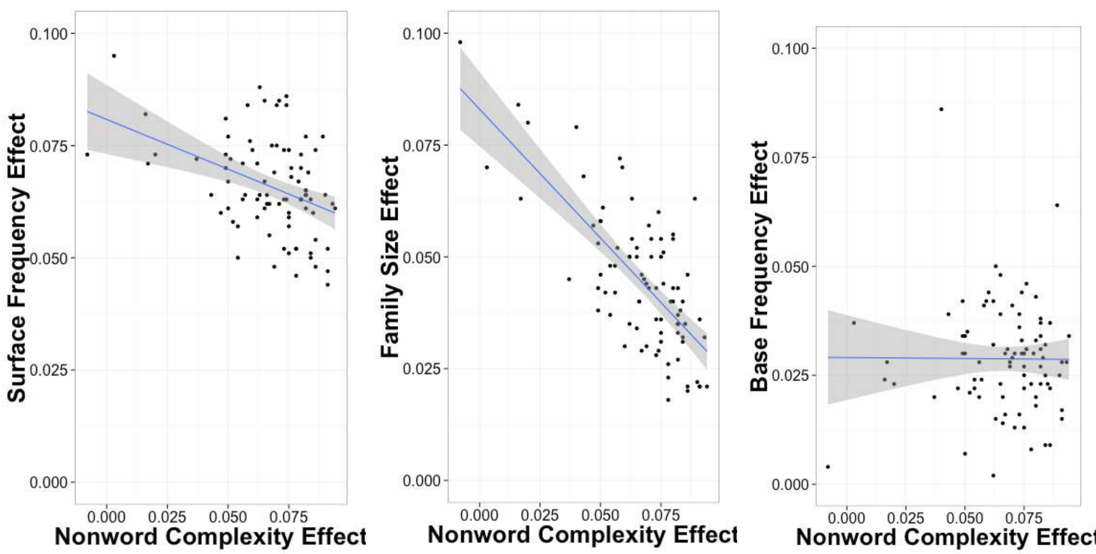



**Figure 5** Scatterplots of relationships between individual word (Surface Frequency, Base Frequency, Family Size) and nonword (nonword complexity in reaction time and error rate) effects extracted from separate nonword and word LME models.

To further explore the relationship between the nonword complexity effect and the word morphological effects (e.g., whether they tap into the same underlying dimension of variation), we examined the patterns of correlation of the morphological effects and the skill measures. However, before we examined the patterns of correlation between the morphological effects and the skill battery, we looked at the relationships between the skill measures. Table

12 displays the correlations between measures in the individual difference battery. First, ART, Vocabulary and Spelling Efficiency, were highly correlated. For example, as individuals spelling score increased, their vocabulary score also increased. The PDE and the CTOPP measures (Nonword Repetition and Blending Words) were also highly positively correlated. For example, as an individual's PDE score increased, their nonword repetition score decreased. Interestingly, ART and spelling, were not strongly related to either the CTOPP or TOWRE measures, but Vocabulary was correlated with all of the measures in the skill battery.

**Table 12. Bivariate Correlations Between Skill Measures**

| | | ART Efficiency | Vocabulary Efficiency | Spelling Efficiency | Sight Word Efficiency | Pseudoword Decoding Effeciency | Nonword Repetition |
|---|---|---|---|---|---|---|---|
| Reading Related Measures | ART Efficiency | | | | | | |
| | Vocabulary Efficiency | .389** | - | | | | |
| | Spelling Efficiency | .429** | .507** | - | | | |
| TOWRE | Sight Word Efficiency | 0.059 | .258* | 0.131 | - | | |
| | Pseudoword Decoding Effeciency | 0.201 | .347** | .314** | .243* | - | |
| CTOPP | Nonword Repetition | 0.173 | .266* | 0.189 | 0.039 | .384** | - |
| | Blending Words | -0.018 | .269* | 0.135 | 0.048 | .395** | .307** |

Table 13 shows bivariate correlations between the morphological effects and individual difference battery. First, correlations between the nonword complexity effect and skill measures were conducted. Vocabulary, ART, spelling, TOWRE, and nonword repetition measures significantly correlated with the nonword complexity effect. All correlations follow the same pattern across the nonword complexity effect and measures of skill. We presented findings from vocabulary efficiency as an example (Figure 6). Individuals with higher vocabulary have larger nonword complexity effects. While, on average, participants display a negative effect of nonword complexity, individuals with low vocabulary had no or very small effects and individuals with high vocabulary had very large effects. The nonword complexity effect in error rate only positively correlated with vocabulary.

**Table 13. Bivariate Correlations Between Nonword and Word Effects in RT and ER and Skill Measures**

| | | Nonword and Word Effects | | | | |
|---|---|---|---|---|---|---|
| | | Error Rate | Reaction Time | | | |
| | | Nonword Complexity | Nonword Complexity | Family Size | Base Frequency | Surface Frequency |
| Reading Related Measures | ART Efficiency | 0.098 | **.375**** | **-.284**** | -0.079 | **-.259*** |
| | Vocabulary Efficiency | **.257*** | **.446**** | **-.371**** | **-.364**** | **-.392**** |
| | Spelling Efficiency | 0.184 | **.420**** | **-.290**** | -0.136 | **-.221*** |
| TOWRE | Sight Word Efficiency | -0.126 | 0.148 | -0.193 | -0.072 | -0.149 |
| | Pseudoword Decoding Effeciency | **.303**** | 0.18 | -0.005 | -0.14 | -0.037 |
| CTOPP | Nonword Repetitons | **.278*** | **.297**** | -0.157 | -0.081 | -0.012 |
| | Blending Words | 0.127 | 0.103 | -0.076 | **-.294**** | -0.202 |

*Note: Skill measures were transformed. Accuracy was divided by speed and inverted to produce a measure of efficiency. Reaction Time transformed with the natural log. Error Rate transformed into Logits using the Empirical Logit transformation (Mirman et al. 2011). Effects are derived from individual random slopes extracted from separate LME Models for Nonwords (RT, ER) and Words (RT)*

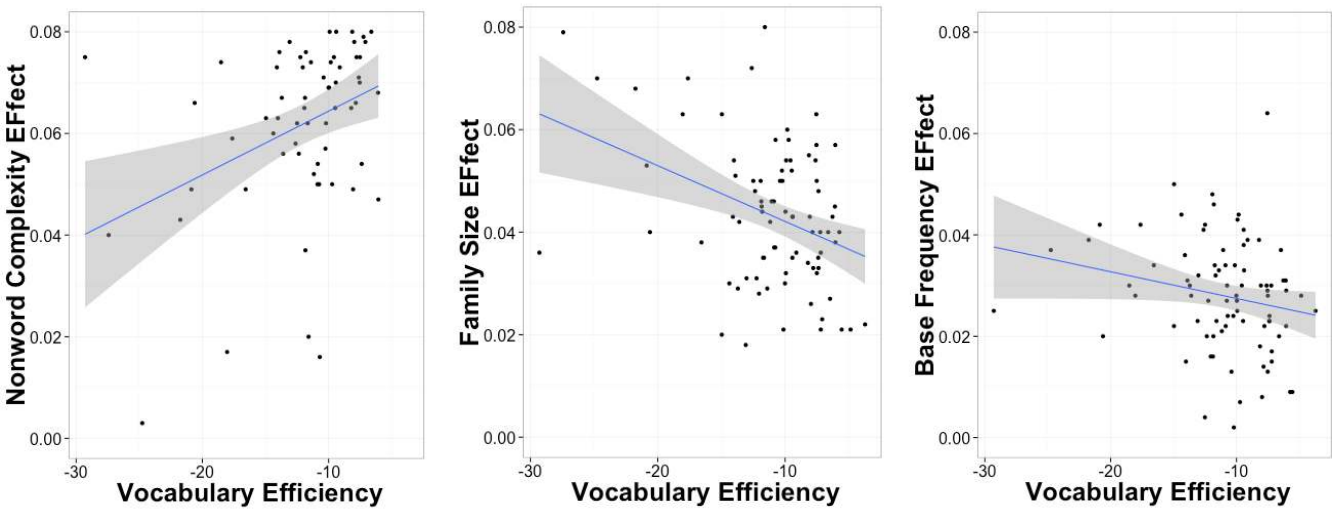



**Figure 6.** Scatterplots of the Nonword Complexity, base frequency, and family size effects and skill respectively. individual random slopes in reaction time and vocabulary.

Next, we examined the patterns of correlation between the word morphological measures and the skill measures. Only vocabulary significantly correlated with both morphological effects. Individuals with higher vocabulary have smaller morphological effects. In order to more thoroughly examine the relationship between base frequency and family size effects, we took a closer look at the distribution of the effects in relation to the skill measures While, on average, participants display a negative effect of both morphological effects, participants with low vocabulary had large effects, but participants with high vocabulary had almost no effect of morphological structure (attenuated in Family Size) (Figure 6b, c). This pattern mirrors the nonword complexity effect, where individuals with high skill have a larger effect and individuals with low skill have almost no effect.

While base frequency and family size pattern in a very similar manner for vocabulary, they are separate predictors of reaction time in words, as the independent significance of each variable (Xu & Taft, 2015), and the modest correlation between base frequency and family size would suggest. We looked at patterns of correlation with the skill measures to further pull these measures apart and compare them to nonword complexity.

First, family size and nonword complexity effects pattern together in terms of correlations with skill measures. In line with the negative correlation between the family size and nonword complexity effect, family size and nonword complexity show opposite patterns of correlation with ART, Vocabulary, and Spelling. For example, as ART efficiency increases, the family size effect decreases and the nonword complexity effect increases. However, the base frequency effect does not correlate strongly with either ART or spelling efficeiency and was instead highly negatively correlated with blending words (CTOPP phonological processing task). While the base frequency effect was related to phonological (form) processing skill, the family size effect was related to ART, a measure of exposure to print. Figure 7 illustrates this mirrored pattern of correlation with ART the nonword complexity and family size effects and the lack of correlation with the base frequency effect. In addition, the nonword complexity effect patterned most closely with family size, but was also correlated with nonword repetition, another measure of phonological processing, similar to blending words. Interestingly, base frequency seems to pattern more with tasks traditionally thought to measure aspects of an individual's reading profile related to form (phonological processing), whereas family size patterns more with tasks traditionally thought to measure aspects of the individual's skills related to meaning/lexical quality (author recognition task, spelling).

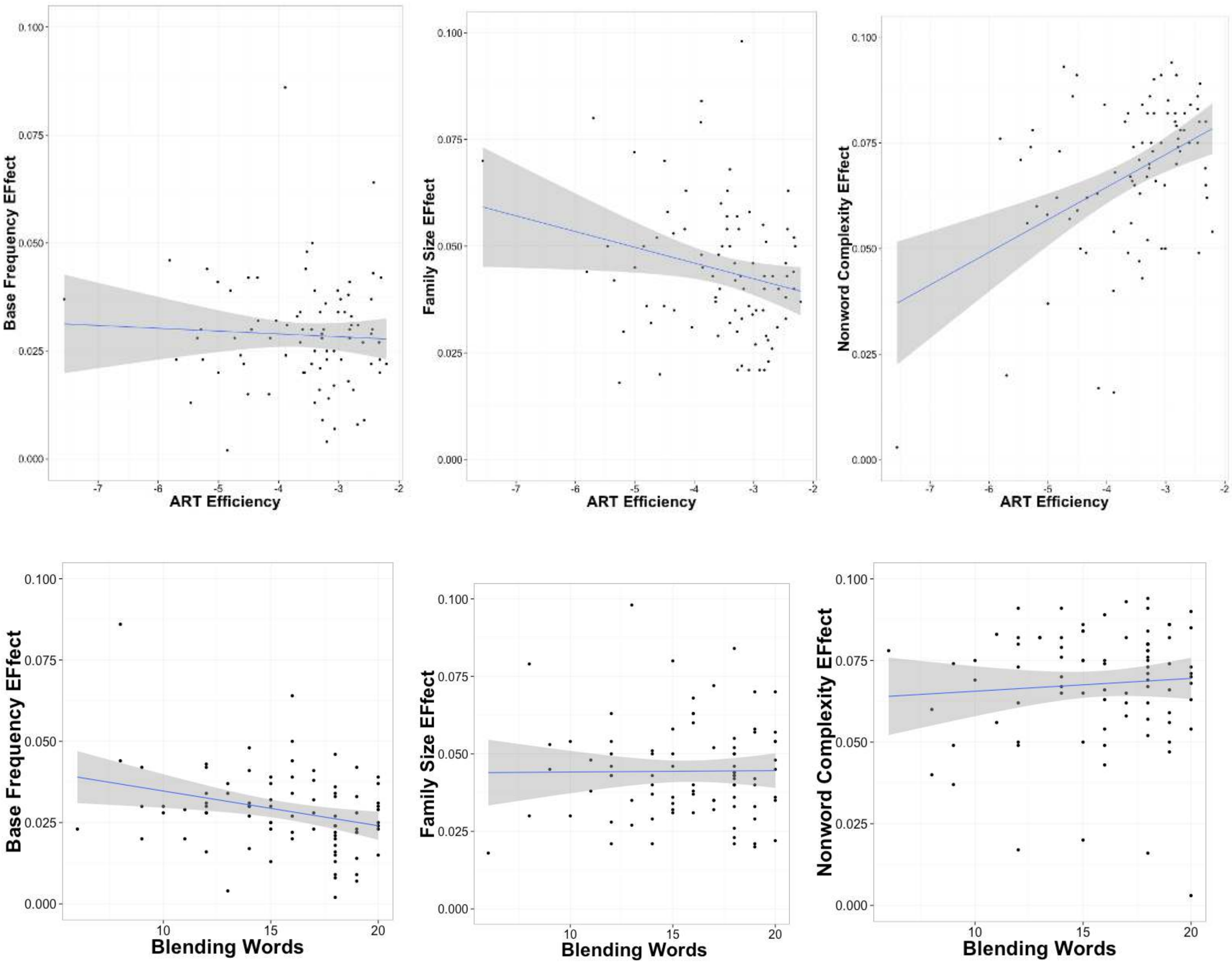



**Figure 7.** Scatterplots of individual base frequency, family size, and nonword complexity effects (left to right) for ART efficiency

As base frequency and family size pattern differentially with aspects of an individual's reading profile associated with phonological processing and lexical quality respectively, we followed the procedure in Andrews et al. (2011) in which vocabulary and spelling were included in a PCA. This approach was valuable to the current study as Andrews was able to use it to show individual differences in morphological priming effects. Component two (most relevant to the current study) reflected unique variance differentiating spelling and vocabulary with overall skill partialled out. Negative individual PCs represented individuals with higher spelling scores relative to vocabulary ("orthographic profile") while individuals with positive individuals PCs represented individuals with higher vocabulary scores relative to spelling ("semantic profile") (Andrews et al. 2011). Individuals with the orthographic profile produced equivalent priming effects for both form and opaque pairs, whereas individuals with the semantic profile produced graded priming for form, opaque, and transparent pairs, which suggested that fine grain differences in an individual's reading profile affect both the sensitivity to and use of morphological structure. By using this technique, we were able to examine these fine grain differences in the use of various sources of morphological information (nonword complexity, family size, base frequency).

Within our data, we included both vocabulary and spelling efficiency in a PCA. The first dimension (PC1) correlated positively with both vocabulary (r =. 86) and spelling (r = .86) and captured 75% of the common variance. Dimension 2 (PC2) on the other hand captured 25% of the common variance and showed opposite patterns of correlation with spelling (r= -.497) and vocabulary (r = .497) similar to Andrews. Table 14 displays the correlations between dimensions 1 and 2 and the nonword and word effects. Unsurprisingly, dimension 1 followed the same pattern of correlation with the nonword and word effects as vocabulary. As individual PCs on dimension 1 increased the nonword complexity effect in reaction time and error rate increased while the family size effect, base frequency effect, and surface frequency effect all decreased. However, dimension 2 only significantly correlated with the base frequency effect (Figure 9). Individuals with more negative individual PCs on dimension 2, (orthographic profile; Andrews et al., 2011) had larger base frequency effects and individuals with more positive individual PCs (semantic profile; Andrews et al., 2011) had smaller base frequency effects.

**Table 14. Bivariate Correlations Nonword/Word Effects and Vocabulary and Spelling Dim 1 and 2**

| | | | Dim 1 | Dim 2 |
|---|---|---|---|---|
| Nonword and Word Effects | RT | Nonword Complexity | **.497**$^{**}$ | 0.026 |
| | | Family Size | **-.379**$^{**}$ | -0.081 |
| | | Base Frequency | **-.288**$^{**}$ | **-.230**$^{*}$ |
| | | Surface Frequency | **-.353**$^{**}$ | -0.172 |
| | ER | Nonword Complexity | **.253**$^{*}$ | 0.073 |

*Note: procedure in Andrews et al. (2011) as described previously. The first dimension (PC1) correlated positively with both vocabulary (r =. 86) and spelling (r = .86) and captured 75% of the common variance. Component 2 (PC2) on the other hand captured 25% of the common variance and showed mirrored patterns of correlation with spelling (r= - .497) and vocabulary (r = .497)*

## Discussion

Experiment 2 examined whether the nonword complexity, family size, and base frequency effects displayed systematic variability and whether these measures of sensitivity to morphological structure in nonwords and words index the same underlying dimension via direct comparisons and patterns of correlation with skill measures. First, each measure of sensitivity to morphological structure showed substantial and systematic variability. In addition, the family size and nonword complexity effects seem to be indexing the same dimension (albeit opposite ends), but the base frequency effect pulls apart from both measures in a systematic way.

Each effect patterned with skill measures in a manner consistent with the literature. First, the nonword complexity effect in both reaction time and error rate and the family size and base frequency effect in reaction time showed large variability across participants. The nonword complexity effect in reaction time, across all indices of individual differences, varied following a simple principle: as skill increased, the nonword complexity effect also increased.

Further, individuals with low skill produced almost no effect of nonword complexity and individuals with high skill produced large effects of nonword complexity. Specifically, as overall reaction time for words and nonwords decreased, the nonword complexity effect increased. In terms of the skill battery, the nonword complexity as skill on the three reading-related measures (ART, Vocabulary, Spelling) increased, the nonword complexity effect increased. The nonword complexity effect also followed this pattern for a measure of phonological processing (Nonword Repetition). The nonword complexity effect in error rate also followed the same principle (as skill increased, the effect increased). However, the nonword complexity effect in error rate increased as overall error rate for only nonwords. In addition, the nonword complexity effect in error rate was only related to one of the reading-related measures (Vocabulary), but was related to both of the nonword phonological and orthographic processing measures (Nonword Repetitions, Pseudoword Decoding).

The family size effect patterned most similarly to the nonword complexity effect in reaction time, albeit in opposite directions. For example, while the nonword complexity effect increased as skill increased across indices of individual differences, the family size effect actually decreased. Further mirroring the nonword complexity effect in reaction time, individuals with high skill actually had very small or no effects of family size, while individuals with low skill were very sensitive to family size. Specifically, as overall reaction time in words and nonwords decreased (faster), the family size effect also decreased. In line with the nonword complexity effect, the family size effect was strongly related to the three reading related measures (ART, Vocabulary, Spelling), but as skill in these measures increased, the family size effect decreased. Consistent with the characterization that the family size effect is semantic in nature (see Baayen, 2014 for review), the family size effect was not related to any of the form based processing measures (Pseudoword Decoding, Blending Words, Nonword Repetitions).

The base frequency effect did not pattern with the family size effect or nonword complexity effect in reaction time, but did pattern somewhat similarly to the nonword complexity effect in error rate. Specifically, the base frequency effect was not related to overall reaction time in nonwords or words. In addition, similar to the nonword complexity effect in error rate, the base frequency effect was also only related to one of the reading related measures, vocabulary. Further, unlike family size and consistent with the characterization of the base frequency effect as an effect related to the form based processing in morphology (e.g., Ford et al., 2010), the base frequency effect was related to word phonological processing (Blending Words) (also similar to the nonword complexity effect in error rate). However, as skill increased on the relevant measures increased, the base frequency effect decreased (similar to the other word effects). In particular, individuals with larger vocabularies and more skill in phonological processing have almost no effect of base frequency, but individuals with less skill in these measures are very sensitive (large effects) to base frequency.

Turning to the similarities between the measures in terms of correlations between the measures and with the skill battery, the nonword complexity effect in reaction time and the family size effect and, to a lesser extent, the nonword complexity effect in error rate and the base frequency effect seem to be related. First, the nonword complexity effect in reaction time and the family size effect were highly negatively correlated. As the nonword complexity effect increased, the family size effect decreased. Consistent with this strong negative correlation, the nonword complexity effect in reaction time and the family size effect patterned very closely across both measures of overall reaction time and all three of the reading related measures (ART, Vocabulary, Spelling), albeit in opposite directions. For example, as vocabulary size increased, the nonword complexity effect increased and the base frequency effect decreased. More specifically, individuals with large vocabularies had almost no effect of family size, but large effects of nonword complexity but individuals with small vocabularies were very sensitive to family size, but not very sensitive to nonword complexity. Also, the nonword complexity effect in error rate and the base frequency effect pattern somewhat similarly, as both were not correlated with ART and spelling, but were instead correlated (also in opposite directions) with vocabulary and form processing related measures (pseudoword decoding and nonword repetition for nonword complexity and blending words for base frequency). For example, as phonological processing skill increased, the nonword complexity effect in error rate increased and the base frequency effect decreased. Lastly, while the base frequency effect and family size pattern with different sets of skill measures, they are both strongly positively correlated with the effect of surface frequency. Similarly, all three effects – nonword complexity (reaction time and error rate), family size, and base frequency, were related to vocabulary size (positive correlation for nonword effects and negative for word effects).

Conversely, while the nonword complexity effect in reaction time and family size seem to be indexing the same dimension of variability, the base frequency effect seems to index a separate dimension. First, examining the main effects analysis, the base frequency effect and the family size effect enter into differential interactions with surface frequency, with base frequency enhancing overall reaction time for low and high surface frequency words and family size enhancing the surface frequency effect. In addition, both the base frequency and family size effects accounted for unique variance in the LME analysis, suggesting that they were separate predictors of reaction time.

Further, turning to the individual difference analysis, including both the base frequency and family size effects as random factors accounted for unique variance, which suggested that the base frequency and family size effects vary across individuals separately. In line with these analyses, the base frequency and family size effects are weakly correlated. Further, the base frequency effect and the nonword complexity effect in reaction time are not correlated. The nonword complexity effect in error rate is not correlated at all with any of the other measures of sensitivity to morphological structure. In terms of the patterns of correlation, while the family size effect and the nonword complexity effect in reaction time pattern very closely to each other, the base frequency effect patterns very differently from family size and nonword complexity (reaction time), despite overlapping correlations with vocabulary. For example, while family size and nonword complexity (reaction time) are strongly correlated with overall reaction time in words and nonwords and with ART and spelling skill, the base frequency effect is only correlated with form based processing measures (word phonological processing). In addition, the base frequency effect varied along dimension 2 of the PCA described in Andrews et al., (2011) which explores the unique variance differentiating vocabulary and spelling skill with overall proficiency martialled out. This allowed us to examine fine grain differences in processing as individuals with the orthographic profile in Andrews et al, (2011) were more affected by obligatory decomposition processes due to the form rather than semantic based processes. Individuals with the orthographic profile (individuals with a higher spelling score relative to vocabulary) in our study, had larger base frequency effects providing further evidence for base frequency as a morphological measure related to form (Ford et al., 2010; Xu & Taft, 2015).

## General Discussion

The current study found that the nonword complexity effect was robust to word context (transparent, mixed) and could therefore be could be examined simultaneously with morphological effects in transparent words (base frequency, family size). In addition, while the main effects of nonword complexity, base frequency, and family size in the group level analysis were consistent with the literature, an important finding was that base frequency and family size account for unique variance and have different patterns of interaction with surface frequency (no interaction with family size). Turning to the individual differences analysis, the nonword complexity effect in reaction time and error rate and the base frequency and family size effects in reaction time show systematic variability in relation to the battery of skill measures and overall reaction time and error rate to nonwords and words. Further, the nonword complexity effect in reaction time and family size effect seem to index opposite ends of the same dimension of variability, while the base frequency effect and nonword complexity effect in error rate index separate dimensions entirely.

**Connection to Previous Literature**

First, in terms of including both morphologically complex nonword and word stimuli, one might suggest that given that our critical stimuli (decomposable nonwords with real world stems) are similar in construction to the Taft (2004) nonword fillers, we might expect to not find base frequency effects and actually replicate the finding of the reverse base frequency effect (medium base/low surface vs. high base/low surface) for low surface frequency words and attenuated base frequency effect in more typical (low base/low surface vs. medium base/low surface) words in the context of decomposable nonwords with real stems. However, only the complex critical nonwords (32 words) follow similar construction principles as the distractor nonwords in Taft (2004) – real world stem plus syntactically correct suffixes (facilitating obligatory decomposition). Both the simple critical nonwords (32 nonwords, “stem” plus no stem – “non-decomposable”) and filler nonwords (126 nonwords) did not follow the “stem+suffix” format of the distractor nonword. In other words, only 32 out of the 190 total nonwords (~17%) and would not produce effects in line with the reverse base frequency effect. The stimuli are in fact more consistent with the nonsense condition in Taft (2004) which generated the classic base frequency effect. Further, the words taken from Ford et al. (2010) included all productive suffixes, while Taft (2004) did not control for this. Taft (2004) also tightly controlled the surface frequency (low/high) and base frequencies (medium/high or medium/low) within relatively small ranges. Ford et al., 2010 words are instead allowed to vary freely.

The nonword complexity effect in reaction time and error rate was in line with Experiment 1 and with previous findings in the literature and replicated the findings in Crepaldi et al., (2010) using the same materials. A unique contribution of this paper was to examine individual differences in sensitivity to morphological structure in nonwords using a stronger manipulation of morphological complexity (decomposable vs. non-decomposable) than Yap et al., (2015) while also conducting a controlled, in-laboratory experiment (Yap et al., (2015) used data from the English Lexicon Project). It is interesting to note, Yap et al., examined individual differences in morphological structure in nonwords via the number of affixes and not the tightly controlled decomposable vs. non-decomposable contrast in Crepaldi et al., (2010). Further, Yap et al., used the automated Affix Detector program as described in Muncer et al., (2013). The program finds morpheme-like elements in a nonword, based on a comprehensive list of affixes listed in Fudge (1984). However, Muncer et al., (2013b), in an analysis of only inflectional affixes using the same program, noted that identifying morphemes solely on the basis of the presence of letter strings that match the list of approved affixes is potentially inaccurate. They provide the example of “s” as an inflectional ending, particularly in words ending in “s” (e.g., in words ending in *is, us, os,* or *ss,* “s” is not necessarily a morpheme). This automated process weakens the overall definition of “number of affixes” in a nonword. In addition, at a theoretical level, the number of affixes effect and the nonword complexity effect may not tap into the same underlying dimensions of variation (similar to base frequency and family size) even though they are based in similar theoretical arguments.

First, complexity in the nonword complexity effect is achieved by attaching a real-world stem with a syntactically corrected suffix (e.g., GASFUL). This is compared both within-subjects and between-subjects with the simple nonword case constructed by attaching the same real-world stem with a letter string with a similar structure to the related suffix (e.g., GASFIL). The resulting nonword is no longer decomposable into stem+affix and therefore does not have morphological structure. Yap et al., reasoned that since the nonword complexity effect points to the automatic decomposition of morphologically complex words and nonwords into morphemes, more morphemes should create longer response latencies and more errors for nonwords. They concluded that the number of affixes in the nonword should have a similar inhibitory affect, as more affixes would generate the need for more decomposition processes. However, the nonwords (even morphologically complex) in their study did not include real world stems. The nonwords also had a mean of 1.1 affixes, meaning that some nonwords included more than 1 affix and there was a range of affixes from no affixes to more than one. Our stimuli had a maximum of one affix. We also examined extremes with no morphological structure (non-decomposable) and both a stem and an affix. Including both the stem and affix, may generate more use of semantic information than simply using affixes with no stem.

Despite differences in the definition of morphological complexity in nonwords, our results are in line with Yap et al., (2015) in which individuals with larger vocabularies had larger morpheme interference effects as indexed by the number of affixes. Overall, there is a strong negative effect of nonword complexity. Interestingly, the nonword complexity effect was related to both reading-related measures (Lexical Quality) and form based (phonological, orthographic) processing measures. In addition, while the nonword complexity effect has been put forth as evidence for automatic decomposition into morphological constituents based on the form of the letter string and before semantic calculations (e.g., Taft, 1975; Crepaldi et al., 2010; Caramazza et al., 1988), it is interesting that the nonword complexity effect in reaction time patterned more closely with the family size effect (purported as related semantic in nature) and not the base frequency effect (purported as related to form based processing).

Now focusing on the effects of sensitivity to morphological structure in words, our results with the base frequency effect in reaction time are in line with Kuperman & Van Dyke (2011). Their skill-based measures examined both comprehension and segmentation (form based processing). Our closest measures to examine overall semantic and form based processing were vocabulary and blending words (phonological processing). Both the overall pattern (as skill increased, the base frequency effect decreased) and fine grain pattern (large facilitative effect for individuals with low skill and no or inhibitory effect for individuals with high skill) lined up with Kuperman & Van Dyke (2011) for both vocabulary (comprehension) and blending words (segmentation).

However, it is interesting to note, Kuperman & Van Dyke (2011) may have used base frequency measure, which encapsulated less morphological information particularly in the case of derivational morphology than the cumulative root frequency used in the current study (for description, see de Jong et al., 2000). Kuperman & Van Dyke, (2011) defined base frequency as the lemma frequency of the base word (summed frequency of the inflectional variants of TEACH – e.g., teach, teaches, teaching - in TEACHER). Lemma frequency has been used mostly for studies of the effect of base frequency in words with inflectional endings, or inflectional morphology (e.g., Baayen, Dijkstra, & Schreuder, 1997, Colé, Beauvillain, & Segui, 1989). However, given Ford et al., (2010) and even Kuperman & Van Dyke, (2011) examined words with derivational endings, the base frequency count which includes the summed lemma frequencies of all related family members is warranted. Additionally, most current examinations of the base frequency effect in derived words describe base frequency as the cumulative lemma frequency (e.g., Taft, 1979; Taft, 2004; Ford et al., 2010; Vannest et al., 2010; Xu & Taft, 2015). The base frequency effect in this context is also a more well defined construct in terms of the relationship with word transparency, decomposability, and suffix productivity, and, more recently, its relationship to the family size effect. However, even given the theoretical differences between the base frequency counts, the two counts of base frequency were highly correlated in the ($r(108)=.955$, $p<.0001$) stimulus set from Ford et al., (2010).

Further, as noted previously, Kuperman & Van Dyke (2011) included both base frequency and family size in their LME analysis, but family size showed less predictive power and was therefore not included in analysis. This is interesting because family size has been theorized to either 1) be the relevant predictor not base frequency (e.g., De Jong et al., 2000) or 2) be a separate predictor of reaction time related to different aspects of morphology (semantics vs. form) (Ford et al., 2010; Xu & Taft, 2015) particularly in transparent words with productive suffixes. In addition, as family size and base frequency effects were separate predictors in our analysis and purportedly related to differential aspects of morphological structure (semantics, form), it would have been interesting to see how family size related to the 17 individual skill measures in Kuperman & Van Dyke (2011). Nonetheless, our family size data were similar to Kuperman & Van Dyke (2011) and our base frequency data in terms of the relationship to vocabulary (e.g., as skill increased, family size decreased). However, as the characterization of family size as an index of sensitivity to semantic structure in morphologically complex words would suggest, the family size effect is not related to form based processing measures (pseudoword decoding, nonword repetition, blending words). The family size effect is instead related to both ART and spelling.

In relation to Andrews et al. (2011), we were also able to examine fine grain differences in sensitivity to and processing of morphologically structured words using their orthographic and semantic profile designations. In Andrews et al., (2011), individuals with a semantic profile seemed to be in line with the graded accounts of morphological priming, while individuals with an orthographic profile had equal priming for form and opaque pairs. This suggested that individuals with the orthographic profile were more affected by obligatory decomposition processes due to the form rather than semantic based processes. In line with Ford et al., (2010), which suggests that base frequency is related to form based morphological processes, individuals with the orthographic profile (more affected by obligatory decomposition) were more sensitive to base frequency.

Turning to Duñabeitia et al., (2014), which also examined early morphological decomposition, albeit via masked transposed letter priming, and used individual differences to adjudicate inconsistencies in the literature (in both morphological and transposed letter priming, some found the effect, some did not). Also as with Andrews et al. (2011), researchers were able to examine fine grain morphological processing (morphological decomposition) through individual differences analysis. Their findings suggest that individuals with faster reading times displayed greater priming for within- than between- morpheme transpositions while individuals with slower readings times showed no difference between the two types of transpositions. In other words, faster readers may be more likely to consistently use decomposition strategies early in processing, while slower readers may not. The closest measure in our study to reading speed (a measure of fluency), was the TOWRE – sight word efficiency and pseudoword decoding efficiency. Interestingly, skill in pseudoword decoding increased, the nonword complexity effect in error rate also increased. As the nonword complexity effect has been put forth as a measure of obligatory form based morphological decomposition, these results seem to be in line with Duñabeitia et al., (2014). For both our study and Duñabeitia et al., (2014), as orthographic to phonological connections (reading speed, decoding efficiency) strengthen, individuals use morphological decomposition strategies to speed word recognition (transposed letter priming), but also cause more errors in nonwords (nonword complexity in error rate).

**Implications for Theories of Morphological Processing**

The nonword complexity, family size, and base frequency effects have been well characterized at the group-level in terms of their relationship to morphological structure. An important contribution of this study is the simultaneous characterization of these effects in terms of individual skill measures and their relationships to each other. The examination of these effects in terms of individuals difference, their relationships to each other, and their patterns of correlation with the with the skill battery have interesting implications for theories of morphological processing. Further, given our previous discussion of how an individual's experience/strength of connections may affect the use of and the sensitivity to the lexical characteristics of a word and the special statistical structure of morphologically related items, individuals should vary systematically in sensitivity to morphological structure related to the form and/or meaning overlap of morphologically structured words. In particular, individual differences measures (reading battery) that index an individuals' form-based processing should vary with morphological measures related to form based aspects of morphology (base frequency) and aspects of the reading profile that index processing related to semantics should also vary with morphological measures that are semantic in nature (family size).

First, from its original discovery in Taft & Forster (1975), the nonword complexity effect has been put forth as a measure of obligatory decomposition of morphologically complex words into its morphological constituents based on orthographic features before semantic processing (see Rastle & Davis, 2008; Rastle et al., 2004 for full description). As morphological structure causes interference for nonwords, which inherently do not have whole-form lexical information, this effect is strong evidence for prelexical decomposition based on orthographic form. Additionally, the purported manipulation is the "decomposability" of the nonword (there is no affix in the simple/non-decomposable case), i.e., the decomposable nonwords with a stem and affix (complex) generate longer response latencies than nonwords that are not decomposable in terms of morphological structure. However, the nonword complexity effect may also be generated by the semantic information encoded in the real-world stems. For example, after this decomposition processes, individuals must then process a real-world stem; which, as our word data suggest, additionally activates morphologically related words.

The family size effect, on the other hand, has been put forth as semantic in nature, particularly in Hebrew (Moscoso del Prado Mart´ın et al. (2005), while the base frequency effect has been put forth as a measure of form related decomposition processes (Ford et al., 2010).. For example, Schreuder and Baayen (1997) found that the removing semantically opaque family members from the count of the Family Size improved correlations with reaction time in visual lexical decision. Further, derivational suffixes, only through the removal of opaque family members were able to obtain a reliable correlation of reaction time in visual lexical decision. Ford et al., (2010) examined both family size and base frequency in the context of words with productive or non-productive suffixes. The family size effect occurred regardless of suffix productivity in line with the characterization of family size as a semantically related morphological effect. Interestingly, the base frequency effect only occurred with words with productive suffixes, suggesting that the base frequency effect was related to the form based morphological decomposition processes. However, some studies suggest that family size and base frequency are in fact indexing the same sensitivity to morphological structure and that family size is the relevant predictor, (Schreuder and Baayen, 1997; Bertram et al., 2000a; De Jong et al., 2000). Our results suggest that this is not the case in transparent words with productive suffixes as family size and base frequency are separate predictors that correlate differentially with aspects of an individual's reading network. In particular, base frequency seems to be most closely related with form based processing ability (phonological processing skill, orthographic profile).

As both the nonword complexity and base frequency effects are purported to be measures of sensitivity to morphological structure based in form and indices of morphological decomposition and family size is purported to be a measure semantic in nature, one might predict that the nonword complexity effect would pattern most similarly to base frequency and not family size. However, the nonword complexity effect in reaction time instead patterns very closely with (mirrors) family size effect (correlations with each other and with skill measures). On the other hand, the base frequency effect does not pattern with the nonword complexity or family size effect in reaction time. Altogether, these data suggest that family size and base frequency are indeed separate predictors of sensitivity to morphological structure, with family size related to semantics and base frequency related to form-based processing.

Ford et al., 2010 and our current study support the conclusion that the family size effect is related to the semantic features of morphology, while the base frequency effect taps into the statistical structure related to form and morphological decomposition. As a result, the relationship between the family size and nonword complexity effects in reaction time could be due to the real-world stem within the decomposable (complex) nonwords. However, in error rate, the nonword complexity effect is correlated with PDE, a measure of O-P processing and the nonword complexity effect in error rate and patterns similarly to the base frequency effect. This suggests that during the lexical decision task, interference from the complex nonword could be caused by not only the prelexical obligatory decomposition of the pseudo-stem and affix but also the real-world stem in the complex nonword co-activating morphologically related family members which generates interference.

Lastly, perhaps the most interesting finding in the current data is the systematic binding of the family size effect (words) and the nonword complexity effect - both in direct correlations (as the family size effect decreases, the nonword complexity effect decreases) and in patterns of correlation (individuals with low skill have large nonword complexity effects and small family size effects). This seems to suggest that the semantically related morphological information that both the nonword complexity and family size effects index, is useful for words at low skill, but does not affect nonword processing. However, once semantic processing skill increases to a certain point, the family size effect goes away, but sensitivity to morphological structure can still be seen in nonword complexity effects.

This is consistent with certain aspects of the Lexical Quality Hypothesis (Perfetti, 2007), in which individuals with worse lexical quality, rely on componential processing, while individuals with higher lexical quality have more automatic full-form lexical representations. Individuals with more full-form lexical representations (higher skill) may have interference from other lexical forms, particularly in morphology where there is both overlap in form and meaning. This would manifest as smaller family size effects for individuals with higher skill. Similarly, the larger number of full-form lexical representations would also create interference for nonword processing and consequently create larger interference effects. However, individuals with lower lexical quality and therefore fewer full-form representations would not have as much interference for both words and nonwords. In addition, these individuals would also be more reliant on componential based processing without the additional interference from other whole-form representations. This lack of interference from other whole-form representations and greater reliance on constituent based processing results in smaller nonword complexity effects and larger family size effects (indicating greater sensitivity to componential aspects of words – morphology).

In order to further examine the relationship between family size and nonword complexity, a future direction could be to examine the family size and base frequency information for the stems used in the nonword complexity effect. In addition, a design including both the number of affixes effect and nonword complexity effect for nonwords and the base frequency and family size effects for words could help further explain the correlation of the nonword complexity effect with the family size effect and not the base frequency effect. In particular, as the Yap et al. (2015) nonwords did not include any stems and the Crepaldi et al. (2010) included complex nonwords with stems, one might see that the number of affixes effect was related to the base frequency effect and the nonword complexity effect was related to the family size effect.

**References**

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.

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

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065-1086.

Baayen, R. H. (2014). Experimental and psycholinguistic approaches to studying derivation. *Handbook of derivational morphology*, 95-117.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*(1), 94-117.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2): Linguistic Data Consortium. *University of Pennsylvania, Philadelphia, PA, USA*.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, *42*(3), 390-405.

Bertram, R., Laine, M., & Virkkala, M. M. (2000). The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology*, *41*(4), 287-296.

Beyersmann, E., McCormick, S. F., & Rastle, K. (2013). Letter transpositions within morphemes and across morpheme boundaries. *The Quarterly Journal of Experimental Psychology*, *66*(12), 2389-2410.

Bradley, D. C. (1979). Lexical representation of derivational relation. In M. Aronoff & M. L. Kean (Eds.), Juncture. Cambridge, MA: MIT Press.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and cognitive processes*, *2*(3-4), 217-227.

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language*, *18*(4), 342.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. Cognition, 28, 297-332. doi:10.1016/

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and language*, *28*(1), 1-13.

Crepaldi, D., Rastle, K., & Davis, C. J. (2010). Morphemes in their place: Evidence for position-specific identification of suffixes. *Memory & Cognition*,*38*(3), 312-321.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology*, *82*(4), 733.

De Jong IV, N. H., Schreuder, R., & Harald Baayen, R. (2000). The morphological family size effect and morphology. *Language and cognitive processes*, *15*(4-5), 329-365.

De Jong, N. H., Schreuder, R., and Baayen, R. H. (2003). Morphological resonance in the mental lexicon. In Baayen, R. H. and Schreuder, R., editors, Morphological structure in language processing, pages 65–88. Mouton de Gruyter, Berlin.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2014). Revisiting letter transpositions within and across morphemic boundaries. *Psychonomic bulletin & review*, *21*(6), 1557-1575.

Duñabeitia, J.A., Perea, M.,& Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho- orthographic decomposition. Cognition, 105, 691–703.

Ford, M. A., Davis, M. H., & Marslen-Wilson, W. D. (2010). Derivational morphology and base morpheme frequency. *Journal of Memory and Language*, *63*(1), 117-130.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time.*Journal of verbal learning and verbal behavior*, *12*(6), 627-635.

Fudge, E. (1984). English word-stress. London: George Allen & Unwin.

Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of memory and language*, *30*(3), 370-384.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological review*, *111*(3), 662.

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, *115*(3), 367-379.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42-73.

Kuperman, V., & Van Dyke, J. A. (2011). Individual differences in visual comprehension of morphological complexity. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 1643-1648).

Lavric, A., Elchlepp, H., & Rastle, K. (2012). Tracking hierarchical processing in morphological decomposition with brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 811.

Lázaro, M. (2012). A Study of Base Frequency in Spanish Skilled and Reading-Disabled Children: All Children Benefit from Morphological Processing in Defining Complex Pseudowords. *Dyslexia*, *18*(2), 130-138.

Mirman, Daniel. *Growth curve analysis and visualization using R*. CRC Press, 2014.

Moscoso del Prado Mart´ın, F., Deutsch, A., Frost, R., Schreuder, R., De Jong, N. H., and Baayen, R. H. (2005). Changing places: A cross-language perspective on frequency and family size in Hebrew and Dutch. Journal of Memory and Language, 53:496–512.

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004). Morphological family size in a morphologically rich language: the case of Finnish compared with Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1271.

Muncer, S. J., Knight, D. C., & Adams, J. W. (2013). Lexical decision and the number of morphemes and affixes. *Scandinavian journal of psychology*,*54*(5), 349-352.

Muncer, S. J., Knight, D., & Adams, J. W. (2014). Bigram frequency, number of syllables and morphemes and their effects on lexical decision and word naming. *Journal of psycholinguistic research*, *43*(3), 241-254.

Nelson, M. J., Brown, J. I., & Denny, M. J. (1960). *The Nelson-Denny Reading Test: Vocabulary, Comprehension, Rate*. Houghton Mifflin.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension.*Scientific studies of reading*, *11*(4), 357-383.

Pinnell, G. S., & Fountas, I. C. (1998). *Word Matters: Teaching Phonics and Spelling in the Reading/Writing Classroom*. Heinemann, 88 Post Rd., W., Box 5007, Westport, CT 06881-5007.

Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Rueckl, J. G., Constable, R. T., ... & Mencl, W. E. (2008). Effects of stimulus difficulty and repetition on printed word identification: An fMRI comparison of nonimpaired and reading-disabled adolescent cohorts. *Journal of Cognitive Neuroscience*, *20*(7), 1146-1160.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*(7-8), 942-971.

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

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

Rueckl, J. G., & Aicher, K. A. (2008). Are CORNER and BROTHER morphologically complex? Not in the long term. *Language and Cognitive Processes*, *23*(7-8), 972-1001.

Rueckl, J. G., & Rimzhim, A. (2011). On the interaction of letter transpositions and morphemic boundaries. *Language and cognitive processes*, *26*(4-6), 482-508.

Sánchez-Gutiérrez, C., & Rastle, K. (2013). Letter transpositions within and across morphemic boundaries: Is there a cross-language difference?.*Psychonomic bulletin & review*, *20*(5), 988-996.

Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., ... & Lyon, G. R. (2003). Neural systems for compensation and persistence: young adult outcome of childhood reading disability. *Biological psychiatry*, *54*(1), 25-33.

Taft, M. (1979). Recognition of affixed words and the word frequency effect.*Memory & Cognition*, *7*(4), 263-272.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology Section A*,*57*(4), 745-765.

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

Torgensen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). Test of word reading efficiency (TOWRE). *Austin, TX: Pro-Ed*.

Vannest, J., Newport, E. L., Newman, A. J., & Bavelier, D. (2011). Interplay between morphology and frequency in lexical access: The case of the base frequency effect. *Brain research*, *1373*, 144-159.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing: CTOPP*. Pro-ed.

Whaley, C. P. (1978). Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 143-154.

Xu, J., & Taft, M. (2015). The effects of semantic transparency and base frequency on the recognition of English complex words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 904.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 53.

Yap, M. J., Sibley, D. E., Balota, D. A., Ratcliff, R., & Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 597.

## Appendix A
## Equations for Person and Item Model – Words

### Error Rate

Level 1 ($Response_{ji}$)

$$\text{logit}(p_{ji}) = \gamma_{ji}$$

Level 2 ($Person_j$ & $Item_i$)

$$\gamma_{ji} = (r_{0101j} + \beta_{001})SurfaceFrequency_i + (r_{0102j} + \beta_{002})BaseFrequency_i + (r_{0101j} + \beta_{003})FamilySize_i + \beta_{004}SurfaceFrequency_i * BaseFrequency_i + \beta_{005}SurfaceFrequency_i * FamilySize_i + \beta_{006}BaseFrequency_i * FamilySize_i + \beta_{006}SurfaceFrequency_i * BaseFrequency_i * FamilySize_i + \pi_{001}PreviousRT_{ji} + \pi_{002}PreviousTrialType_{ji} + r_{010j} + r_{001i}$$

### Reaction Time

($Person_j$ & $Item_i$)

$$\gamma_{ji} = (r_{0101j} + \beta_{001})SurfaceFrequency_i + (r_{0102j} + \beta_{002})BaseFrequency_i + (r_{0101j} + \beta_{003})FamilySize_i + \beta_{004}SurfaceFrequency_i * BaseFrequency_i + \beta_{005}SurfaceFrequency_i * FamilySize_i + \beta_{006}BaseFrequency_i * FamilySize_i + \beta_{006}SurfaceFrequency_i * BaseFrequency_i * FamilySize_i + \pi_{001}PreviousRT_{ji} + \pi_{002}PreviousTrialType_{ji} + r_{010j} + r_{001i}$$

## Appendix B
## Equations for Person and Item Model - Nonwords

### Error Rate

Level 1 ($Response_{ji}$)

$$\text{logit}(p_{ji}) = \gamma_{ji}$$

Level 2 ($Person_j$ & $Item_i$)

$$\gamma_{ji} = (r_{0101j} + \beta_{001})NonwordComplexity + \pi_{001}PreviousRT_{ji} + \pi_{002}TrialOrder_{ji} + r_{010j} + r_{001i}$$

### Reaction Time

($Person_j$ & $Item_i$)

$$\gamma_{ji} = (r_{0101j} + \beta_{001})NonwordComplexity + \pi_{001}PreviousRT_{ji} + \pi_{002}TrialOrder_{ji} + r_{010j} + r_{001i}$$