uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Glenn Thompson
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Psychology)
GRADE / DEGREE

Department of Psychology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Towards a Process-Based Characterization of Syllable Effects in Visual Word Recognition

TITRE DE LA THÈSE / TITLE OF THESIS

Alain Desrochers
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Sylvain Chartier

Charles Collin

Steve Lupker (University of Western Ontario)

Pierre Mercier

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Towards a Process-Based Characterization of Syllable Effects in Visual Word Recognition

Glenn L. Thompson

University of Ottawa

Thesis Supervisor: Alain Desrochers, PhD

List of Examiners:

Sylvain Chartier, PhD

Charles Collins, PhD

Pierre Mercier, PhD

Stephen Lupker, PhD (External)

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the PhD degree in Experimental Psychology

School of Psychology

Faculty of Social Sciences

University of Ottawa

Your file Votre référence
ISBN: 978-0-494-69112-0
Our file Notre référence
ISBN: 978-0-494-69112-0

Foreword

It would like to begin this foreword with a technical note. The document was prepared in 'article' format, which means that there is more redundancy in the content than would be expected in a traditional monograph. This characteristic, in part, explains the length of this (brick of a) dissertation. I deserve some of the blame and credit I suppose.

This work obviously represents several years of effort on my part. I did not do it alone. It would not have been possible without the support that I received from various sources. I would like to express my thanks here for the assistance I received.

During my doctoral studies, I was fortunate enough to be awarded masters- and doctoral-level scholarships by the Natural Sciences and Engineering Research Council (NSERC). I am grateful for this support as it allowed me to pursue my studies while helping to support my family.

On a more personal level, I would like to thank the members of the Cognitive Psychology of Language Laboratory. At the risk of forgetting someone, I want to specifically thank the following people for their help in recruiting participants and/or running participants: Nathalie Ricard, Lisa Mask, Ashley Lemieux, Marianne Chevrier, Michelle Zenko, Sabrina Fréchette, Julie Noël, and Véronic Quann. They all endured my remorseless teasing with good humour and aplomb. For that, they are to be commended. I am also grateful to the members of my examining board for having carefully read this manuscript and provided me with constructive comments (the members of my committee are listed on the title page of this document).

A number of great professors helped shape my thinking over the years. Tay Wilson and Micheal Persinger at Laurentian University taught me to expect more from myself and how to use statistics. Cynthia Whissell, also of Laurentian, got me interested in psychology (taught me intro) and in psycholinguistics specifically. These great professors ensured that I was more than ready for graduate school. I would like to thank Claude Lamontagne at the University of Ottawa for widening my perspectives on philosophy of science and perception.

At this point, I must of course devote at least a paragraph to my thesis supervisor and mentor Alain Desrochers. He actively collaborated with me on the conceptual and empirical work presented in this dissertation (as indicated in the author's notes) and he was quite a demanding/generous editor when it came to the writing. More generally, he smoothed out my rough edges as much as he could over the years since I began working in his lab. It could not have been easy. Thankfully, Alain is kind and patient and always generous with his time. I recognized these qualities in him almost immediately, but I did not appreciate their true value until I learned first hand of his integrity, intelligence, erudition, and wisdom. I am deeply indebted to him for nearly everything that I am able to do well as a researcher. You would think that this would make up for how resentful I am that his command of the English language is superior to mine (despite the fact he's a late bilingual and English is his second language). You would be wrong. Nevertheless, I'm proud to count myself among the disciples of such a first-rate researcher. More importantly, I'm proud to consider him a friend.

Finally, I would like to say a word of thanks to my family. First, I want to thank my loving wife Rachel, without whose love and support I might have graduated two years earlier (wuv you, bug). My life would have been emptier without you though. Je tiens aussi à remercier mon fils Parker pour sa grande patience. Papa travail beaucoup quand il voudrait biens jouer avec toi. Je t'aime de tout mon cœur, duders. To my unborn daughter: thank you for waiting until daddy was finished school before arriving. I love you already and I can't wait to get to know you.

To mom and dad: My entire life you've loved me unconditionally, while caring enough to make sure I grew up right and stayed in school. You gave me a home and you gave me the resources to pursue my academic goals. You were with me from the start. As far as I'm concerned, this doctorate is as much yours as it is mine. I dedicate it to you.

Abstract

Existing evidence on the role of the syllable in visual word recognition is insufficient to inform the extention of models of single-syllable word reading to the more general case of multi-syllable words. Open questions include what syllable units are relevant (e.g., Max Onset, Max Coda), when and why are they relevant, and for whom? A common paradigm for addressing this issue is the word-splitting manipulation in visual lexical decision, where a visual boundary is introduced within words; this boundary is either consistent or inconsistent with that specified by a theoretical syllable unit. If an advantage is observed for the consistent condition, then a 'syllable effect' has been obtained. Recent work has focused on identifying variables that moderate the effect of such word-splitting manipulations in an attempt to explain the empirical inconsistency in this literature (Chen & Vaid, 2007; Taft, 2001, 2002). This line of enquiry was pursued in two studies employing the word-splitting paradigm in lexical decision. The first study reports a series of three experiments (N=48 x 3) where a combination of random coefficient analysis and multi-level modeling was used to examine the joint contribution of syllabic complexity, lexical frequency, and the tendency of participants to rely on phonology. The results indicate that participants who are most sensitive to homophone interference tend to prefer the 'phonological' Max Onset syllable with low-frequency words, situating the locus of this effect in lexical phonology. A follow-up study (N=122) using the same analytical technique with a broad set of indicators provides some corroboration of this effect, and additional information suggesting that apparent patterns of syllable preference arise from multiple causes, some of which depend on Print Exposure and/or participant response speed. The results highlight the need to disentangle response speed from other participant characteristics in predicting syllable preference.

Table of Contents

Chapter 1

Towards a Process-Based Characterization of Syllable Effects in Visual Word Recognition

Abstract

Most computational models of reading and word recognition suffer from an important limitation: they can only process a small pool of monosyllabic words (Exceptions include ad hoc models, such as Ans, Carbonnel, & Valdois, 1998). The task of extending these models to handle multi-syllable words is hindered by uncertainty regarding the way polysyllabicity should be managed within the next generation of models. This state of affairs is due in large part to inconsistent empirical support for the syllable as a functional unit. The present review summarizes the syllable-effect literature with the aim of identifying critical open questions that may serve to elucidate the empirical and theoretical status of syllable effects in the visual word recognition literature. Inter alia, it is concluded that individual differences provide an informative way of addressing the empirical inconsistency associated with syllable effects, but that such research is most useful when explicitly linked to process-based variables derived from computational models.

Introduction

Alphabetic languages explicitly represent the fundamental contrastive units of speech (i.e. phonemes) in print through the use of letters and letter combinations called graphemes (Coltheart, 1978; Henderson, 1985). This alphabetic principle (Rozin & Gleitman, 1977) has been taken as evidence that reading processes are based primarily on grapheme-to-phoneme translation and that they are parasitic on pre-existing knowledge of the oral language for the extraction of meaning (for elaboration of the 'Strong Phonology' hypothesis, see Van Orden, Pennington, & Stone, 1990; Frost, 1998). Another possibility, which happens to be more consistent with the existing body of evidence, is that the cognitive system is actually quite flexible in how it exploits printed text. For instance, despite the fact that alphabetic systems were designed to represent the segmental units of speech, skilled-readers appear to develop direct associations between print and semantics as their competency develops in an effort to speed up the recognition process (Harm & Seidenberg, 2004; for reviews and additional evidence, see, Coltheart, 1978; Shelton & Caramazza, 1999; Taft & van Graan, 1998). Similarly, the cognitive system may detect correspondences between print (orthography) and the other major bodies of knowledge (i.e., phonology, semantics) using letter-cluster types other than the grapheme (e.g., letter-clusters of varying size; rimes), to overcome limitations in the consistency of statistical mappings for example (see 'grain-size' theory; Ziegler & Goswami, 2005 and the discussion of the context-size issue in Plaut, McClelland, Seidenberg, & Patterson, 1996). The bits of information that are supposed to be associated with one another across bodies of knowledge such that otherwise distinct symbols may be perceived as a unit (e.g., the letters *th* in the word BATH) are called functional units (or 'relational units', Venezky, 1970) and they are the cornerstone of any adequate theory of the reading process.

The nature of the functional units employed by the reading system is not self-evident. Consequently, theoretical statements regarding the role of a particular functional unit must be justified using a combination of logical, computational, and especially empirical evidence (Grainger & Jacobs, 1998; Coltheart, 1978; Coltheart, Rastle, Langdon, Perry, & Ziegler, 2001). When a given unit receives strong empirical support, it provides a useful constraint on the development of theoretical models because, whether theorists choose to (a) explicitly represent a given unit in their model via symbolic coding or (b) endow their models with the capacity for discovering a particular class of units within an adaptive and graded system, an adequate model of the reading process would be expected to simulate a stable pattern of empirical results (for discussion, see Coltheart et al., 2001; Seidenberg & Plaut, 2006; Perry, Ziegler, & Zorzi, 2007). However, when empirical evidence for or against a given class of functional unit is insufficient, as is the case when the associated literature is inconsistent, the assumptions of theoretical models go unsupported and unchallenged, which is a state of affairs that hinders theoretical progress. The present paper concerns itself with the syllable as a functional unit in visual word recognition.

Currently, major computational models of skilled-reading are designed to account for the reading and recognition of a small pool of 2,800 words that are both monosyllabic and monomorphemic (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004).[1] This class of words is the most frequently used in the language according to token frequency counts, but comprises a tiny fraction of all words in the English language according to type frequency and a mere ten percent of an adult's vocabulary (for vocabulary-size estimates, see Nation & Waring, 1997). The specificity of existing models raises the possibility that their assumptions do not apply to the processing of all words (Chateau & Jared, 2003) and it excludes a number of phenomena specific to multisyllable (e.g., stress placement; Rastle & Coltheart, 2000) and multimorphemic words (morphological decomposition; Chialant & Caramazza, 1995; Taft, 1991, 1994) that may prove useful in the task of adjudicating among competing theories. Not surprisingly, major theoretical works by Coltheart and colleagues (2001), Plaut and colleagues (1996), Harm & Seidenberg (2004) and Perry and colleagues (2007) have all pointed to multisyllable words as a logical next step in theoretical development.

This theoretical demand lacks a sufficient empirical base, however. The information required to make sound theoretical generalizations to the case of multi-syllabic words is incomplete, because (a) such words are chronically under-studied (as first noted by Jared & Seidenberg, 1990; and reiterated more than a decade later by Chateau & Jared, 2003; Duncan & Seymour, 2003) and (b) the experimental studies that do exist have not converged on a stable pattern of results (see Chateau & Jared, 2003; Chen & Vaid, 2007; Taft, 1987, 2001). Taken together, these factors have contributed to the dearth of solid evidence for syllable effects and slow progress in the theoretical account of complex word reading. Relevant open questions include the nature of syllable representation and the locus of syllable effects within a general theoretical framework for visual word recognition. As Harm & Seidenberg (2004, p. 62) note:

> *How are multisyllabic words read? Complex words could be processed as wholes (...) or in parts (...) [that] could be syllables or morphemes or clumps of adjacent letters that sometimes cross structural boundaries. These issues have not been resolved by behavioral research. If there were better information about how complex words are processed, it could be used to guide the development of a model.*

The development of such a model would, of course, support increased interest in experimenting with complex stimuli, which would in turn provide a richer database to guide model development.

Recent research on the syllable effects have yielded inconsistent results and led to an impasse in theoretical development. Correlational studies examining the influence of syllable length have consistently indicated that words with additional syllables take longer to name (Jared & Seidenberg, 1990) and recognize (New, Ferrand, Pallier, & Brysbaert, 2006). However, such findings are difficult to interpret given that length in syllables is known to be confounded with other lexical attributes, including the number of phonemes in a word. Thus, any effect of syllable length may actually be an effect related to the number of phonemes (Jared & Seidenberg, 1990). Experimental research has not produced evidence that is any more convincing. Typically, such studies introduce an artificial boundary within words that is either consistent or inconsistent with some definition of the syllable (see below for a more detailed discussion). If the syllable-consistent condition in word-splitting experiments results in faster reaction times (or greater accuracy), then a 'syllable effect' has been observed. Evidence that has thus far been collected in this way has been mixed (Chen & Vaid, 2007; Taft, 1987, 2001).

In a step towards resolving this issue, Taft (2001) investigated the possibility that individual differences in reader skill play a role in determining the experimental manifestation of syllable-effect phenomena. In a series of three experiments (see also Experiment 1, Taft, 2002), he found that syllable effects vary systematically with reader skill such that high- and low-skill readers produce qualitatively different patterns of results. A more detailed discussion of these findings and the earlier empirical work is presented below. At present, we simply want to highlight the fact that Taft's findings have important practical implications for the interpretation of conflicting reports in the literature, but that their theoretical implications are not as clear.

In practical terms, the results suggest that syllable effects should not be dismissed as unreliable (e.g., the result of Type-I errors and the file-drawer effect). Instead, it appears that a failure to consider reader skill as a relevant factor produced varying patterns of results depending on whether chance yielded a sample that was of high, low, or medium skill. In the unselected case, the most common result would be a heterogenous sample that is of moderate skill on average and which produces a null syllable effect (for review and discussion, see Chen & Vaid, 2007; Taft, 2001, 2002). In other words, this finding is important in that it (a) provides a general way of resolving the apparent inconsistencies in the literature and (b) improves our ability to manipulate experimental conditions so as to replicate syllable effects. In contrast, the theoretical implications of Taft's results are not as readily determined. The reason for this is that reader skill, defined as performance on a speeded comprehension test (as seen in several papers, Chateau & Jared, 2000; Sears, Siakaluk, & Buchanan, 2008; Taft, 2001), is a concept so general, depending on any number of component sub-skills, that it has no clear referent within extant computational models of the word recognition process (i.e., a specific aspect of architecture, the default value of a parameter or a combination of these properties). Thus, Taft's results do not speak directly to the critical issue of the locus of syllable effects within models of reading and, by extension, how computational models should address the syllable as a functional unit when they are generalized to account for the processing of words with more than one syllable. As will be argued, the more interesting questions from a theoretical standpoint involve identifying the properties of the word recognition system that covary with skill, and determining whether estimates of such theory-based variables are predictive of syllable preference.

Purpose

The present document reviews the syllable effect literature with the aim of summarizing what is currently known and relating this information to existing theoretical models while paying special attention to the dimensions (differences in architecture, parameters) that are supposed to model individual differences among readers. The discussion addresses itself explicitly to computational theories of the reading and recognition of individually presented monosyllabic words. This decision is based on the twin premises that a) a stronger link to a computational theoretical context can guide empirical work amidst inconsistent results and b) theory development should be a cumulative enterprise such that any model equipped to handle multi-syllable words must also account for experimental effects related to single-syllable words (from this perspective, a multi-syllable model like that of Ans et al., 1998, does not offer a solid base). Ways of measuring individual differences along the dimensions suggested by extant models of monosyllabic word reading are discussed.

The first section of the paper is devoted to a review of the empirical literature on syllable effects in visual word recognition. This review draws primarily from research using the two 'gold standard' experimental tasks for studying the word recognition and reading system: naming and lexical decision (Balota et al., 2004; for fMRI evidence that the two tasks tap one and the same system, see Carreiras, Mechelli, Estevez, & Price, 2007). This discussion sketches the inconsistency in the literature across decades of research, even among methodologically similar studies, and culminates with a description of the experiments reported by Taft (2001, 2002), which point to individual differences in reader skill as the culprit. Finally, the theoretical ambiguities associated with Taft's findings are discussed in greater detail. It is argued that defining individual differences in terms of the parameters of computational models is a more useful approach.

The second section is devoted to a discussion of the word recognition literature, specifically the assumptions underlying current computational models. Some of these assumptions are common to all or most models, while others are model-specific. The general and model-specific assumptions are each presented to prepare the ground for a discussion of how individual differences in syllable processing might be conceptualized in terms of these theories.

In the final section, a path for future research is sketched that proposes to build upon the work of Taft (2001, 2002) by predicting syllable preference based on more fine grained estimates of individual difference. The purpose of this path of enquiry is not to immediately contrast extant models, but rather to constrain future extensions to the domain of multi-syllable word reading. Accordingly, the overarching goal is to relate syllable effects to a limited number of theoretical 'individual difference' parameters that are common to both major models of reading, which will as a consequence (a) refine present understanding of the conditions under which syllable effects emerge and (b) support model-building decisions at the critical stage of their generalization to cover multi-syllable words.

The Syllable Effect in Visual Word Recognition

The notion of a syllable as a functional unit has its origin in the study of oral language processing. Because an important part of the reading system is the mapping from print-to-sound, researchers have explored the possibility that the letter-clusters representing the orthographic counterpart of the syllable in speech may be a relevant functional unit (e.g., seminal work by Spoehr & Smith, 1973, 1975) in the same way that the letters representing a phoneme appear relevant (i.e., graphemes effects; Rey, Ziegler, & Jacob, 2000). A parallel line of work has developed the idea that syllable-like units may exist in the orthographic domain, but as the result of factors governing the organization of orthographic knowledge (e.g., structure of the orthographic lexicon or, alternatively, meaning retrieval processes) as opposed to decoding processes (Forster &

Taft, 1976; Taft, 1979). Both possibilities are considered in what follows.

This section begins with an exposition of the evidence for and against the syllable as a functional unit of oral language in English on the one hand and of reading on the other. This discussion reveals that the findings in both domains are remarkably inconsistent, even when identical methodologies are used. Among other things, this suggests that some attribute of participants or items may be moderating the disputed effects. The section concludes with a detailed examination of the findings reported by Taft (2001, 2002) suggesting that individual differences associated with reader skill may account for inconsistent result, at least in the visual word recognition literature.

Syllable in Oral Language

The monosyllabic focus of reading research has produced a widely held view among researchers that phonological representation consists of phoneme strings. This view is consistent with the idea that pressure to identify words quickly causes the phonological system to focus on the minimally contrastive unit of the language: the phoneme (minimality constraint; Frost, 1998). As Ashby & Rayner (2004) & Ashby & Martin (2008) have argued, this perspective ignores an enormous amount of linguistic evidence for the hierarchical, multi-layered and multi-dimensional nature of phonological representation. For example, phonology/phonetics can be described as the combination of three layers of information: melodic, skeletal, and prosodic (Bybee, 2001; Clements & Keyser, 1983; Selkirk, 1982). According to this conception, phoneme strings are specified at the melodic level, vowel and consonant identities at the more abstract skeletal level, and the syllable structure/identity at the still more abstract prosodic level. All models of monosyllabic word reading in English represent the melodic level and some even encode vowel and consonant identities (e.g., Harm & Seidenberg, 2004; Plaut et al., 1996), but as yet none has had to manage the level of representation that is specific to multi-syllable words, which is to say prosody. The role of prosody (the syllable) in oral language is now considered.

Theoretical status of syllable units. The syllable is a fundamental unit for the description of phonology in the field of linguistics (Blevins, 1995; Hooper, 1972; Selkirk, 1982). By definition, a syllable is a unified speech stream, a series of phonemes that are pronounced together. The syllabic structure of a word thus organizes phonemes into pronunciation units, though the attribution of this function to a particular locus within models of speech production is a matter of contention, with some theories positing an abstract phonological-level representation (e.g., Ferrand, Segui, & Grainger, 1996) and others a representation within a phonetic syllabary at a more concrete output-level syllabary (Levelt, Roelofs, & Meyer, 1999). Regardless of how it is achieved, an important consequence of syllabic unification is that it creates a coarticulation context that influences the acoustic properties of individual phonemes (Lindblom & Studdert-Kennedy, 1967; Krakow, 1999; Uhry & Ehri, 1999) in a way that can make them more difficult to identify for listeners. This form of 'invariance' is taken as strong evidence for the role of the syllable in language processing (Cole & Scott, 1974; Massaro, 1974). For instance, the computational challenge that invariance presents to recognition processes can be solved by assuming the existence of a syllabic level of representation in models of auditory word recognition (i.e. a syllabary; Norris & Cutler, 1985; for a non-syllabic solution, see Norris, 1994). It has also been argued that syllables are a useful cue for initiating lexical search from a continuous speech-stream input (Cutler & Norris, 1988).

According to Treiman (1986), syllables are composed of the following structural elements: a mandatory nucleus (a vowel, as in the monosyllabic article A) that can be bookended by either an onset (an initial consonant or consonant cluster, as in the word

BRA), a coda (a final consonant or consonant cluster, as in the word AT), or both (as in the word BRAT). For monosyllabic words, this description of syllable structure is relatively unproblematic, but with multisyllable words an ambiguity arises concerning the way the consonant(s) at the boundary between two syllables (medial consonant or consonant cluster) should be coded. Specifically, a system for syllabifying multi-syllable words needs some way of deciding whether a medial consonant or consonant cluster (ND in THUNDER) should be interpreted as the coda of the first syllable (THUND-ER), the onset of the second (THU-NDER), or whether it should be divided between the two (THUN-DER) or belong to both simultaneously (ambisyllabicity: THUND-NDER). Thus, the widely held view that the syllable is a functional unit in phonology belies the fact that there is substantial disagreement over some very fundamental questions, such as the segmentation of words into syllabic units (for a brief review, see Content, Kearns, & Frauenfelder, 2001). On the one hand, determining the number of syllables in a word is as simple as counting the number of vowels it contains, as each of these forms the nucleus of a different syllable. It is the boundary between syllables that is not as easily determined.

Syllabification Heuristics. The typical approach to specifying syllable boundaries is to describe the processing of all words in terms of a simple heuristic or rule, which may then be constrained in various ways if required. The descriptive power of such rules is especially evident in languages where syllable structure is nearly perfectly regular, such as Spanish (Cutler, Mehler, Norris, & Segui, 1986) or Japanese (Taft, 2002). In such languages, the division of words according to a Maximal Onset principle (Carr, 2000; Spencer, 1996) is relatively uncontroversial (for discussion, see Taft, 2001). The Max Onset rule is also the most common description of English syllabification (Carr, 2000; Pulgram, 1970), despite complications like variable stress placement and complex syllable structure (Delatre, 1965).

The *Max-Onset principle* states that the words should be partitioned into syllables so as to maximize the number of onsets (consonants or consonant clusters at the onset of a syllable), within the constraints of a language's phonotactics (e.g., some consonants cannot legally be combined as onsets). In English, a clear example is the word NAVY whose Max-Onset boundary lies between the first vowel (a long vowel) and the second consonant, NA-VY. For a word like THUNDER, the consonants ND cannot legally form an onset, therefore the word is syllabified as THUN-DER. The complexities of the English language bring a number of phonotactic constraints to be superimposed on the Max-Onset principle, and a coherent description of these rules has yet to be provided. Some constraints are relatively clear-cut (Anderson & Jones, 1974), such as the case of words with a stressed first syllable containing a short vowel, which is a phoneme that cannot occupy a syllable-final position (e.g., syllabified as MEL-ON as opposed to Max-Onset rule ME-LON). Other constraints seem to operate more probabilistically. For example, stressed syllables appear to attract intervocalic consonants (consonants occuring between two vowels) regardless of the characteristics of the phoneme (Treiman & Danis, 1988), in violation of the Max-Onset principle (e.g. lem/on vs de/mand). Similarly, the phonetic properties of a consonant also seem to influence whether it tends to be bound to the first or second syllable (for discussion of this and other empirically relevant factors, see Treiman, Gross, and Cwikiel-Glavin, 1992; Treiman, Straub, and Lavery, 1994; Treiman & Zukowski, 1990). For such items, the principle of Max-Onset can be maintained by assuming that the medial consonant belongs to both the first syllable and the second syllable, a state that is called *ambisyllabicity* (Kahn, 1980; for counter-arguments, see Selkirk, 1982). Overall, it is clear that the issue of syllable identity in the oral domain, at least in English (and in

French, see Content, Kearns, & Frauenfelder, 2001), is not settled. For this reason, the Max-Onset principle should be considered at best a useful approximation of syllabification behavior in spoken language for English that is accurate for a subset of unambiguous cases (e.g., intervocalic/medial consonant clusters; words with a long first vowel; Taft, 1987, 2001).

The Syllable as a Functional Unit in Oral Language. Research on spoken language processing in English suggests that the syllable is a unit of production but not of reception. In recognition experiments based on the sequence monitoring paradigm, where participants must detect whether a fragment (either CV or CVC) is present in a target word (syllabified either CV-CVC or CVC-VC), the key findings suggest that English speakers are insensitive to syllable boundaries (Bradley, Sanchez-Casas & Garcia-Albea, 2003; Cutler, Mehler, Norris, & Segui, 1986) whereas Spanish speakers, for example, produce the classic interaction between fragment and word type that is indicative of syllable sensitivity (Bradley et al. 2003). In contrast, studies of non-reading language production are more consistent with the idea of the syllable as a functional unit. In English, picture naming latencies appear to increase with the length of words in syllables (Bates et al., 2003; Santiago, MacKay, Palma, & Rho, 2000). What is more, length in syllables is known to affect eye movement latencies during object recognition and visual search tasks (Zelinsky & Murphy, 2000). The latter finding is attributable to the additional time and/or cognitive resources required to evoke multi-syllable object names and to keep them active in short-term memory. On the whole, these findings suggest a role for syllable units in the production of oral language.

Syllable in Reading and Visual Word Recognition

In highly 'syllabic' languages like Spanish, the syllable has been shown to be relevant both the visual word recognition and reading aloud (Carreiras, Alvarez, & de Vega, 1993; Ferrand, Segui, & Grainger, 1996). In contrast, the body of evidence relating the syllable to visual word recognition and reading is relatively thin for skilled readers of English (though it is widely cited as a relevant unit in developing readers, Ehri, 1992; Frith, 1985; Seymour, 1999). The review to follow begins with the most obvious point of contact with spoken language: oral reading.

Syllable in Naming. In the word naming literature, syllable-based processing has been examined by manipulating lexical attributes that are associated with syllable structure, including syllable length and syllable frequency. For example, Jared & Seidenberg (1990) found that the length of words in syllables is associated with longer naming latencies for nonwords and low-frequency words even when length in phonemes is equivalent. More recently, words beginning with high-frequency syllables (token counts) were associated with faster naming latencies than those beginning with low-frequency syllables in a regression analysis of 3029 items from the English Lexicon project database that controlled for various confounding factors such as word frequency and orthographic N (based on token frequency; Macizo & Van Petten, 2007). Taken together, both findings are consistent with the idea that the syllable is involved in oral reading and yet they cannot be considered hard evidence of the syllable's status as a functional unit. The reason for this is that both length in syllables and syllable frequency are confounded with the distributional properties of items. Specifically, syllable length is confounded with the number of vowels and vowels are the main source of print-to-sound inconsistency in English (Jared & Seidenberg, 1990). Further, syllable frequency (especially token frequency) is likely confounded with estimates of a vowel's print-to-sound consistency, such as the ratio of a word's 'friends' to its 'enemies' (Jared, McCrae, & Seidenberg, 1990; not statistically controlled by Macizo & Van Petten, 2007). In contrast, direct experimental manipulations, using fragment priming for example, control for such confounds through the use of counter-balancing schemes.

In one such study, Ferrand, Segui, & Humphreys (1997) primed target words with orthographically matched fragments (masked) that either respected or violated their syllable boundary. These authors found a reliable priming effect for the syllable-consistent condition, but this effect has not been replicated in follow-up studies (Schiller, 1999; 2000). In more recent work, Ashby & Martin (2008, Experiment 2) primed target words with orthographically matched fragments that either respected or violated their syllable boundary. They found that parafoveally presented primes resulted in faster naming latencies when they did not violate the syllable boundary. This effect was not attributable to cohort activation because the advantage was maintained when compared to words that shared an initial tri-gram, but possessed a different syllable structure (e.g., gender, genocide). Thus, this latter finding constitutes strong evidence that the phonological syllable defined by Max-Onset division is relevant to oral reading. However, two aspects of the study suggest its limited applicability to the processing of individually presented words: the use of a priming manipulation and the fact that primes were presented parafoveally. Both characteristics suggest the syllable effect being reported is related to short-term memory processing during the naturalistic reading of text, which is a domain that is admittedly important, but that normally falls outside the scope of computational models of single-word lexical processing (e.g. Seidenberg & McClelland, 1989).

Syllable in Visual Word Recognition. Visual word recognition per se is typically defined as performance on the lexical decision task, but can also be examined using the duration of fixations in an eye-tracking study (e.g., using fixation durations as a dependent variable) or other tasks (e.g., letter identification). In lexical decision, Macizo

& Van Petten (2007) found that words with frequent word-initial syllable were recognized faster than words beginning with low-frequency syllables. This finding contrasts with the inhibitory effects of syllable-frequency that is observed in languages with more transparent syllables (for a review, see Macizo & Van Petten, 2007), which the authors interpreted in terms of a longer time-course for the activation of phonological information in English. Another possible interpretation of this discrepancy across languages is that the effect in English has a completely different basis (e.g., a confound with print-to-sound consistency or sound-to-print consistency, which can both influence visual lexical decision performance; Ziegler, Montant, & Jacobs, 1997). This same limitation infirms the observation that length in syllables is positively associated with reaction time in lexical decision, even after controlling statistically for length in letters and orthographic neighborhood (New et al., 2006).

Inconsistent with the findings of Macizo & Van Petten (2007) and New et al. (2006), Ferrand and colleagues (1997, Experiment 2) reported a failure to obtain syllable effects in lexical decision, in this case using a masked fragment priming paradigm. This null effect was later replicated by Ashby & Rayner (2004, Experiment 1: foveal presentation) in an eye-tracking study, finding that fixations were shorter with larger fragments regardless of whether the syllable boundary was violated. However, when primes were presented parafoveally, which is something that can only be reliably achieved with eye-tracking equipment, a reliable syllable-effect was obtained (Ashby & Rayner, 2004; Experiment 2). The latter finding was replicated by Ashby & Martin (2008; Experiment 1) using parafoveal presentation of primes in a lexical decision task (i.e., with response time and accuracy as dependent measures). Again, the specificity of the effect to priming and parafoveal presentation suggests a basis within the short-term memory systems that support the fluent reading of text.

More directly relevant to visual word recognition per se is the finding by Printzmetal, Treiman, and Rho (1986) that identification of a target letter's colour is more accurate when it is consistent with the other letters in the same syllable. Similarly, the syllable boundary appeared to be protective against color confusion with adjacent letters. In response, Seidenberg (1987) presented evidence that the true source of the effect is the statistical boundary that is created when two letters are infrequently encountered next to each other (bigram frequency trough), but the bigram-frequency trough hypothesis has since been disconfirmed (Rapp, 1992; see also Carreiras, Alvarez, & de Vega, 1993). Assuming that it is valid, the original finding strongly suggests that suprasegmental units (units larger than a letter, smaller than a word) are functionally relevant in word recognition.

'BOSS' or Max Coda Segmentation. Using a simple unmasked priming manipulation, Taft (1979) also failed to find a Max-Onset syllable priming effect in lexical decision with uncontrolled, presumably foveal, presentation. Importantly, Taft's study was not a simple test of the idea that the phonological syllable is relevant to visual word recognition. Rather, the explicit purpose of his study was to contrast two alternative definitions of the syllable that agree on the segmentation of some words and disagree for others. Specifically, Taft proposed that visually presented words are segmented so as to be consistent with their Basic Orthographic Syllable Structure (BOSS). The *BOSS* is defined as the word-initial onset and nucleus of a word plus any additional consonants that would form an attested word ending, or violate morphemic boundaries. It has since been recast in phonological terms as conforming to the Maximal Coda principle (Taft & Radeau, 1995), which states that words are segmented so as to maximize the number and size of codas in a word (e.g., NAV-Y instead of NA-VY; THUND-ER as opposed to THUN-DER). Like the Max Onset principle discussed

above, Max Coda division is constrained by things like legality (i.e., the rule is adjusted if it creates an illegal coda, such as PR for example).

The rationale behind the 'BOSS unit' or 'Max-Coda syllable' is primarily empirical, arising originally from the observation that polysyllabic monomorpheme words can be primed with fragments that bear a superficial resemblance to morpheme stems (Taft & Forster, 1976; see also the more recent study by Rastle, Davis, & New, 2004). This observation led to the proposal that the orthographic system tends to represent words in a decomposed manner according to some combination of their morphological and orthographic properties. According to this proposal, the Max Coda syllable (BOSS) is a morphographic unit that should be implicated in the retrieval of semantic information and the organization of the orthographic lexicon (see Taft, 1994). This idea has not been applied consistently though as the Max Coda syllable has been evoked as a unit that is relevant to pronunciation and stress-placement (Taft, 1992). This inconsistency may be a symptom of the lack of theoretical constraint associated with the absence of a computational model for multi-syllabic words.

Max Onset vs Max Coda in Lexical Decision. Using a simple unmasked fragment priming paradigm, Taft (1979, 1987) found that words were identified faster when preceded by a fragment that coincided with the Max Coda boundary relative to words preceded by fragments consistent with the Max Onset boundary or a control condition. As noted above, the priming methodology has the disadvantage of implicating memory processes that may be more telling of the mechanisms presiding the reading of continuous text than those governing word recognition per se. A technique for manipulating the relevance of sublexical units that is perhaps more useful in this respect is word-splitting. In a word-splitting paradigm, words are divided visually by introducing a mid-word change in case (e.g. THUNder ;Taft, 1979), foreign characters (e.g., THUN/DER ; Katz & Baldasare, 1983; Lima & Pollatsek, 1983; Taft, 1979; 1987), spaces (e.g. THUN DER;Taft, 2001, 2002), or by creating distinct units using colour (e.g., Rouibah & Taft, 2001). The logic behind this manipulation is that if the boundary created by the manipulation coincides with a cognitively relevant boundary, then a reliable effect (facilitory or inhibitory) relative to a comparison condition should be obtained. An advantage of such designs is that the assignment of words to word-splitting conditions is counter-balanced across participants, which means that words act as their own controls in such designs, thereby neutralizing the multifarious confounds associated with between-word comparisons (e.g., syllable frequency effects).

Using this methodology, Taft (1979, 1987) reported an advantage in lexical decision of the Max Coda defined syllable over that defined by Max Onset in English. Specifically, a processing advantage was observed when words were artificially divided at a boundary coinciding with that of the Max Coda syllable in comparison with other locations. Various alternative explanations for the Max Coda advantage were tested and it was found that the Max Coda advantage is not attributable to the size of the word-initial unit (a word-initial Max Coda is always equal to or larger/longer than Max Onset), either in terms of orthographic overlap (Taft, 1979; See also, Chen & Vaid, 2007) or its potential for uniquely identifying a word (i.e., narrowing the list of potential lexical candidates; Taft, 1987). Given these results, Taft concluded that the word-initial Max Coda syllable (BOSS) is a functional unit of word recognition.

This conclusion has not gone unchallenged as several conflicting experiments have been reported that employed virtually identical item-sets and procedures to those used by Taft (Taft, 1979 vs. Jordan, 1986 and Lima & Pollatsek, 1983 and Katz & Baldasare, 1983; Taft, 1987 vs. Seidenberg, 1987). The variable results that were obtained (Null results, Max Onset advantage) despite the apparent methodological consistency suggest that some sort of sampling issue is at work. Specifically, the syllable effect in lexical decision using the word-splitting technique is either (a) the result of a combination of Type-I errors and the file-drawer effect, (b) real, but dependent on some unmeasured participant- or item-attribute, or (c) a combination of those factors.

The Syllable Effect and Reader Skill. Taft (2001) observed that he had replicated the Max Coda advantage in both Australia and the United states in his earlier work (discounting a dialect-based explanation), but that in doing so he had enlisted the participation of a special population of readers that were likely more proficient than the average undergraduate-level reader (i.e., graduate students; colleagues). He reasoned that relative syllable preference might be skill-dependent, such that high-skill readers are biased towards 'efficient' Max Coda division (associated with direct recognition based on orthography) whereas low-skill readers may prefer the phonologically defined Max Onset division.

Indeed, Taft (2001, Experiments 1 & 2; 2002, Experiment 1) found that reader skill, operationally defined as performance on a timed comprehension task similar to the Nelson-Denny test (Brown, Bennett, & Hanna, 1981), was associated with a participant's relative preference for Max-Onset over Max-Coda division in the order of .25 to .30 (Pearson r). Specifically, as reader skill increased, the Max-Onset preference reversed to become a Max Coda advantage, suggesting that the Max Coda unit is functionally relevant for skilled readers, while the Max-Onset unit is operational within unskilled readers. An unselected group of participants would therefore normally produce a null effect and sampling strategies that bias participant selection towards one end of the skill spectrum could yield the corresponding syllable-effect pattern. In sum, the pattern of results reported by Taft (2001) provides a plausible explanation for the inconsistency that had plagued the literature.

Interestingly, Taft's finding is consistent with results reported by Katz & Baldasare (1983), which demonstrated that lower-skill/degraded letter presentation is associated with Max-Onset preference. Specifically, unskilled children and adults under conditions of degraded letter presentation (taken as an experimental simulation of less efficient orthographic processing) showed preference for Max Onset division. Under neutral conditions, the adults showed no bias towards syllabic or non-syllabic segmentation. A study by Butler, Jared, & Hains (1984) also found a relationship between reader skill and sensitivity to syllable structure, but is difficult to interpret here since it was not designed to contrast Max-Onset and Max-Coda syllabification. Since Taft (2001)'s paper was published, a study using SAT scores as an estimate of reading skill reported a Max Coda preference for all skill-levels, and therefore an absence of reversal to Max Onset preference. This finding is actually consistent with Taft's work because, as Chen & Vaid (2007) admit, even their 'poor readers' were above the SAT average. In sum, there is reason to believe that reader skill, more specifically a cognitively-based concomitant of reader skill, moderates syllable preference.

In developing a theoretical rationale for skill-dependent syllable preference, Taft proposed a number of possible explanations. Perhaps the most convincing of these is the hypothesis that low-skill participants are more reliant on phonology and therefore more sensitive to phonologically defined units (e.g., Max Onset syllable). This possibility is consistent with the finding that skilled readers appear to be less sensitive to phonology during the reading of text and in controlled experimental tasks (reviews and results reported by, Jared, Levy, & Rayner, 1999; Lewellen, Goldinger, Pisoni, & Greene, 1993; Sears, Siakaluk, Chow, & Buchanan, 2008). In contrast, high-skill participants may be relatively efficient at using orthographic (Sears et al., 2008) and morphological information, and therefore prefer a word-initial unit that preserves word-stem information (i.e., Max Coda segmentation). This possibility is considered in more detail below, but for now it is interesting to note that Taft's results do not exclude the possibility that skilled readers simply represent the phonological syllable differently (e.g., Taft, 1992). This ambiguity renders all theoretical interpretations of these findings in terms of the locus of syllable effects highly speculative and begs a replication of these results using an estimate of phonological reliance (see below). In closing, Taft reported that processing speed, as indicated by the overall reaction time of participants, was not a reliable predictor of syllable preference. Therefore, the link between syllable effects and reader skill is not reducible to simple processing speed.

The Syllable Effect and Item Properties. Participant characteristics are not alone in moderating syllable preference. Using the split-word paradigm, syllable preference is known to interact with some item characteristics as well. Any pursuit of the relationship between skill and syllable preference must adequately control or manipulate these factors where possible. Known item-characteristic moderators include the phonetic properties of the phonemes straddling the syllabic boundary (Chateau & Jared, 2003; Taft, 2002), the syllabic structure of the item (i.e., Consonant-vowel syllables vs. medial consonant cluster; post-hoc analysis by Taft, 2002), and perhaps most importantly the lexical frequency of the carrier word (Chen & Vaid, 2007; see also, Jared & Seidenberg, 1990; Macizo & Van Petten, 2007). Taft (2001, Experiment 3) had identified the boundedness of the Max Coda unit (i.e., whether it occurs in one or many words) as an additional factor, but Chen & Vaid (2007) demonstrated that this effect is neutralized when frequency is properly controlled.

Summary. A review of the syllable effect in the visual word recognition literature indicates that such effects have been historically difficult to replicate. More recent work casts a positive light on this empirical inconsistency: participant and item characteristics moderate the target effect. Among other things, this means that syllable effects should be predictably obtained with proper control and/or manipulation of these factors. Notably, though item characteristics like phonetic properties and lexical frequency are represented in many models of word recognition (see below), reader skill is not, at least not directly. We now turn to a discussion of these computational models and how they represent words and individual difference. This discussion will be followed by a section dealing with the selection of process-based variables that will help pin down the locus of syllable effects in these models.

Current State of Visual Word Recognition

The field of word recognition investigates a circumscribed problem-space that is nested at once within the fields of psycholinguistics and human memory processing (Seidenberg & McClelland, 1989). Specifically, this field concerns itself with explaining the acquisition and representation of the functional architecture (Coltheart, 1987) within the broader cognitive system that presides over the recognition, oral reading, and comprehension of individually presented words. Thus construed, the phenomena under investigation are assumed to interact with other components of linguistic knowledge (e.g., syntax, discourse) and to be governed by many of the same principles as other memory systems.

This literature has yielded a large body of empirical work from which a number of widely accepted assumptions have been derived about the cognitive system that supports word recognition and reading processes (for a perspective that differs somewhat from the mainstream, see Murray & Forster, 2004). Taken together, these assumptions form a very general framework for understanding reading that includes more specific theoretical models as special cases. First, this general framework is developed and then a detailed examination is presented of two more specific theoretical camps: that of the dual route cascaded model (Coltheart, 1978; Coltheart et al., 1993; Coltheart et al., 2001; and, arguably, Perry et al., 2007) and that of the PDP triangle-model (Seidenberg & McClelland, 1989; Plaut et al., 1996; Harm & Seidenberg, 1999; Harm & Seidenberg, 2004). In presenting these models, special attention is paid to the architectural details of the models, especially as they relate to representing individual differences.

The Reading System: Some Basic Assumptions

Developmentally (i.e., ontogenetically), reading skills are preceded by the acquisition of oral language, which implies, among other things (e.g., knowledge of grammar), detailed knowledge of the relationship between spoken words and their meaning. Literacy acquisition can be described as the process of relating knowledge of print to these two preexisting bodies of knowledge (for discussion, see, Frost, 1998; Harm & Seidenberg, 2004). Consistent with this simplified view is the basic observation that skilled-readers can do at least two things: convert print into sound (i.e., read aloud by retrieving phonological information based on orthographic cues) and extract meaning from printed words (e.g., retrieve semantic information in response to orthographic cues). In sum, the most basic description of the reading system has at its core a minimum of three bodies of knowledge and their specified interrelationships: orthography for representing knowledge of written words, phonology for representing knowledge of speech sounds, and semantics for representing knowledge of meaning. The relationship between these types of knowledge is presented schematically in Figure 1.

Uncontroversial Processing Streams or "Routes"

Figure 1 implies a number of general characteristics about the reading system. In the first place, it can be seen that the bodies of knowledge are connected to one another. These connections represent the associations that are there to be acquired by developing readers and subsequently exploited during online processing. In assuming that readers develop all these connections to some degree (Frost, 1998; Harm & Seidenberg, 2004), the schematic implies redundancy in the system for both the extraction of meaning and reading aloud. Specifically, reading aloud can be accomplished either directly via the orthography-to-phonology connection or indirectly by way of semantics (orthography-to-semantics-to-phonology). Similarly, comprehension can be achieved directly via the orthography-to-semantics connection or indirectly by way of phonology (orthography-to-phonology-to-semantics). Importantly, the *direct* and *indirect* processing streams are not typically viewed as mutually exclusive options: each route makes a contribution to processing in what is largely conceived as a cooperative division of labor, though direct routes are often assumed to be more efficient (Coltheart et al., 2001; Harm & Seidenberg, 2004). Among other sources of evidence, this general framework is consistent with the performance of patients with various kinds of acquired language impairments (for reviews, see Coltheart et al., 2001; Harm & Seidenberg, 2004; Perry et al., 2007).

Interactivity and Interactive Activation

The other assumption implied by the schematic presented in Figure 1 is that of interactivity and its ubiquity throughout the reading system (for a discussion of the interactivity principle, see McClelland, 1987). The interactivity principle was initially proposed to account for context effects in letter-recognition studies (McClelland & Rumelhart, 1981), whereby an accuracy and response time advantage is observed for letters embedded in a lexical context (i.e., words). In this case, it is assumed that letters cue the retrieval of orthographic knowledge, which in turn supports the recognition of letters through increased feedback ‘activation’ of letter-recognition units (for a discussion of the context-effect studies and the principle of activation more generally, see McClelland, 1979; McClelland & Rumelhart, 1981).

A number of concepts/assumptions have been developed to support this conception of interactivity (commonly called ‘Interactive Activation’), including ‘detectors’ and ‘activation’. To begin, it is generally conceded that input to the system is processed initially by all-or-none perceptual detectors, which in turn feed activation into the functional architecture whose purpose is to funnel this activation to appropriate representational units according to various assumptions, including interactivity. Within this context, the concept of activation has been proposed by way of an analogy to neural electrical signal transmission in the brain wherein the representational units of a computational model are thought to map loosely onto the functions of neuronal aggregates, as opposed to individual neurons. According to such representational schemes, each unit in the system has an attribute called activation, the level of which determines things like its availability for decision-making processes and its capacity for exciting or inhibiting other representational units, which is done continually/continuously (i.e., cascaded manner, as opposed to threshold-contingent transmission). Variants of this interactive-activation assumption have since been invoked to account for a number of phenomena that imply feedback from one body of knowledge to another, such as the influence of semantics (Imageability Effect: James, 1975; Balota et al., 2004; Polysemy Effect; Pexman & Lupker, 1999) and phonology (Homophone Effect: Pexman, Lupker, & Jared, 2001; Pexman, Lupker, & Reggin, 2002; and Feedback Consistency Effects: Stone, Vanhoy, & Van Orden, 1997; Ziegler, Montant, & Jacobs, 1997) in a task that only explicitly requires knowledge of orthography, visual lexical decision. Finally, an aspect of interactivity that may not be obvious from the schematic is that it seems to be automatic in skilled-readers, largely beyond their executive control (McClelland, 1987).

The General Framework and Specific Models

As mentioned earlier, the framework presented in Figure 1 is intended to be general enough to include major models of reading as special cases. For instance, it is worth noting that its graphical depiction is similar in many respects to that of the triangle-model (Seidenberg & McClelland, 1989; see below for a detailed discussion). The main difference between the two is that the triangle-model framework makes a number of additional assumptions regarding how the connections are developed and used by readers (e.g., hidden units). Similarly, the dual route cascaded (DRC) model framework proposed by Coltheart and colleagues (2001) -- and discussed below-- is a special case of the general model in Figure 1. The key distinction is that DRC and other dual route models propose that print-to-sound connectivity is split into two parallel processing streams, one lexical and the other non-lexical, which technically makes DRC a 'triple route' model of word reading (2 phonological routes, 1 semantic route). A more detailed examination of these theoretical complications and their broader implications is now presented.

Computational Theories of Reading

Recently, a particular class of theoretical model has risen to prominence in the cognitive sciences: the computational model. First introduced to a wide audience in the cognition literature of the 1960s (e.g., Miller, Galanter, & Pribram, 1960), computational models are 'computable' theories that have been expressed in the form of executable computer programs. As such, these models have the advantage of being fully specified, which means that all their assumptions are necessarily explicitly stated.

Perhaps more importantly, the implications of these assumptions can be readily determined by running simulations for the purpose of generating experimental predictions (e.g., Coltheart et al, 2001; Harm & Seidenberg, 2004). By comparing the simulation results to human-generated data, it is possible to ascertain the validity of a particular set of assumptions for the description of normal reading performance or an attested pattern of deficits (though there has been debate over what it means to empirically validate a model, see Seidenberg & Plaut, 2006). For these reasons and others, the computational approach is clearly superior to the verbal and box-diagram methods that preceded it (for discussion, see Seidenberg, 1992; Colheart et al., 2001; Perry et al., 2007).

The discussion to follow addresses itself specifically to two classes of model that are committed to the computational approach: the dual route model and the triangle-model. The following aspects of each model are discussed: (a) their theoretical antecedents, (b) the principles that guide their development, (c) the details of their most recent implementations, (d) the role that functional units play in their respective contexts, and (e) the possibilities for modeling individual differences in a population of readers. This discussion is intended to provide a theoretical background to the problem under study as well as some perspective on what the generalization to multi-syllable words might involve and what role a syllable unit might play in such models. It will be argued that individual differences in process-based effects are potentially useful in pinning down the locus of syllable effects.

Dual-Route Framework

The DRC framework is an approach to studying reading processes that was first described verbally and with box –and-arrow diagrams (Coltheart, 1978). Later, the print-to-sound component of the model was implemented in successive computational instantiations (Coltheart et al., 1993; Coltheart et al., 2001). The description to follow draws most heavily from the arguments and discussion presented in the 2001 paper, but it is intended both as a description of the specific implementation and of the more general theory on which it is based. In keeping with this spirit, elements of a newer dual route model (Perry et al., 2007) are also discussed where they contrast with the assumptions of the 2001 DRC model.[2]

Theoretical Antecedents. The developers of the most recent instantiation of the DRC framework (Coltheart et al., 2001) present their model as the synthesis of two distinct theoretical frameworks: the logogen model (Morton, 1969; 1980) and the Interactive Activation and Competition (IAC) model (McClelland & Rumelhart, 1981; McClelland, 1991; and the additional mechanisms for performing lexical decisions proposed by Grainger & Jacobs, 1996). More specifically, the macro architecture of the DRC model is explicitly adapted from the reading-system component of the logogen model, while most of the assumptions relating to online processing, including the attributes of its basic representational units, are drawn from the IAC framework. In other words, the logogen and IAC models can be said to have contributed the macro- and micro-level structure of the DRC model respectively. The common core of these models is a commitment to modularity (for reviews in support of the proposed modules, see Ellis & Young, 1988; Shallice, 1988; Shelton & Caramazza, 1999; Coltheart et al., 2001; Perry et al., 2007) and to the symbolic representation of linguistic information (e.g., letter units corresponding to letters; word units corresponding to words). The IAC model contributes the assumption of interactivity and widespread parallel processing, while the logogen model contributes the non-interactive and serial print-to-sound decoding mechanism that distinguish DRC from most PDP models.

The 'DRC' approach to model building. Coltheart and colleagues (2001) identified two fundamental principles that describe their approach to model building. The first, a view first articulated in the field by Jacobs & Grainger (1994) and reiterated by Perry and colleagues (2007), is that theoretical progress is promoted through adherence to the principle of 'nested modeling'. Nested modeling refers to the notion that a new computational model should match the descriptive adequacy of its predecessors in addition to accounting for new phenomena. Because the critical elements of both the logogen and IAC models were incorporated or nested within DRC, Coltheart and colleagues have argued that it matches and extends the adequacy of these models as explanations of reading and word recognition phenomena, though critics have pointed out that they have been less than systematic in verifying this claim (Seidenberg & Plaut, 2006; see Perry et al., 2007, for a more systematic approach). The second model-building principle identified by Coltheart and colleagues asserts that it is preferable for theorists to make explicit hypotheses concerning the functional architecture of their models rather than delegate this task to a learning algorithm whose limitations may prevent it from discovering the 'true' structure of the system (Coltheart et al., 2001, p.204-206). Coltheart and colleagues (2001) explicitly contrast their position with the supposedly atheoretical methods of 'New Behaviorism' adopted by the proponents of the triangle-model (for a vigorous response to such charges, see Seidenberg, 1992; Seidenberg & Plaut, 2006).

Overview of the DRC model. Like the logogen model, the synthetic DRC model proposes a bipartite print-to-sound decoding apparatus comprised of parallel and partially interacting routes to pronunciation, one lexical and the other non-lexical. The first 'lexical' mechanism is often referred to as 'addressed phonology' because it is based on the recognition of unique letter sequences corresponding to individual words. As such, it is well-suited to the processing of atypical words (e.g. words with irregular pronunciations) and words that are highly familiar (which are easier to retrieve as a unit, Adams, 1991), as well as the retrieval of semantics via the unimplemented 'indirect' route to oral reading. The second 'non-lexical' print-to-sound strategy is not specific to any word as it is based on generalized knowledge of the correspondence between the basic units of orthography and of speech: graphemes and phonemes. It is specialized for reading novel (unknown) words that are 'typical' in that their pronunciation can be inferred based on those of words that are already known. This strategy is associated with the term 'assembled phonology', which refers to its supposed construction of a complete phonological representation online. Together, the assumption of parallel print-to-sound connections, one lexical and one non-lexical, can account for certain benchmark effects observed in skilled readers along with many patterns of neurological impairment (for a list, see Coltheart et al., 2001; Perry et al., 2007). Further, the addition of certain assumptions about the developmental relationship between these two processes, as yet unimplemented computationally, allow the more general framework to explain certain developmental phenomena, such as the two major profiles of developmental dyslexia (Coltheart et al., 2001; Jackson & Coltheart, 2001; Desrochers & Thompson, 2008).

The two print-to-sound routes comprise a number of sequentially ordered modules, which is a characteristic inherited from both the logogen and IAC models. Another aspect of DRC's theoretical heritage is that each module contains explicitly defined symbolic units or 'locally represented' units. These basic processing units are considered symbolic because they are abstract representations with readily identifiable referents (for arguments in support of the use of symbolic representation in computational modeling, see Grainger & Jacobs, 1998; Page, 2000). For example, the letter-detection module contains position-specific 'letter units' corresponding to each letter in the language; the orthographic lexicon contains a word unit for each visual word-pattern known to the system; the phonological lexicon contains a unit for each word known in oral language; and the phoneme output buffer contains position-specific units corresponding to each phoneme attested in the language of interest. In addition, the non-lexical route contains a module that represents the graphemes of the language as a set of conditions within a coherent system for applying actions rules for transforming a string of symbols at the letter-level into the increased activation of units in the phoneme output buffer (for a different implementation of the non-lexical route that relies on activation within a 'grapheme' level of representation, see Perry et al., 2007). Thus, the DRC model explicitly represents many of the functional units of language within its architecture. If syllable effects can be pinned down to a given locus within the word recognition system, it is reasonable to expect that the syllable will also be symbolically represented. That is, unless syllable effects can be captured as an emergent property of the system (emergent representations in DRC are discussed below).

Online Interaction Between Lexical and Non-Lexical Streams. Based on this description it can be seen that the two routes are functionally segregated with the exception of two points of interaction, located at the first and final levels of the reading process respectively. Orthographic input is initially processed by feature detectors, but the first site of inter-mingling between the two processing streams occurs when activation reaches the letter-detection level, which feeds both the lexical and non-lexical routes. The second site of interaction is the phoneme output buffer, which receives activation from both decoding mechanisms simultaneously, and ultimately generates the overt output for the model when an arbitrary criterion has been reached. An implication of this property is that both routes always contribute to naming and word recognition; the relative contribution of each is a matter of degree, depending on the parameters of the model (Coltheart et al., 2001) and the properties of the letter-string being processed (e.g., such shifts in emphasis allow DRC to reproduce grapheme size, lexicality, and serial position effects; Coltheart & Rastle, 1994; Weekes, 1997; Rastle & Coltheart, 1998, 1999; Hino & Lupker, 2000). The model simulates both accuracy (i.e., it provides a phonemic output) and latency data, the latter of which is computed as the number of processing cycles required to reach either the naming criterion or that for lexical decision (based on activation in the orthographic lexicon). We now turn to a detailed examination of the lexical and non-lexical routes, which are each discussed in turn.

Operation of the lexical route. Within the lexical route, letter-detection proceeds in parallel across all serial-positions resulting in the cascaded activation of a series of modules: the orthographic lexicon, the phonological lexicon, and finally the phoneme output buffer. As noted above, there is also an as yet unimplemented semantic pathway that is directly connected to both the orthographic and phonological lexicon. Throughout, the process is governed by various processing assumptions, including interactivity and cascaded activation transmission.[3] An important implication these processing assumptions, especially that of cascaded activation transmission, is that units within a module that are similar in some way may act as a group (cohort) to influence performance, which is an example of distributed knowledge in DRC.

Perhaps the best known example of a cohort effect is the influence of orthographic neighborhood on word recognition performance (e.g. in lexical decision; Grainger & Jacobs, 1996). An orthographic neighborhood refers to a group of words whose orthography overlaps in some way (e.g. they share all their positional letters except one; Coltheart, Davelaar, Jonasson, & Besner, 1977). If the number of neighbors or 'neighborhood density' is important enough, then DRC predicts that the additional activation in the orthographic lexicon will generally improve naming performance (e.g.

Laxon, Masterson, & Moran, 1994; Peereman et Content, 1995, 1997) and influence lexical decision performance in various ways depending on variables such as the relative frequency of the target and its neighbors, and the lexicality of the letter-string (e.g. Andrews, 1992; Carreiras, Perea, & Grainger, 1997; Sears et al., 2008). In sum, orthographic neighborhood constitutes a type of emergent functional unit that manifests itself within any level of representation in the model when enough representational overlap is distributed across modules (for brief a discussion of localist vs distributed representation in DRC, see Seidenberg & Plaut, 2006, Note 3).

Operation of the non-lexical route. Unlike its lexical counterpart, the non-lexical mechanism in DRC (a) analyses letter-sequences serially one at a time rather than in parallel and (b) begins activating phonemes as graphemes are identified by the transcoding process which operates according to condition-action rules as opposed to competition and activation. The contition-action rules are implemented in a system of grapheme-to-phoneme correspondence relations (also known as GPC rules). As implemented, the graphemes definitions in the model do not possess the property of activation, which means that there is no feedback (i.e., interactivity) in the non-lexical route, but this is not a strong assumption, according to its authors. In fact, a recent implementation of the DRC conceptualization possesses a non-lexical decoding mechanism that incorporates both serial processing and interactive activation transmission (Perry et al., 2007). The main advantage of this revision is that it allows the model to account for the graded effects of grapheme consistency (e.g., Zevin & Seidenberg, 2006).

The assumptions of all-or-nothing rule application and serial processing have a number of implications for non-lexical processing in DRC. First, implicit functional units of the type that emerge in the lexical route are not admissible because activation is absent in the rule application process (though it is present in Perry et al., 2007). Second, the serial mechanism implies that the left-most graphemes within a letter-string are privileged relative to the others who receive activation only later, if at all. Thus, any property of letters or graphemes that is likely to cause the non-lexical route to generate an incorrect output (e.g., irregularity; Hino & Lupker, 2000) has a greater effect when it occurs early in a word. Another aspect of the serial-processing hypothesis is that the non-lexical mechanism works with incomplete data until the entire letter-string is processed. The result is that even an optimally configured rule application system generates an erroneous phoneme output much of the time. For example, the identity of multi-letter graphemes (e.g. th) is ambiguous until all its letters are revealed (see the DRC explanation for 'grapheme-size' or 'whammy' effects, Coltheart & Rastle, 1998).

Functional Unity in DRC. Because of the pervasive modularity of the DRC model, the representation of units tends to be localized and so it has been argued that these can be tweaked without directly affecting unrelated components of the system (Coltheart, 1999). However, the impact of such changes on a non-linear system like that of DRC requires simulations to be fully evaluated. The same property causes there to be two types of functional units in DRC. The first type is explicitly encoded into the system by modelers. Examples include the letter units and orthographic word units in the model, and the conditions associated with the application of non-lexical GPC rules. The second type of unit emerges during online processing due to the confluence of a number of factors. An example in the lexical route is the cohort effect relating to orthographic neighborhood. Within the non-lexical route, an example is the effect of grapheme size which arises from the serial left-to-right application of its decoding mechanism. The type of functional units that are available within the DRC framework have implications for the way in which it can be modified to account for the role of the syllable, if any, within the system. For example, changes to the explicit coding in the system may be required to account for syllable effects; or a multisyllable-compatible version of the model may account for the empirically attested syllable effects as an emergent property of its dynamics (e.g., the way words are represented and compete with each other during online processing). Importantly, irrespective of how such effects are supposed to arise, the macro-architecture of the model is stable across implementations, which means that localizing the source of functional unity to a particular processing stream (lexical-semantic, lexical, and non-lexical) will be useful when multi-syllable generalizations to the model are undertaken, and tests of the adequacy of various forms of representation performed.

Individual Differences in DRC. DRC is intended as a static model of the average and mature reader, and therefore its potential for explaining effects relating to individual differences within populations of normal readers has yet to be fully explored, though individual differences arising from abnormalities in the system have been investigated within the context of explaining patterns of acquired dyslexia by testing strategic lesions to the model (Coltheart et al., 1993; Coltheart et al., 2001). For the purpose of modeling individual differences in normal readers, several aspects of the model could be modified in principled ways, including aspects of its structure (e.g., the number and type of units represented; their connectivity) and the parameters governing its online functioning (e.g., the strength of feedback from phonology to orthography).

One such strategy would be to adjust the number of units in the orthographic and phonological lexicons in an attempt to model the effects of differences in vocabulary size, which is related to reading skill (Butler et al., 1984). Such an adjustment could have a pervasive effect on both lexical and non-lexical processing streams as the capacity of the system for formulating adequate GPC rules is constrained by the content of the vocabulary available to the system (Desrochers & Thompson, 2008). In addition to modifying the content of an existing module, new modules could be added in order to evaluate the hypothesis that groups of participants may be distinguished based on the modules their reading system employs, and the symbolic units within those modules. Indeed, this is the approach that Taft (1994, 2001) has taken in proposing a morpho-syllabic decomposition level of representation in between the letter-level and the orthographic lexicon that varies depending on reader skill (Taft, 2001). Other researchers have suggested that syllables may be represented symbolically in the non-lexical route as part of the transcoding mechanism (Ans, Carbonel, & Valdois, 1998), which is an idea that can account of individual differences in syllable preference by assuming that the default or dominant method of extracting phonology from print varies with reading skill.

A different type of strategy for modeling individual differences would involve systematically adjusting the more dynamic properties of the model, including the parameters controlling strength of connections among individual modules, in ways that are motivated by hypotheses as to how readers differ from each other. Currently, the space available to such parameters has been explored primarily with the limited aim of finding a single set of parameters that allows the model to simulate a large number of benchmark effects (Coltheart et al., 2001). However, for parameters that have a meaningful interpretation (e.g. route activation strength), the adjustment process may be informative in its own right as an exploration of how changes in a given parameter covary with changes in the magnitude of predicted effects. Such patterns of covariance constitute predictions about the way empirical estimates of parameter strength in a given reader should be related to that reader's sensitivity to a set of linguistic variables.

This way of modeling individual differences has been applied in a limited way to such phenomena as cross-language differences in the magnitude of certain benchmark effects (e.g., lexical frequency; phonological irregularity) when DRC has been extended to cover languages other than English (e.g., see the rationales that have been developed for parameter changes in the German and French instantiations; Ziegler, Perry, & Coltheart, 2000; Ziegler, Perry, & Coltheart, 2003). However, it could be applied more broadly in an attempt to move beyond modeling of the 'average' reader inferred from factorial ANOVA experiments and towards a model-based account of why some readers are more sensitive to a given benchmark experimental manipulation than others. Moreover, such a strategy may be useful for examining inter-participant variability in studies examining the functional unity of letter-clusters in word recognition (e.g., syllable effects), especially since the strength of functional unity is often explicitly linked to model parameters like feedback strength in interactive activation models.

Triangle-Model

The triangle-model framework, which is sometimes identified with the terms connectionism and parallel distributed processing (PDP) in the word-recognition literature, originated with a paper by Seidenberg & McClelland (1989). This paper presented a general framework for single word reading and an implementation of one of its components, the association between print and sound that supports oral reading. Various other computational instances of the framework have improved on the original (Plaut et al., 1996) and/or modeled other processes, such as the acquisition of phoneme representations (Harm & Seidenberg, 1999) and the interplay of orthography, phonology, and semantics in a full implementation of the 'triangle' (Harm & Seidenberg, 2004). In contrast to the DRC model, the triangle-model framework advances the idea that a single mechanism suffices to account for print-to-sound translation. Another major difference between the triangle-framework and DRC model is that the functional architecture of the reading system is conceived as a distributed network of weighted connections among pools of units rather than a series of modules that harbor symbolic units. Finally, the triangle-model is a computational model of reading acquisition as much as it is a model of the static reading system, whereas DRC is primarily conceived as a model of a static mature reader.

Theoretical Antecedents. According to its main proponent (Seidenberg, 1992), the triangle-model framework is an extension of the 'learnability approach' to studying reading acquisition that was developed by a number of linguists and psycholinguists in the 1970s and 1980s (Pinker, 1979; Wexler & Culicover, 1980). This framework attempts to explain the process of language acquisition in terms of four main aspects: (a) the starting state of the system, (b) the ultimate stable state that the system must achieve (e.g., mastery of a language's grammar), (c) the inputs to the cognitive system that are supplied by the environment, and (d) the system's capacity for learning. Seidenberg (1992) argues that the connectionist approach advocated within the triangle-model framework extends this basic approach by contributing a substantive theory of representation and learning, a formal way of expressing hypotheses concerning various aspects of learnability and their interplay across development.

The Connectionist Approach. Consistent with the idea that it is an extension of the learnability approach, the 'theoretical' component of connectionist modeling involves making various statements about the initial state of the reading system, such as the nature of input and output units (e.g., quality of phonological output units, Harm & Seidenberg, 1999), the parameters governing the learning process (e.g., learning algorithm; constraints on such algorithms like 'greed' for activation, Harm & Seidenberg, 2004), and the input available to the system (e.g., the impact of various training regimens, Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004; Powell, Plaut, & Funnell, 2006) and subsequently evaluating whether the steady-state that is achieved by the model (a) approximates the performance of the specific population of readers being modeled (e.g. normal readers or dyslexic readers) or (b) informs the process of relating computational functions to their brain bases (Seidenberg, 2005). This technique has provided new insights into the causes of reading difficulties, (e.g., the limitations of the original 1989 model have been reinterpreted as a restricted model of phonological dyslexia, Harm & Seidenberg, 1999). Modelers from this school are constrained by a number of assumptions about how the cognitive system works (e.g., GRAIN: Graded, Random, Adaptive, Interactive, Nonlinear; McClelland, 1991, 1993; Plaut et al., 1996; Seidenberg, 2005), including the principles that processing is in parallel within a given pool of units and knowledge is represented in a distributed fashion throughout a network, which means that the same processing units a) encode the input-output mappings of words that have been encountered before and b) venture generalizations in response to novel letter-strings.

Connectionist Architecture(s). The issue of architecture within connectionist models is complex. In the first place, the initial state of a connectionist network, its network architecture, can be distinguished from the functional architecture that is developed with learning and which serves as the skilled-reading analogue of the DRC model (Coltheart et al., 2001). That these two aspects can be dissociated is easily demonstrated by the fact that identical network architectures that are subjected to different training histories can achieve vastly different functional architectures (Hutzler et al., 2004; Powell et al., 2006), and changes in network architecture can generate different functional architectures despite identical training histories (Harm & Seidenberg, 1999, 2004). Moreover, it is possible to distinguish between the simple case of an input-output mapping network (e.g., Seidenberg & McClelland, 1989) and more complex networks where different kinds of input-output networks are forced to interact with one another (Harm & Seidenberg, 1999, 2004). Connectionist architecture at both levels of analysis is considered in turn. Since the functional architectures of such models depend on learning processes, this aspect of the model is discussed along with strictly architectural elements.

A key feature of 'connectionist' models is that the associations among domains of knowledge are modeled as interconnected pools of neuron-like units that learn input-output associations (Seidenberg, 2005). Within a network architecture (i.e., initial state), two broad types of unit-pools are distinguished: those that represent an input or an output (i.e., domain of knowledge: e.g., letter forms; phonemes) and the 'hidden units' that mediate their association. The hidden units are justified sometimes on computational grounds by virtue of the fact that they are several times more powerful than simple input-output networks (Hinton, McClelland, & Rumelhart, 1986; but see Harm & Seidenberg, 2004 and Perry et al., 2007 for arguments against universal use hidden units). Whether a network employs hidden units or not, the system learns (i.e., acquires its functional architecture) by gradually adjusting the weighted connections among the pools of units in the network architecture in response to training that associates specific inputs with their corresponding outputs (e.g., letter form; phoneme form) and it does this according to a set of rules for performing such operations (e.g., learning algorithm like backpropagation or the delta rule).

In more complex systems, such as the full implementation of the triangle-model framework presented by Harm & Seidenberg (2004), the acquisition of a given input-output mapping is influenced by the fact that some or all of the same units are simultaneously acquiring another set of associations. For example, Harm & Seidenberg

(2004) found that the task of learning to retrieve word meanings in reading was divided amongst the indirect 'semantically mediated' route and the direct Orthographic-Semantics pathway. Among other things, the relative importance of the two routes in the retrieval of semantics was determined by pressure to respond quickly (i.e., the direct route is faster, and as a result retrieves meaning for most words, especially those of high frequency), greed for activation (e.g., both the direct and indirect routes are implicated in the processing of most words), and the distributional properties of words (e.g., words with inconsistent pronunciations are read primarily via the direct route). An important implication of these simulations is that the dominance of a particular processing route is determined by both experience (e.g., amount of training, with more training resulting in increased dominance of the direct route to semantics) and the properties of individual items (consistency of a given input-output mapping, such as spelling consistency, encouraging reliance on the alternate route). This observation suggests that if syllable effects are route-dependent, they may vary as a function of both participant- and item-attributes.

Operation of PDP models. Mirroring the description of its functional architecture, the online operation of PDP models can be described at a minimum of two levels of analysis. Again, the simplest level concerns the computation of an output based on a single input pattern (e.g., phonology based on orthography). In fact, early PDP models (Seidenberg & McClelland, 1989; Plaut et al., 1996) were based entirely on the mapping of a single input domain to an output domain. For simple input-output relations, the weighted connections of a PDP model allow it to generate a pattern of activation in response to an input pattern. This pattern determines the identity of the output units that get activated, and therefore the response of the model (e.g., pronunciation in the case of a naming model). The accuracy of the model can therefore be compared against human data. What is more, response time has also been modeled in earlier models using a rough measure of the network's certainty (e.g., Mean summed error) and, in more recent implementations, using the settling times of an attractor network (Harm & Seidenberg, 1999, 2004).

In more complex architectures (Harm & Seidenberg, 2004), the dynamic properties of attractor networks allow the time-course of activation flow to be modeled. For example, simulations reveal that the number of input-output steps in a processing stream (1 in a direct connection; 2 if it is indirect) is inversely related to response speed such that the most direct association is most efficient and continues to improve with training long after the indirect route has reached asymptote. Nevertheless, parallel processing streams tend to cooperate in generating a response pattern (e.g., both direct and indirect routes help the output in the pool of semantic units to settle on a stable response) and tend to compensate for each other's weakness on an item-by-item basis, whether during online processing or while learning.

Functional Unity and Individual Differences in PDP models. The issues of functional unity and individual differences in PDP models are dealt with together in this section because in practice they are modeled using the same three general strategies. The only exception is that functional unity is strongly influenced by the distributional properties of words, which are largely out of a modeler's control. For example, the vowel-consonant unit in monosyllabic words happens to be more predictive of pronunciation than the grapheme, which suggests that a connectionist model and actual readers would learn to treat these orthographic 'rimes' as a unit (see, Treiman, Mullennix, Bijeljac_Babic, & Richmond-Welty, 1995; Kessler & Treiman, 1997). That said, PDP theorists can control various aspects of their model in order to capture empirically attested functional units, individual differences, and any interaction between these aspects.

For instance, two network architectures with identical starting states can be made to differ in their ultimate stable state (i.e., mature functional architecture) by manipulating factors relating to training. A network could be specifically trained to learn grapheme mappings, which would have an impact on the representation patterns in its network, the main effect of which is to improve novel word or nonword naming (Powell et al., 2006). The idea here would be to model differences in training that are observed across groups of readers with the object of simulating the features that distinguish such groups on reading tasks (e.g., across educational systems; Hutzler et al., 2004). For present purposes, an interesting feature of the Harm & Seidenberg (2004) 'comprehension' network in this regard is that increased training— which is analogous to print exposure and/or skill— results in increased reliance on the direct semantic pathway (which continues gains in processing speed as the phonologically mediated route reaches an asymptote), suggesting that better (more experienced) readers rely more on semantics during reading or at least have that option available to them when trying to avoid the negative influence of incorrect phonological information. This prediction is consistent with the observation that skilled readers are less sensitive to the phonological properties of words during reading (Jared, Levy, & Rayner, 1999; Lewellen, Goldinger, Pisoni, & Greene, 1993; Unsworth & Pexman, 2003) and the finding that extensive reading experience (i.e., print exposure) increases efficiency in the use of orthographic patterns (Chateau & Jared, 2000; Sears, Siakaluk, Chow, & Buchanan, 2008; Stanovich & West, 1989). This shift in emphasis with reading skill and how it might be estimated is taken up again in the final section of the paper.

Another method of modeling individual differences would be to keep the training regimen constant while varying the starting state of the network (network architecture), with the aim of evaluating the impact of decisions about the starting state of a network on the stable state that it ultimately achieves. For example, the nature of the input representations could be modified in various ways by changing the symbolic unit that constitutes the input (e.g., letters or graphemes), for example, or by changing the way spatial position is coded (i.e., highly distributed relative position coding, as in wickelfeatures, Seidenberg & McClelland, 1989; grouping of letter inputs based on their relative position within syllabic structure; Plaut et al., 1996). As stated earlier, such manipulations have provided the basis for a PDP account of phonological types of developmental dyslexia (Harm & Seidenberg, 1999). More directly relevant to the issue of functional unity in multi-syllable words is the fact that recent implementations of the triangle model code letter input in terms of their position within syllable structure (Plaut et al., 1996; Harm & Seidenberg, 1999, 2004; the same is true of the Perry et al., 2007 model). Obviously, additional assumptions or modifications to the existing coding system are required in order to handle multisyllabicity, specifically the issue of segmenting words into pronunciation units. Ideally, the nature and locus of such processes will be constrained by relevant data, such as interactions with effects specific to a given processing stream.

Finally, various properties of a trained model can be adjusted to model acquired reading impairments or strategic emphasis. For instance, a mature network might be lesioned by removing certain connections in order to simulate the effects of acquired reading impairments. What is more, the strength of activation transmission can be adjusted for various types of connection (e.g., feedback connections) so as to simulate, among other things, strategic emphasis or de-emphasis of certain types of knowledge. To the extent that they are based on the same general model of reading, it may be possible to directly compare the DRC and triangle models based on the predictions that they make regarding the impact of varying the value of activation strength parameters related to the fundamental semantic and phonological processing streams and to pit these predictions against the performance of actual readers. Such comparisons are impossible until it is known what patterns the results reveal.

Summary. In this section of the paper, an attempt has been made to sketch a general framework for understanding word recognition processing and two specific implementations of this general framework have been presented. The structural features of these models have been highlighted and the philosophical positions of the two theoretical camps have been presented so that future extensions to the models might better be anticipated. The possibilities for implementing functional unity have been discussed, both in terms of explicit representation and emergent properties. DRC tends to emphasize the former while PDP models emphasize the latter type of representation, and yet neither possibility is excluded in either.

With respect to individual differences, both models assume parameters related to connection strength, both 'feedforward' and 'feedback', whose psychological correlates would be assumed to vary across readers. Further, both models assume that the processing strength of different routes is independent in the sense that certain pools of information may be more important than others during processing. Within the more general framework presented in Figure 1, it can be seen that the processing streams common to both models are those of orthography-to-phonology and orthography-to-semantics. Characterizing item-level effects (frequency of occurrence, spelling-sound consistency, syllable effects) in terms of their relationship to these two processing streams would provide a useful way of establishing the construct validity of certain controversial effects and of further constraining models of reading with an under-exploited set of empirical parameters: the way the magnitude of a given experimental effect correlates with another effect across a population of participants.

In the present case, a critical question related to syllable-effects concern their relationship to individual differences in general and their locus within models of reading in particular. The interaction between skill and syllable effect patterns reported by Taft (2001, 2002) is intriguing, and spurs speculation that process-based covariates of reading skill, which is to say variable efficiency in the use of phonology, semantics and orthography, may be among the cognitive moderators. A number of specific predictions arise from the discussion to follow, which considers ways of operationalizing individual differences in cognitive terms. If these predictions hold, then the empirical inconsistency in the literature will have been explained in a more detailed way and theorists will have been provided with critical information about the role of the syllable in multisyllabic word reading.

Towards A Process-Based Characterization

A critical question in the word recognition literature, as noted earlier, concerns the status of the syllable as a functional unit. Regrettably, the relevant literature is presently quite inconsistent, which is the issue addressed by Taft (2001, 2002). His key finding, that skill moderates syllable preference, explains the inconsistency of previous results and provides an important first step towards articulating an individual-differences-based account of syllable effects in visual word recognition. The next obvious step is to refine the theoretical interpretation of Taft's results by examining promising hypotheses (e.g., a shift away from phonology as a primary source of information with increasing skill). Ideally, such research would speak directly to computational models of reading by operationally defining individual differences in ways that have a relatively clear theoretical interpretation. What follows is the description of a research program that aims to do just that.

Process-Based Individual Differences

If a general measure of reading skill is too coarse grained to inform theoretical models of reading, the question then becomes: what aspects of computational models are likely to vary with reader skill and therefore be a plausible source of moderating effects reported by Taft (2001, 2002)? Perhaps the most obvious explanation, differences in pure processing speed, was investigated by Taft, who reported a non-significant correlation with overall reaction time. This null effect is important in that response delays are associated with greater involvement of semantic and phonological information (Seidenberg, Petersen, Plaut, & MacDonald, 1996). Thus, it would seem that a more nuanced explanation is warranted.

As an explicitly developmental model, the triangle framework is the most obvious source of insight into the relationship between reader skill and an activation-based model. Indeed, Harm & Seidenberg (2004) report simulations whereby greater reading experience is associated with correspondingly greater involvement of direct semantic processing during comprehension over indirect phonologically-mediated comprehension. This result is consistent with the finding that skilled or experienced readers are less sensitive to phonology during the reading of text and in controlled experimental tasks (review and results reported by, Jared, Levy, & Rayner, 1999; Lewellen et al., 1993; Sears et al., 2008). It also suggests a theoretical mechanism underlying the phonological-shift hypothesis that Taft (2001) evoked in explaining his results. According to this account, skilled or experienced readers de-emphasize phonologically mediated processing in favor of the direct orthography-to-semantics connection. Such readers may then draw upon a richer source of information when making decisions about words and when reading words aloud. They are therefore accordingly less sensitive to a variety of factors that make words more difficult to recognize, especially those related to phonology (e.g., homophony; phonological syllable). In contrast, low-skill readers rely strongly on phonological information by necessity and are therefore especially sensitive to the phonological attributes of stimuli. They may also be less sensitive to the semantic attributes of words, all other things being equal, since their direct connections between orthography and phonology are less well developed. In sum, skilled-readers may make very rapid use of phonology (e.g., Ashby, Sanders, & Kingston, 2007), but they also have rapid access to other sources of information, such as semantics. In contrast, low-skill readers are much less efficient at retrieving all kinds of information including orthography, phonology, and semantics (Sears et al., 2008; Unsworth & Pexman, 2003), but are forced to rely primarily on phonology by default. When it comes to syllable effects, this analysis predicts that low-skill readers would prefer a phonological syllabification, while high-skill readers would be more likely to demonstrate preference for other types of units, such as a Max Coda orthographic syllable.

Phonological Emphasis Hypothesis. A stringent test of the phonology-emphasis hypothesis would require that the reliance on phonology be estimated for each participant so that its potential moderating effects on patterns of syllable preference can be evaluated. If the hypothesis is confirmed, syllable preference will depend on the degree to which a participant relies on phonology. This dependency could take two forms: (a) a symmetrical effect (high-phonology participants prefer Max Onset; low-phonology participants prefer Max Coda) or (b) an asymmetrical effect (high-phonology preference for Max Onset; Null effect for low-phonology participants). In contrast, if the Max Coda syllable (i.e., BOSS) is merely a different kind of phonological syllable, then no relationship should be observed, all other things being equal. A third possibility is that is that both the Max Onset and Max Coda syllable are units of orthographic access, in which case no relationship should be observed with phonological reliance. The latter idea is broadly consistent with Taft's conception of syllable units as gateways for lexical access (Taft, 1979, 1994, 2001) and relies implicitly on the idea that phonological structures shape orthographic representations during development, rather than on the notion of functional unity arising from the retrieval of phonological information in an interactive-activation-based model. With respect to computational theory, the first of these hypotheses is arguably the simplest to capture because it corresponds to the parameter or groups of parameters governing the activation strength of the orthography-to-phonology processing stream.

Unlike the PDP networks, DRC is not developmental, but it does possess parameters that can be used to simulate differences in phonological emphasis. Particularly relevant is the fact that DRC proposes two different types of phonology, one lexical and the other non-lexical. Accordingly, two distinct estimates of phonological reliance in visual word recognition are proposed; though all phonological influences on lexical decision in DRC are assumed to be lexically mediated (i.e., via feedback from the phonological lexicon to the orthographic lexicon, which is the seat of decision-making processes in this task). Operationally, the regularity and homophony effect in lexical decision would provide estimates of the strength of non-lexical and lexical phonology respectively (Unsworth & Pexman, 2003). The homophony effect may be especially useful in that it has proven reliable across studies and experimental conditions (Pexman & Lupker, 1999; Pexman, Lupker, Jared, 2001; Pexman, Lupker, & Reggin, 2002) and it has consistently been interpreted as being the result of feedback from phonology to the orthographic domain. Within the triangle-model framework, sensitivity to homophony is simply interpreted without 'lexical' qualification as the tendency to rely on phonological information.

Obverse of Phonological Emphasis? If not phonology, what do low-phonology participants rely on? One possibility, derived from the DRC perspective, is that the processing focus shifts to a purely orthographic level of representation, the orthographic lexicon (Coltheart et al., 2001). In this case, a Max Coda preference might correlate with an estimate of orthographic sensitivity, like the orthographic N effect. Specifically, reading skill is known to be associated with insensitivity to orthographic N (Chateau & Jared, 2000; Sears, Siakaluk, Chow, & Buchanan, 2008). The idea here is that stronger orthographic representations for such readers result in better resistance to interference. A second possibility is that readers shift from a phonological emphasis to a strategy that exploits both phonologically-mediated and direct semantic information, with the direct connection to semantics dominating most of the time (Harm & Seidenberg, 2004). One way of estimating reliance on direct semantics would be to estimate a participant's sensitivity to a semantic variable like imageability (James, 1975) or polysemy (Pexman & Lupker, 1999). Should tests of the phonological-emphasis hypothesis reveal an asymmetrical effect, it may be possible to complete the pattern that was observed with reader skill (Taft, 2001) by testing the combined moderating potential of phonological emphasis and either semantic reliance or sensitivity to orthographic N. In the latter case, "poor" readers, or at least readers with relatively limited exposure to print, are more sensitive to the influence of N, where low N is a relatively greater hindrance with words and high N is a relatively greater hindrance for correctly rejecting nonwords (Sears et al., 2008). Strain & Herdman (1999) reported stronger imageability effects for low-skill participants with low-frequency irregular words, but the nature of this relationship may reverse after controlling statistically for overall reaction time and/or phonological reliance.

Dual Skill Components. The account just derived from the Harm & Seidenberg

(2004) model makes a number of predictions about the relationship between reliance on phonology and reliance on semantics. First, after controlling for overall reaction time (Seidenberg et al., 1996), the two should arguably be negatively related to one another. At the very least, the relationship between these attributes is an interesting empirical issue. Second, the relationship between these variables and syllable preference should be dissociable such that high-semantics participants prefer the Max-Coda analysis and high-phonology participants prefer a Max-Onset analysis. More nuanced tests might also be considered whereby the ratio of phonological to semantic reliance (i.e., their relative importance) or their interaction is actually most predictive of syllable preference. For instance, an equal reliance on both semantic and phonological information may yield a null effect when Max Coda and Max Onset conditions are compared. If used together, estimates of phonological and semantic reliance (or sensitivity) may permit the locus of syllable effects to be pinned down within the general reading and word recognition framework presented in Figure 1 and by extension the DRC and triangle-model frameworks.

A similar investigation of the role played by orthographic sensitivity, as indexed by Orhographic N, in determining syllable preference could be undertaken. The use of semantic sensitivity as a predictor of Max Coda preference assumes that it is an emergent property of the process of retrieving semantic information. This is indeed what would be predicted if morphological effects had a semantic basis, but the Max Coda unit is assumed to be represented independently of semantics, though it may act as an efficient access unit (Forster & Taft, 1976; Taft, 2001, 2003) and therefore result in greater semantic involvement. Nevertheless, if the orthographic interpretation of Max Coda preference is correct, then the orthographic N effect may be a better predictor of Max Coda preference than semantic reliance. In fact, this is the result that one would predict given the finding that morphological units provide equivalent priming regardless of whether they are semantically empty (Rastle, Davis, & New, 2004). It is advisable then that both semantic and orthographic sensitivity be considered jointly in pursuing this line of enquiry.

A Separate Issue: Representation

A contentious issue concerns the attribution of functional unity to explicit representation or implicit representation. For instance, it is possible that the syllable is perceived as a unit because letter-clusters act as a cue for the retrieval of syllable-units. Alternatively, functional unity might arise from the fact that syllables provide a useful context from which to constrain grapheme pronunciation, for example (Chateau & Jared, 2003). If, as in the latter example, syllable effects were dependent on some aspect of the distributional property of words, then this would count as evidence against the notion that syllables are somehow explicitly represented in models of reading (e.g., bigram frequency trough hypothesis; Seidenberg, 1987; spelling-sound consistency in naming; Chateau & Jared, 2003). In comparing two syllabification heuristics, the work of Taft (2001, 2002) and the research strategy proposed here do not address this issue directly. In other words, the results of such experiments may be due to either explicit representation of the syllable or, perhaps less likely, to a confound with the distributional properties of an input-output domain. A direct contrast of these two ideas would require an analysis of the distributional properties of the input-output domain in question (e.g., orthography-phonology; orthography-semantics). Given that it is unclear which domain is giving rising to the syllable preference effects in lexical decision (especially the Max Coda syllable), it is perhaps best to determine which input-output domains are relevant to the question at hand, as the research program outlined in this article proposes to do, before exploring such finer grained hypotheses further.

Conclusions

Extant computational models of reading can only handle single-syllable words. Inconsistency in the syllable-effect literature is currently an obstacle to the principled generalization of such models to the more common case: multi-syllable words. Taft (2001) suggested an explanation for this inconsistency based on reader skill. It was argued that this type of association is not particularly informative for model building, because of the cognitively amorphous nature of such measures. Instead, a new approach is advocated that exploits an under-used source of empirical constraint: the magnitude of experimental effects across participants, and their inter-relationships. This approach would link syllable preference to individual differences that are interpretable in terms of computational theories. Of particular interest to determining the locus of syllable effects in visual word recognition are the relative strength of semantic and especially phonological feedback across participants, and their possible role in mediating syllable preference. Ways of estimating these variables were proposed and hypotheses were derived from extant models. It is hoped that further enquiry into this area will yield answers to support theorists at critical stages in model development.

References

Adams, M.J. (1991). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Anderson, J. & Jones, C. (1974). Three theses concerning phonological representations. *Journal of Linguistics, 10,* 1-26.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18,* 234-254.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105,* 678-723.

Ashby, J. & Martin, A.E. (2008). Prosodic phonological representations early in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance,34,* 224-236.

Ashby, J. & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes, 19,* 391-426.

Ashby, J., Sanders, L.D., & Kingston, J. (2007). Skilled readers begin processing sub-phonemic features by 80 ms during visual word recognition: Evidence from ERPs. *Biological Psychology,* doi:10.1016/j.biopsycho.2008.03.009.

Balota, D.A., Cortese, M.J., Sergent-Marshall, S.D., Spieler, D.H., & Yap, M.J. (2004). Visual word recognition of single-syllable words, *Journal of Experimental Psychology: General, 133,* 283-316.

Balota, D., Yap, M., Cortese, M., Hutchison, K., Kessler, B., Loftus, B., et al. (2004).The English lexicon project, *Behavior Research Methods, 39,* 445-459.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A. et al. (2003). Timed picture naming in seven languages. *Pscychonomic Bulleting & Review, 10,* 344-380.

Blevins, J. (1995). The syllable in phonological theory. In J.A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206-244). Cambridge, MA: Blackwell.

Brand, M., Rey, A., & Peereman, R. (2003). Where is the syllable priming effect in visual word recognition? *Journal of Memory and Language, 48,* 435-443.

Brown, J.I., Bennett, J.M., & Hanna, G. (1981). *The Nelson-Denny reading test.* Chicago: The Riverside Publishing Company.

Butler, B.E., Jared, D., & Hains, S. (1984). Reading skill and the use of orthographic knowledge by mature readers. *Psychological Research, 46,* 337-353.

Bybee, J. (2001*). Phonology and language use.* Cambridge, England: Cambridge University Press.

Carr, P. (2000). *English phonetics and phonology: An introduction.* Oxford: Blackwell.

Carreiras, M., Alvarez, C.J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32,* 766-780.

Carreira, M., Mechelli, A. Estevez, A., & Price, C.J. (2007). Brain activation for lexical decision and reading aloud: Two sides of the same coin? *Journal of Cognitive Neuroscience, 19,* 433-44.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 857-871.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28,* 143-153.

Chateau, D., & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language, 48*, 255-280.

Chen, H. & Vaid, J. (2007). Word frequency modulates the Basic Orthographic Syllabic Structure (BOSS) effect in English polysyllable word recognition. *Language and Cognitive Processes, 22*, 58-82.

Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 55-76). Hilsdale, NJ: Erlbaum.

Clements, G.N., & Keyser, S.J. (1983). *CV Phonology*. Cambridge, MA: MIT Press.

Cole, R. & Scott, B. (1974). Toward a theory of speech perception. *Psychological Review, 81*, 348-374.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.

Coltheart, M. (1987). Functional architecture of the language-processing system. In Coltheart, M., Sartori, G., and Job, R. (Eds.), *The Cognitive Neuropsychology of Language*. London: Lawrence Erlbaum Associates.

Coltheart, M. (1999). Modularity and cognition, Trends in Cognitive Science, 3, 115-120.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: evidence for dual-route models of reading. *Journal of Experimental Psychology: Human, Perception and Performance, 20*, 1197-1211.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589-608.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 534-555). New

York: Academic Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Content, A., Kearns, R., & Frauenfelder, U. (2001). Boundaries versus onsets in syllabic segmentation, *Journal of Memory and Language, 45*, 177-199.

Cortese, M., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words, *Behavior Research Methods, 36*, 384-387.

Cutler, A. (1981). Making up materials is a confounded nuisance, or: Will we be able to run any psycholinguistic experiments at all in 1990? *Cognition,* 10, 65-70.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English, *Journal of Memory and Language, 25*, 385-400.

Delatre, P. (1965). *Comparing the phonetic features of English, French, German and Spanish: an interim report.* Heidelberg: Julius Groos Verlag.

Desrochers, A. & Thompson, G. (2008). De l'orthographe à la lecture orale. In A. Desrochers, F. Martineau & Y.C. Morin (Eds.). *Orthographe française : évolution et pratique*. Ottawa: Éditions David.

Duncan, L., & Seymour, P. H.K. (2003). How do children read multisyllabic words? Some preliminary observations. *Journal of Research in Reading, 26*, 101-120.

Ehri, L.C. (1992). Reconceptualizing the devleoopment of sight word reading and its relationship to recoding. In P. Gough, L.C. Ehri & R. Treiman (Eds.), *Reading* *acquisition.* (pp. 107-143). Hillsdale, NJ: Erlbaum.

Ellis, A., & Young, A. (1988). *Human cognitive neuropsychology.* London: Erlbaum.

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: the role of syllabic units. *Journal of Memory and Language, 35*, 708-723.

Ferrand, L., Segui, J., & Humphreys, G. (1997). The syllable's role in word naming. *Memory & Cognition, 25*, 458-470.

Forster, K.I. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory & Cognition, 28*, 1109-1115.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K.E. Patterson, J.C. Marshall & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* (pp. 301-330), Hillsdale, NJ: Erlbaum.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin, 123*, 71-99.

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

Grainger, J., & Jacobs, A. (1998). On localist connectionism and psychological science. In J.Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition.* (pp. 1-38). Mahwah, NJ: Erlbaum.

Harm, M. & Seidenberg, M. (1999). Reading acquisition, phonology, and dyslexia: Insights from a connectionist model. *Psychological Review, 106*, 491-528.

Harm, M. & Seidenberg, M. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Henderson, L. (1985). On the use of the term "grapheme". *Language and Cognitive Processes*, 2, 135-148.

Hino, Y. & Lupker, S. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 166-183.

Hinton, G. E., McClelland, J.L., Rumelhart, D.E. (1986). Distributed representations. In D.E. Rumelhart, & J.L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol, 1, pp. 77-108). Cambridge, MA: MIT Press.

Hox, J. (2002). *Multi-level analysis: Techniques and Applications.* London: Lawrence Erlbaum.

Hooper, J. (1972). The syllable in phonological theory. *Language*, *48*, 525-540.

Hutzler, F., Zielger, J.C., Perry, C., Wimmer, H. & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition, 91*, 273-296.

Jackson, N. E., & Coltheart, M. (2001). *Routes to reading success and failure.* New York: Psychology Press.

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

James, C.T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception & Performance, 1*, 130-136.

Jared, D., Levy, B., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General, 128*, 219-264.

Jared, D., McRae, K., & Seidenberg, M. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29*, 687-715.

Jared, D., & Seidenberg, M. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 92-105.

Jordan, T. (1986). Testing the BOSS hypothesis: Evidence for position-insensitive orthographic priming in the lexical decision task. *Memory & Cognition, 14*, 523-532.

Kahn, D. (1980). *Syllable-based generalizations in English phonology*. New York: Garland.

Katz, L., & Baldasare, J. (1983). Syllable coding in printed-word recognition by children and adults. *Journal of Educational Psychology, 75*, 245-256.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295-311.

Krakow, R. (1999). Physiological organization of syllables: A review. *Journal of Phonetics, 27*, 23-54.

Laxon, V., Masterson, J., & Moran, R. (1994). Are children's representations of words distributed ? Effects of orthographic neighbourhood size, consistency and regularity of naming, *Language and Cognitive Processes, 9*, 5-7.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Lewellen, M.J., Goldinger, S.D., Pisoni,D.B., & Greene, B.G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and categorization. *Journal of Experimental Psychology: General, 122*, 316-330.

Lima, S. D., & Pollatsek, A. (1983). Lexical access via orthographic code? The Basic

Orthographic Syllabic Structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior, 22*, 310-332.

Lindblom, B., & Studdert-Kennedy, M. (1967). On the role of formant transitions in vowel recognition. *The Journal of the Acoustical Society of America, 42*, 830-843.

Lorch, R. & Myers, J. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149-157.

Massaro, D. (1974). Percpetual units in speech recognition. *Journal of Experimental Psychology, 102*, 199-208.

Macizo, P. & Van Petten, C. (2007). Syllable frequency in lexical decision and naming of English words. *Reading and Writing, 20*, 295-331.

McClelland, J. (1979). On the time relations of mental processes: An examination of systems of processing in cascade. *Psychological Review, 86*, 287-330.

McClelland, J. (1987). The case of interactionism in language processing. In M Coltheart (Ed.), *Attention and Performance XII: The psychology of reading.* London: Erlbaum.

McClelland, J. (1991). Stochastic interactive proceses and the effect of context on perception. *Cognitive Psychology, 23*, 1-44.

McClelland, J. (1993). The GRAIN model: A framework for modeling the dynamics of information processing. In D.E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 655-688). Hillsdale, NJ:Erlbaum.

McClelland, J., & Rumelhart, D. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings.

Psychological Review, 88, 375- 410.

Miller, G., Galanter, E., & Pribam, K. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart, & Winston.

Morton, J. (1969). The Interaction of information in word recogniton. Psychological Review, 76, 165-178.

Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive processes in spelling*. London: Academic Press.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111*, 721-756.

Nation, P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.) *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). New York: Cambridge University Press.

New, B., Ferrand, L., Pallier, C., Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13*, 45-52.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189-234.

Norris, D., & Cutler, A. (1985). Juncture detection. *Linguistics, 23*, 689-705.

Page, M. (2000). Connectionist modeling in psychology: A localist manifesto. *Behavioral and Brain Sciences, 23*, 443-512.

Peereman, R. & Content, A. (1995). The neighborhood size effect in naming: lexical activation or sublexical correspondences, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 409-421.

Peereman, R. & Content, A. (1997). Orthographic and phonological neighborhoods in naming: Not all neighbors are equally influential in orthographic space. *Journal* *of Memory and Language*, *37*, 382-410.

Perry, C., Ziegler, J., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*, 273-315.

Pexman, P., & Lupker, S., (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects. *Canadian Journal of Experimental Psychology*, *53*, 323-334.

Pexman, P., Lupker, S., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 139-156.

Pexman, P., Lupker, S., & Reggin, L. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 572-584.

Pinker, S. (1979). Formal models of language learning. *Cognition*, *7*, 217-283.

Plaut, D., McClelland, J., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading : Computational principles in quasiregular domains. *Psychological Review*, *103*, 56-115.

Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a reading unit. *Journal of Memory and Language*, *25*, 461-475.

Powell, D., Plaut, D., & Funnell, E. (2006). Does the PMSP connectionist model of single word reading learn to read in the same way as a child? *Journal of Research in Reading*, *29*, 229-250.

Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus.* The Hague: Mouton.

Raajimakers, J. (2003). A further look at the "Language-as-fixed-Effect Fallacy". *Canadian Journal of Experimental Psychology*, *57*, 141-151.

Rapp, B. (1992). The nature of sublexical orthographic organization: the bigram trough hypothesis examined. *Journal of Memory and Language, 31*, 33-53.

Rastle, K. & Coltheart, M. (1998). Whammy and double whammy: the effect of length on nonword reading. *Psychonomic Bulletin and Review, 5*, 277-282.

Rastle, K. & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 482-503.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language, 42*, 342-364.

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

Rey, A., Zielger, J., & Jacobs, A. (2000). Graphemes are perceptual reading units. *Cognition, 75*, B1-B12.

Rouibahh, A., & Taft, M. (2001). The role of syllabic structure in French visual word recognition. *Memory & Cognition, 29*, 373-381.

Rozin, P. & Gleitman, L. (1977). The structure and acquisition of reading II: The reading process and the acquisition of the reading principle. In A. Reber & D. Scarborough (Eds.) *Toward a Psychology of Reading* (pp. 55-141). Hillsdale, NJ: Erlsbaum.

Santiago, J., MacKay, D.G., & Palma, A. & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes, 15*, 1-44.

Schiller, N. (1999). Masked syllable priming of English nouns. *Brain and Language, 68*, 300-305.

Schiller, N. (2000). Single word production in English. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 512-528.

Sears, C.R., Siakaluk, P.D., Chow,V.C., & Buchanan, L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research* (Online March 15th, 2008).

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention & Performance XII* (pp.245-263). Hillsdale, NJ: Erlbaum.

Seidenberg, M.S. (1992). Connectionism without tears. In S. Davis (Ed.) *Connectionism: Theory & Practice* (pp. 84-122). New York, NY: Oxford University Press.

Seidenberg, M.S. (1993). Connectionist models and cognitive theory. *Psychological Science, 4*, 225-235.

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science, 14*, 238-242.

Seidenberg, M. S., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Seidenberg, M.S., Petersen, A., Plaut, D., & MacDonald, M.C. (1996). Pseudohomophone effects and models of word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 48-62.

Seidenberg, M. S., & Plaut, D. (2006). Progress in understanding word reading : Data fitting versus theory building. In Andres, S. (Ed.), *From inkmarks to ideas: Current issues in lexical processing*. Hove, UK: Psychology Press.

Selkirk, E. (1982). The syllable. In H. Van der Hulst & N. smith (Eds.), *The structure of* *phonological representations* (pp. 337-383). Dordrecht: Foris.

Seymour, P. (1999). Cognitive architecture of early reading. In I. Lundberg, F.E. Tonnessen & I. Austad (Eds.), *Dyslexial: Advances in Theory and Practice* (pp. 59-73). Dordrecht: Kluwer.

Shallice, T. (1988). *From neuropsychology to mental scturcutre*. Cambridge, England: Cambridge University Press.

Shelton, J., & Caramazza, A. (1999). Deficits in lexical semantic processing: Implications for models of normal language. *Psychonomic Bulletin & Review, 6*, 5-27.

Spencer, A. (1996). *Phonology*. Oxford, UK: Blackwell.

Spoehr, K.T., & Smith, E.E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology, 5*, 71-89.

Spoehr, K.T., & Smith, E.E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance, 104*, 21-34.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402-433.

Stone, G.O., Vanhoy, M., & Van Orden, G.C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36*, 36, 337-359.

Taft, M. (1979). Lexical Access via an orthographic Code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behaviour, 18*, 21-39.

Taft, M. (1987). Morphographic processing: The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and performance, XII: Reading* (pp. 265-299). Hillsdale, NJ:

Erlbaum.

Taft, M. (1991). *Reading and the Mental Lexicon.* Hove: Lawrence Erlbaum Associates Ltd.

Taft, M. (1992). The body of the BOSS: Subsyllabic units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1004-1014.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, *9*, 271-294.

Taft, M. (2001). Processing of orthographic structure by adults of differing reading ability. *Language & speech, 44*, 351-376.

Taft, M. (2002). Orthographic processing of polysyllabic words by native and nonnative English speakers. *Brain & Language, 81*, 532-544.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink. & D. Sandra (Eds.) *Reading complex words*. Amsterdam: Kluwer.

Taft, M., & Forster, K. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15*, 607-620.

Taft, M., & Radeau, M. (1995). The influence of the phonological characteristics of a language on the functional units of reading: A study in French. *Canadian Journal of Experimental Psychology*, *49*, 330-346.

Taft, M. & van Graan, F. (1998). Lack of phonological mediation in a semantic categorization task. *Journal of Memory and Language*, *38*, 203-224.

Thompson, G.L. (2008). Eliminating aggregation bias in experimental research: Random coefficient analysis as an alternative to performing a 'by-subjects' and/or 'by-items' ANOVA. *Tutorials in Quantitative Methods for Psychology*, *4*, 21-34.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, *25*, 476-491.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language*, *27*, 87– 104.

Treiman, R., Gross, J., & Cwikiel-Glavin, A. (1992). The syllabification of /s / clusters in English. *Journal of Phonetics*, *20*, 383–402.

Treiman, R., Mullennix, J., Bijeljac-Babic, R. & Richmond-Welty, E.D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*, 107-136.

Treiman, R., Straub, K., & Lavery, P. (1994). Syllabification of bisyllabic nonwords: Evidence from short-term memory errors. *Language and Speech*, *37*, 45 –60.

Treiman, R. & Zukowski, A. (1990). Towards an understanding of English syllabification. *Journal of Memory and Language*, *29*, 66-85.

Unsworth, S.J., & Pexman, P. (2003). The impact of reader skill on phonological processing in visual word recognition. *Quarterly Journal of Experimental Psychology*. *56A*, 63-81.

Uhry, J. & Ehri, L. (1999). Ease of segmenting two- and three-phoneme words in kindergarten: Rime cohesion or vowel salience? *Journal of Educational Psychology*, *91*, 594-603.

Van Orden, G., Pennington, B., & Stone, G. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488-522.

Venezky, R. L. (1970). *The structure of English orthography*. The Hague: Mouton.

Weekes, B. (1997). Differential effects of number of letters on word and nonword naming latency. The *Quarterly Journal of Experimental Psychology, 50A*, 439-456.

Wexler, K. and P. Culicover (1980). *Formal Principles of Language Acquisition*, Cambridge, MA: MIT Press.

Zelinsky, G.J., & Murphy, G.L. (2000). Synchronizing visual and language processing: An effect of object name length on eye movements, *Psychological Science*, *11*, 125-131.

Zevin, J.D. & Seidenberg, M.S. (2006) Simulating consistency effects and individual differences in nonword naming. *Journal of Memory and Language*, *54*, 145-160.

Ziegler, J.C., & Goswami, U. (2005). Reading acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory, *Psychological Bulletin, 131*, 3-29.

Ziegler, J.C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language*, *37*, 533-554.

Ziegler, J.C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology, 12*, 413-430.

Ziegler, J.C., Perry, C., & Colhteart, M. (2003). Speed of Lexical and Nonlexical Processing in French: The case of the regularity effect. *Psychonomic Bulletin & Review, 10*, 947-953.

Authors Note

Glenn L. Thompson & Alain Desrochers, School of Psychology, University of Ottawa, Ottawa, Ontario, Canada.

The research reported in this article was supported by the National Sciences and Engineering Council of Canada (NSERC) in the form of Master- and Doctoral-level scholarships awarded to the first author, as well as NSERC grants awarded to the second author.

Correspondence should be addressed to Glenn Thompson at the School of Psychology, University of Ottawa, 145 Jean-Jacques Lussier, P.O. Box 450, Station A, Ottawa, Canada, K1N 6N5, or e-mail: GlennLThompson@gmail.com.

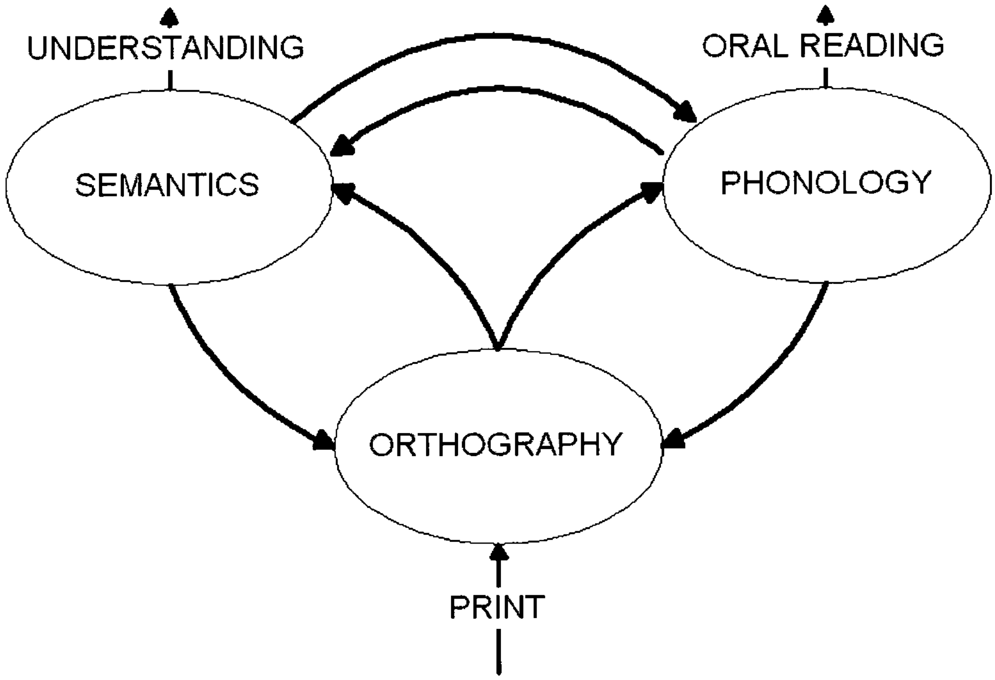
Notes

[1] The DRC model actually reads all monosyllabic words (approximately 7000), including the roughly 3000 items that are polymorphemic. The importance of this generalization is open to question given that semantics (morphology pertains to the relationship between semantics and both orthography and sound) are not implemented in the model.

[2] Perry and colleagues (2007) frame the origin of their model in another way, but it is arguably a straight-forward modification of the DRC model (the lexical route is essentially identical) that incorporates a grapheme-level of representation in the non-lexical route and an adaptive distributive network associating this grapheme-level with the units of the phoneme-level of representation. This model is not fully adaptive in that grapheme knowledge is built into the model, which is another way in which their model is similar to DRC.

[3] Cascaded activation can be contrasted with threshold-based transmission, which assumes that activation transmission between levels of representation is limited to a single unit: the first to reach its firing threshold (see for e.g. the logogen model; Morton, 1980). Cascaded transmission derives from the view that the units and activation of computational models reflect the action of neuronal aggregates whose behavior is not thresholded though their constituent elements may be.

List of Figures


UNDERSTANDING
ORAL READING
SEMANTICS
PHONOLOGY
ORTHOGRAPHY
PRINT

Chapter 2

Characterizing the Syllable Effect in Visual Word Recognition:

A Novel Methodology & New Evidence

Abstract

Thompson & Desrochers (2009a) argued that the most important question facing theorists in the visual word recognition literature concerns the role of the syllable. Most computational models of reading and word recognition can only process a small pool of monosyllabic words. Extending these models to the case of multi-syllable words requires additional empirical information. For instance, what syllable units are relevant (e.g., Max Onset, Max Coda), when and why are they relevant, and for whom? A research program is described that proposes to address these issues. First, a novel methodology is proposed that is suitable for examining the relationship between participant attributes and syllable preference: multi-level and random coefficient analysis. Second, two series of experiments are described that address whether reliance on phonology, reliance on semantics, and general 'orthographic' reading are related syllable preference. The research program, which comprises the first author's doctoral dissertation, aims to improve upon current understanding of syllable effects and the role they should play in the next generation of theoretical models.

Keywords: Multi-level modeling; Reading; Visual Word Recognition; Syllable; Functional Unit; Lexical Processing

Introduction

In their review of the syllable-effect literature, Thompson & Desrochers (2009a) observed that most computational models of visual word recognition are monosyllabic, and that major theoretical papers in the area have consistently emphasized the need to better understand the role of the syllable. Additional information on this score would inform the principled development of more general models that are capable of handling multi-syllable words. Perhaps the most commonly used method for experimentally examining the syllable in visual word recognition is the word-splitting paradigm. According to this method, words are presented so that they are divided somehow (e.g., case manipulation, insertion of a foreign character). This division creates an artificial boundary within the word. If this boundary coincides with a psychologically real boundary, then the logic of the paradigm assumes that word recognition should be facilitated (i.e., faster reaction time, greater accuracy) relative to comparison conditions (e.g., boundary + 1 letter). When the boundaries being examined are syllable boundaries, this facilitation is known as a syllable effect. Thus far, the literature has focused on contrasting two syllabic units: the Max Onset syllable (dubbed the 'phonological' syllable) and the Max Coda syllable (dubbed the 'Basic Orthographic Syllable Structure' or BOSS; Taft, 1979). Both units are rules or heuristics for dividing words into syllables. The rules agree for some words, but can be constrasted for others, such as when a cluster of consonants occurs between vowels (e.g., THUN/DER versus THUND/ER).

As Thompson & Desrochers (2009a) observed, the research results that pertain to the syllable effect in the scientific literature are inconsistent, with more recent work focusing on the moderating role of individual differences (e.g., reader skill; Taft, 2001, 2002) and lexical frequency (Chen & Vaid, 2007) as explanations for these inconsistent findings. Despite the utility of reader skill in predicting syllable preference (low-skill readers prefer Max Onset division, high-skill readers Max Coda), Thompson & Desrochers (2009a) argued that it is not a particularly informative construct, at least not with respect to computational theory. In its stead, they proposed that process-based variables would provide information that is more theoretically relevant. Specifically, the relative value of semantic, orthographic, and phonological reliance in predicting syllable preference could provide clues as to the localization and function of the syllable within the word recognition system. If syllable preference depends on any of these variables, it suggests that they share some communality in the cognitive system. What follows is a description of a research program that was undertaken for following through on the general strategy suggested by Thompson & Desrochers (2009a).

The strategy that we implemented is described in three sections. First, we briefly describe the advantages of using a multi-level strategy for examining interactions among participant- and item-level characteristics. These advantages are described in more detail in Thompson (2008), where a procedure for performing random coefficient analysis in SPSS is described. This procedure was employed in the analysis of empirical data reported in the two papers to follow. In the first of these, we describe the logic behind an initial study designed to address the issue of whether Max Onset syllable preference is related to phonological reliance. The justification, methodology, and results of this study are described in greater detail in Thompson & Desrochers (2009b). In the second, we describe the motivation behind a more synthetic study that examined simultaneously various participant-characteristics, including print exposure, phonological reliance, semantic reliance, and various indices of orthographic sensitivity. This second study is described in detail by Thompson & Desrochers (2009c). Finally, the contribution of this series of experiments is described briefly here. These considerations are presented in greater detail in the General Discussion.

Multi-level Modeling

Traditional methods of analyzing psycholinguistic data have been criticized for a number of years (Clark, 1973), and this criticism has become more heated with the increased availability of software for multi-level modeling (Baayen, Davidson, & Bates, 2008; Locker, Hoffman, & Bovaird, 2007). In a nutshell, psycholinguistic research is a special case because it typically draws samples from two distinct populations: participants and items (i.e., words). As a result, psycholinguistic experiments yield datasets that are structurally complex: they contain clusters of correlated observations. Each participant responds to multiple items within each experimental condition, and each item is responded to by multiple participants. Under a standard least-squares regression model, such data violates the assumption of independent observation. One strategy for sidestepping this issue involves simplifying the dataset by averaging over items (or participants) in what has been called an 'aggregation strategy'. An important problem with this strategy is that it tends to underestimate the uncertainty of coefficient estimates (i.e., standard errors; Thompson, 2008). As a result, analyses conducted in this way are much more susceptible to Type-I error inflation.

An alternative strategy that is not subject to the same limitation is the use of multi-level modeling (Baayen et al., 2008; Locker et al., 2007; Thompson, 2008). The key advantage to this strategy is that it allows participant and item variability to be modeled simultaneously, eliminating the need to test both participant and item effects separately. Additional advantages to this strategy include the possibility of using continuous predictors instead of simplified binary factors (e.g., high and low frequency; For a discussion of variable factorization, see Desrochers, Thompson, & Fréchette, 2009), testing interactions between continuous predictors, and testing cross-level interactions between continuous predictors (e.g., word and participant variables). These advantages are discussed by Thompson (2008). Further advantages include greater flexibility with respect to the parameters relating to the variance-covariance matrix (e.g., the addition of random slopes, covariance among random effects, autoregressive variance components).

With specific reference to the problem of syllable effects in visual word recognition, multi-level analysis is interesting in that it provides a way to estimate individual differences in sensitivity to a variety of word recognition difficulty factors. Such estimates are obtained by estimating distinct regression models for each participant (Thompson, 2008). For some participants a given variable may be relevant (e.g., statistically and practically significant) and for others not. This variability, effectively individual differences in the magnitude of beta coefficients, may be used to predict syllable preference in a separate multi-level analysis. Indeed, this strategy was used in the two empirical studies described below.

Syllable Preference and Phonological Reliance

Taft (2001) reported evidence that syllable preference depends on reader skill. In his discussion, Taft speculated that individual differences in phonological reliance may be responsible for his observed patterns of syllable preference. According to this view, poor readers would place disproportionate emphasis on phonology, and therefore prefer the 'phonological' syllable. In contrast, good readers have the possibility of consistently recognizing words purely on the basis of orthography. Given this ability, reliance on phonology would be minimal, and so preference of Taft's orthographic syllable would be prevalent. These two patterns would correspond to Max Onset and Max Coda syllable preference demonstrated respectively by poor and good readers (Taft, 2001, 2002).

Based on this analysis, readers who rely more heavily on phonology could be supposed to be more sensitive to aspects of phonology that make words more difficult to recognize: phonological difficulty factors such as regularity and homophony (Thompson & Desrochers, 2009b). To the extent that this is true, it is possible to directly test Taft's interpretation of his results. If Taft is correct, then participants who are highly sensitive to such difficulty factors should prefer Max Onset division. If such preference is non-lexical, then it may be specifically associated with sensitivity to regularity. In contrast, if it is lexical, then it should be specifically correlated with sensitivity to homophony. By simultaneously considering both sensitivity to regularity and homophony, the general phonological-reliance hypothesis can be evaluated as well the more fine-grained lexical/non-lexical distinction. It is worth noting that if the phonological sensitivity is positively associated with syllable preference, then this would be inconsistent with the view that both the phonological and orthographic syllable exist in orthographic syllabaries as units of lexical access (see Taft, 1979, 2001). The goal of the first series of experiments was to examine this idea more closely (Thompson & Desrochers, 2009b).

Thompson & Desrochers (2009b) modeled individual differences in syllable preference in two steps. First, the sensitivity of each participant to regularity and homophony in visual lexical decision was estimated using random coefficient analysis. Second, these estimates of sensitivity were employed as participant-level predictors of syllable preference in a word-splitting paradigm. The results promised to directly test Taft's suggestion that syllable preference depends on phonological reliance, and to do so using 'process-based' variables (Thompson & Desrochers, 2009a). Two versions of the phonological reliance hypothesis were considered: a) a strong form of the hypothesis whereby heavy phonological reliance would be positively associated with

Max Onset preference and negatively associated with Max Coda preference, and b) a weak, asymmetrical form whereby heavy phonological reliance would only be predictive of Max Onset preference. A third possibility, a null result, was deemed to be consistent with the idea that phonological reliance does not distinguish between Max Onset and Max Coda preference. Two explanations for such a result are possible: a) the relevance of both types of syllable is an emergent property of phonological feedback or b) both types of syllable are units of orthographic access. Being a null result, it would not of course provide any direct evidence for either possibility. The results that were actually obtained are described in Thompson & Desrochers (2009b). The central finding was that homophone sensitivity was positively related to Max Onset syllable preference with low-frequency words, suggesting that such preference in visual lexical decision arises from feedback lexical phonology. Max Coda preference did not emerge consistently. Thus, it was the predicted asymmetrical pattern that was observed.

Syllable Preference and Strong Orthography

A limitation of the previously described study was that only one kind of participant variable was considered: phonological reliance. Words, however, are multi-dimensional stimuli. Semantic and orthographic attributes are also known to affect the speed and accuracy with which words are recognized. What is more, sensitivity to such attributes or 'difficulty factors' is known to be related to reading skill. It was hypothesized by Thompson & Desrochers (2009b) that the asymmetrical pattern they obtained in contrasting Max Onset and Max Coda syllable preference was due to the fact they lacked a direct predictor of a Max Coda advantage. A broad set of difficulty-factor indicators that tap an assortment of constructs (e.g., semantics, orthography) would provide the best chance of identifying a reliable process-based associate of Max Coda preference. When considered at once with phonological indicators, it may be possible to achieve positive prediction of both Max Onset and Max Coda preference, as in by Taft (2001). For instance, a general variance component might emerge representing a general insensitivity to attributes that make letter-strings more difficult to discriminate. Such a factor might complete the picture, yielding symmetrical prediction of syllable preference.

Thompson & Desrochers (2009c) proceeded to estimate sensitivity to a variety of phonological, orthographic, and semantic factors. To make a closer link with Taft (2001)'s work, a proxy for reading skill was also obtained, print exposure (as estimated by the Author Recognition questionnaire, Stanovich & West, 1989). This more exploratory correlational approach seemed justified given the inconsistency of more high-constraint experimental research. It seemed wise to take a step back to gain a wider view of the problem, acknowledging the uncertainty in this area.

In the first place, sensitivity to the word recognition difficulty factors was estimated. A principal components analysis was conducted to reduce the number of variables to a series of orthogonal predictors. It was hypothesized that high-skill would be associated with a general insensitivity to difficulty factors, which would in turn be associated with syllable preference. However, a general sensitivity-skill variance component was not obtained. Rather, major sources of variance were distributed across six components. The pattern of results indicates that a) apparent syllable preference arises from multiple sources, and b) these sources do not provide clear general support for the Phonology-Max Onset hypothesis or the Skill-Insensitivity hypothesis. Response speed emerged as a moderator of some of these relationships, indicating that response speed is an area where greater experimental control might be attempted when examining syllable effects. Manipulating participant instructions, for example, to emphasize speed or accuracy could reduce instability in the emergence of syllable effects. For a more detailed discussion of these results, see Thompson & Desrochers (2009c).

Theoretical and Methodological Contributions

The research reported in this dissertation contributes to the existing literature in at least three ways. These contributions are considered briefly here, and taken up again in the General Discussion of this dissertation document. First, it acts as a demonstration of how random coefficient analysis and multi-level modeling generally can be used to estimate the relationship among 'process-based' variables, and their value as predictors of other types of behavior. The power of multi-level models arises from their flexibility with respect to modeling the covariance structure of complex databases like those produced by psycholinguistic experiments. These methods provide a more comprehensive empirical model of data than is obtained using typical methods, which is interesting as a general strategy in that an adequate model of word recognition should be able to explain the covariance among difficulty-factor sensitivities (i.e., estimated using random coefficient analysis), as well as their relationship to other behaviors such as syllable preference. Second, it provides a more complete account of the conditions under which syllable effects are observed. Thus, the discipline is closer to achieving empirical stability in the syllable effect literature, a stability that is necessary if theoretical progress is to be made. Finally, a fairly strong theoretical implication of the reported results is that Max Onset preference is directly related to lexical feedback phonology for low-frequency words. This result provides a useful constraint for theorists (e.g., Coltheart et al., 2001; Harm & Seidenberg, 2004) as they generalize their models to multi-syllable words, in that it situates some aspect of syllable preference to feedback phonology. More generally, it would seem that apparent syllable preference in visual word recognition is determined by a number of tendencies, not all of which are consistent with a locus for Max Onset preference in feedback phonology or for Max

Coda preference with high-skill orthographic processing.

Conclusion

The research program described above had two main objectives: a) adapt random coefficient analysis to the purpose of estimating sensitivity to difficulty factors in word recognition and b) employ this methodology to address open questions relating to the role of the syllable in visual word recognition. In fact, random coefficient analysis proved to be a useful way to model individual differences in sensitivity, both in terms of reaction time and accuracy. From a substantive standpoint, individual differences in syllable preference were associated with lexical feedback phonology in an initial series of experiments. In follow up work, there was a little evidence of a general 'insensitivity' variance component. Rather, individual differences seemed spread out over multiple variance components. The picture presented by these results is complex, but provides an indication of the participant characteristics that are implicated in the emergence of apparent syllable preference.

References

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*, 390-412.

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Desrochers, A., Thompson, G. L., & Fréchette, S. (in press). The Impact of Information Technology on the Selection of Lexical Items for Research in Human Experimental Psychology. In F. Columbus (Ed.), *Advances in Psychological Research*. Hauppauge, NY: Nova Science Publishers, Inc.

Harm, M. & Seidenberg, M. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Locker, L.Jr., Hoffman, L., & Bovaird, J.A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, *39*, 723-730.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 402-433.

Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language & Speech*, *44*, 351-376.

Taft, M. (2002). Orthographic processing of polysyllabic words by native and nonnative English speakers. *Brain & Language, 81*, 532-544.

Thompson, G.L. (2008). Eliminating aggregation bias in experimental research: Random coefficient analysis as an alternative to performing a 'by-subjects' and/or 'by-items' ANOVA. *Tutorials in Quantitative Methods for Psychology, 4*, 21-34.

Thompson, G.L., & Desrochers, A. (2009a). Towards a Process-Based Characterization of Syllable Effects in Visual Word Recognition. Manuscript submitted for publication.

Thompson, G.L., & Desrochers, A. (2009b). Sensitivity to Homophone Interference as Correlate of Syllable Preference in Visual Word Recognition. Manuscript submitted for publication.

Thompson, G.L., & Desrochers, A. (2009c). Individual Differences and Syllable Preference: Distal and Proximal Correlates. Manuscript submitted for publication.

Authors Note

Glenn L. Thompson & Alain Desrochers, School of Psychology, University of Ottawa, Ottawa, Ontario, Canada.

The research reported in this article was supported by the National Sciences and Engineering Council of Canada (NSERC) in the form of Master- and Doctoral-level scholarships awarded to the first author, as well as NSERC grants awarded to the second author.

Correspondence should be addressed to Glenn Thompson at the School of Psychology, University of Ottawa, 145 Jean-Jacques Lussier, P.O. Box 450, Station A, Ottawa, Canada, K1N 6N5, or e-mail: GlennLThompson@gmail.com.

Chapter 3

Eliminating Aggregation Bias in Experimental Research: Random Coefficient Analysis as an Alternative to Performing a 'by-subjects' and/or 'by-items' ANOVA

Abstract

Experimental psychologists routinely simplify the structure of their data by computing means for each experimental condition so that the basic assumptions of regression/ANOVA are satisfied. Typically, these means represent the performance (e.g. reaction time or RT) of a participant over several items that share some target characteristic (e.g. Mean RT for high-frequency words). Regrettably, analyses based on such aggregated data are biased toward rejection of the null hypothesis, inflating Type-I error beyond the nominal level. A preferable strategy for analyzing such data is random coefficient analysis (RCA), which can be performed using a simple method proposed by Lorch & Myers (1990). An easy to use SPSS implementation of this method is presented using a concrete example. In addition, a technique for evaluating the magnitude of potential aggregation bias in a dataset is demonstrated. Finally, suggestions are offered concerning the reporting of RCA results in empirical articles.

Introduction

Researchers routinely transform their data in order to satisfy the assumptions of statistical analyses (e.g. regression analysis). For example, log, reciprocal, and square-root transformations are all used to correct the shape of empirical distributions so that the assumption of normality (Gaussian distribution) is satisfied. Such considerations also guide how raw data is prepared for analysis. For instance, the regression/ANOVA approach that is taught in undergraduate and graduate-level statistics classes requires that each data point be independently collected or at least uncorrelated. In practical terms, this means that each participant must contribute a single data-point to an analysis. This assumption is violated in many situations such as when a repeated measures design is used. The solution in the special case of repeated measures is to extract the offending 'correlated' between-subject variance prior to analysis (e.g. repeated-measures ANOVA; correlated t-test). A limitation of this strategy is that it is only applicable to cases where researchers are interested in comparing the repeated observations or matched observations. This approach is not applicable to cases where participants generate a number of responses that is greater than the number of experimental conditions. Often, this is the case when the independent variable (IV) is the property of items or a collection of items, such as a comparison of reaction times (RTs) in response to high- and low-frequency words. In such cases, the data points or observations contributed by each subject are clusters of inter-dependent scores within the dataset as a whole that are often summarized by a mean to satisfy the assumption of independence (e.g. mean RT in response to high- or low-frequency words). Such clustering is a consequence of a dataset structure that is said to be complex, hierarchical, or multi-level.

The purpose of this paper is to present an SPSS macro for analyzing multi-level data using the Random Coefficient method proposed by Lorch & Myers (1990). First, the case is made for abandoning a commonly used strategy for accommodating multi-level data, namely aggregation (i.e. computing subject- or item-type means for each experimental condition), in favour of Random Coefficient Analysis (RCA). This discussion is limited to a relatively non-technical summary of previous work. Readers interested in more technical details, such as mathematical proofs, may consult Lorch & Myers (1990) and the other relevant sources cited below. The development of the justification behind using RCA is followed by (a) instructions for using the SPSS syntax provided here to perform RCA, (b) a description of how the results of such analyses are interpreted using a concrete example, (c) demonstration of a method for evaluating the magnitude of potential bias in a dataset, and (d) suggestions on how to report a RCA in an empirical article.

Strategies for Analyzing Hierarchically Structured Data

It is easiest to introduce the concept of hierarchical structure with a concrete example. As discussed above, hierarchically structured datasets contain clusters of inter-dependent observations that are caused by the presence of multiple levels of analysis (e.g. participant, item). For instance, a researcher may examine whether a group of participants reads frequently used words aloud more rapidly than words that are used infrequently in print. A sample of high- and low-frequency words (e.g. 20 each) is collected and each participant (e.g. N = 40) generates a single response for each word. The comparison of interest is high- vs low-frequency words on RT. Each item is associated with 40 responses and each participant produced 20 responses per condition. Clearly, whether the dataset is examined from the item or the subject perspective, it contains correlated observations that must be accommodated somehow.

Cohen, Cohen, West, & Aiken (2003) identify three possible ways of dealing with this situation, which they refer to as the "clustering" problem (p. 539). The simplest strategy is called *disaggregation*, which amounts to ignoring the correlated responses in the dataset despite the fact it violates a fundamental assumption of regression (i.e. independent observation). The problems with this strategy are so obvious that they will not be considered further here. A much more common strategy involves replacing correlated observations with an estimate of central tendency like the mean so that the dataset satisfies the assumption of independent observation or, stated differently, the assumption of uncorrelated observation (Hox, 2002). This strategy is called *aggregation* and it yields results that are subject to important, but oft-ignored, conceptual limitations (e.g. ecological fallacy; Robinson, 1950) and a host of concomitant statistical problems that are discussed below. A third strategy, sometimes called *Random Coefficient Analysis* (i.e. RCA), involves first analyzing the data within individual participants, and then determining whether the magnitude of this within-subject effect differs significantly from zero on average for the sample of participants. In what follows, the drawbacks of the aggregation strategy are presented and then the superior alternative RCA is discussed.

The Aggregation Strategy

The habitual way of dealing with the type of hierarchically structured data discussed here is aggregation (Hox, 2002). Aggregation involves computing an estimate of central tendency to summarize multiple scores at one level of analysis (e.g. items) with a single observation at another level of analysis (e.g. participants). For example, the responses of each participant might be averaged over items within cells of the experimental design (averaging over subjects is also an option, but bias remains an issue).[1] For the scenario developed earlier, this procedure would result in two observations per participant: mean RT for high- and low-frequency words. The 'by-subjects' solution yields data that can be submitted to a repeated-measures analysis because the structure of the dataset has been simplified to reflect the comparison of interest, which is high vs low frequency words. However, while accurately estimating cell means, aggregation decreases the complexity of datasets at the expense of (i) forcing researchers to choose between performing a by-subjects analysis or a by-items analysis, (ii) decreasing the accuracy of population variability estimates, and (iii) inflating the probability of spuriously rejecting the null hypothesis (Lorch & Myers, 1990; Raudenbush & Bryk, 2002; Raaijmakers, 2003).

Type-I error inflation arises in the by-subjects frequency example because the treatment effect is confounded with the degree to which the treatment effect varies across participants. Thus, statistically significant effects that are observed with ANOVA using the aggregate strategy may be due to a participant by experimental-effect interaction rather than the experimental effect per se (Lorch & Myers, 1990). To understand why this is so, the logic of the F-test (ANOVA, Regression) needs to be understood.

The reasoning behind the F-ratio test is simple. If an estimate of overall variance is large enough relative to an estimate of error variance, then the null hypothesis of no experimental effect is rejected (Howell, 2002, pp. 324-325). The estimate of overall variance is called the mean square treatment (MS_{effect}) and the estimate of error variance is called the mean square error (MS_{error}). The MS_{effect} estimate comprises two general components: (a) variance caused by the experimental effect and (b) error variance. In contrast, the MS_{error} variance estimate is comprised solely of error variance. If MS_{effect} and MS_{error} are equal, then F is equal to 1, and there is obviously no treatment effect. If MS_{effect} is greater than MS_{error}, then the probability of obtaining the resulting F value – assuming the null hypothesis of no effect is true – is determined using its degrees of freedom and the known F-ratio probability distribution. If this probability is small enough (say .05), we reject the null hypothesis in favor of the alternative: the treatment has an effect. The logic of the F-ratio test, of course, only holds if the error variance component within MS_{effect} is comparable to that represented by MS_{error}.

Regrettably, the aggregation strategy produces MS_{effect} and MS_{error} terms that contain qualitatively different error estimates. The violation of the F-ratio's logic is apparent when one examines the MS_{effect} (1) and the MS_{error} (2) terms that result from aggregation:

(1) MS_{effect} = [Treatment effect] + [(Subject · Treatment Interaction error) + (Residual error)]

(2) MS_{error} = [Residual error]

An F-ratio based on such a mean square error does not disentangle the contribution of the experimental effect (i.e. the linear component) and the degree to which the treatment effect varies across participants (i.e. the non-linear participant by experimental-effect interaction). Thus, a statistically significant effect could be due to one of three things: (a) a significant experimental effect, (b) significant variation in the treatment effect across participants, or (c) both of these things. The ambiguity is caused by the absence of the participant by treatment interaction in the MS_{error} term. This source of error must be present in the denominator of the F-ratio to statistically control for its presence within the numerator.

An obvious solution to this problem would be to generate an F-ratio based on an MS_{error} term that includes both the 'residual error' and the 'participant by treatment interaction' components. This more appropriate error term can be obtained by analyzing the effect of item variables within each subject separately and then testing whether estimates for these effects differ reliably from zero, on average, across participants. The MS_{error} that is produced using such a strategy effectively isolates the treatment effect when the F-ratio is computed (3).

(3) MS_{error} = [(Subject · Treatment error) + (Residual error)]

Among other things, this strategy avoids the biases inherent in the aggregation strategy, and it is less awkward to apply because the statistical analysis is tailored to the dataset rather than vice-versa.

Random Coefficient Analysis (RCA)

Lorch & Myers (1990) recommended the use of a simple RCA procedure for analyzing the type of multi-level data that is common in experimental research.[2] Within this context, the term random refers to the fact that RCA examines the effect of IVs on the dependent variable (DV) indirectly via the values of unstandardized beta coefficients that are sampled 'at random' from a probability distribution for each participant (for elaboration on this use of the term 'random', see Cohen et al, 2003, p. 544).

RCA is a two-step procedure for evaluating the reliability of effects in hierarchically structured designs that has been used only sporadically by cognitive scientists (e.g., Borowsky, Owen, & Masson, 2002; Chateau & Jared, 2003) despite the fact that it estimates the statistical significance of experimental effects more accurately than does aggregation. At step one, the analysis begins with the assessment of item-level effects within each participant (e.g. item-characteristics predicting subject responses). For step two, the statistical significance of the item-level effects across participants is assessed using standard tests like the single-sample t-test and possibly ANOVA/regression. In other words, the influence of subject-level variables can be evaluated at this point (e.g. individual differences like gender). The combination of steps 1 and 2 are essentially equivalent to a least-squares estimated hierarchical linear or multi-level model (Hox, 2002; Raudenbush & Bryk, 2002). Such analyses can be applied to many types of hierarchically structured data (e.g. students nested within schools; children nested within families), but only the special case of items nested within participants will be considered here. In what follows, steps 1 and 2 of the Lorch & Myers (1990) method for RCA are each described in turn.

The first step in performing an RCA involves performing an ordinary regression within each participant. This level of analysis can be considered the item-level or level-1. Each regression at the item-level involves the prediction of a DV (e.g. RT) on the basis of a set of predictors that can be item attributes or other types of variables (see Hox, 2002). These predictors can be both main effects and interactions involving categorical (e.g. ANOVA design; Cohen et al., 2003, pp.302-308) and/or continuous variables (pp. 255-300). The N regressions performed during the first step of a RCA yield N unstandardized beta coefficients for each item-level IV. These beta coefficients serve as data at the level of participants in the second step of RCA (level-2).

For the second step, the beta coefficients from step one can be used to answer at least two kinds of level-2 statistical questions (i.e. participant-level questions). First, a researcher might be interested in determining whether estimates of an item-level effect, which are represented by beta coefficients computed for each participant, are significantly different from zero for the sample on average. Alternatively, a researcher might be more interested in evaluating the effect of individual differences by, for example, comparing groups of participants to each other. Both types of comparisons are possible so long as the same type of regression analysis is performed within each participant (i.e. same predictor variables).

The first type of participant-level test is performed by comparing a collection of beta coefficients to the value of zero using a single-sample t-test. A statistically significant result indicates that, providing that the null hypothesis is true, the probability of observing an average beta coefficient as big as this or bigger for the sample of participants is less than the nominal alpha level. The second kind of participant-level test evaluates the relationship between participant variables (e.g. gender) and item-level parameters. To test whether a participant-level variable has a direct effect on the DV (main effect), its association with the item-level intercepts is evaluated. The intercept is useful here as it is a baseline value on the DV for each participant with which a predictor may or may not be associated. Of some use is the fact that, when predictors are centered (see below), regression equation intercepts can be interpreted as an unweighted mean for the participant on the DV. To test whether a participant-level variable interacts with an item-level predictor by modulating its effect on the DV, a regression predicting the item-level unstandardized beta coefficients is performed. The nature of the observed relationship depends on whether the participant-level variable is associated with an increase or a decrease in the absolute magnitude of the item-level beta coefficients (increasing or decreasing the strength of the effect) and whether the direction of the effect is reversed.

In summary, RCA represents the complexity of hierarchical datasets in a single, if multi-step, procedure without introducing the bias associated with aggregation. The advantages of RCA are many and they include (a) unbiased estimation of error variance that maintains the nominal Type-I error rate, (b) synthesis of 'by-subject' and 'by-item' analyses within a single statistical test without artificially inflating Type-II error (for a discussion of the limitations associated with other strategies for combining subject and item analyses, see Note 2; for a more detailed discussion, see Raaijmakers, 2003), (c) outputs that facilitate use of alternative forms of data presentation (e.g. confidence intervals for main effect and interaction slopes; for recommendations, see Loftus, 1996; for formula and other details, see Loftus & Masson, 1994; Masson & Loftus, 2003), and (d) the possibility of using continuous predictor variables so as to avoid the loss of power associated with imposing an artificial dichotomy on the predictor to satisfy the requirements of a by-subjects ANOVA with aggregated data (on the cost of dichotomization, see Cohen, 1983; Donner & Eliasziw, 1994; Hunter & Schmidt, 2004, p. 210). Finally, the procedure does not require major leaps in conceptual and mathematical understanding for the typical researcher because it is based on a straight-forward combination of statistical techniques that are covered in undergraduate-level statistics courses (regression, t-test).

The simplicity of RCA is advantageous as the strengths and limitations of the statistical procedures on which it is based are well-studied and familiar to most researchers. The responsible application of RCA involves, among other things, ensuring that the assumptions of regression and of t-test analyses are satisfied for the data to which they are applied. For example, the assumptions of regression must be verified within each participant to ensure the validity all inferences (for a detailed treatment see Cohen et al., 2003; for an introductory treatment see Tabachnick & Fidell, 2001). Similarly, care should be taken to plan studies that are likely to have sufficient power for detecting meaningful effects at each step in the analysis. Formulas for power calculation are widely available for regression and t-tests.[3]

When should RCA be applied? RCA is appropriate when data are hierarchically structured and the fundamental assumptions of regression are satisfied (for its application to the binary case, see Myers & Broyles, 2000). Many, if not most, experiments reported in the cognition literature that examine the effect of item-attribute variables across participants meet these criteria. Under certain circumstances, RCA may prove to be useful as a tool for verifying results obtained using an aggregation strategy (e.g. it is reported alongside conventional analyses by Chateau & Jared, 2003). In principle though, RCA can and should replace more commonly-used techniques like aggregation when it is appropriate (see above), unless the total number of participants is very small (e.g. N < 10), in which case a procedure known as the *fixed effects approach to clustering* should be employed (see Note 2).

More advanced procedures for estimating parameters (e.g. maximum likelihood) and adjusting parameters (e.g. Empirical Bayes estimation) within a random coefficient framework are available with sophisticated specialized programs like HLM (e.g. Raudenbush, Bryk, & Congdon, 2004) or MLwin (e.g. Rasbash et al., 2000), and also in SPSS. However, the simple RCA method described here and developed by Lorch & Myers (1990) is a viable alternative to aggregation for researchers without the background necessary for using more advanced techniques effectively (for readable introductions to such analyses see, Hox, 2002; Raudenbush & Bryk, 2002).

A Macro for RCA

It is possible to perform the RCA described above using the SPSS drop-down menus (i.e. with a mouse), but this procedure is time-consuming and prone to errors. A more efficient and reliable strategy is to run RCA analyses using a user-supplied program called a macro. In the appendix, the macro syntax for performing RCA as well as some syntax for executing the macros is presented. The functions performed by this syntax are described in what follows. In the final section, this macro is applied to some realistic data in the hopes of facilitating understanding of RCA in general and the macro in particular.

RCA in SPSS

To perform a RCA, the relationship between the DV (e.g. RT) and the IVs must be summarized by beta-weights for each participant. In the appendix, the macro named 'RCAsetup' performs this function. To use this macro, the raw data file must contain a variable that identifies each participant uniquely (i.e. an ID variable), one variable or column for each item-level independent variable (IV), and a variable/column for the dependent variable (DV). Essentially, the data file should be structured so that each row represents a single experimental trial. The datum (DV value) for a given trial is generated by a participant (Subject ID) in response to an item with particular properties (IV 1, IV 2, etc…). For example, the first trial in a data set may contain an RT of 566, a subject ID of 1, and values for the IVs of .5 (e.g. for 'high frequency') and -.5 (e.g. for 'low imageability'). If necessary, an IV representing the interaction between two IVs can be created by computing a variable representing the product of the two variables/vectors (for important cautions when using continuous IVs, such as the need to center predictors, see, Cohen et al., 2003, pp. 255-300).

If the data file is structured appropriately, executing the macro should generate a new data file containing one row with an intercept and a series of unstandardized beta-weights for each participant. The number of beta-weight variables in the new file should be equal to the number of first-order predictors (e.g. item-level variables like frequency or imageability). In the new file, the initial ID field is preserved but all other variables in the original file will be absent.[4]

The second macro tests the mean value of each beta-weight variable against 0 (i.e. a one-sample t-test). This macro generates an output containing descriptive statistics for the beta-weight variables (Mean, Standard Deviation, Standard Error), 95 percent confidence intervals for the average beta-weights, and summary statistics for a single-group t-tests, which is equivalent to a repeated-measures test of the differences between conditions with unbiased error terms. If an ANOVA or a Regression is desired using participant-level IVs, then it can be performed through the drop-down menus in the usual manner using the intercepts (CONST_) as a DV to test for a main effect, or using the beta-weights as DV to test for interactions between IVs (item-level by participant-level interaction).

To use these macros effectively, they should be executed in isolation. This can be accomplished by selecting the relevant syntax, and selecting 'run current' from the right-click menu. Executing this syntax loads the macros into SPSS memory. Once the macros are in memory, separate syntax must be provided for analyzing data with the macros (i.e. macro calls). This syntax must begin with the macro name and be followed by a list of variables to be included in the analysis. For the first macro, this list of variables must be provided in a specific order: ID variable name in rounded brackets, followed by the DV name in rounded brackets, and finally a list of item-level IVs in rounded brackets. An example of syntax for executing the macros is provided in the Appendix, but the variable names that correspond to those in your database must replace the default names. A description of how to analyze the dataset that is available for download with this article using the supplied macro syntax is presented in the following section.

An Example to Try with the RCA Macro

This section begins with a description of how to use the RCA syntax to analyze the data provided. A detailed description of how to interpret the results generated by this analysis is then undertaken, which is followed by suggestions for effect size estimation in RCA, and a comparison of the RCA results with the results of an analysis based on aggregated data.

Replication of the example reported here requires the use of two files that are available for download with this article. The first is an SPSS data file labeled *Thompson.sav*. The second is an SPSS syntax file labeled *Thompson.sps*. The data file contains four variables (columns): participant ID number, the DV (RT in milliseconds, ms, for a specific item) and the IVs Imageability (coded as - .5 = low-imageability, + .5

= high-imageability), Frequency (coded as - .5 = low-frequency, + .5 = high-frequency), and the interaction between the two, which was obtained by the following SPSS command:

*COMPUTE fxi = freq*imag.*
EXECUTE.

The data are structured so as to allow the regression equivalent to ANOVA to be performed within each participant. The logic behind using the values .5 and -.5 to denote membership within levels of the IVs is explained in the section below entitled 'Coding issues'.

To perform Step one of RCA on these data, both the data file and the syntax file identified above must be opened in SPSS using *File/Open/Data* and *File/Open/Syntax* from the drop-down menu. Begin by examining the syntax file. It contains two types of lines: those that begin with the character * are dedicated to comments explaining the syntax, which are ignored by SPSS; those that do not begin with this character contain active syntax that is interpreted by SPSS when executed. Initially, the RCA macros must be loaded into memory. To do this, select the block of text containing the macro syntax, right-click the mouse, and then click the 'run current' option. The macro is now available to be called upon like any other syntax command. To call the first macro into action, select the syntax beginning with the word 'RCAsetup'. Executing this syntax causes SPSS to perform a standard regression analysis within each participant individually, and then to create and open a data file named *betas.sav* containing a row of unstandardized beta coefficients for each participant. Examine the output file that is generated for error messages and then close the output file without saving. If there are no problems, execute the next line of syntax to call the second macro into action and test whether the item-level effects are significantly different from zero for the sample of participants (e.g. for frequency, imageability, and frequency by imageability).

Coding the predictors. In this example, we are analyzing ANOVA type data (categorical predictors, continuous DV) using a regression approach. In order to produce meaningful unstandardized beta coefficients, the predictors, in this case frequency and imageability, must be given appropriate values. Cohen and colleagues (2003) suggest a number of methods for coding categorical variables that produce equivalent overall regression equations, but different unstandardized beta coefficients. Arguably the simplest of these strategies is to "dummy code" the IVs assigning the value of 0 to one group and the value of 1 to the other.[5] However, in most cases it is desirable to center predictors so that 0 represents the average value of each predictor. Centering predictors prior to running regression produces beta coefficients that represent the effect of a predictor averaged over levels of the other predictors included in the analysis, which is useful since that is precisely the type of effect that an F-test in an factorial ANOVA table evaluates. Similarly, centering predictors causes the intercept to be equal to the value of the DV when all predictors are average (i.e. 0), which makes it the unweighted participant mean on the DV across all predictors.

In the present case, the item-level IVs are centered round the value zero because the low-frequency and high-imageability items are coded as -.5 while the high-frequency and high-imageability items are coded as +.5. We use the absolute value of .5 so that the difference between groups is equal to 1 [.5 – (-.5) = 1], which is important to ensure that the unstandardized beta coefficients is easy to interpret. If the difference between codes was a value other than 1, the beta coefficient would not represent the average difference between levels of a main effect. This is true because beta coefficients represent the average increase in DV associated with a 1-unit increase in the predictor.

Interpreting the results. The *Thompson.sav* file contains data from 64 participants (mixed condition; Thompson & Desrochers, 2003) that were modified slightly by adding a small non-zero value that was sampled from a normal distribution to each observation. The original data were taken from an experiment examining the influence on lexical decision performance of lexical frequency (i.e. the frequency of occurrence of a word in a corpus of text) and imageability (i.e. the ease with which participants evoke a mental image in response to a word). Visual lexical decision is a task that requires participants to discriminate between real words and nonsense words that are presented one at a time on a computer screen by button press. The DV in the data file is reaction time (RT) in milliseconds (ms). Each participant made twenty-five responses to words per experimental condition (e.g. 25 highly imageable words of low-frequency, 25 highly imageable words of high-frequency, etc...) for a total of 100 observations per participants minus the data for incorrect responses, which were discarded prior to analysis. Participants made an equal number of responses to nonsense words and these were also discarded.

As noted above, regression analyses were performed separately for each participant (i.e. step one). Interpretation of an average beta coefficient for a sample of participants is similar to the interpretation of beta coefficients generated by a more conventional analysis. If a beta coefficient is interpreted as the average x-unit increase in the DV associated with a 1-unit increase in the IV, then the average beta for a sample of participants is interpreted as the mean average x-unit increase in the DV associated with a 1-unit increased in the IV for the sample. From an ANOVA perspective, an average beta is simply the average mean difference between conditions for the sample.

For the present example, interpretation of the average beta coefficients is relatively easy. Because the main effects discussed here only have two levels (coded -.5, +.5), an average beta coefficient is equal to the average difference in ms across participants between the two conditions and also the average effect of the IV in ms. If the procedure described above was executed correctly, the results should indicate that the average beta coefficients for the Frequency effect, the Imageability effect, and their interaction are -129.13, -53.33, and 84.43 respectively. The single-sample t-tests indicate that all three effects are significantly different from zero. Thus, we have observed statistically significant main effects for frequency and imageability, and a significant interaction between the two. These significant effects are interpreted as follows. The direction of the frequency effect (negative) indicates that high frequency words (coded +.5) are read aloud 129.13 ms (Standard Error = 9.27) more rapidly than low-frequency words (coded -.5), $t(63) = 13.93, p < .001$. The direction of the Imageability effect (negative) indicates that high-imageability words (coded +.5) are recognized 53.33 ms (Standard Error = 7.13) faster than low-imageability words (coded -.5), $t(63) = -7.48, p < .001$. These two main effects are qualified by a statistically significant interaction, average unstandardized beta = 84.43 (Standard Error = 14.34), $t(63) = 5.89, p < .001$. The signs of the main effects (both negative) and the interaction (positive) indicate that the effect of one IV is reduced as the value of the other increases. Decomposing the interaction so that it can be fully interpreted requires a bit more work.

Simple effects testing. Statistically significant interactions are the statistical justification for examining the statistical significance of one IV within levels of another. The reason for this is clear in RCA as the unstandardized beta coefficient for the interaction is literally the average difference between the simple effects (or simple slopes) of one IV across levels of the other for the sample. Because the difference between simple effects is significant here, we know that the effect of imageability is statistically different depending on the frequency of the associated words. We now might want to determine more precisely what the nature of the imageability effect is within levels of frequency. For example, is the effect of imageability reversed as frequency-level changes? It is possible to answer this question using the simple slope tests that are available in the literature (Aiken & West, 1991; Cohen et al, 2003). However, an easier way to perform simple effects testing with categorical predictors is to re-run the analysis that tests the effect of only one of the IVs (e.g. imageability) twice: once using only low-frequency words and once using only high-frequency words.

To perform simple effects testing with the example data, open the original data file (*Thompson.sav*) and then execute the first block of simple effects syntax. Then, open the original data file again (without saving the version that is already open) and execute the second block of simple effects syntax (see *Thompson.sps*). This procedure will re-execute the original analysis twice, once with high-frequency words and once with low-frequency words. If performed correctly, the results should indicate that the effect of imageability is statistically significant for low-frequency words only. The average beta value for the imageability effect within the low-frequency condition is -95.54 ms (Standard Error = 11.72), $t(63) = 8.15$, $p < .001$. In contrast, the average unstandardized beta coefficient for high-frequency words is only -11.11 ms (Standard Error = 8.19), $t(63) = 1.36$, $p = .18$. Examination of the 95 % (within-subject) confidence interval for this non-significant difference, which is provided in the output, indicates that the data are consistent with both a relatively large high-imageability advantage over the low-imageability condition (lower-bound for the difference between conditions: -32.88 ms) and a small effect that is of about the same magnitude as the observed difference in the opposite direction (upper-bound for the difference between conditions: + 10.65 ms). In other words, the evidence for an effect of imageability with high-frequency words, in either direction, is weak to say the least. Note that the difference between the simple slopes is equal to the average beta coefficient for the interaction that was obtained in the overall analysis, $-95.54 - (-11.11) = 84.43$.

Calculating effect-size in RCA. Lorch & Myers (1990) did not recommend an estimate of effect size for the RCA technique described here. Reaction time (e.g. in ms) is a DV that has an intuitive meaning and therefore standardized measures of effect size are less relevant than they otherwise might be. However, if standardized estimates of effect size are desired, it is possible to compute an appropriate within-subjects Cohen's d (d_{Cohen}) by dividing the average beta coefficient value for the sample of participants, which can be considered an average difference between conditions, by its standard deviation, which can be considered the standard deviation for the difference between conditions (for more details, see Cohen, 1988; Howell, 2002, pp. 235-236). This estimate of effect-size can be conceived as an expression of the experimental effect (absolute difference between conditions) in standard deviation units (i.e. a standardized effect or standardized difference). For more information on d_{Cohen} consult the sources noted above or a good textbook. To explore the implications of different effect sizes on things like distribution overlap try the program g*power 3, which is freely available for download (Faul, Erdfelder, & Buchner, in press).

Readers that prefer thinking about effect size in terms of percent of variance explained can calculate partial-η^2, which is the estimate of observed effect size produced by SPSS for repeated-measure designs, for RCA using sum of squares that are calculated in the manner described below. For any given IV, partial-η^2 is the proportion of variance in the DV left unexplained by the other IVs in the analysis that is accounted for by that variable (analogous to a partial correlation, or more accurately partial-r^2). Using within-subject SS, the formula for this calculation is: partial-$\eta^2 = SS_{effect} / (SS_{effect} + SS_{error})$. The d_{Cohen} and partial-η^2 values for the RCA along with those obtained by analyzing the aggregated data are reported in Table 1.

Practical Importance of Aggregation Bias

At this point, one might wonder what the importance of aggregation bias is in practice. Lorch & Myers (1990) examined this issue through simulations, but empirical study of this issue in experimental psychology (esp. cognitive psychology) is hard to find. The purpose of this section is to provide two types of evidence designed to convince skeptical readers of the practical importance of aggregation bias. First, the results of the Lorch & Myers simulations are briefly reviewed. These results demonstrate that in principle the magnitude of aggregation bias can be quite large. Then, researchers are provided with a method to estimate the potential bias in their own data. By providing researchers with this method, it is hoped that the problem presented by aggregation bias will be more difficult to ignore. The method is demonstrated using the data described above, but this demonstration is not intended as a general test of the practical importance of aggregation bias nor should it be interpreted as such.

Review of Simulation Results

Lorch & Myers (1990) demonstrated that the magnitude of aggregation bias, which they operationalized as Type-I error inflation in their simulations, depends on at least three factors that seem to interact synergistically. In other words, the effect of one of these factors is magnified as the magnitude of the others increases. The first and most obvious factor is the variance of between-condition differences in the population (i.e. the population variance for the average beta coefficients computed in RCA). The larger the variance is in the population, the greater the potential magnitude of aggregation bias. We can extend the implications of this finding a bit by calling upon what is known about sampling error, which causes over- or under-estimation of the population variance from sample to sample. Because the magnitude of aggregation bias depends on this variance, it should also vary from sample to sample even if all other factors are kept constant. These fluctuations will be especially important when sample sizes (number of participants) are small (see the law of large numbers; central limit theorem). The two other factors examined by Lorch & Myers (1990) were (a) the number of items per participant and (b) the inter-item correlation within experimental conditions. Both factors are positively associated with aggregation bias, which is a somewhat counter-intuitive finding in that both factors are positively associated with a score's — in this case a mean's — reliability. In fact, 'improvement' on these two parameters can actually degrade the quality of the analysis in situations where RCA is appropriate and aggregation is used instead. According to the results reported by Lorch & Myers (1990), even with a conservative estimate of variability (.25), inter-item correlation (.20) and the use of only 10 items per participant, aggregation bias can inflate the nominal alpha level from .05 to .79 (their Table 2)! Again, sample to sample fluctuations in reliability (for a discussion of this issue, see Thompson & Vacha-Haase, 2000) could cause additional variation in the extent of aggregation bias across experiments, even those with identical methodologies.

In addition to the factors examined by Lorch & Myers, many other methodological factors (e.g. the magnitude of real experimental effects, the number of participants, design complexity) could directly or indirectly determine the extent of bias. Though we know some things about what tends to increase or decrease the amount of potential bias in a given dataset, it is currently impossible to know what the practical importance of aggregation bias will be for any given experiment. Ultimately, it is best to avoid aggregation bias altogether by using a bias-free technique like RCA.

Assessing the Magnitude of Aggregation Bias with Real Data

The demonstration provided by Lorch & Myers (1990) should be sufficient to convince researchers of the practical importance of aggregation bias. Nevertheless, researchers may be more motivated to change their ways if they are able to evaluate the extent of potential aggregation bias within their own data. Further, in the absence of formal meta-analyses examining the issue, informal tests of aggregation bias that are conducted by researchers could increase awareness of the problem and its potential magnitude for specific types of research. For these reasons, a method is demonstrated for comparing RCA results to those produced by analyzing aggregated data. A secondary benefit of providing this demonstration is that it requires the conversion of RCA statistics into a form that is familiar to many researchers: the ANOVA summary table. Among other things, this transformation allows the easy computation of effect-size estimates like partial-η^2.

To perform the comparison, two types of values were obtained: (a) ANOVA and effect size estimates from an analysis using the aggregation method and (b) ANOVA and effect size estimates computed based on the RCA statistics from the example reported above. The results for the analysis of aggregated data were obtained by running the syntax in the file *Thompson_aggr.sps* with the following data file open: *Thompson_aggr.sav*. This data file was produced by transforming the original (*Thompson.sav*) using the *aggregation* and the *restructure* functions in SPSS. In contrast, the standard ANOVA counterparts to the RCA statistics reported above were obtained through calculations that were made without the use of SPSS. This conversion process must be explained in detail to be easily understood and replicated. First, the sum of squares effect (SS_{effect}) and the sum of squares error (SS_{error}) are computed based on the unstandardized beta values and their respective standard deviations. To compute the SS_{effect}, the unstandardized beta coefficients is squared and then multiplied by N. The exception is the interaction coefficient, which is divided by two before squaring.[6] To obtain the SS_{error} for a given effect, the standard deviation of its coefficient is squared and then multiplied by N. Again, the exception is the interaction SS_{error}, which requires that the standard deviation be divided by two before squaring. The other F-test statistics are derived from the resulting sums of squares values in the usual manner. First, each sum of squares value is divided by its degree(s) of freedom to produce corresponding mean-square values. Second, F-ratios are computed by dividing each MS_{effect} estimate by its associated MS_{error} term. The ANOVA statistics for the RCA and Aggregation analysis are reported for easy comparison in Table 1.

The first thing to note about Table 1 is that the MS_{effect} is identical for both analytical strategies. The hand calculation of the RCA values introduced a little rounding error, but otherwise the comparison is consistent with formal demonstrations that RCA and aggregation produce equivalent estimates of the absolute magnitude of treatment effects (i.e. estimates of means and therefore differences between means; Lorch & Myers, 1990). The second thing to note about the table is the difference between the MS_{error} values that are produced by the two strategies. For the main effects and the interaction, the error term is larger in the RCA than it is in the aggregation analysis. This result was expected because RCA produces relatively unbiased error estimates (Lorch & Myers, 1990), which are larger and therefore result in standardized effect-size measures that are smaller than those obtained from aggregated data (e.g. d_{Cohen}).

It is clear that aggregation bias is present in the data. The magnitude of this bias can be examined through standardized statistics like the F values (same df throughout) and the effect size estimates. Note that in Table 1 the differences between the F-ratios across aggregation and RCA strategies range from .54 to 3.04. Given the overall size of the experimental effects, these differences might be considered small. The effect size estimates would seem to support that interpretation because the observed aggregation bias is limited to about 1 percent of partial variance explained (frequency effect, the interaction). Nevertheless, the aggregation bias observed here could mean the difference between a significant result and a null result with a smaller sample size or when examining less powerful effects, which is important if only because statistical significance plays a role in determining whether a study is published.

Reporting RCA Results

At present, there are no official norms for reporting the result of RCA. In principle, the results can be reported as a regression analysis or as an ANOVA analysis as long as use of the technique is acknowledged. A regression analysis can vary in complexity depending on the number of predictors involved. When the number of item-level predictors is large, the results of the analysis can be reported in a table as with any other regression analysis, with proper acknowledgement that the results were obtained using Lorch & Myer's method (for an example, see Chateau & Jared, 2003). If the regression analysis is simply an implementation of an ANOVA design (3 or fewer IVs), then its results can be reported within the body of the text in a manner similar to that used to report ANOVA analyses in the literature (see the example below).

Interestingly, the output generated by the RCA macro presented here lends itself well to reporting within-subject confidence intervals of different types as well as within-subject d_{Cohen} effect sizes (for a critique of the dominant hypothesis testing philosophy, see Loftus, 1996). Within-subject confidence intervals (95%) for main effects and interactions are provided automatically in the output (i.e. for the average beta coefficients). Further, within-subject confidence intervals can be computed for individual cell means by first transforming the standard error for the average beta coefficients (i.e. difference between means) into the standard errors for the means themselves (Note: the value will be the same for both means) by the following formula: $SE_{mean} = SE_{difference} / \sqrt{2}$ (for additional formula and recommendations for calculating confidence intervals for various designs and comparisons, see Loftus & Masson, 1994; Masson & Loftus, 2003).

The italicized paragraphs that follow demonstrate how the RCA results that are discussed above might be reported in-text within the results section of an empirical article. Estimates of effect size are not reported directly, but d_{Cohen}, for example, can be computed using the reported information in the manner described above (simply convert the standard error into a standard deviation, $SD = SE \cdot \sqrt{n}$. Similarly, the average beta coefficients are not reported because they can be derived from the reported t-values and their standard errors ($b = t \cdot SE$) or from the difference between the reported means. The example is intended to reflect a style that is typical of articles reporting repeated-measures ANOVA results in the field of cognitive psychology.

> *The frequency by imageability (2 x 2) design was analyzed using random coefficient analysis (Lorch & Myers, 1990). Random coefficient analysis (RCA) is a multi-level regression technique that produces unbiased error term estimates, unlike ANOVAs based on aggregated data. Within RCA, the magnitude of an experimental effect is first estimated within each participant and then the hypothesis that these within-subject effects are significantly different from zero for the sample is tested. The tests of main effects and interaction are reported as one-sample t-tests, which in this case are equivalent to correlated t-tests of the difference between conditions, because that is how such effects are evaluated in RCA.*
>
> *The results indicate that the latency advantage for high-frequency words over low-frequency words (578.42 vs 707.55 ms) is significant for the sample of participants, $t(63) = 13.93$, SE difference = 9.27, $p < .001$. Similarly, the advantage of high-imageability words over low-imageability words (616.32 vs*

669.65 ms) is significant across participants, t (63) = -7.48, SE difference = 7.13, p < .001. These two main effects are qualified by a statistically significant interaction, t (63) = 5.89, SE difference = 14.34, p < .001. Decomposition of the interaction confirmed that high-imageability words were associated with faster responses than the low-imageability words (659.78 vs 755.33 ms), t (63) = 8.15, SE difference = 11.72, p < .001. In contrast, the effect of imageability fell short of significance for high-frequency words (572.86 vs 583.78 ms), t (63) = 1.36, SE difference = 8.19, p = .18. The 95 percent confidence interval for this non-significant difference indicates that the data are consistent with both a relatively large high-imageability word advantage over the low-imageability condition (lower-bound difference: -32.88 ms) and a small effect that is of about the same magnitude as the observed difference (i.e. -11.72), but in the opposite direction (upper-bound difference: + 10.65 ms). In other words, the results do not support the idea that imageability exerts a meaningful effect on reaction times when words are also high-frequency.

Conclusion

The numerous flaws of the aggregation strategy that is widely applied by cognitive psychologists and experimental psychologists more generally were reviewed. In its stead, it has been proposed that a procedure sometimes referred to as random coefficient analysis should be used to test the effect of item-attribute variables (Lorch & Myers, 1990). The simple RCA procedure proposed by Lorch & Myers was described in general terms, and then an easy to use program for performing random coefficient analysis in SPSS (version 11 or better) was presented. The operation of this program, called a macro, was explained in terms of a concrete example using data that is available for download with this article. The results obtained from analysis of this data were interpreted in detail and suggestions were offered for reporting such results in empirical articles. Finally, a method for evaluating the potential magnitude of aggregation bias for any dataset was presented so that researchers can better appreciate the consequences of choosing to report analyses based on aggregated data.

In closing, RCA can and should be applied in most cases when a repeated measures design is used to examine item-attribute effects. More generally, RCA is preferable to aggregation whenever multi-level data are involved (Hox, 2002). Whether a traditional hypothesis testing approach is adopted or more informative confidence intervals are used, it is important that error variance be accurately estimated because otherwise the validity of effect size estimation (e.g. power- and meta-analyses) and hypothesis testing are negatively affected. In a discipline where the difference between $p = .03$ and $p = .06$ can mean the difference between a manuscript's publication and its rejection, the use of strategies like aggregation that are known to bias error estimation and therefore p-values is hard to justify, no matter what the practical importance of such bias may be.

Program Availability

A syntax file containing both macros is available directly from the author (GlennLThompson@gmail.com) or online (http://www.geocities.com/glennleothompson /OriginalCode.html). The code may also be typed into an SPSS syntax editor from the Appendix or downloaded from the journal's website, where the data reported in this article may also be found.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Borowsky, R., Owen, W., & Masson, M. (2002). Diagnostics of phonological lexical processing: Pseudohomophone naming advantages, disadvantages, and base-word frequency effects. *Memory & Cognition, 30*, 969-987.

Campbell, J.D. & Thompson, V.A. (2002). More power to you: Simple power calculations for treatment effects with one degree of freedom. *Behavior Research Methods, Instruments, & Computers, 34*, 332-337.

Chateau, D., & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language, 48*, 255-280.

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

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlational analysis for the behavioral sciences.* (Third ed.) New Jersey: Lawrence Erlbaum Associates.

Donner, A. & Eliasziw, M. (1994). Statistical implications of the choice between a dichotomous or continuous trait in studies of inter-observer agreement. *Biometrics, 50*, 550-555.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (in press). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods.*

Green, S. B. (1991). How many subjects does it take to do a regression analysis?

Multivariate Behavioral Research, 26, 449-510.

Howell, D.C. (2002). *Statistical Methods for Psychology, Fifth Edition.* Pacific Grove, CA: Wadsworth.

Hox, J. (2002). *Multi-level analysis: Techniques and Applications.* London: Lawrence Erlbaum.

Hunter, J.E. & Schmidt, F. L. (2004). *Methods of Meta-Analysis, Second Edition.* London: Sage Publications.

Levine, J. (1997). Overcoming feelings of powerlessness in "Aging" researchers: A primer on statistical power in analysis of variance designs. *Psychology and Aging, 12*, 84-106.

Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science*, 161-171.

Loftus, G.R. & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476-490.

Lorch, R. & Myers, J. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 149-157.

Masson, M.E. & Loftus, G. R. (2003). Using confidence intervals for graphically-based data interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203-220.

Myers, L. & Broyles, S. (2000). Regression coefficient analysis for correlated binomial outcomes. *Journal of Applied Statistics, 27*, 217-234.

Raaijmakers, J. (2003). A further look at the "Language-as-Fixed-Effect Fallacy". *Canadian Journal of Experimental Psychology, 57*, 141-151.

Rasbash, J., Browne, W., Healy, M., Cameron, B., & Charlton, C. (2000). MLwin. Multilevel Models Project Institute of Education.

Raudenbush, S. & Bryk, A. (2002). *Hierarchical linear models.* Thousand Oaks: Sage Publications.

Raudenbush, S., Bryk, A., & Congdon, R., (2004). HLM for Windows (6.00). HLM Software.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*, 351-357.

Tabachnick, B., & Fidell, L. (2001). *Using Multivariate Statistics, Fourth Edition.* Boston: Allyn and Bacon.

Thalheimer, W., & Cook, S. (2002, August). *How to calculate effect sizes from published research articles: A simplified methodology.* Retrieved January 30th, 2007 from http://work-learning.com/effect_sizes.htm.

Thompson, G. & Desrochers, A. (2003, June). The role of word imageability in visual word recognition. Poster presented at the 13th Annual Meeting of the Canadian Society for Brain, Behaviour and Cognitive Science (BBCS), Hamilton, Ontario, Canada.

Thompson, B. & Vacha-Haase, T. (2000). Psychometrics is datametrics: the test is not reliable. *Educational and Psychological Measurement, 60*, 174-195.

Author Note

Glenn L. Thompson, School of Psychology, University of Ottawa, Ottawa, Ontario, Canada.

This work was supported by two Natural Sciences and Engineering Council of Canada (NSERC) scholarships. I would like to thank my doctoral supervisor Alain Desrochers for agreeing to let me use some of our data here in slightly modified form as an example. I would also like to thank two anonymous reviewers for their useful comments on a previous draft of this paper. Finally, I would like to thank Denis Cousineau for his helpful suggestions concerning the text and the macro code.

Correspondence concerning this article should be addressed to Glenn Thompson, School of Psychology, University of Ottawa, 145 Jean-Jacques Lussier, Box 450, Station A, Ottawa, Canada, K1N 6N5 or e-mail (GlennLThompson@gmail.com).

Notes

[1] Raaijmakers (2003) recommends analyzing aggregated data by-subjects or by-items, but not both unless there is a special reason for doing so. Statistics that are available for combining the F-tests produced by both analyses are too conservative in most cases (i.e. they inflate Type-II error rates). The RCA method presented here renders the point moot, however, as it effectively combines both by-subjects and by-items analyses.

[2] An analytical strategy that accomplishes the same thing as RCA is called the *fixed-effects approach to clustering* (Snijders & Boskers, 1999, as cited in Cohen et al, 2003, p. 541; Presented as an alternative way of implementing RCA in Lorch & Myers, 1990). This strategy requires that the raw data file (as defined in the body of the text) be analyzed with a single regression equation. The experimental effect and error variance (e.g. participant by treatment interactions) are disentangled by entering a series of dummy variables in the regression equation along with the predictor variables. The simple between-participant differences are statistically controlled by coding the identity of each participant using N-1 dummy variables. These variables remove the same variance that is associated with the main effect of participants (i.e. the individual differences) in a standard repeated-measures analysis. To control for the error variance associated with aggregation bias, the product of each dummy variable with the predictor variable(s) is entered into the regression equation (i.e. the participant by treatment interaction).

While this strategy and the RCA method presented in the body of this paper accomplish the same goal, there are reasons for choosing one over the other as circumstances dictate (Cohen et al, 2003, pp. 565-566). The fixed-effects approach is appropriate when the cluster has a substantive meaning, but less useful for cases where the grouping is simply a random sample across which a researcher wishes to generalize experimental effects. For example, a social-psychology project where participants are clustered within ethnic neighborhood may be an appropriate case. In the case of item-attribute effects, the participants are not meaningful per se except in so far as they are associated with participant-attributes (e.g. gender). Therefore, the type of research discussed here (item-attribute effects within and across participants) is more appropriately submitted to an RCA. The only exception is the case where there are fewer than 10 participants, in which case the fixed-effects method is recommended.

[3] Green (1991) proposed formula for estimating power and sample size for regression analyses. To ensure adequate power for the item-level coefficients (e.g. minimum of .80), the following formula can be applied for estimating N: $N \geq (8 / f^2) + (m - 1)$, where N is the number of observations required, m is the number of beta coefficients to be estimated, f^2 is equal to $r^2 / (1 - r^2)$, and r^2 is the expected effect size (ω^2 or another adjusted "percent explained" effect-size estimator may be used in place of r^2). For t-tests, Campbell and Thompson (2002) present a simple technique for computing effect sizes, power, and required N to achieve a given level of power for tests with 1 degree of freedom (for other cases, see Levine, 1997). First, the expected effect size (Cohen's d) is estimated based on past research (See Thalheimer & Cook, 2002, for many simple formulas for calculating observed d in published research using commonly reported statistics like MS_{error}) or the values proposed by Cohen (1988) for small (.20), medium (.50), and large (.80) effects. For both regression and t-tests, special circumstances that would reduce power like non-normal distributions within participants or unreliable measures would require that the sample size be increased to maintain the stated level of power (Tabachnick & Fidell, 2001, p. 117).

[4] If applicable, participant-level variables may be recovered from the original file by using the Data/Merge Files/Add Variables from the drop-down windows. From the resulting window, simply select the original data file and choose to import the participant-level variable of interest. The two files can be matched using the participant ID field. Once participant-level variables are in the same file as the unstandardized beta coefficients, level-2 hypotheses can be tested in the usual way using intercepts as DV to test main effects and the average beta coefficients as DV to test for interactions (see text).

[5] If there are more that two levels (*g*) per predictor, the categorical variable is represented by g-1 dummy variables (e.g. 0 vs 1) or contrast code variables (.5 vs -.5). For details on how to devise coding schemes for regressions with categorical independent variables (especially orthogonal coding schemes), see Cohen et al (2003), Tabachnick & Fidell (2001, pp. 149-150), or another good statistics textbook.

[6] Without getting into too many details, we have to divide the interaction beta coefficient by two because the difference between .025 and -.025 is .5 rather than 1. To be able to compute a SS_{effect} that is comparable to (a) the two main effects and (b) the estimate produced by an ANOVA using aggregated data, the 'squared deviations' must be put on the same basis. Since beta coefficients reflect the average increase in the DV associated with a 1-unit increase in x, the scale of the interaction coefficient is effectively double that of the effect used to calculate the SS_{effect} for the aggregation ANOVA. Dividing the interaction coefficient by two puts all effects on the same basis again.

Table 1

Magnitude of Aggregation Bias: Comparing ANOVA Statistics and Effect Size Estimates for the Aggregation and RCA Strategies

Strategy		MS_{effect}	SS_{error}	MS_{error}	F-ratio	d_{Cohen}	Partial η^2
Aggregation (A)							
	Frequency (F)	1067243.74	346104.96	5493.73	194.27	-1.76	0.76
	Imageability (I)	182013.45	205091.60	3255.42	55.91	-0.94	0.47
	F x I	114060.36	207184.95	3288.65	34.68	1.48	0.36
RCA							
	F	1067242.72	351601.56	5580.98	191.23	-1.74	0.75
	I	182013.50	208300.96	3306.36	55.05	-0.94	0.47
	F x I	114060.47	210460.74	3340.65	34.14	1.47	0.35
Bias (A – RCA)							
	F	1.02	-5496.60	-87.25	3.04	-0.02	0.01
	I	-0.05	-3209.36	-50.94	0.86	0.00	0.00
	F x I	-0.11	-3275.78	-52.00	0.54	0.01	0.01

Note. The observed (retrospective) power for all tests is effectively 1. MS_{effect} is equal to SS_{effect} because all treatment effects have a single degree of freedom. Partial η^2 reflects the proportion of variance explained by a predictor after between-subject variance and the variance attributable to the other predictors has been removed. The symbol d_{Cohen} is an estimate of effect size that expresses the mean difference between conditions in standard deviation units. All A and RCA effects in the example data are statistically significant for degrees of freedom (1, 63) at $p < .01$.

Appendix

This appendix contains SPSS syntax (version 11 or later) for the two macros that were discussed within the main body of the text (see the downloadable syntax file named: *Thompson.sps*). The macro named 'RCAsetup' performs step one of RCA, which involves running regression analyses within each participant and saving the resulting intercepts and unstandardized beta coefficients to a new datafile. The macro named 'RCAtest' performs step two, which involves testing whether the unstandardized beta coefficients are significantly different from zero for the sample of participants. In what follows, each block of syntax (italicized) is presented within its own section and is preceded by a short description.

Macro for Step One: RCAsetup

The macro called RCAsetup accepts three variable-name parameters as input. The first is the name of the participant identification variable. The second is the name of the dependent variable. The third is a list of the names of the independent variables. If it is executed with an appropriately structured data file open, the program creates a data file called *betas.sav* containing the participant identification variable and an intercept as well as unstandardized beta coefficient(s) for each participant. A commented version of the syntax is available in the downloadable SPSS syntax file. Example syntax for the macro is provided in the final section of the appendix, but the default variable names must be replaced.

```
DEFINE RCAsetup (!positional !enclose ( '(', ')' )/
          !positional !enclose ( '(',')' )/
          !positional !enclose ( '(',')' )).

SORT CASES BY !1 .
SPLIT FILE BY !1 .
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA
```

```
 /NOORIGIN
 /DEPENDENT !2
 /METHOD=ENTER !3
 /OUTFILE=COVB('C:\temp1.sav') .
SPLIT FILE OFF.

GET FILE = 'c:\temp1.sav'.
SELECT IF (rowtype_ = 'EST').
SAVE OUTFILE='C:\temp2.sav'
 /DROP=DEPVAR_ ROWTYPE_ VARNAME_
 /COMPRESSED.

GET FILE = 'C:\temp2.sav'.
EXECUTE.
SORT CASES BY !1 .
CASESTOVARS
 /ID = !1
 /GROUPBY = VARIABLE .
SAVE OUTFILE='C:\betas.sav'.
GET FILE = 'C:\betas.sav'.
EXECUTE.

ERASE FILE= 'c :\temp1.sav'.
ERASE FILE= 'c:\temp2.sav'.
EXECUTE.
!ENDDEFINE.
```

Macro for Step Two: RCAtest

The macro called RCAtest accepts a single variable-name input parameter: a list of the item-level independent variables. If it is executed with the file called *betas.sav* open, it produces as output a t-test for each variable that evaluates the hypothesis that the associated effect is statistically significant for the sample of participants. A commented version of the syntax is available in the downloadable SPSS syntax file. Example syntax for the macro is provided in the following section, but the default variable names must be replaced with more appropriate ones.

```
DEFINE RCAtest (!positional !enclose ( '(', ')' )).
T-TEST
 /TESTVAL = 0
 /MISSING = ANALYSIS
 /VARIABLES = !1
```

```
 /CRITERIA = CI(.95) .
!ENDDEFINE.
```

Calling the Macros: Example syntax

What follows is example syntax for calling the macros reported above. In order for the syntax to work, the macros they refer to (RCAsetup, RCAtest) must have been previously loaded into memory. Macros are loaded into memory by selecting the associated syntax and executing it. To adapt the example below to a particular case, simply replace the default variable names (ID, DV, IV1, and IV2) with appropriate variables from the dataset to be analyzed.

```
RCAsetup (ID) (DV) (IV1 IV2).
RCAtest (IV1 IV2).
```

Chapter 4

Sensitivity to Homophone Interference as Correlate of Syllable Preference in Visual Word Recognition

Abstract
Taft (2001) found that preference for one method of syllabifying words over another, namely segmentation by maximal-onset (Max-Onset) or maximal-coda (Max-Coda) rule, depends on reading skill. A limitation of this finding is that 'skill' has no clear referent in current computational models of reading. With respect to resolving issues like the locus of syllable effects, process-based variables, like sensitivity to feedback phonology, are potentially more informative. Accordingly, the results of three lexical decision experiments are reported (3 x N = 48) that associate syllable preference in a word-splitting paradigm with the magnitude of the homophone interference and regularity effects. Participants who were sensitive to homophone interference tended to respond faster when low-frequency words were divided at the Max-Onset boundary. In contrast, phonologically insensitive participants preferred the Max-Coda boundary with some low-frequency words when the background contrasted poorly with the word-initial syllables. The results suggest a developmental shift away from a phonology-dominated processing strategy towards one that is more heavily reliant on orthography. The observed individual differences could be accounted for by assuming that this shift is in various stages of completion within the skilled-reading population.

Introduction

Extant theoretical models of the visual word recognition system have been developed to account for the reading and recognition of 2, 800 monosyllabic words in English (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 2004). The theoretical focus then has been on a relatively small subset of all words: monosyllables, which comprise only ten percent of an adult's vocabulary (Nation & Waring, 1997). If computational models of lexical processing are to be extended to account for the remaining ninety percent of words, it is critical that the status of sublexical functional units like the syllable be resolved (for discussion, see Rastle & Coltheart, 2000; Coltheart et al., 2001; Chateau & Jared, 2003; Harm & Seidenberg, 2004). Regrettably, multi-syllable words are chronically under-studied (for discussion, see Chateau & Jared, 2003) and direct examination of the syllable as a functional unit has yielded inconsistent results (for reviews, see Chen & Vaid, 2007; Chetail & Mathey, 2009; Taft, 2001).

Phonology: Simple or multi-layered?

A commonly held belief in the word recognition literature is that the minimality constraint (i.e., the necessity of uniquely identifying words) tends to produce phonological representations that are articulated at the level of the phoneme (Frost, 1998), which is the minimally contrastive unit of the English language. On the other hand, linguistic theory ascribes to phonology a hierarchical, multi-layered and multi-dimensional representation, which has been described as consisting of three levels: the melodic, the skeletal, and the prosodic (Bybee, 2001; Clements & Keyser, 1983; Selkirk, 1982). According to this conception, phoneme identities are specified at the melodic level, consonant and vowel identities are specified at the skeletal level, and the syllable is presumed to reside at the prosodic level of representation. Current computational theories of reading and word recognition represent either the melodic level (Coltheart et al., 2001) or both the melodic and skeletal level (Plaut, McClelland, Seidenberg, & Patterson, 1996). Prosody, however, is thought to be a property of multi-syllable words, and so it has escaped the attention of monosyllabic models.

Prosody or 'What is a syllable?'

While the syllable is a fundamental unit for the description of phonology in the field of linguistics (Blevins, 1995; Hooper, 1972; Selkirk, 1982), its status in the psycholinguistic literature is still an open question (see below). By definition, a syllable is a unified speech stream, a series of phonemes that are pronounced together. Linguists have tried to describe the way syllables are formed using simple rules or heuristics. The most commonly evoked heuristic in describing English syllabification is the Maximal Onset (Max Onset) principle (Carr, 2000; Pulgram, 1970).

Regardless of the heuristic employed, syllable units are thought to comprise three main components (Treiman, 1986): (1) a mandatory nucleus which is always a vowel (e.g., *A*), (2) an optional onset which is always a consonant (*TO*) or consonant cluster (*TREE*), and (3) an optional coda which is always a consonant (*AT*) or consonant cluster (*ART*). Syllabification of speech into syllables is based on the attribution of phonemes to these roles. For example, the Max-Onset principle of syllabification states that words are partitioned into syllables so as to maximize the number and size of onsets in a word while respecting phonotactic constraints. According to this rule, the word *NAVY* would be partitioned into two syllable units: *NA-VY*. Similarly, the word *THUNDER* is partitioned into two units as follows *THUN-DER*. Notice that *ND* form an illegal onset in English and therefore the *N* is treated as the coda of the first syllable. Finally, it is worth noting that the Max-Onset principle is violated for many ambiguous cases (Anderson & Jones, 1974; Treiman & Zukowski, 1990) and so it should be considered merely an approximation of syllabification behavior.

Empirical Evidence for the Syllable

Researchers studying the status of the syllable as a functional unit in English have attacked the problem in a number of ways using the two 'gold standard' word recognition tasks: naming and lexical decision (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). For example, Jared & Seidenberg (1990) found that low-frequency words are read aloud more rapidly when they have fewer syllables, all other things being equal, including length in letters. Further, the frequency of the syllable units themselves is also known to affect lexical processing. Macizo & Van Petten (2007) discovered in their analysis of two-syllable items from the English Lexicon project that words with a high-frequency syllables are read aloud more rapidly and, perhaps more surprisingly, are recognized more quickly in visual lexical decision, which is a task that does not explicitly require phonological recoding. Taken together, these studies are suggestive of a role for the syllable as a functional unit, but the possibility of a confound with some distributional property of words cannot be ruled out (e.g., length in syllables is confounded with number of phonemes; syllable frequency may be confounded with print-to-sound consistency, etc…). Other kinds of manipulation, such as priming, avoid such ambiguity by capitalizing on the possibility of using counter-balancing, and items as their own controls.

In a masked-priming study, Ferrand, Segui, & Humphreys (1997) primed target words with orthographically matched fragments that either matched or violated their syllable boundary. These authors reported a naming advantage for the syllable-match condition and no effect in the lexical decision task. This finding was interpreted as strong evidence for the role of the syllable in naming, but not in word recognition per se. However, Schiller (1999, 2000) has failed to replicate syllable preference, reporting results that support instead the view that what matters in fragment priming is the degree of segmental overlap (i.e., more overlapping letters, greater priming). Using a simple (non-masked) fragment-priming paradigm, Taft (1979) also failed to find evidence in support of the Max-Onset syllable as a functional unit. Instead Taft found that a different type of syllable unit, one defined by the maximal coda (Max Coda) principle (see Taft & Radeau, 1995), produced a processing advantage relative to comparison conditions. In contrast to the Max Onset syllable, Max Coda division maximizes the number and size of codas in a word while respecting phonotactic and morphological constraints. It is important to note that this definition only conflicts with the Max Onset syllable for some words (see Taft, 1987, 2001). Following Taft & Forster (1976), Taft (1979) ascribed a morphological and orthographic function to the Max Coda unit, which contrasts with the primarily phonological basis of Max Onset division (though see Taft, 1992, for a phonological interpretation of Max-Coda preference; see also, Chateau & Jared, 2003).

In a more recent study, Ashby & Martin (2008) found a Max Onset syllable advantage in both naming and lexical decision when masked fragment primes were presented parafoveally. However, the specificity of this finding to parafoveal presentation and the priming manipulation suggests a locus for such syllable effects in the memory processes that support the fluent reading of continuous text rather than an individually presented word.

Word-splitting literature

A more direct way of studying the role of the syllable in visual word recognition is to employ the word-splitting paradigm, whereby stimuli are presented in a segmented form. This manipulation can be achieved by a mid-word change in case (Taft, 1979), introduction of a foreign character (e.g., /; Katz & Baldasare, 1983; Lima & Pollatsek, 1983; Taft, 1979, 1987) or spaces (Taft, 2001, 2002), or creation of distinct letter groupings using colour (e.g., Rouibah & Taft, 2001). In an experiment examining the importance of syllabic units, for example, the word THUNDER might be presented with a character inserted at the boundary of the phonologically defined syllable (i.e., *THUN/DER*) or at the boundary of the orthographically/morphologically defined syllable (i.e., *THUND/ER*). As discussed above, the former is defined operationally as segmentation according to the maximal onset principle (Max Onset; Pulgram, 1970; Spencer, 1996) and the latter is defined as segmentation according to the maximal coda principle (Max Coda, Taft & Radeau, 1995; see also the definition of the 'BOSS' unit, Taft, 1979). As with the priming methodology, the multifarious confounds associated with between-word comparisons are neutralized in the word-splitting paradigm by counterbalancing the association between word list and experimental condition across participants. Assuming such counterbalancing has equalized all other factors, an advantage of one type of segmentation over the other may be taken as evidence for the relevance of that unit during processing.

A history of inconsistent findings. For the last 20 years, research based on the word-splitting paradigm in lexical decision has yielded contradictory findings (for reviews, see Chen & Vaid, 2007; Taft, 2001), with some sources supporting the role of word-initial letter-chunks defined by bigram-frequency troughs (Seidenberg, 1987; but see Rapp, 1992; and also, Carreiras, Alvarez, de Vega, 1993), others the role of the Max-Coda Syllable (Rouibahh & Taft, 2001; Taft, 1979, 1987), and still others Max-Onset syllable preference or a general preference for syllabic division (Jordan, 1986; Katz & Baldasare, 1983; Lima & Pollatsek, 1983). Such variability across studies, especially when using identical stimuli (for e.g. Taft, 1979 vs Lima & Pollatsek, 1983, Taft, 1979 vs Kats & Baldasare, 1983; Taft, 1987 vs Seidenberg, 1987), has led some observers to term syllable effects 'unreliable' (e.g. Brand, Rey, & Peereman, 2003). In other words, the causes underlying the inconsistency in the literature are largely unknown.

Individual differences moderate syllable preference. Acknowledging the inconsistency in the literature, Taft (2001) proposed that fluctuation in the reading skill of participants across studies was responsible for the conflicting published reports (see also, Taft, 2002). Indeed, he found that the effect of dividing words at different locations varied as a function of reading skill, which he defined operationally as accuracy on a timed reading comprehension test similar to the Nelson-Denny (Taft, 2001). Specifically, skilled readers responded more rapidly under Max-Coda segmentation and less-skilled participants responded more rapidly when the same words were split according to the Max-Onset principle. The former result was replicated subsequently with a sample of high-skill readers (better than average verbal SAT scores) in more recent work (Chen & Vaid, 2007). Thus it would seem that the inconsistencies in the literature are due, at least in part, to sample-to-sample variability in reading skill. An important limitation of this finding, however, is that a concept like reading skill does not have a clear referent in models of visual word recognition, which means that it is of limited utility in informing the generalization of current models of word recognition to the case of multi-syllabic words. For instance, the moderating role of skill does not speak directly to the issue of the locus within such models of syllable effects.

Extrapolating from his 'skill' data, Taft (2001) offered a number of process-based speculations for the observed pattern of results. Perhaps the most convincing of these is the idea that the low-skill sample preferentially relied on phonological information during word recognition. Low-skill participants tend to also have less exposure to print (Lewellen, Goldinger, Pisoni, and Greene, 1993; Stanovich & West, 1989; Unsworth & Pexman, 2003). This constellation of reader attributes is associated with a heightened sensitivity to the phonological characteristics of words, both at the lexical and sub-lexical level (Chateau & Jared, 2000; Jared, Levy, & Rayner, 1999; Sears, Siakaluk, Chow, & Buchanan, 2008; Jared & Seidenberg, 1991; Unsworth & Pexman, 2003). This greater sensitivity is thought to reflect the fact that low-skill/low-print-exposure readers rely strongly on phonology at the expense of other types of information in visual word recognition out of necessity (Sears et al., 2008), but this does not imply that low-skill participants are any better at using phonology, quite the opposite (see Seidenberg & McClelland, 1989; Strain & Herdman, 1999). In fact, skilled readers appear to be more efficient at using all types of information, including phonology and orthography (Chateau & Jared, 2000; Sears et al., 2008). Returning to the issue of syllable preference in word recognition, it could be surmised that low-skill participants prefer Max-Onset segmentation because it is a 'phonological' unit and phonology tends to be a relatively strong determinant of their performance when compared to other domains of knowledge like orthography or semantics. If correct, this places the locus of Max Onset preference in the phonological domain. Depending on one's preferred theory of visual lexical decision making, greater reliance on phonology could mean stronger influence of feedback from phonology on the activation of orthographic units (e.g., models that place the seat of decision-making in the orthographic domain; Coltheart et al., 2001; Grainger & Jacobs, 1993), the direct use of phonological information in making visual lexical decisions, or both (e.g., models of lexical decision that posit flexible use of all types of linguistic information; Ratcliff, Gomez, & McKoon, 2004).

With respect to the preference of high-skill participants for Max-Coda analysis, Taft (2001) evoked the idea of a level of representation populated by syllable-like units whose characteristics are determined by pressure for efficient orthography-based lexico-semantic retrieval. The main support for this interpretation derives from the result of his third experiment, wherein the Max Coda preference was found to be conditional on whether the word-initial Max-Coda unit was bounded to a single word. If a word-initial Max-Coda unit was unbounded, which is to say it was common to multiple semantically-related words, then no preference was observed. This effect was attributed to competition created by the activation of multiple semantic representations. According to this interpretation, Taft's results provide some indirect support for the idea that there is a connection between word-initial Max-Coda units (presumably orthographic) and the semantic system. This interpretation differs somewhat from the original conceptualization of the 'BOSS' unit as serving the efficient retrieval of orthographic lexical representations (Taft, 1979).

Word Attributes moderate syllable preference. A study published subsequent to Taft's 2001 paper has failed to replicate the results of his third experiment (Chen & Vaid, 2007). The manipulation of boundedness in the original stimulus list is demonstrably confounded with lexical frequency such that bounded items are of lower frequency. When frequency is properly controlled, the dependency between syllable preference and boundedness disappears (Chen & Vaid, 2007). Lack of a boundedness effect is consistent with results reported by Rastle, Davis, and New (2004) indicating that Max-Coda fragments offer equivalent masked-priming regardless of semantic attributes (e.g., transparent stems vs semantically vacuous apparent stems). These findings suggest that Max Coda preference is a property of orthographic lexical organization rather that semantic retrieval per se.

Moreover, Chen and Vaid (2007) found that Max-Coda preference in a high-skill sample depends on lexical frequency such that it is contingent on the carrier word being of low frequency. This finding is broadly consistent with the observation that semantic and morphological characteristics have their most obvious effects on low frequency words (Alegre & Gordon, 1999; Becker, 1979; Meunier & Segui, 1999; Tse & Neely, 2007). Thus, it has been shown that Max Coda syllable preference is observed with low frequency words but not high frequency words. It remains to be seen whether Max Onset syllable preference is also frequency dependent. Indeed, this is what one would expect based on the general finding that the manifestation of phonological effects in lexical decision is more commonly observed with low-frequency words (Kerswell, Siakaluk, Pexman, Sears, & Owen, 2007; Lacruz & Folk, 2004; Paap & Noel, 1991; Seidenberg, 1985). In fact, because the magnitude of the frequency effect is known to depend on participant characteristics like exposure to print (Chateau & Jared, 2000; Sears et al., 2008), it is possible that sample-to-sample variability in this characteristic has contributed to the historically unstable word-splitting results.

Process-based individual differences. As stated earlier, variables like reading skill or print exposure are external to the machinery of the reading system, which means that there is no clear conceptual relationship between the parameters of theoretical models and these estimates of individual difference. In fact, the task of empirically linking these variables to more detailed 'process-based' effects has been an object of active research (Chateau & Jared, 2000; Sears et al., 2008; Lewellen et al., 1993; Stanovich & West, 1989). The term *process-based effect* is used here to refer to an experimental manipulation that has a relatively straight-forward explanation in terms of the parameters of existing models. For example, the homophone interference effect in lexical decision is thought to be caused by spurious feedback from the phonological to the orthographic domain (Pexman, Lupker, & Jared, 2001; Pexman, Lupker, & Reggin, 2002). The idea here is that a low frequency word like HARE activates its phonological representation, which in turn activates via feedback its relatively high-frequency competitor HAIR at the orthographic level. The activation of two competing responses at the orthographic level for such items increases response latencies and lowers accuracy. The explanation offered by a given model for this effect may vary, but what remains constant is that an adequate model will offer a precise explanation, linked to specific parameters. In other words, the processing that underlies the effect is relatively well-understood in model-specific terms. Another interesting process-based variable is the regularity effect in lexical decision, which is attributed to competitive interference in the phonological domain between the correct irregular response and the incorrect 'regularized' response (Pexman et al., 2002), and is only observed with low-skill participants (Unsworth & Pexman, 2003). An interesting property of this effect is that its magnitude is, among other things, thought to vary as a function of the extent to which phonological processes are involved in decision making (Unsworth & Pexman, 2003).

Such process-based variables are interesting generally as an under-exploited source of constraint on computational models of reading. An adequate model of reading, for example, would have a set of parameters that allows it to simulate the covariance of various process-based effects across participants. For instance, the size of the frequency effect and the orthographic neighborhood effect vary together across participants (Sears et al., 2008), then a computational model should be able to explain why this is the case by demonstrating that the pattern of covariance can be simulated by varying appropriate model parameters (e.g., by implementing multiple versions of a model to simulate individual differences). With respect to the issue of syllable preference, process-based effects are useful in that they can serve to pin down the locus of syllabic effects. The homophone interference effect, for example, is an index of how sensitive a participant is to feedback from phonology in the lexical decision task. On the other hand, due to the nature of the print-to-sound mappings involved, regularity effects are more appropriately interpreted in terms of the degree to which phonological information is relied upon (but see Unsworth & Pexman, 2003). In sum, if Max Onset preference is linked to degree of sensitivity to the homophone interference effect, then it can be attributed to the phonological domain via feedback from phonology. If, on the other hand, Max Onset preference is linked to individual differences in the regularity effect, then Max Onset preference will be more appropriately interpreted in terms of more direct reliance on phonological information which is unmediated by feedback.

Purpose

The purpose of the present research is to expand on the work of Taft (2001, 2002) by associating syllable preference with a participant attribute that is more directly tied to the online processing that is presumed to occur by extant computational models (e.g. Coltheart et al, 2001; Harm & Seidenberg, 2004): sensitivity to homophony and regularity during visual word recognition. Using a mixed-model statistical approach (Baayen, Davidson, & Bates, 2008; Jaeger, 2008; Locker, Hoffman, & Bovaird, 2007), we report a series of three experiments designed to address this issue. Among other things, mixed-models have the advantage of simultaneously representing item- and participant-attributes in a single analysis.

In terms of word-attribute effects, these experiments constitute an attempt at replicating what might now be considered the dominant finding in the word-splitting literature in visual word recognition, which is an advantage of the Max-Coda syllable over the Max-Onset syllable (e.g. Chen & Vaid, 2007; Taft, 1979, 1987, 2001). Previous work has suggested that lexical frequency moderates syllable preference (Chen

& Vaid, 2007), which is a relationship that is verified for the first time here. Finally, syllable complexity (i.e., presence or absence of an inter-vocalic consonant cluster) is sometimes cited anecdotally as a potential moderator (Taft, 2002), and so it was also included as an item-level predictor. The purpose of representing such a detailed word-attribute structure was to capture any potential interactions with the phonological participant attributes considered here.

At the participant-level, the degree to which participants rely on phonology, estimated here via the homophone interference effect and the regularity effect in lexical decision (Pexman et al., 2002; see the Method section for details), was used to predict word-splitting task performance. In Experiment 1, the feedback phonology indicator, homophone interference, is employed as a predictor in a standard word-splitting paradigm. In Experiments 2 and 3, both homophone interference and the regularity effect are estimated for each participant and used as predictors. If the phonological interpretation of Max Onset preference is correct, then sensitivity to phonological variables should be a reliable predictor such that participants who rely heavily on phonology should prefer the Max-Onset syllable and participants who do not rely on phonology prefer the Max-Coda syllable. A weaker 'asymmetrical' result is also consistent with this idea whereby a Max-Onset advantage for high-phonology participants and a null effect for low-phonology participants is observed. An interesting aspect to Experiments 2 and 3 is the contrasting theoretical interpretations of the homophone and regularity effects, which allow reliance on phonology to be distinguished from reliance on feedback phonology. Finally, cross-level interactions involving word- and participant-attributes are expected. Specifically, it is anticipated that manifestation of syllable preference one way or the other will be observed with low-frequency words. Taft (2002) has noted that syllable preference effects are more common with syllabically complex stimuli. Thus, it is possible that the observed effects will be limited to a particular stimulus type.

Experiment 1

At the item-level, Experiment 1 was intended as a basic replication of classic word-splitting experiments (e.g., Taft, 1979). At the participant level, it was intended as a test of the idea that sensitivity to feedback phonology (i.e., homophone interference) is predictive of syllable preference in a word-splitting task. Both lexical frequency and syllabic complexity were considered potential moderators of this relationship.

Method

Participants. Participants were 48 volunteer psychology students (22 men, 26 women) at the University of Ottawa, Canada. Mean age was 20.19 years with a 95 % Confidence Interval (*CI*) of 18.95 to 21.44. The sample comprised 29 first and second year students as well as 19 upper-year students and graduate students. Mean years of University study was 2.04 with a 95 % CI of 1.72 to 2.36. All participants reported normal or corrected to normal vision and English as their mother tongue.

Materials. A total of 120 polysyllabic experimental items were selected for the word-splitting task from the CELEX database (Baayen, Piepenbrock, Gulikers, 1995). According to frequency estimates obtained from the English Lexicon Project database (Balota et al., 2007), the mean log frequency (out of 131 million) for this sample of morphologically simple (i.e. monomorphemic) words was 6.63 (SD=1.63). For a summary of other item-characteristics, see Table 1.

Following Taft (2001, 2002)'s item selection rules, words that began with a prefix or a semantically transparent stem were excluded (e.g. *blunder* is acceptable, but not *virus* because of the existence of *viral*; but see Chen & Vaid, 2007). Half the items were syllabically simple and the other half were syllabically complex (i.e. possessed a medial consonant cluster). Critically, all the words allowed the Max-Onset and Max-Coda syllabifications rules to be contrasted (e.g. possessed a medial consonant cluster, see Taft, 2001). In all cases, the Max-Onset segmentation was consistent with the phonotactically-guided syllabified pronunciation provided by the Merriam-Webster's online dictionary (2004). Thus, the Max-Onset syllable might be considered the phonological syllable for the purpose of this experiment, which is not necessarily always the case (see Taft, 2002) because the maximal onset principle is constrained by the phonotactics of a language, on which linguists sometimes disagree (Carr, 2000). It is important to note at this point that the distinction between simple and complex words is confounded with length in letters of the word-initial unit. Thus, any effects associated with this manipulation may be attributed to either unit-length or complexity or some combination of the two. An equal number of nonwords were produced that matched the words on length in letters and CVC structure. These were created by modifying two letters in each word, while preserving the number of phonemes. A complete list of the 240 items used in the word-splitting task in Experiment 1 is presented in the Appendix A (words) and Appendix B (nonwords).

The homophone task was created using 72 items from the study reported by Pexman and colleagues (2002). This set of items has consistently generated homophone interference effects across several experiments (e.g., Pexman et al., 2001; Pexman et al, 2002; Unsworth & Pexman, 2003). The items included 4 practice words, 17 homophone-control word pairs (matched on frequency, length, and orthographic neighborhood), and 34 nonword filler items (for the complete list, see Appendix C). According to Pexman et al (2002), all of the homophones are low in frequency and share their pronunciation with a frequently occurring word that was not presented as an item, both of which are necessary conditions for the effect to be observed. Reader skill or exposure to print was not measured here directly, but Unsworth & Pexman (2003) have demonstrated that performance on the homophone task depends to some extent at least on reader skill. Low-skill participants make significantly more errors in response to homophones, and there was a corresponding trend in RT.

Equipment. The experiment was controlled using E-Prime 1.1 software (Psychology Software Tools, 2001). Hardware included an IBM compatible computer with a 17-inch 50 Hertz monitor operating with 1024 x 758 resolution. Participants indicated their response by pressing the far-left or far-right button of a five-button box response box (Psychology Software Tools, Model 200A).

Procedure. Participants were tested individually in a dimly-lighted room. They were asked to make judgments concerning the lexical status of letter-strings as quickly as possible while not exceeding an error rate of approximately ten percent. Judgments were made by pressing one of two buttons, where half the participants made positive word responses using their right hand and the other half with their left.

Participants completed two separate lexical decision tasks in a single 30 minute session. The first of these was the word-splitting task, which consisted of three blocks of trials within which a total of 270 words and nonwords were presented. The sequence began with a thirty-item practice block (15 words, 15 nonwords), which was then followed by two experimental blocks of 120 items each (60 words, 60 nonwords). Each experimental block began with four practice items (2 words, 2 nonwords), responses to which were discarded from the statistical analysis. Prior to experimentation, the word items were assigned to separate blocks for presentation so that various factors were equivalent (e.g., number of words, the experimental and control variables noted below) and the order in which the blocks were presented was counterbalanced across participants to control far any potential remaining carry-over effects. In contrast, the nonword items were drawn randomly without replacement from a single master list for each participant, with the constraint that an equal number of nonwords were presented in both blocks. In all cases, the order of item presentation was randomized within blocks for each participant.

For the word-splitting manipulation, a visual boundary was created for each item at one of three locations: the Max-Onset (*cro/cus*), the Max-Coda (*croc/us*), or the Max-Coda+1 (*crocu/s*) boundary. Following Rouibah & Taft (2001), this visual segmentation was created using colour (blue/red) against a white background rather than by dramatically altering the appearance of letter-strings through the introduction of a foreign character (e.g. /, #, or a few spaces). The following controls were in place to isolate the effect of word-splitting from other lexical characteristics. First, the 120 words were divided into three lists that were equated as much as possible on frequency, length in letters, number of phonemes, orthographic neighborhood, and bigram frequency (see Table 1). For presentation, these lists were assigned to one of three word-splitting conditions in such a way that the list-appearance association was counterbalanced across participants. In contrast, the nonwords were not assigned to a specific stimulus block. Instead, they were assigned randomly to one of the three word-splitting conditions for each participant, with the constraint that the three word-splitting conditions were represented equally in each block.

Each trial of the word-splitting task consisted of the presentation of a fixation mark (+) for 500 ms at the centre of the screen and presentation of an item for 500 ms. After 500 ms, the item was replaced with a blank screen pending a response from the participant. A 250 ms delay followed the participant's response prior to the start of the next trial. Items were presented in color (blue/red) against a white background.

The second task consisted of a single block of 74 trials. The block began with 6 practice items which were followed by 68 randomized experimental items, half of which were words and half nonwords. The task was intended to estimate the magnitude of homophone interference for each participant and was performed second to avoid biasing performance on the word-splitting task, which is less robust across studies than the homophone effect. Items for this task were presented in black, in lower-case, against a white background. Standard lexical decision instructions were used. For each trial, item presentation was preceded by a fixation mark (500 ms).

Results & Discussion

The data were analyzed using the statistical software R, Version 2.8.1, and the associated package for testing mixed-model effects *lme4*, as reviewed by Baayen et al., (2008) and Jaeger (2008). The analyses focused on two behavioral outcomes: response latencies for correctly identified words (RT) and error rate (ER) across all words. The RT data were analyzed via a linear mixed-model with restricted maximum likelihood estimation (Baayen et al., 2008), while the error data were modeled using mixed-model logistic regression for binomial outcomes (Jaeger, 2008). Both approaches are demonstrably superior to the way psycholinguistic data are more commonly analyzed, achieving excellent power without compromising the nominal Type-I error rate. Of currently available mixed-model programs, the algorithm in R is especially useful in that it generates accurate parameter estimates even in the presence of missing data (Baayen et al., 2008). The latter feature is useful given that trial-level missing data are ubiquitous in reaction time experiments due various factors (Miller, 1991).

A regression approach was adopted throughout because it admits continuous predictors, unlike factorial ANOVAs. Note that since it is unnecessary to compute 'mean RTs' with mixed-models, the main justification for the artificial dichotomization of continuous variables like lexical frequency disappears. As argued in the introduction, in the absence of such motivations, the preservation of statistical power is paramount. Following the conventional strategy (analogous to Type III sum of squares ANOVA, Cohen et al, 2003, p. 362), the analyses that follow consider only the unique variance associated with each predictor over and above the contribution of the other predictors in a particular model. The results of the word-splitting task are reported first followed by the results of the phonological task.

For Experiment 1, two sources of data were analyzed: homophone task and word-splitting task performance. First, the homophone interference effect was evaluated for the group. This analysis was intended to verify previous reports of homophone interference and also to describe the performance of the group on this measure. Subsequently, the statistical model of homophone task performance was re-estimated as a random slope or random coefficient analysis model (RCA; Lorch & Myers, 1990; Thompson, 2008). The purpose of the latter analysis was to provide custom descriptions of each participant's sensitivity to homophony in lexical decision. This indicator reflects the extent to which participants rely on phonological feedback at the expense of orthographic knowledge in word recognition. Second, reaction time and accuracy data on the word-splitting task were fitted by a model including word-attribute predictors (e.g., word-splitting condition; lexical frequency), participant-attribute predictors (e.g., magnitude of the homophone interference effect), and potential interactions across levels of analysis.

Prior to analysis, responses were screened on reaction time such that responses faster than 150 milliseconds (ms) were eliminated along with those falling further than 3.5 standard deviations above the participant mean. This procedure resulted in a loss of less than 3 percent of data points.

Use of the regression approach requires a number of decisions with respect to the way variables are coded. For example, all predictors are normally centered to facilitate the interpretation of regression coefficients and their associated significance tests and to avoid statistical problems when testing interactions (i.e., non-essential multi-colinearity). With dichotomous predictors, groups are normally coded so that the difference between conditions is equal to 1. For the phonological task, the sole item-level variable was homophone status (Homophone vs Orthographic Control, coded as +.5 and -.5, respectively). For the word-splitting task, the item-level variables were: log frequency (centered), syllable structure (simple vs complex, coded as -.5 and +.5, respectively), and the position of the color boundary (Max-Onset, Max-Coda, comparison condition). The word-splitting effect was represented by two orthogonal contrasts, the first comparing Max-Onset and Max-Coda segmentation to each other (-.5 vs +.5, henceforth 'Contrast A') and the second contrasting a linear combination of the Max-Onset and Max-Coda conditions with a third comparison condition, Max-Coda plus one letter (-.25 & -.25 vs .5, henceforth 'Contrast B'). All contrast codes were orthogonal according to accepted criteria (Cohen et al., 2003, p. 333).

Phonological Task. Mixed-models with random effects for items and participants were fitted to the RT and error data for the homophone task. The RT analysis ($N = 48$, Items = 34, Observations = 1425) estimated the variance components to be 3873.58 and 916.15 for the random effects of participant and item, with residual variance estimated at 9561.21. The fixed effect of homophone status was estimated at b = 27.34 ms, SE = 11.62, with a sample intercept of 571.52. Markov Chain Monte Carlo (*MCMC*), a bootstrapping procedure for small samples, revealed that the homophone effect is statistically significant, $p = .02$, Highest Posterior Density (*HPD*) 95 percent Confidence Interval (*CI*) of 5.55 to 49.47. The error analysis ($N = 48$, Items = 34, Observations = 1563) estimated the random effect variance at .78 for participants and

.77 for items (residual variance is unestimated in such log linear models). The logit of the intercept and homophone effect were -2.95, $SE = 0.23$, $Z = -12.90$, $p < .001$ and 0.80, $SE = .37$, $Z = 2.17$, $p = .03$ respectively. The test of the intercept indicates that on average errors were unlikely for any given item, as the logit corresponds to odds of .05 for making a mistake. This result is hardly surprisingly given that errors, by design, are rare events in RT experiments with minimal speed-accuracy trade-off. The direction of the homophone effect coefficient indicates that participants were more likely to get homophone items wrong. When the logit is converted into an odds ratio, it is revealed that homophones are 2.23 times more likely to be responded to incorrectly than the orthographic controls. The average estimated probability of making a mistake on any given trial was .085, or an error rate of 8.5 percent. In sum, the results confirm the stability of the homophone interference effect in lexical decision (Pexman et al., 2001; Pexman et al, 2002; Unsworth & Pexman, 2003). The fact that an effect was observed of homophony on ER suggests that the present sample is relatively 'low-skill' given that previous work has shown that this effect is observed with low-skill participants (Unsworth & Pexman, 2003).

Word-Splitting Task. The preceding models of RT and ER data were re-estimated, this time as an RCA model. The purpose of this analysis was to provide custom estimates of homophone interference effect for each participant, which were to be used as predictors of word-splitting task performance. It happens that only the slopes from the RT analysis were useful. In the interest of economy, only the results obtained using the RT estimates as predictors are reported here. It is worth noting that it is ER and not RT which was significantly associated with estimates of reader skill in previous work (Unsworth & Pexman, 2003), and yet ER was not useful in predicting responses to the word-splitting manipulation.

When discussing the word-splitting task results, the predictor corresponding to the random coefficients for homophone interference is henceforth referred to as 'homophone sensitivity' to avoid confusion. The mixed-model of word-splitting task RT data is reported in Table 2 and the ER in Table 3. A summary of the decisions taken regarding random effects is presented first. This is followed by a report of the item-level results and the participant-level results, including the main effect of Homophone Sensitivity and all cross-level interactions.

With respect to random effects, log-likelihood tests, Aikake's Information Criterion (AIC), and Bayesian Information Criterion (BIC) attested the utility of random intercept parameters for both items and participants in both the RT and ER analysis. In addition, a random slopes parameter for syllable preference (see Contrast A below) was found to significantly improve the fit of a model with direct effects only, $\chi^2(2) = 20.29$, $p < .001$. However, when the full model including interactions with Homophone interference and other word-level variables was tested in a subsequent step the Contrast A random slope parameter became redundant. Removing this parameter from the final model did not significantly affect model fit. This result attests the fact that random effect variability in syllable preference is well accounted for by the fixed-effect variables included in the model.

At the item level, log frequency was the only variable to have a direct effect, both on response latencies and errors. The coefficient for the RT analysis indicates that for every one unit increase in log frequency a corresponding 84.91 ms decrease in response latency was observed. The sign of the logit coefficient produced by the ER analysis indicates that the odds of making a mistake decrease with increasing frequency. Taking the exponential of the logit reveals an odds ratio of .12, which corresponds to an 88 percent decrease in the odds of making an error with every one unit increase in log frequency. A significant two-way interaction was observed between syllabic complexity and the critical word-splitting contrast involving Max Onset and Max Coda division: Contrast A. Post-hoc decomposition of the interaction indicates that the effect of word splitting is null for simple words, $b = -6.00$ ms, $SE = 8.66$, MCMC $p = .48$, *HPD* 95 percent *CI* -23.10 to 10.50, and statistically significant in favor of Max Onset segmentation for complex words, $b = 20.65$ ms, $SE = 8.71$, *MCMC* $p = .02$, *HPD* 95 percent *CI* 3.39 to 36.99. The results replicate the standard frequency effect, and indicate a Max Onset preference for a sub-set of words. While this interaction was qualified by the higher-order interactions discussed below, it seems to call into question the wisdom of treating both types of items as if they are similar (see Taft's work). Further, the observation of a Max-Onset advantage with complex words seems to run counter to the preponderance of Max-Coda effects in the word-splitting literature, but it is consistent with some conflicting reports (e.g., Lima & Pollatsek, 1983). If the presence of an overall effect of homophony on ER is taken as a manifestation of 'poor skill' (Unsworth & Pexman, 2003), then this Max Onset advantage would seem to confirm Taft's suggestion that participants who rely heavily on phonology in visual word recognition prefer words to be divided at the Max Onset syllable.

At the participant-level, Homophone Sensitivity was found to have a direct effect on RT in the word-splitting task whereby for every 1 ms increase in homophone interference there was a corresponding 1.64 ms increase overall RT on the word-splitting task. This direct effect was qualified by an interaction with frequency, $b = -0.68$, $SE = 0.17$, *MCMC* $p < .001$, *HPD* 95 percent *CI* -1.03 to -0.33. Post-hoc decomposition of this interaction revealed that the effect of frequency is greater when participants rely heavily on phonology (+1 Standard Deviation in Homophone Sensitivity), $b = -107.98$, $SE = 17.02$, MCMC $p < .001$, *HPD* 95 percent *CI* -138 to -

75.88, than when they do not (at -1 *SD*), $b = -61.15$, $SE = 16.99$, *MCMC* $p < .001$, HPD 95 percent *CI* -92.77 to -30.31. Slower responses as well as greater sensitivity to phonological effects and lexical frequency are associated with low reader skill and poor print exposure (Chateau & Jared, 2000; Sears et al., 2008; Unsworth & Pexman, 2003). The direct effect of homophone sensitivity in addition to the interaction observed here are consistent with this general observation, and support the idea that the homophone effect on lexical decision RT reported here is part of a constellation of individual-difference effects comprising print-exposure and reading skill (for mixed empirical support of this idea, see, Unsworth & Pexman, 2003).

The focus of the present analysis was cross-level interaction tests involving Contrast A, the comparison of Max-Onset and Max-Coda presentation. As can be seen in Table 2, none of the interactions is statistically significant at the .05 alpha level. We decided against a strict application of this decision rule for a few reasons. In the first place, the present analysis is based on an analytical technique that faithfully preserves the nominal Type-I error rate (Baayen et al., 2008), unlike standard analyses which claim the value .05 while actually being much more liberal. Second, we are reporting confidence intervals for the estimated coefficients, which are more informative than p-values and allow a more flexible interpretation of results, especially in cases where low power is an issue (Loftus, 1996; Masson & Loftus, 2003). Finally, the present analysis was considered exploratory, and any results were to be replicated in any case.

With this approach, two interesting interactions are suggested by the data. First, there was a marginally significant three-way interaction between the Homophone Sensitivity, Frequency, and Contrast A predictors. A four-way interaction involving Syllable Complexity seemingly qualified this effect. The observed *p*-value for both effects was .07. Post-hoc decomposition of the higher-order four-way interaction indicated that the three-way interaction was statistically significant for complex syllable words only, $b = -1.57$, $SE = 0.71$, *MCMC* $p = .03$, *HPD* 95 percent *CI* -3.00 to -0.22. For simple syllable words, there was no significant effect, $b = -.009$, $SE = 0.52$, *MCMC* $p = .99$, *HPD* 95 percent *CI* -1.07 to -0.99.

When the three-way interaction within complex words was decomposed, it was discovered that high-sensitivity participants (+1 SD) showed an interaction between Frequency and Contrast A, $b = -78.44$, $SE = 33.93$, *MCMC* $p = .02$, *HPD* 95 percent *CI* -144.51 to -12.39, while low-sensitivity participants did not, $b = 30.02$, $SE = 34.53$, *MCMC* $p = .37$, *HPD* 95 percent *CI* -36.49 to 97.67. Breaking down the Frequency by Contrast A interaction for high-sensitivity participants revealed a Max Onset advantage for low-frequency words only, $b = 57.30$, $SE = 17.09$, *MCMC* $p < .001$, *HPD* 95 percent *CI* 23.17 to 90.64. For high-frequency words, high-sensitivity participants showed no syllable preference either way, $b = -10$, $SE = 20.93$, MCMC $p = .63$, *HPD* 95 percent *CI* -52.31 to 30.45.

Residuals & other issues. The RT residuals indicated that the model described longer latencies less well. To investigate the impact of these scores, a filter was applied to the residuals whereby scores exceeding an absolute value of 2 were excluded from the analysis, which resulted in the exclusion of approximately four percent of datapoints. Under these conditions, a main effect of Contrast A, a Max Onset advantage, was detected, b = 8.87, $SE = 3.98$, *MCMC* $p = .02$, *HPD* 95 % *CI* 1.43 to 16.76, and its two-way interaction with Complexity fell short of significance, $b = 10.90$, $SE = 7.95$, *MCMC* $p = .19$, *HPD* 95 % *CI* -5.08 to 26.37. The marginally significant four-way interaction observed above remained relatively robust to the screen on RT, $b = -0.92$, $SE = 0.61$, *MCMC* $p = .14$, *HPD* 95 % *CI* -2.14 to 0.27, but the non-significant trend toward a three-way interaction involving Homophone Sensitivity, Frequency, and

Contrast A all but disappeared, $b = 0.04$, $SE = 0.61$, $MCMC\ p = .88$, HPD 95 % CI -0.54 to 0.65. Taken together, these results indicate that the presence of outliers in the solution cannot be evoked as an explanation for the observed pattern of results. The most notable change was a generalization of the Contrast A by Complexity interaction to a straight Max Onset advantage.

In addition it is worth noting that the inclusion of cross-level interaction terms involving Contrast B does not affect the observed pattern of results. This result is inconsistent with the idea that the Contrast A advantage and its associated interactions are reducible to a 'unit size' effect. If the most important factor was the size of the word-initial unit, then Contrast B would show a parallel series of effects. Unit-size based explanations for Max-Onset preference have been rejected in other work as well (see Taft, 1987, 2002).

Summary. In sum, the results of Experiment 1 provide qualified support for the asymmetrical pattern evoked in the introduction: phonological reliance seems to be predictive of Max Onset preference, but not of Max Coda preference. The fact that this pattern applies here only to low-frequency words is consistent with similar reports that the frequency of a carrier word moderates syllable effects (Chen & Vaid, 2007; Jared & Seidenberg, 1990). The apparent relevance of syllable complexity has important methodological implications as this factor is routinely ignored in the formal construction of stimulus lists.

Experiment 2

Experiment 1 provided preliminary support for the idea that homophone sensitivity and syllable preference are related. Post-hoc decomposition of the four-way interaction revealed that Max Onset preference was observed for complex items only, when words were low-frequency and when participants relied relatively heavily on phonology. A limitation of our demonstration is that the test of the four-way interaction falls short of statistical significance, which casts doubt on the validity of the post-hoc decomposition, regardless of whether it yielded a relatively clear pattern. The present experiment served to replicate Experiment 1 with increased power, which was achieved by doubling the number of items from 120 to 240.

Another goal of this experiment was to verify whether another indicator of phonological reliance, the magnitude of the regularity effect across participants, was predictive of syllable preference. The DRC model (Coltheart et al., 2001) attributes regularity effects to the non-lexical route (and its influence on lexical processing) and homophone interference specifically to the lexical route. If the pattern observed with the regularity effect mirrors that of the homophone effect, then syllable preference will have been linked to non-lexical phonology, according to a DRC conceptualization. Another difference is that regularity effects in lexical decision are not thought to be based on feedback from phonology (Pexman et al., 2001). When regularity effects occur (Unsworth & Pexman, 2003), the nature of the conflict created by irregularity suggests that they are the result of direct reliance on phonological processing in the decision-making process as opposed to feedback phonology.

Method

Participants. Participants were 48 psychology students (4 men, 44 women) at the University of Ottawa. Mean age was 20.62 years with a 95 % *CI* of 20.08 to 21.16. The sample comprised 16 introductory psychology students and 32 upper-year psychology students. Mean years of University study was 2.66 with a 95 % *CI* of 2.36 to 2.98. All participants reported normal or corrected to normal vision and English as their mother tongue. About 30 percent of the sample was drawn from a subject pool and received course credit in exchange for their participation.

Materials. For Experiment 2, 240 words were sampled from the Celex database to act as experimental items. Forty percent of the stimuli were carried over from Experiment 1. Not all of the items from the first experiment were used, however, as the item-selection criteria were tightened, especially for the simple syllable structure words whose initial syllable had to contain a long vowel in Experiment 2 (as defined by the Celex database). Otherwise, the item selection procedure remained the same. The log frequency for the Experiment 2 word-splitting items was 6.00 (SD = 1.72). For a summary of the other characteristics of these items, see Table 1. As for the phonological task, the same Pexman et al (2002) items as in Experiment 1 were used, with the addition of the irregular items (n=10) and regular controls (n=10) used by the same authors (items listed in Appendix C).

Procedure. The instructions for Experiment 2 were identical to those given in Experiment 1, as was the strategy for dissociating item-list and word-splitting condition. That is, three lists were created, which were assigned to word-splitting condition according to a counterbalancing scheme across participants (see Table 1 for list characteristics). In addition, the way the items were organized for presentation was identical. The main differences between Experiment 1 and Experiment 2 lie in the number of items used and the fact that the items were presented in blocks according to frequency. Contextual blocking effects are always a possibility in psycholinguistic research (Lupker, Brown & Columbo, 1997; Rastle, Kinoshita, Lupker, & Coltheart, 2003), therefore researchers are left to decide whether they want to enhance or degrade the strength of a given effect. In this experiment, we elected to maximize the potential frequency effect, which is considered a potential moderator of the relationship between phonological processing and detection of syllable preference effects. In any case, blocking effects in lexical decision are typically weak and require the manipulation of both word and nonwords attributes (Rastle et al., 2003). The items from Taft (2001)'s experiment 3 were also presented as a separate block of items, either immediately before or immediately after the other item blocks were presented (after the 30-item practice list). These items were excluded from the present analysis, as they will form the object of a separate paper investigating the moderating role of boundedness specifically.

Results & Discussion

Except for the differences described above, the variables used in Experiments 1 and 2 were identical. The analyses were based on the same mixed-model procedure, the variables were coded the same way, and the same strategy for screening outliers was employed. The outlier detection procedure resulted in a loss of less than two percent of data points.

Phonological Tasks. Mixed-models with random effects for items and participants were fitted to the RT and error data for the phonological effects: the homophone interference and regularity effects. The homophone interference effect is evaluated first and the regularity effect second.

The RT analysis of the homophone effect (N = 48, Items = 34, Observations = 1548) estimated the variance components to be 5377.3 and 1828.7 for the random effects of participant and item, with residual variance estimated at 19821.8. The fixed effect of homophone status on RT was estimated at b = 33.25 ms, SE = 11.62, with a sample intercept of 596.23, which can be considered a sort of grand mean. The *MCMC* bootstrapping procedure revealed that the homophone effect is statistically significant, p = .04, 95 % Confidence Interval (*CI*) of 1.26 to 64.39. The error analysis (N = 48, Items = 34, Observations = 1630) estimated the random effect variance at .42 for participants and .74 for items. The logit of the intercept and homophone effect were -3.46, SE = 0.22, Z = -15.62, p < .001 and 0.40, SE = .40, Z = 1.01, p = .31, respectively. The mean estimated probability of making a mistake with the homophone items was .049. Taken together, the result confirms the homophone effect on RT that was observed in Experiment 1 and in previous work (Pexman et al., 2001; Pexman et al, 2002; Unsworth & Pexman, 2003), but the effect is less strong here and was not corroborated by ER. Inter alia, this suggests that the present sample of participants is more skilled overall than in Experiment 1, which concords with the fact that a greater number of upper-year students and graduate students were included.

The RT analysis of the regularity effect ($N = 48$, Items = 20, Observations = 926) estimated the variance components to be 6394 and 1268.6 for the random effects of participant and item, with residual variance estimated at 31801. The fixed effect of Regularity was estimated at $b = 0.90$ ms, $SE = 18.55$, with a sample intercept of 576.01. The *MCMC* bootstrapping procedure confirmed the apparent null effect, $p = .94$, with a 95 % *CI* of -37.31 to 37.85. The error analysis ($N = 48$, Items = 34, Observations = 960) estimated the random effect variance at near zero for participants and .72 for items. The logit of the ER intercept and regularity effect were -3.62, $SE = 0.27$, $Z = -13.27$, $p < .001$ and 0.03, $SE = .55$, $Z = 0.06$, $p = .95$, respectively. The mean estimated probability of making an error on any given trial was .033. Thus, a null regularity effect was observed here, which is consistent with the idea that the present sample is relatively highly skilled (Unsworth & Pexman, 2003). If this supposition is correct, then a null effect of word-splitting at the item level is expected.

Word-Splitting Task. As in Experiment 1, the preceding models were re-estimated, this time using an RCA approach. The purpose of this analysis was to provide custom estimates of the phonological effects for each participant, which were to be used as predictors of word-splitting task performance. This procedure generated four participant-level variables (2 phonological variables by 2 behavioral indicators, RT and

ER). The predictors were not significantly correlated with each other, with the exception of Homophone ER and Regularity RT, $r = .32, p = .03$. Based on the Unsworth & Pexman (2003) analysis, this relationship would seemingly reflect common 'skill' variance. It is perhaps surprising then that only the slopes from the homophone RT analysis were predictive of performance. If general reading skill is a determining factor in syllable preference, then one would expect its strong phonological concomitants to strongly associate with syllable preference. In the present experiment, this did not appear to be the case. At any rate, the fact that regularity predictors were not useful isolates the predictive value of phonological reliance to the relative importance of 'feedback' from phonology to orthography in determining responses.

In the interest of economy, only the results using the homophone RT estimates as predictors are reported here. The mixed-models of word-splitting task performance in terms of RT and error rate are reported in Tables 4 and 5 respectively. Again, random intercept parameters significantly improved model fit and continued to do so even after all fixed-effects were added. The same was not true of the random slope parameter for syllable preference, which became redundant with the addition of interaction effects. In what follows, the item-level variables are reported first while the participant-level phonological reliance and cross-level interactions are reported second.

With respect to item-attributes, a glance at Tables 4 and 5 indicates that the only lexical characteristic to have a direct effect on performance was lexical frequency. The RT analysis indicates a 34.84 ms decrease in response latency for every log (base 10) unit of lexical frequency. The ER analysis corroborates this effect as the odds of making a mistake decrease by 44 percent for every 1 unit increase in log frequency. It is apparent that the frequency effect is not as strong with the current sample of items and participants as it was in Experiment 1, which is yet another confirmation that the present sample is relatively highly skilled. Further, in contrast to Experiment 1, an interaction was not observed between Complexity and Contrast A. This discrepancy is to be expected given that the present sample is more highly skilled (and therefore less likely to show an overall Max Onset preference), and is actually quite consistent with the historically inconsistency of the effect.

At the participant level, the magnitude of the homophone interference effect was again found to be directly related to word-splitting RT. Specifically, for every 1 ms increase in homophone interference, there was a corresponding 0.64 ms increase in overall reaction time. As with frequency, its direct effect was weaker in Experiment 2 than in Experiment 1. Somewhat counter intuitively, the effect of Homophony Sensitivity on ER was observed here where it was absent in Experiment 1. The effect reflects a slightly higher likelihood of making a mistake on the word-splitting task with every 1 ms increase in the homophone interference effect, with odds of 1.004.

Table 4 indicates that the interaction between homophone effect and lexical frequency that was reported for the RT analysis in Experiment 1 was replicated here. The signs of the main effects and interaction indicate that the pattern is the same as that observed in Experiment 1. Post-hoc decomposition was therefore not undertaken, especially since this effect was qualified by a high-order interaction: a three-way interaction, which was detected among homophone sensitivity and the item-level variables lexical frequency and syllabic complexity. For simple words, the two-way interaction between Homophone Sensitivity and Frequency was statistically significant, $b = -0.14$, $SE = 0.03$, *MCMC* $p < .001$, *HPD* 95 % *CI* -0.21 to -0.09. For complex words, this interaction was smaller but still statistically significant, $b = -0.06$, $SE = 0.04$, *MCMC* $p = .02$, *HPD* 95 % *CI* -0.12 to -0.008. Thus, for complex words, the magnitude of the frequency effect depended less strongly on Homophone Sensitivity.

Table 4 confirms the marginal four-way interaction that was observed in the RT analysis in Experiment 1. All the relevant effects were identical in sign, but in this case the four-way interaction was statistically significant. An interesting difference with Experiment 1 is that the hint of a three-way interaction between Homophone Sensitivity, Frequency, and Contrast A is absent. This may be attributable to the fact that the nature of the stimuli in the 'simple' condition was changed slightly. Also, the ER analysis reported in Table 5 indicates a significant four-way interaction which was not present in Experiment 1.

Post-hoc decomposition of the four-way interaction on RTs confirms that the three-way interaction among Homophone Sensitivity, Frequency, and Contrast A falls short of statistical significance for simple words, $b = 0.05$, $SE = 0.07$, $MCMC\ p = .50$, HPD 95 % CI -0.09 to 0.20, but is statistical significant for complex words, $b = -0.20$, $SE = 0.07$, $MCMC\ p = .005$, HPD 95 % CI -0.34 to -0.06. For high-sensitivity participants (+1 SD), the two-way interaction between Frequency and Contrast A is statistically significant, $b = -15.02$, $SE = 5.29$, $MCMC\ p = .004$, HPD 95 % CI -25.13 to -4.46, while for low-sensitivity participants (-1 SD) it is not, $b = 5.72$, $SE = 9.56$, $MCMC\ p < .67$, HPD 95 % CI -3.95 to 16.37. The source of the Frequency by Contrast A interaction for high-sensitivity participants is a significant Max-Onset advantage for low frequency words, 34.77, $SE = 12.82$, $MCMC\ p = .005$, HPD 95 % CI 8.74 to 59.49, and a corresponding null effect for high-frequency words, $b = -15.48$, $SE = 12.01$, $MCMC\ p = .20$, HPD 95 % CI -38.94 to 8.76. In sum, the pattern of the four-way interaction is identical to that in Experiment 1, with the exception that it was statistically significant due, at least in part, to increased power resulting from the use of nearly double the number items in Experiment 2.

Post-hoc decomposition of the four-way interaction observed with ER suggested the possibility of a response bias in the data. The three-way interaction involving Homophone Sensitivity, Frequency, and Contrast A was not significant for simple words, *logit* = -.001, *SE* = 0.001, *Z* = -0.90, *p* = .37, and was marginally so for complex words, *logit* = 0.002, *SE* = 0.002, *Z* = -1.90, *p* = .056. For high-sensitivity participants (+1 *SD*), the two-way interaction between Frequency and Contrast A (complex words) was marginal, *logit* = 0.15, *SE* = 0.09, *Z* = 1.66, *p* = .09, while for low-sensitivity participants (-1 *SD*), it was clearly non-significant, *logit* = -0.10, *SE* = 0.09, *Z* = 1.02, *p* = .31. When the two-way interaction for high-sensitivity participants was broken down further, a null effect is observed for high-frequency words, *logit* = 0.06, 0.24, *Z* = 0.28, *p* = 0.78. In contrast, a significant Max Coda advantage is observed for low-frequency words, *logit* = -0.45, *SE* = 0.19, *Z* = -2.44, *p* = .01. In the RT analysis, it was a Max Onset advantage that was observed in this condition. The complementary pattern suggests the possibility of some sort of response bias (speed-accuracy trade-off) whereby high-sensitivity participants tended to respond to low frequency words quickly (and inaccurately) when they are divided at the Max Onset boundary, and slowly and accurately when they are divided at the Max Coda boundary.

Speed-Accuracy Bias? To test whether a response bias could explain the RT effect, we entered the estimated probability of making a mistake on any given word trial from the ER analysis as a predictor of trial-level RTs, along with the other predictor variables. If the RT effect is reducible to a speed-accuracy trade-off, then the four-way interaction in the RT analysis should cease to be significant when the estimated error probabilities are entered into the equation, or should at least be dramatically reduced. This is true because the error probabilities are the result of a model that contains the same predictors as the RT analysis. If high-sensitivity participants are biased towards errors with certain types of stimuli, then this will be represented in the fitted error probabilities. If this bias is responsible in whole are in part for any RT effects, especially the 4-way interaction in question, then controlling for this variable will eliminate or reduce the observed effect.

Accordingly, the fitted trial-level error probabilities were entered as predictors of RT in a model that was otherwise identical to that reported in Table 4. As expected, the estimated probability of making a mistake on a given word-participant combination (i.e. trial) was highly predictive of RTs, $b = 218.65$, $SE = 19.31$, *MCMC* $p = .0001$, *HPD* 95 % *CI* 186.81 to 262.99. The regression coefficient represents the increase in RTs associated with 1-unit increase in error probability (i.e., zero probability vs 1.0 probable). Inclusion of this predictor did not, however, eliminate or dramatically reduce the 4-way interaction with RTs, which in fact became slightly stronger, $b = -0.32$, $SE = 0.10$, *MCMC* $p = .0008$, *HPD* 95 % *CI* -0.52 to -0.12, suggesting some mild suppression. The RT frequency effect was reduced by the inclusion of trial-level error probability as a predictor, but remained significant, $b = 21.64$, $SE = 2.30$, *MCMC* $p = .0001$, *HPD* 95 % *CI* -25.66 to -17.04. Similarly, the Homophone Sensitivity effect on word-splitting RTs was reduced slightly in this analysis, but remained statistically reliable, $b = .56$, $SE = 0.21$, *MCMC* $p = .008$, *HPD* 95 % *CI* 0.18 to 0.94. The same was true of its interaction with Frequency, b = -0.09, $SE = 0.02$, *MCMC* $p = .0001$, *HPD* 95 % *CI* -0.13 to -0.05. Finally, the three-way interaction between Homophone Sensitivity, Complexity and Frequency nearly dropped from significance, $b = 0.09$, $SE = 0.03$, *MCMC* $p = .047$, *HPD* 95 % *CI* 0.004 to 0.17. In sum, none of the RT effects reported in Table 4 are reducible to a trade-off between response speed and accuracy, and importantly the 4-way interaction was shown to be unrelated.

Residuals & other issues. Inspection of the RT residuals indicated some misfit with the longer latencies. To investigate whether or not these influenced the results, the data were reanalyzed without the scores that exceeded 2 Standard Deviations above and below the regression line. This screening procedure again resulted in the loss of approximately four percent of data points. Analysis of the re-estimated model indicated that the observed pattern was nearly identical. Inter alia, the four-way interaction was reduced but remained significant, $b = -0.21$, $SE = 0.29$, $MCMC\ p = .04$, HPD 95 % CI -0.41 to 0.004. Note also that the results are unaffected by the inclusion of cross-level interaction terms involving Contrast B. Thus, the cross-level interactions are unlikely to be simple 'unit size' effects. Unit-size based explanations for Max-Onset preference have been ruled out in other work as well (see Taft, 2002).

Summary. Experiment 2 confirms the result of Experiment 1: participants who are highly sensitivity to feedback phonology, operationalized here as the magnitude of the homophone interference effect, show a preference on RT for Max Onset division with some low frequency words. This preference is not reducible to a weak but significant mirror-image effect on ER in the opposite direction. Overall, the results extend those of Taft (2001, 2002) and Chen & Vaid (2007) by linking word-splitting syllable preference to an indicator of phonological reliance that is process-based.

Experiment 3

Experiments 1 & 2 establish that a four-way interaction is observed with the present set of data whether words are split so as to emphasize the word-initial segment. This interaction is caused by a Max Onset advantage for certain words (complex, low frequency). In contrast, sensitivity to feedback phonology was not predictive of Max Coda preference on RT. It is possible that the way words were segmented was not remarkable enough to produce syllable preference for low-sensitivity participants. Accordingly, a more extreme colour-division manipulation was attempted that de-emphasized the word-initial unit by making its letters more difficult to recognize. If such a manipulation were successful, it would disrupt the normal processing of the homophone-insensitive participants sufficiently for a syllable effect to be observed. Specifically, it was hypothesized that a Max Coda effect would be observed for such participants despite the fact that words divided in this way necessarily have fewer letters that are easy to identify relative to the Max Onset condition. Similarly, if the Max Coda syllable is unrelated to word recognition for high-sensitivity participants, then a straight unit-size advantage should be observed whereby smaller units (i.e., Max Onset syllable) are preferred to larger units (e.g., Max Coda syllable). If, in contrast the Max Coda syllable is characteristic of processing for high-sensitivity participants, then a Max Coda advantage may be observed despite the fact it results in a greater proportion of hard to discriminate letters for each letter string. The bias against Max Coda preference inherent in the methodology means that this analysis constitutes a strong test of the functional utility of this unit.

Accordingly, in Experiment 3 the items are again split by color, but the word-initial segment is de-emphasized by presenting its letters in yellow against the same white background as before. If an advantage is observed for Max Coda division in this context, then it would have to be considered strong support for this word-initial segment.

Method

Participants. Participants were 48 psychology students (7 men, 41 women) at the University of Ottawa. Mean age was 20.11 years, with a 95 % *CI* of 18.74 to 21.57. The sample comprised 33 introductory psychology students and 15 upper-year students. Mean years of University study was 1.92 with a 95 % *CI* of 1.59 to 2.25. All participants reported normal or corrected to normal vision and English as their mother tongue.

Materials. Exactly the same set of words and nonwords as in Experiment 2.

Procedure. The difference between Experiment 2 and Experiment 3 is in the trial-level procedure. First, the word-splitting was performed using yellow for the first letter chunk and red for the final letter-chunk, in contrast to Experiment 2 which used blue and red respectively. The fact that items were presented against a white background provided a poor contrast for the word-initial letter-chunk, which was presented in yellow. Not surprisingly, this manipulation increased the difficulty of the task substantially. To accommodate this reality, the experimental procedure was altered so that items remained visible until a response was collected. Further, in Experiment 3, mistakes were signaled to participants with an auditory signal (short beep) as an additional strategy for keeping error rates down.

Results & Discussion

This experiment was structured identically to the previous experiment. Accordingly, the data were analyzed according to the same analytical strategy employed in Experiment 2. The outlier delection procedure resulted in a loss of less than two percent of data points.

Phonological tasks. A mixed-model of the Homophone task yielded estimates of item, subject, and residual variance of 2334.5, 18285.2, and 47784.9, respectively, and 1522 observations across 48 participants, and 34 items. The intercept was estimated at 614.33, *HPD* 574 to 651 while the effect of Homophone Interference was estimated at 42.32 ms, $SE = 20.02$, $MCMC\ p = .04$, *HPD* 95 % *CI* 2.04 to 81.75. The mixed-effects logistic regression model of ER estimated variance in the logit at 0.53 and 0.56 for items and participants, respectively, based on 1618 observations, 48 participants, and 34 words. The intercept, $logit = -3.28$, $SE = 0.21$, $Z = -15.67$, $p < .001$, indicated that participants were significantly more likely to respond correctly than make a mistake.

The mean estimated probability of making a mistake on any given trial was 0.056 (i.e., error rate). The test of the Homophone interference effect on ER was statistically significant, $logit = 0.76$, $SE = 0.35$, $Z = 2.17$, $p = .03$. Thus, the robust effect of homophone status on RT was replicated again, and the corresponding effect on ER re-emerged here where it was absent in Experiment 2, suggesting that the present sample is of lower-skill given that the experimental materials and procedure for this take were identical to those used in Experiment 1.

The irregular items and their regular controls were submitted to a similar mixed-model analysis of RT and ER. The RT model estimated the item, participant, and residual variance at 772.52, 9556.15, and 22333.08 respectively, based on 914 observations over 48 participants and 20 items. The intercept was estimated at 577.82 ms, $SE = 16.19$, *HPD* 95 % $CI = 549.75$ to 604.8. The test of the regularity effect on RT was not statistically significant, $b = -6.96$, $SE = 15.89$, *MCMC* $p = .65$, *HPD* 95 % $CI =$ -39.59 to 24.6. The mixed logistic model of ER estimated the item and participant variance in the logit at 0.52 and 0.34, respectively, based on 954 observations over 48 participants and 20 items. The intercept was estimated at -3.53, $SE = 0.26$, $Z = -13.59$, $p < .001$. The test of the regularity effect was not statistically significant, $logit = -0.25$, $SE = 0.49$, $Z = -0.51$, $p = .61$. This lack of statistical significant replicates the finding obtained in Experiment 2, but in the present case it is more puzzling. Unsworth & Pexman (2003) found that the Homophone effect on ER and the Regularity effect on RT were both characteristic of low-skill processing. The dissociation of the two effects reported here suggests that two effects are not as strongly related as previous work suggested.

Word-splitting tasks. As with the preceding two-experiments, a random slopes analysis was conducted to produce customized phonological effects for four participant-level variables (2 phonological tasks by 2 dependent variables, RT and ER, respectively). In this analysis, the homophone and regularity effects on RT were found to be negatively correlated, $r = -.40, p = .005$. This result was unexpected and contrasts with that obtained in the previous experiment where the same correlation was non-significant, $r = -.15, p > .05$. Recall that the phonological tasks were administered last, which means that the disruptive manner in which word-splitting boundaries were introduced may have influenced the behavior of participants on the phonological tasks. The shift from a null to significant negative relationship across experiments which was observed between the two RT effects suggests that a) the relationship between the RT effects for homophony and regularity is context dependent and b) this contextual strategic emphasis involves a shift away from a mode of processing that is highly sensitive to regularity (relatively speaking) to one that is relatively sensitive to homophony, or vice versa. The roughly 10 ms increase in the magnitude of the homophone interference effect from Experiment 1 to Experiment 2 suggests that it is predominantly the former. The motivation for this hypothetical shift in emphasis may be related to the fact that regularity effects in lexical decision are more prominent when stimulus quality is degraded (Hino & Lupker, 1996).

It was unclear what effect if any this seeming shift should have on the relationship between these factors and syllable preference. To investigate the matter further, the four estimates were entered into models predicting word-splitting task performance. As in the previous experiments, only the RT coefficients for the Homophone task were predictive of performance. Accordingly, only the analyses using the RT 'Homophone sensitivity' coefficients as participant-level predictors are reported. The mixed-models of word-splitting task performance in terms of RT and error rate are reported in Tables 6 and 7 respectively. Again, random intercept parameters significantly improved model fit and continued to do so even after all fixed-effects were added. The same was not true of the random slope parameter for syllable preference, which became redundant with the addition of interaction effects. The latter indicates that there is consistently substantial variability in syllable preference and that this variability is accounted for by the fixed effects included in the final model. As above, the item-level effects are discussed first, followed by the participant-level phonological reliance and cross-level interactions.

The word-splitting manipulation was more drastic in this experiment, and this is reflected in the direct effect of both Contrast A and B on RT. These effects are best interpreted in terms of the proportion of letters for each condition that were difficult to discrimination: the larger the proportion, the worse the performance. In other words, Max Onset division responses are fastest because the fewest letters were colored yellow against a white background. In contrast, Max Coda+1 responses were slowest because more letters were colored yellow. More interesting is how these effects interact with Homophone Sensitivity (more on this below). Lexical frequency was found to affect both RT and ER. This effect was qualified by an interaction between frequency and complexity, on both RT and ER. In both cases, the pattern of the interaction is synergistic such that the frequency effect is greatest with complex words. For complex words, the difficult-to-discriminate word-initial segment was always longer than the corresponding segment for simple words and it would seem that this property was particularly disruptive to the processing of low-frequency words. While stimulus quality effects tend to be additive with word frequency in lexical decision (Yap, Balota, Tse, & Besner, 2008), in the present case the stimulus was only partially degraded, which is arguably more similar to other ways of disrupting the familiar form of stimuli, like case-mixing, that do interact with frequency (Besner & McCann, 1987).

At the participant-level, the direct effect of Homophone Sensitivity on word-splitting RT is remarkable by its absence. It did, however, interact with several lexical characteristics. The two-way interaction between Frequency and Homophone sensitivity observed in Experiments 1 & 2 was replicated. Most interesting within the context of this paper, is the two-way interaction between homophone sensitivity and Contrast A. The pattern of the interaction is synergistic, indicating that on average the Max Onset advantage gets larger as Homophone sensitivity increases. This effect is consistent with the four-way interactions observed in Experiments 1 & 2, only it is more general. Note that including the two-way interaction between Homophone sensitivity and Contrast B does not affect the results, which means that an obvious alternative explanation, a simple unit-size effect, is less likely.

This two-way interaction is of course qualified by two three-way interactions and the four-way interaction. The signs of the three-way interactions indicate that the preceding two-way between Homophone Sensitivity and Contrast A is reduced as frequency increases and when words are syllabically complex. The former is consistent with the previous experiments, but the latter is not. A syllable preference is being demonstrated here with the syllabically simple words. A full account of the results requires a decomposition of the four-way interaction.

For simple words, the three-way interaction involving Homophone Sensitivity, Frequency, and Contrast A is statistically significant, $b = -0.54$, $SE = .09$, $MCMC\ p < .001$, HPD 95 % CI -0.73 to -0.36, but for complex words it is not, $b = -0.01$, $SE = 0.09$, $MCMC\ p = .91$, HPD 95 % CI -0.18 to 0.16. It is readily evident then that the results of the present experiment do not mirror those of the previous two: it is the simple words rather than the complex words that are showing a Homophone-Sensitivity-dependent pattern. Complex words are longer, have slightly higher frequency, and the difference between Max Coda and Max Onset division can be more than one letter (as opposed to words in the simple syllable condition). It is possible that the emphasis and de-emphasis of the word-initial unit interacts with one of these factors so that syllable preference is conditional on the method of word-splitting.

When the three-way interaction with simple words was explored further, it was discovered that, for high-sensitivity participants (+1 *SD*), the interaction between Frequency and Contrast A was significant and negative, $b = -32.13$, *MCMC* $p = .0002$, *HPD* 95 % *CI* -49.36 to -15.56. In contrast, for low-sensitivity participants (-1 *SD*), the effect was significant and positive, $b = 37.33$, $SE = 11.32$, *MCMC* $p < .001$, *HPD* 95 % *CI* 21.12 to 54.04. Obviously, the contrasting direction of the coefficients suggests that the high- and low-sensitivity participants produced a qualitatively different pattern of results. For high-sensitivity participants, a large Max Onset advantage is observed for low-frequency words, $b = 140.50$, $SE = 19.78$, *MCMC* $p = .0001$, *HPD* 95 % *CI* 102.95 to 178.62, and a null effect for high-frequency words, $b = 32.80$, $SE = 20.47$, *MCMC* $p = .12$, *HPD* 95 % *CI* -8.98 to 72.02. This pattern is the same as in Experiments 1 & 2, except in this case it is with simple words. For low-sensitivity participants, a statistically significant Max Coda advantage was observed with low-frequency words, $b = 71.36$, $SE = 19.88$, *MCMC* $p = .0004$, *HPD* 95 % *CI* -110.05 to -32.42, and a statistically significant Max Onset advantage for high-frequency words, $b = 53.74$, $SE = 20.17$, *MCMC* $p = .008$, *HPD* 95 % *CI* 13.74 to 93.45. The former effect was predicted based on the analysis developed in the introduction. The latter effect, however, was not, and is not readily interpretable as a 'syllable' effect, which is unlikely to be observed with high-frequency words (see Introduction).

One possible scenario is that the syllable preference effect is time-course dependent such that low-sensitivity participants show preference for the phonological unit early (Max Onset) and the morphographic unit later (Max Coda). A problem with this explanation is that orthography usually precedes phonology when the time-course of activation is examined (Ferrand & Grainger, 1993, 1994), and in this case preference for the phonological unit came first (i.e., with the quickly recognized high-frequency words). This problem might be resolved, however, by assuming that the Max Coda unit is not an orthographic unit but rather the result of semantic feedback. Semantic priming, for example, normally takes longer to manifest than either orthographic priming or phonological priming (Rossell, Price, & Nobre, 2003; but for an example of rapid semantic priming, see Perea & Gotor, 1997). Of course, it is unlikely that syllable effects would be observed at all with high-frequency words (see Introduction). Thus, a non-syllable explanation may be in order. For instance, it is perhaps more plausible that low-sensitivity participants were better able to take advantage of the smaller proportion of low-contrast letters when words were high-frequency based on their familiarity. If this is the case, the advantage observed in this condition would have less to do with syllable preference and more to do with the quality of the orthographic information available to the participant and the ability of the participant to make use of this information. If this line of reasoning is pursued, it is evident that the more remarkable finding is the syllable preference effect which was observed with low-frequency words, because this condition actually benefited from the larger low-contrast word-initial letter cluster. The expected disadvantage associated with this condition was seemingly more than compensated by congruency with Max Coda division. In sum, the more extreme visual segmentation implemented in Experiment 3 influenced the processing of low-sensitivity participants, thereby revealing a connection between insensitivity to homophone interference and preference for Max Coda division with low-frequency words.

Residuals & other issues. Again, inspection of the residuals indicated that longer latencies were not fitted particularly well. To investigate whether or not these were a determining factor in producing the observed pattern of results, the analysis was carried without the scores exceeding 2 standard deviations above and below the regression line. Analysis of the re-estimated model indicated that the observed pattern was again nearly identical. Inter alia, the four-way interaction was reduced but remained significant, b = 0.29, SE = 0.29, $MCMC\ p$ = .002, HPD 95 % CI 0.11 to 0.48. Note that the participant-level effects were not influence by the inclusion of interactions involving Contrast B, suggesting that unit-size cannot account for the associated moderated word-splitting effects.

General Discussion

The aim of the present research was to verify whether process-based phonological variables were predictive of participant syllable-preference. With some qualifications, the answer to this question was positive in all three experiments. Specifically, participants who were sensitive to feedback phonology (i.e., showed a large homophone interference effect) tended to prefer the 'phonological' syllable defined by the Max Onset rule with low frequency words. This pattern of syllable preference with low-frequency words was qualified by an apparent interaction between the experimental word-splitting manipulation and syllabic complexity. Specifically, when the letters of the word initial segment were highly contrastive against the background (Experiments 1 & 2), the syllable effects were observed with syllabically complex words. In contrast, when the word-initial segment contrasted poorly with the background (Experiment 3), the observed pattern of effects was obtained with syllabically simple words. For this experiment, it was hypothesized that the more extreme technique for splitting the words would successfully disrupt the normal responding of the low-sensitivity participants. Indeed, when the word-initial unit contrasted poorly against the background, the asymmetrical results reported in experiment 1 & 2 was mirrored by a low-sensitivity participant preference for Max Coda division with low-frequency words. Taken together, these findings extend previous work by linking previous individual-difference-based accounts of syllable preference (i.e., Taft, 2001, 2002; Chen & Vaid, 2007) to a process-based phonological variable that is meaningfully interpreted in terms of extant computational models. We now turn to more specific questions related to the results.

Basic Item-Level Effects

The standard frequency effect (Monsell, 1991; Norris, 2006) was obtained despite the introduction of an artificial boundary with the stimuli. This suggests that at a certain level the introduction of even a highly disruptive artificial boundary (as in Experiment 3) did not qualitatively disrupt normal word recognition processing. Second, the homophone interference effect was replicated in all three experiments, which provides independent confirmation of the work of Pexman and colleagues (Pexman et al., 2001; Pexman et al, 2002; Unsworth & Pexman, 2003). Overall, the instability of word-splitting effects was generally confirmed, with a direct effect of word-splitting observed in Experiment 1 but not 2 (Experiment 3 was not a fair item-level test). Various indicators suggested that the sample from Experiment 1 was relatively less skilled (i.e., fewer years of University, homophone effect on error rate), which would account for the word-level syllable effect observed with this sample (Taft, 2001). Importantly, it was discovered that the interaction among item- and participant-level variables could account for such variability in syllable preference.

Individual Differences and Syllable Preference

The central question of the present research concerned whether syllable preference was reliably associated with estimates of phonological reliance. Overall, syllable preference varied consistently with sensitivity to feedback phonology (i.e., homophone interference) rather than general phonological reliance (i.e., regularity effect). Whereas Taft (2001) reported a symmetrical interaction whereby high-skill participants preferred Max Coda division and low-skill participants preferred Max Onset division, the present research only managed to consistently replicate the second half of that pattern using an estimate of phonological reliance. This finding suggests that insensitivity to the homophone effect is not a good predictor of Max Coda preference, which was only obtained here under low-contrast conditions. It remains to be seen whether other indictors might more reliably associate with a Max Coda advantage. Possible process-based candidates include indicators of sensitivity to semantic attributes (e.g., imageability; James, 1975) and orthographic attributes (e.g., orthographic N, Coltheart et al., 1977).

Chen & Vaid (2007) found that Max Coda preference was observed only for low frequency words. Experiment 3 confirms this finding while extending such preference to participants who are insensitive to homophony. Moreover, all three experiments extend the frequency-dependence of the syllable effect to include Max Onset syllable preference. Overall, the observed patterns of syllable preference were quite consistent, but depended on syllabic complexity in an inconsistent way depending on the way stimuli were divided. It is unclear why syllable preference effects would be observed with complex stimuli in the first two experiments and with simple stimuli in the final experiment. Syllabically complex stimuli tended to be longer in terms of number of letters. This property may have interacted with the type of visual segmentation to generate observable effects.

Despite the various ways in which the direct relationship between phonology and syllable preference is qualified, the findings reported here allow a certain kind of account of Taft (2001, 2002)'s results to be discounted. Specifically, it does not appear to be the case that syllabification is an orthographic process, as least not Max Onset syllabification. Rather, Max Onset syllable preference is associated with overall reliance on phonological feedback in lexical decision. This suggests that Max Onset preference is an emergent property of feedback from the phonological domain. This finding provides a potentially valuable constraint for the extension of existing computational models to multi-syllable words. In the DRC model, for example, the homophone effect is necessarily lexical feedback. Within such models, the syllable boundary could either be represented directly in the lexicon (e.g., via abstract syllable structure) or indirectly via the activation of lexical cohorts that share the same syllable (e.g., via feedback from the phonetic processing level). Presumably a PDP model would have to encode elements of syllabic structure in its input (e.g., cvc structure) and output units at the very least, if not devote an entire pool of units to representing the syllable. However the syllable is represented, the effect should manifests itself behaviorally when feedback is strong enough and/or when other sources of information for making a lexical decision are deemphasized. Note that the present results do not prove conclusively that low-sensitivity readers fail to represent the Max Onset syllable, only that such preference is not observed empirically. This null result raises the questions: a) what information are low-sensitivity participants relying on if not phonology and b) what accounts for their shift away from phonology or, perhaps more accurately, their recruitment of other types of information to support the word recognition process.

Phonology and Semantics: Shifting Emphasis

In regards to the first question, there are at least two possibilities given existing frameworks for understanding word recognition (e.g. Coltheart et al., 2001; Harm &

Seidenberg, 2004): orthography and semantics. Taft (2001) operationally defined skill in terms of comprehension ability on a speed task. Good performance on such a task will positively correlate with rapid recognition of orthographic patterns and efficient retrieval of semantic information. The so-called 'good readers' may prefer Max Coda division either because they syllabify words prior to recognition of orthographic patterns (Taft, 1979, 2001) and/or because Max Coda units emerge as a property of the interaction between orthographic processes and semantic retrieval (Taft, 2001). Experiment 3 of this paper demonstrated the participants who are relatively insensitive to phonological feedback prefer Max Coda division. This finding is consistent with the idea that the Max Coda syllable is not a phonological unit in lexical decision (though it may be in naming, Taft, 1992; possibly as an epiphenomenon, Chateau & Jared, 2003), but does address what its possible correlates are. Future work may link Max Coda preference more specifically to an alternative mode of processing.

At least one prominent model of word reading and recognition (i.e. Harm & Seidenberg, 2004; see also early expressions of similar ideas, Seidenberg & McClelland, 1989; Doctor & Coltheart, 1980) has modeled a developmental shift from heavy reliance on phonology in reading to reliance on other modes of reading, specifically direct orthography-to-semantics processing, both feedforward and feedback. Harm & Seidenberg (2004) note that the primary purpose of reading is the extraction of meaning, and for this purpose the indirect route (orthography-to-phonology-to-meaning) is slower than processing involving direct connections between orthography and semantics. With sufficient training, a network that is motivated to be as efficient as possible (i.e., to be 'greedy for activation') will shift the balance of processing from a phonology-based system to one that incorporates a heavy orthographic-semantic component. If we assume that this shift is in different stages of progression across a population of adult readers, then the pattern of syllable preference results observed here may be explained. Participants whose shift to direct semantic retrieval is relatively incomplete will be more sensitive to the phonological characteristics of words, including their (phonological) syllabic structure. Conversely, participants whose shift to direct semantic retrieval is relatively complete are correspondingly insensitive to phonological characteristics. Indeed, their relatively multi-faceted system may be generally less vulnerable to many kinds of inefficiencies whether they are orthographic, semantic, or phonological (Chateau & Jared, 2003; Sears et al., 2008; Unsworth & Pexman, 2003). Future research using a broader set of indicators (orthographic, semantic) may shed more light onto how the balance of different types of information conspires to determine observed syllable preference. Bowers & Kirby (2007) report the results of an intervention demonstrating that even among skilled readers there is room for substantial improvement in morphological awareness (awareness of the relationship between orthographic patterns and semantics), suggesting that enough individual differences exist along this dimension to produce concomitant variation in 'orthographic' syllable preference. It is possible that an intervention (e.g., morphological training) could induce a corresponding shift from Max Onset to Max Coda preference.

Conclusions

The present research extends the results reported by Taft (2001) by establishing that sensitivity to feedback phonology during visual word recognition is associated with observable patterns of syllable preference. For low frequency words, participants who rely on phonology tend to perform better when words are segmented at the boundary defined by maximal onset segmentation. Further, under some conditions, participants who are relatively insensitive to phonology tend to make faster responses to low-frequency words when the word-initial syllable is taken to end at the boundary defined by Max-Coda rule. These findings are cautiously interpreted within a framework that assumes a gradual shift from a word recognition strategy that recruits activation heavily from phonology (leading to Max-Onset preference) to a more diverse strategy that can lead to other patterns of results (e.g. preference for the more morpho-semantically consistent Max-Coda syllable) that is in various stages of progression across the population of University-level readers.

References

Alegre, M. & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41-61.

Anderson, J. & Jones, C. (1974). Three theses concerning phonological representations. *Journal of Linguistics*, *10*, 1-26.

Ashby, J. & Martin, A.E. (2008). Prosodic phonological representations early in visual word recognition. Journal of Experimental Psychology: Human Perception & Performance, *34*, 224-236.

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*, 390-412.

Baayen, R.H., Piepenbrook, R., Gulikers, L. (1995). *The CELEX Lexical Database*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia.

Balota, D.A., Cortese, M.J., Sergent-Marshall, S.D., Spieler, D.H., & Yap, M.J. (2004). Visual word recognition of single-syllable words, *Journal of Experimental Psychology: General*, *133*, 283-316.

Balota, D. A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J. J., Nelson, D.L., Simpsons, G.B., and Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445-459.

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

Brand, M., Rey, A., & Peereman, R. (2003). Where is the syllable priming effect in visual word recognition? *Journal of Memory and Language*, *48*, 435-443.

Blevins, J. (1995). The syllable in phonological theory. In J.A. Goldsmith (Ed.), *The* *handbook of phonological theory* (pp. 206-244). Cambridge, MA: Blackwell.

Bowers, P. & Kirby, J. (2007, June). Morphology in Reading and Spelling. Paper presented at the Sixth Annual Meeting of the Canadian Language and Literacy Research Network.

Bybee, J. (2001*). Phonology and language use.* Cambridge, England: Cambridge University Press.

Carr, P. (2000). *English phonetics and phonology: an introduction.* Oxford: Blackwell.

Carreiras, M., Alvarez, C.J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32*, 766-780.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28*, 143-153.

Chateau, D. & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language*, *48*, 255-280.

Chen, H.-C. & Vaid, J. (2007). Word frequency interacts with the BOSS in orthographic processing of English polysyllable words. *Language and Cognitive Processes, 22,* 52-82.

Chetail, F., & Mathey, S. (2009). Syllabic priming in lexical decision and naming tasks: The syllable congruency effect re-examined in French. *Canadian Journal of Experimental Psychology, 63*, 40-48.

Clements, G.N., & Keyser, S.J. (1983). *CV Phonology*. Cambridge, MA: MIT Press.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlational analysis for the behavioral sciences.* (Third ed.) New Jersey: Lawrence Erlbaum Associates.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud.

Psychological Review, 108, 204-256.

Doctor, E.A., and Coltheart, M. (1980). Phonological recoding in children's reading for meaning. *Memory and Cognition, 80*, 195-209.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31*, 119-122.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology* , *47*A, 365-382.

Ferrand, L., Segui, J., & Humphreys, G. (1997). The syllable's role in word naming. *Memory & Cognition, 25*, 458-470.

Grainger, J., & Jacobs, A.M. (1993). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518-565.

Harm, M. & Seidenberg, M. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Hino, Y., & Lupker, S.J. (1996). Effects of polysemy in lexical decision and naming : an alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1331-1356.

Hooper, J. (1972). The syllable in phonological theory. *Language, 48*, 525-540.

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434-446.

James, C.T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 104*, 130–136.

Jared, D., Levy, B., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General, 128*, 219-264.

Jared, D., & Seidenberg, M. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 92-105.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120*, 358-394.

Jordan, T. (1986). Testing the BOSS hypothesis: Evidence for position-insensitive orthographic priming in the lexical decision task. *Memory & Cognition, 14*, 523-532.

Katz, L., & Baldasare, J. (1983). Syllable coding in printed-word recognition by children and adults. *Journal of Educational Psychology, 75*, 245-256.

Kerswell, L., Siakaluk, P.D., Pexman, P.M., Sears, C.R., & Owen, W.J. (2007). Homophone effects in visual word recognition depend on homophone type and task demands. *Canadian Journal of Experimental Psychology, 61*, 322-327.

Lewellen, M.J., Goldinger, S.D., Pisoni, D.B., & Greene, B.G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization, *Journal of Experimental Psychology: General, 122*, 316-330.

Lima, S. D., & Pollatsek, A. (1983). Lexical access via orthographic code? The Basic Orthographic Syllabic Structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior, 22*, 310-332.

Locker, L., Hoffman, L., & Bovaird, J. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research*

Methods, 39, 723-730.

Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science*, 161-171.

Loftus, G.R. & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476-490.

Lorch, R. & Myers, J. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 149-157.

Lupker, S. J., Brown, P., & Columbo, L. (1997). Strategic control in a naming: changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 570-590.

Macizo, P., & Van Petten, C. (2007). Syllable frequency in lexical decision and naming of English words. *Reading & Writing, 20*, 295-331.

Merriam-Webster Online. *Searchable dictionary.* Retrieved in the fall of 2003, from http://www.m-w.com/

Meunier, F., & Segui, J. (1999). Morphological priming effect: The role of surface frequency. *Brain and Language, 68*, 54-60.

Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *Quarterly Journal of Experimental Psychology, 43*A, 907-912.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148-197). London, England: Lawrence Erlbaum Associates.

Nation, P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.) *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). New York: Cambridge University Press.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113*, 327-357.

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short SOAs in lexical decision and naming. *Cognition, 62,* 223-240.

Pexman, P.M., Lupker, S.J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 139-156.

Pexman, P.M., Lupker, S.J., & Reggin, L.D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 572-584.

Plaut, D., McClelland, J., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading : Computational principles in quasiregular domains. *Psychological Review*, *103*, 56-115.

Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus*. The Hague: Mouton.

Rapp, B. (1992). The nature of sublexical orthographic organization: the bigram trough hypothesis examined. *Journal of Memory and Language, 31*, 33-53.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language*, *42*, 342-364.

Rastle, K., Kinoshita, S., Lupker, S. J. (2003). Cross-task strategic effects. *Memory and Cognition,* 31, 867-876.

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

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

Rossell, S. L., Price, C.J., and Nobre, A.C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia, 41*, 550-64.

Rouibahh, A., & Taft, M. (2001). The role of syllabic structure in French visual word recognition. *Memory & Cognition, 29*, 373-381.

Schiller, N. (1999). Masked syllable priming of English nouns. *Brain and Language, 68*, 300-305.

Schiller, N. (2000). Single word production in English. *Journal of Experimental Psychology:* Learning, Memory, and Cognition, 26, 512-528.

Sears, C. R., Siakaluk, P. D., Chow, V. C., & Buchanan, L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research, 37*, 269-291.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention & Performance XII* (pp.245-263). Hillsdale, NJ: Erlbaum.

Seidenberg, M. S., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Selkirk, E. (1982). The syllable. In H. Van der Hulst & N. smith (Eds.), *The structure of phonological representations* (pp. 337-383). Dordrecht: Foris.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402-433.

Spencer, A. (1996). *Phonology.* Oxford, U.K.: Blackwell Publishers.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53*, 347-359.

Taft, M. (1979). Lexical Access via an orthographic Code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behaviour, 18*, 21-39.

Taft, M. (1987). Morphographic processing: The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and performance, XII: Reading* (pp. 265-299). Hillsdale, NJ: Erlbaum.

Taft, M. (1992). The body of the BOSS: Subsyllabic units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1004-1014.

Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language & Speech, 44*, 351-376.

Taft, M. (2002). Orthographic processing of polysyllabic words by native and nonnative English speakers. *Brain & Language, 81*, 532-544.

Taft, M., & Forster, K. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15*, 607-620.

Taft, M. & Radeau, M. (1995). The influence of the phonological characteristics of a language on the functional units of reading: A study in French. *Canadian Journal of Experimental Psychology*, *49*, 330-346.

Tse, C-S. & Neely, J.H. (2007). Semantic priming from letter-searched primes occurs for low- but not high-frequency targets: Automatic semantic access may not be a myth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1143-1161.

Thompson, G. L. (2008). Eliminating aggregation bias in experimental research: Random coefficient analysis as an alternative to performing a 'by-subjects' and/or 'by-items' ANOVA. *Tutorials in Quantitative Methods for Psychology, 4,* 21-34.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of* Memory and Language, 25, 476-491.

Treiman, R. & Zukowski, A. (1990). Towards an understanding of English syllabification. *Journal of Memory and Language*, *29*, 66-85.

Unsworth, S.J. & Pexman, P.M. (2003). The impact of reader skill on phonological processing in visual word recognition. *The Quarterly Journal of Experimental Psychology*, *56A*, 63-81.

Van Orden, G.C. (1987). A rows is a rose : spelling, sound and reading. *Memory and Cognition*, *15*, 181-198.

Yap, M. J., Balota, D. A., Tse, C-S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: Evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34*, 495-513.

Appendix A

List of word-splitting task words for Experiment 1

Complex Words					
List A		List B		List C	
turban	tambourine	vortex	vampire	yonder	turmoil
tampon	timber	curfew	sultan	gargoyle	cactus
quagmire	hernia	wampum	snorkel	dolphin	dorsal
morphine	kernel	vendetta	morsel	garlic	tantamount
custard	gorgon	septic	harbinger	tertiary	tundra
bastion	dalmatian	gargantuan	magnum	sphagnum	symposium
maestro	vulpine	tungsten	pectoral	sardine	blunder
dandelion	herpes	balsam	lantern	vulture	swastika
gospel	cauldron	walnut	zombie	torpedo	bongos
versus	circus	tantrum	nectar	seldom	nostril
Simple Words					
List A		List B		List C	
stadium	mosaic	lemur	lesion	crouton	feral
cypress	junior	bohemian	rhubarb	creature	tirade
tulip	viper	typhoid	tunic	nubile	bonus
gonad	radar	locust	juvenile	serum	cedar
ceiling	blatant	slogan	sucrose	femur	cider
molar	zebra	bovine	jujube	glucose	sodium
diesel	jujitsu	mogul	zodiac	vinyl	zucchini
kibosh	cadence	stipend	brochure	vitamin	libation
cucumber	scrotum	crisis	lilac	feline	tiger
suture	cougar	muzak	mucus	crocus	juniper

Appendix B

List of word-splitting task nonwords for Experiment 1

Complex Nonwords					
lindainite	tisten	tampine	torvix	polthen	jander
loastate	santen	tulpan	kurtij	slunden	gastoile
jastua	mestin	slirkel	sabtem	dersat	nurtual
pispol	polmosion	porponsein	fectin	tondomeint	bympetium
dergen	cortist	lirbonter	tuskis	lartik	tirteira
larthine	birneon	semnis	torset	bastri	sphignot
milpone	leustro	terptiral	langsten	borpeto	sordite
tantotion	horpet	bonwern	bilsan	snastita	multune
keistran	dornil	zalboe	wastud	tandis	salnom
tercus	vernun	pyndyssa	nendron	nesdrol	lundas
Simple Nonwords					
tiekel	plodeum	siteon	lomin	trucese	rodir
luniot	tiprett	thutart	pahitian	torate	jughtana
ceitock	glidint	tokip	pithoid	cidut	labite
todor	strigum	tavetone	pekist	fomar	simam
jutip	vepar	satrane	slukin	jitoner	majut
lotra	tomar	bababe	katoit	tracut	gilan
tatitra	topeec	nitine	nigyl	treipure	zinet
bamonce	biroth	brathire	trakend	lobasion	bikanon
gytod	lulonter	tojak	srosis	miper	pamile
saiger	tasare	wikis	vavuc	prewdon	makial

Appendix C

Phonological indicator stimuli borrowed from Pexman et al (2002).

Homophone	Homophone Controls	Regular	Regular controls
blew	baked	bowl	beam
bored	bleed	worm	Yell
brake	boil	wool	wick
coarse	cheese	doll	Stab
feat	deed	warn	wink
hare	flip	pint	dock
haul	hack	deaf	dusk
hire	heap	bush	Rust
ladder	hoop	comb	Sank
leased	locate	wasp	hunt
maid	loomed		
mane	maze		
mined	mess		
mourning	mounting		
reed	mused		
reel	rude		
sole	seal		

Appendix D

List of word-splitting task words for Experiments 2A & 3A

Complex syllable structure

List A		List B		List C	
balsam	barbecue	arsonist	cauldron	bonkers	biscuit
bamboozle	census	bambino	charlatan	brontosaurus	blunder
burlap	gargantuan	bantam	dolphin	calvary	boulder
comport	hamper	banter	garlic	constrict	complexion
corvette	hurdle	bolster	garnish	corsage	compound
gander	martini	canker	garter	culvert	garden
gorgon	mercenary	charnel	gender	gargoyle	harlot
harbinger	mercury	dactyl	harbour	garland	kerchief
holster	mundane	garner	harness	gypsum	kernel
kleptomaniac	pectoral	garnet	hectic	kumquat	martial
lactose	sardonic	jamboree	lactic	minstrel	martyr
larceny	septic	magnum	marquis	mongoose	pilgrim
mongrel	sermon	marten	mermaid	mordant	salmon
nuptial	swastika	numskull	morphine	nectarine	scepter
pilfer	turban	pontoon	purple	peptic	scoundrel
purloin	turmoil	quagmire	serpent	sphagnum	sordid
tambourine	vendetta	sardine	sponsor	tampon	sultan
tundra	vernacular	snorkel	strychnine	turquoise	turpentine
varsity	versus	turgid	tertiary	vandal	turtle
yonder	vortex	varmint	walnut	wampum	vector

Simple syllable structure

List A		List B		List C	
boron	auburn	bebop	beagle	auger	boudoir
brouhaha	beacon	butane	beaver	caudal	caucus
bucolic	boutique	coupon	boulevard	cubit	cedar
cougar	brunette	couture	cupid	eructation	crucial
crouton	cereal	crusade	deacon	fecund	cucumber
culotte	easel	cuneiform	fever	femur	epoch
cuticle	ecumenical	demon	frugal	goulash	faucet
feline	edict	fluvial	glucose	gubernatorial	genius
glaucoma	evil	geezer	humus	jujube	harem
humerus	helium	jihad	Kleenex	lemur	hubris
lather	hooligan	jukebox	lethal	lesion	Jupiter
muzak	junior	julep	moron	mucous	juvenile
nougat	juniper	kudos	peacock	mural	lager
oriole	lucid	lupin	plenary	orangutan	legion
pecan	lugubrious	munificent	pupil	pugilist	mausoleum
rubric	maharajah	munition	putative	quorum	raucous
sequin	paucity	oodles	souvenir	slalom	rhubarb
suture	treason	rhesus	thorax	sumac	specious
velum	tunic	tautology	Venus	troubadour	tulip
yuletide	vehement	toucan	weasel	venal	zebra

Appendix E

List of word-splitting task nonwords for Experiments 2A & 3A

Complex Syllable Structure

bactam	gurnor	mentaid	pormant	shirnel	trintisairas
bardut	gypsum	mertacilur	porquoise	snorkel	trunder
bartenue	hantic	mertono	quastire	sornin	tulvern
bentoral	herbaur	milgrem	ramtal	sorthine	tumksitt
bertiary	hindle	mirtury	ramzoodle	spensot	tundra
bontir	homstens	misten	ranzer	sphignit	turgot
burmoit	horbanger	mortax	recturine	strechtite	turnat
buscout	horlin	mulster	relmon	swontifa	ursonant
caindron	jombosee	mumbaurune	sactik	talphin	vantor
calcasy	kontars	muptiel	samner	tanton	vensity
cermint	kornot	musquam	samter	tasmess	vindessa
cinmer	kundul	namdine	santle	tercetary	voulter
cirvunne	lensas	narteny	santuno	terpend	watnum
conbletian	lintase	nomteind	sarmin	therleten	wenper
fardonit	lomtort	nultam	sarquit	tirsage	wenteen
garland	marner	nustip	sarsoybe	tognot	wursas
gergentuin	masdyr	paptuc	sartyt	tonstrint	yanden
gingin	mastian	pentic	scawndril	tonter	zlintanetail
girlec	mendone	pershief	scintre	tontrel	zurlian
gurnesh	mengeete	pimpom	senstrel	tortantine	zurtle

Simple Syllable Structure

aumer	ewtres	jovenate	paucus	sobot	trugel
auturg	fashit	jutin	peenate	soulath	truzial
beasock	ferane	kelun	penion	soupom	tubilist
beazer	fivop	lagabriant	Pleebex	srilade	tulebide
beedair	flupose	lamid	pobative	stenious	tunolic
bezan	funep	lenut	pougat	sunitícant	tuvous
breagon	geetigan	leshal	putid	taitonagy	usil
broutadair	geterbatunial	lurac	putune	taliper	Vanum
bunauform	gloutona	mabataree	quofim	tarel	vearox
bunibion	goomit	moutan	rablic	taucat	veasen
cager	goozir	nausolian	rhaberb	tekund	vehamund
cautire	grunesse	nenot	rhisis	temal	venur
celon	guneas	ninibe	roubanir	tequan	verean
cumonne	haliun	nooletard	ruferus	tohad	vetra
degiot	heatle	nubic	rupal	torun	voutoque
denar	humad	nulib	rutacle	totamber	vudis
eabal	hutrig	ocumanigal	setine	toumar	vurat
epack	irestation	oliale	shorex	trenary	waumous
eringonan	japior	Pabiger	slumial	troonana	weavon
etind	jatebot	pahnal	snarom	troufon	wutak

Authors Note

Glenn L. Thompson and Alain Desrochers, School of Psychology, University of Ottawa, Ottawa, Ontario, Canada.

The research reported in this article was supported by doctoral scholarships awarded to the first author from the Natural Sciences and Engineering Research Council (NSERC) and by a research grant from the same agency awarded to the second author.

Correspondence should be addressed to Glenn Thompson at the School of Psychology, University of Ottawa, 145 Jean-Jacques Lussier, P.O. Box 450, Station A, Ottawa, Canada, K1N 6N5, or e-mail: GlennLThompson@gmail.com.

Notes

[2] Logically, interaction terms that are included in the model must be accompanied by all associated lower-order effects. Otherwise, the interaction effect is confounded with its main effects.

Table 1. Stimulus list characteristics for word-splitting tasks, experiments 1, 2, and 3

Exp 1	Simple Syllable Structure, n = 60					Complex Syllable Structure, n = 60				
List	N Let	N Phon	Log F	Orth N	Bi-Freq	N Let	N Phon	Log F	Orth N	Bi-Freq
A, n=40	6.6(1.10)	5.75(1.33)	6.48(1.66)	.70(1.17)	3843(1270)	6.55 (1.27)	5.6 (1.14)	7.07(1.86)	.50(1.00)	3746(1367)
B, n=40	6.5(1.23)	6.15(1.38)	6.39(1.35)	.45(0.76)	3015(1398)	6.55 (1.10)	5.75 (0.97)	6.33(1.71)	.40(0.99)	3215(1721)
C, n=40	6.6(0.99)	5.65(0.99)	6.70(1.75)	.65(0.99)	3709(1514)	6.6 (1.43)	5.9 (1.45)	6.86(1.43)	.95(1.99)	3908(1452)
NW, n=120	6.15(1.02)	5.85(1.25)	NA	.20(1.05)	1715(706)	7.00 (1.19)	5.75 (1.20)	NA	.18(.70)	1989(781)
Exp 2 & 3	Simple Syllable Structure, n = 120					Complex Syllable Structure, n = 120				
List	N Let	N Phon	Log F	Orth N	Bi-Freq	N Let	N Phon	Log F	Orth N	Bi-Freq
A, n=80	6.58(1.39)	5.72(1.51)	5.55(1.89)	.68(1.42)	3423(1298)	7.27(1.45)	6.55(1.61)	5.89(1.60)	.85(1.69)	3818(1222)
B, n=80	6.38(1.33)	5.82(1.48)	6.35(1.48)	.80(1.70)	3020(1294)	6.90(0.96)	6.02(0.89)	6.38(1.55)	1.28(2.28)	4010(1498)
C, n=80	6.55(1.77)	5.69(1.70)	5.72(1.78)	1.03(2.08)	3171(1336)	7.28(1.39)	6.23(1.42)	6.15(1.91)	0.93(1.52)	3446(1200)
NW, n=240	6.51(1.50)	5.74(1.56)	NA	.28 (0.98)	1748 (688)	7.15 (1.29)	6.27(1.34)	NA	.25 (0.64)	2007 (763)

Note. According to a counterbalancing scheme over participants, lists A, B, and C were assigned to Max-Onset, Max-Coda or Max Coda+1. Lists D and E were assigned to either Max-Onset or Max-Coda. Parentheses denote standard deviations. N Let and N Phon refer to number of letters and phonemes. Log F,Orth N, and Bi-Freq refer to the log HAL frequency, orthographic N, and mean bigram frequency (Balota et al, 2007).

Table 2. Mixed-model regression of Experiment 1 lexical decision latencies as a function of item and participant characteristics.

				HPD 95% *CI*[b]		Random Effects
	b	*SE*	*t*[a]	Lower	Upper	
Intercept	631.27	22.57		597.38	665.05	
Word-Level						**S^2=3886.4 S=62.34**
Contrast A (1A)	7.32	6.14	1.19	-5.26	19.15	
Contrast B (1B)	1.95	7.10	0.28	-11.99	15.38	
Log F (2)	-84.91	15.90	5.34†	-112.50	-54.14	
Complexity (3)	-2.16	12.76	0.17	-25.02	21.49	
1A x 2	-9.94	14.97	0.87	-40.06	19.65	
1A x 3	26.65	12.28	2.17*	1.91	50.10	
1B x 2	-32.37	17.14	1.89[m]	-66.42	2.07	
1B x 3	12.29	14.20	0.87	-16.38	39.65	
2 x 3	1.55	31.80	0.05	-55.66	61.86	
1A x 2 x 3	-30.13	29.99	1.01	-88.01	29.16	
1B x 2 x 3	-0.54	34.34	0.02	-67.77	68.33	
Participant-Level						**S^2=22510 S=150.03**
Homophone Effect (4)	1.64	0.62	2.61†	0.70	2.55	
4 x 1A	-0.06	.18	0.34	-0.42	0.29	
4 x 2	-0.68	0.17	3.88†	-1.03	-0.33	
4 x 3	0.12	.14	0.86	-0.16	0.40	
4 x 1A x 2	-0.79	.44	1.80[m]	-1.67	0.04	
4 x 1A x 3	.29	.36	0.82	-0.40	0.99	
4 x 2 x 3	0.13	.35	0.37	-0.62	0.78	
4 x 1A x 2 x 3	-1.58	.088	1.77[m]	-3.27	0.11	
Residual						**S^2=28193 S=167.91**

Note. Results based on 4666 datapoints, which are based on N = 48 and Items = 120. [a] Statistical significance determined using the default values of the Markov chain Monte Carlo (*MCMC*) algorithm, R version 2.8.1 (for description, see Baayen, Davidson, & Bates, 2008). [b] Confidence interval values generated using *MCMC* Highest Posterior Density procedure. The inclusion of the four-way interaction involving Contrast B does not affect the four-way interaction reported here. [m] $p < .10$, *$p < .05$, †$p < .001$.

Table 3. Mixed-model logistic regression of Experiment 1 Error rates as a function of item and participant characteristics.

	logit	*SE*	*Z*	Random Effects
Intercept	-2.33	0.20	-11.60	
Word-Level				S^2=2.43 S=0.87
Contrast A (1A)	-0.06	0.11	-.56	
Contrast B (1B)	0.03	0.13	0.24	
Log F (2)	-2.14	0.42	-5.13†	
Complexity (3)	-0.11	0.31	-0.35	
1A x 2	-0.13	0.38	-0.35	
1A x 3	-0.17	0.23	-0.74	
1B x 2	-0.31	0.44	-0.72	
1B x 3	0.13	0.27	0.50	
2 x 3	0.73	0.83	0.88	
1A x 2 x 3	-0.72	0.76	-0.95	
1B x 2 x 3	1.11	0.88	1.26	
Participant-Level				S^2=0.76 S=0.87
Homophone Effect (4)	0.003	0.004	0.81	
4 x 1A	-0.002	0.003	-0.65	
4 x 2	-0.007	0.004	-1.61	
4 x 3	0.002	.003	0.97	
4 x 1A x 2	0.01	0.01	1.21	
4 x 1A x 3	-0.01	0.007	-1.49	
4 x 2 x 3	0.01	0.009	1.44	
4 x 1A x 2 x 3	-0.007	0.02	-0.31	

Note. Mean estimated probability of making an error was .18 (i.e., error rate). Results based on 5,760 data points, which are based on N = 48 and Items = 120. $^{m}p < .10$, $^{*}p < .05$, $^{\dagger}p < .001$

Table 4. Mixed-model regression of Experiment 2 lexical-decision latencies as a function of item and participant characteristics.

	b	SE	t[a]	*HPD* 95% *CI*[b] Lower	Upper	Random Effects
Intercept	645.54	11.78		623.19	666.41	
Word-Level						**S^2=3766.4 S=61.37**
Contrast A (1A)	0.85	4.48	0.19	-7.56	9.95	
Contrast B (1B)	2.97	5.13	0.58	-6.99	13.17	
Log F (2)	-34.84	2.58	13.51†	-39.49	-30.18	
Complexity (3)	-4.81	8.85	0.54	-20.30	11.16	
1A x 2	-1.60	2.67	0.60	-6.70	3.79	
1A x 3	4.16	8.96	0.46	-13.83	21.49	
1B x 2	3.16	3.10	1.02	-3.18	9.06	
1B x 3	-3.44	10.27	0.33	-23.52	16.54	
2 x 3	5.62	5.15	1.09	-4.06	14.62	
1A x 2 x 3	-6.11	5.33	1.15	-16.87	3.98	
1B x 2 x 3	1.92	6.17	0.31	-10.18	13.95	
Participant-Level						**S^2=5720.4 S=75.63**
Homophone Effect (4)	0.64	0.21	3.05†	0.26	1.02	
4 x 1A	0.14	0.09	1.63	-0.02	0.31	
4 x 2	-0.10	0.02	5.03†	-0.15	-0.07	
4 x 3	-0.06	0.07	0.84	-0.20	0.08	
4 x 1A x 2	-0.07	0.05	1.45	-0.19	0.02	
4 x 1A x 3	-0.02	0.17	0.14	-0.36	0.33	
4 x 2 x 3	0.08	0.04	2.02*	0.003	0.16	
4 x 1A x 2 x 3	-0.25	0.10	-2.39**	-0.45	-0.04	
Residual						**S^2=25138.6 S=158.55**

Note. Results based on 7,708 data points, which are based on N = 48 and Items = 239. [a] Statistical significance determined using the default values of the Markov chain Monte Carlo (*MCMC*) algorithm, R version 2.8.1 (for description, see Baayen, Davidson, & Bates, 2008). [b] Confidence interval values generated using *MCMC* Highest Posterior Density procedure. The inclusion of the four-way interaction involving Contrast B does not affect the four-way interaction reported here. $*p < .05$, $**p < .01$, $^{\dagger}p < .001$.

Table 5. Mixed-model logistic regression of Experiment 2 Error rates as a function of item and participant characteristics.

	logit	*SE*	*Z*	Random Effects
Intercept	-1.83	0.13	-14.26†	
Word-Level				S^2=1.79 S=1.34
Contrast A (1A)	-0.08	0.07	-1.05	
Contrast B (1B)	-0.09	0.09	-1.04	
Log F (2)	-0.58	0.06	-9.99†	
Complexity (3)	-0.23	0.19	-1.20	
1A x 2	0.01	0.04	0.33	
1A x 3	-0.10	0.15	-0.65	
1B x 2	-0.06	0.05	-1.08	
1B x 3	-0.06	0.17	-0.35	
2 x 3	0.15	0.11	1.32	
1A x 2 x 3	0.03	0.09	0.28	
1B x 2 x 3	-0.11	0.11	-1.02	
Participant-Level				S^2=0.36 S=0.60
Homophone Effect (4)	0.004	0.002	2.16*	
4 x 1A				
4 x 2	-0.0001	0.0003	-0.53	
4 x 3	-0.0004	0.001	-0.26	
4 x 1A x 2	0.0007	0.0009	0.78	
4 x 1A x 3	-0.002	0.003	-0.68	
4 x 2 x 3	-0.0003	0.0007	-0.42	
4 x 1A x 2 x 3	0.003	0.002	2.01*	

Note. Mean estimated probability of making an error was .24 (e.i., error rate). Results based on 10,141 datapoints, including $N = 48$ and Items = 240. $^{m}p < .10$, $*p < .05$, $^{\dagger}p < .001$.

Table 6. Mixed-model regression of Experiment 3 lexical-decision latencies as a function of item and participant characteristics.

				HPD 95% *CI*[b]		Random Effects
	b	*SE*	*t*[a]	Lower	Upper	
Intercept	819.54	25.72	31.87	776.22	862.42	
Word-Level						**S^2=9223.6 S=96.04**
Contrast A (1A)	38.63	6.89	5.61†	25.40	52.10	
Contrast B (1B)	66.66	7.95	8.39†	51.44	82.52	
Log F (2)	-48.94	4.06	12.06†	-56.25	-41.44	
Complexity (3)	-15.19	13.90	1.09	-37.17	10.98	
1A x 2	-1.34	4.12	0.33	-9.47	6.47	
1A x 3	-0.58	13.78	0.04	-27.18	26.52	
1B x 2	-13.37	4.78	2.80**	-22.46	-3.96	
1B x 3	24.32	15.89	1.53	57.74	0.13	
2 x 3	6.99	8.11	0.86	-7.78	21.19	
1A x 2 x 3	-7.87	8.20	0.96	-24.04	7.96	
1B x 2 x 3	-12.11	9.52	1.27	-29.97	6.97	
Participant-Level						**S^2=29442 S=171.59**
Homophone Effect (4)	0.36	0.38	0.96	-0.27	0.98	
4 x 1A	0.36	0.11	3.40†	0.15	0.57	
4 x 2	0.29	0.09	3.36*	-0.10	-0.002	
4 x 3	-0.05	0.03	-2.00†	0.12	0.46	
4 x 1A x 2	-0.27	0.06	-4.27†	-0.40	-0.14	
4 x 1A x 3	-0.75	0.21	-3.49†	-1.19	-0.34	
4 x 2 x 3	-0.04	0.05	-0.78	-0.14	0.06	
4 x 1A x 2 x 3	0.53	0.13	4.10†	0.28	0.79	
Residual						**S^2=69334.6 S=263.31**

Note. Results based on 8,913 data points, which are based on $N = 48$ and Items = 237. [a] Statistical significance determined using the default values of the Markov chain Monte Carlo (*MCMC*) algorithm, R version 2.8.1 (for description, see Baayen, Davidson, & Bates, 2008). [b] Confidence interval values generated using *MCMC* Highest Posterior Density procedure. The inclusion of the four-way interaction involving Contrast B does not affect the four-way interaction reported here. $*p < .05$, $**p < .01$, $^{\dagger}p < .001$.

Table 7. Mixed-model logistic regression of Experiment 3 Error rates as a function of item and participant characteristics.

	logit	*SE*	*Z*	Random Effects
Intercept	-1.87	1.09	-17.16†	
Word-Level				S^2=1.39 S=1.18
Contrast A (1A)	0.04	0.07	0.63	
Contrast B (1B)	-0.05	0.08	-0.65	
Log F (2)	-.53	0.05	-10.06†	
Complexity (3)	-.50	0.16	-2.38	
1A x 2	0.02	0.04	0.59	
1A x 3	0.05	0.13	0.32	
1B x 2	-0.13	0.05	-2.64**	
1B x 3	0.04	0.02	0.25	
2 x 3	0.06	0.10	0.58	
1A x 2 x 3	-0.07	0.08	-0.86	
1B x 2 x 3	-.11	0.09	-1.10	
Participant-Level				S^2=0.23 S=0.48
Homophone Effect (4)	0.0008	0.001	-0.76	
4 x 1A	0.002	0.001	2.33*	
4 x 2	0.0004	0.0003	1.48	
4 x 3	0.0006	0.0008	-0.70	
4 x 1A x 2	0.000004	0.0006	0.06	
4 x 1A x 3	-0.001	0.002	-0.66	
4 x 2 x 3	-0.00003	0.0005	-0.63	
4 x 1A x 2 x 3	0.0002	0.001		

Note. Mean estimated probability of making an error was .21 (e.i., error rate). Results based on 11,376 data points, including $N = 48$ and Items = 237. $^{m}p < .10$, $*p < .05$, $^{\dagger}p < .001$.

Chapter 5

Individual Differences and Syllable Preference: Distal and Proximal Correlates

Abstract

Taft (1979) proposed that all skilled readers recognize words using a unit that is now known as the Max Coda syllable. More recently work suggests that this unit is only relevant for some people (i.e., good comprehenders; Taft, 2001, 2002; those who are insensitive to the homophone effect, Thompson & Desrochers, 2009b) with some words (low-frequency words; Chen & Vaid, 2007). A multi-level mixed-model analysis was carried out (N=122) using a broad set of item- and participant-level indicators in predicting lexical decision performance in a word-splitting paradigm. The central question addressed was whether insensitivity to various linguistic attributes that make word recognition more or less difficulty would be associated with Max Coda preference. The main results indicate that a) participant sensitivity could be reduced to six independent variance components (only two of which were related to print exposure) and b) the relationship of these components to syllable preference presents a relatively complex picture which implicates as moderators both lexical frequency and participant response speed. The results highlight the need to provide more detailed descriptions of participant samples when studying syllable effects using the word-splitting paradigm.

Introduction

A major source of debate in the visual word recognition literature (for reviews, see Coltheart, Rastle, Langdon, Perry, & Ziegler, 2001; Frost, 1998) has concerned the extent to which lexical memory depends upon the use of integrated units like lexical representations (e.g., addressed phonology) or sub-lexical units (e.g., assembled phonology). Major computational models of visual words recognition are designed to process monosyllabic words exclusively, which means that so far the universe of sub-lexical units has been restricted to letters and graphemes. A logical next step for models of reading is an extension to handle multi-syllable words (Coltheart et al., 2001; Harm & Seidenberg, 2004). With this upcoming extension comes a substantial expansion in the number of possible types of sub-lexical units that can be evoked. In this vein, at least two non-morphological supra-segmental units (larger than a grapheme, smaller than a word) have been proposed: the phonological syllable and the basic orthographic syllable structure (or 'BOSS', Taft, 1979). These units have been associated with the domains of phonology and orthography respectively: phonological syllables are speech units (i.e., phonemes that are pronounced together), while the BOSS is taken to be a word retrieval unit. In more recent work (Chen & Vaid, 2007; Taft, 2001), these types have been captured by the following heuristics for dividing words at their syllable boundaries: Max Onset principle and the Max Coda principle (for elaboration on the nature of these heuristics, see Rouibah & Taft, 2001).

Regrettably, multi-syllable words have been relatively understudied (Chateau & Jared, 2003) and the evidence that does exist for and against these two syllabic definitions has been largely inconsistent since Taft (1979)'s seminal paper (for reviews, see Chen &

Vaid, 2007, Taft, 2001, Thompson & Desrochers, 2009a, 2009b). More recent work has begun to clarify the conditions under which a given unit appears to be implicated in visual word recognition by examining the moderating role of both item and participant characteristics (Chen & Vaid, 2007; Taft, 2001, 2002; Thompson & Desrochers, 2009b). Theoretical progress on the issue of multi-syllable words depends on the resolution of these issues. The present work is an attempt to replicate and extend this body of work.

The rationale of the study is introduced as follows. First, the main paradigm used to study syllable effects in visual word recognition is introduced: word-splitting. Second, the known moderators of word-splitting syllable effects are reviewed. Finally, the specific research questions addressed by the present work are introduced by identifying questions left open by the existing literature.

Word-Splitting Paradigm

A commonly used method of contrasting these two units experimentally has been the word-splitting task in lexical decision (for a reviews, see Chen & Vaid, 2007). The validity of this paradigm hinges on the following logic. Words are presented so that they are divided in a manner that is consistent with either Max Onset segmentation, Max Coda segmentation, or a comparison condition. It is assumed that if a given artificially created boundary coincides with a boundary that is relevant to the cognitive system, then a word should be recognized more rapidly than when it is divided elsewhere (Taft, 1979). For example, if words are identified more rapidly and/or accurately when divided at their Max Onset boundary, then it is concluded that the Max Onset boundary is relevant to the word recognition process. Conversely, if words are identified more efficiently when divided at their Max Coda boundary, then it is the Max Coda unit that would seem to be relevant.

Either possible outcome is called a syllable effect: preference for one syllable-like unit over another which is not reducible to length in letters. The word-splitting boundary can be created using a variety of methods, including the introduction of spaces between letters or a foreign character (Katz & Baldasare, 1983), or simply by distinguishing letter groups by case (Taft, 1979) or the colour of letters (e.g., Rouibah & Taft, 2001; Thompson & Desrochers, 2009b). Fragment priming has also been used with similar intent (Jordan, 1986; Lima & Pollatsek, 1983; Taft, 1987).

Taft (1979) used this logic in designing his experiments and consistently obtained an advantage for the BOSS unit (i.e., the Max Coda unit). Other researchers have had difficulty replicating these findings (Jordan, 1986; Katz & Baldasare, 1983; Lima & Pollatsek, 1983; Seidenberg, 1987). We now consider variables that have been shown to moderate word-splitting syllable effects, thereby providing an explanation for the instability of such effects.

Syllable Effects & Lexical Frequency

As Chen & Vaid (2007) have argued, word frequency is a commonly reported moderator of sub-lexical effects. It is therefore surprising that it had not been considered more systematically as a potential moderator of word-splitting syllable effects (though such a relationship has been reported in English using methods other than word-splitting in the naming task, Jared & Seidenberg, 1990). In their review of the literature, Chen & Vaid remarked that previous word-splitting studies generated larger and more consistent Max Coda syllable advantages with relatively low-frequency words. They proceeded to demonstrate this relationship formally in a series of three experiments, finding that the Max Coda advantage was indeed specific to low frequency words. The pattern of results was the same whether they selected for reading skill (based on SAT scores) or not, though even their low-skill participants were above average verbal SAT scorers. Thompson & Desrochers (2009b, Experiment 3) reported a similar relation except that their results were specific to a sub-sample of participants who were insensitive to homophony.

Max Onset preference has been similarly been associated primarily with low-frequency words. For a sub-sample of participants, Thompson & Desrochers (2009b) observed a Max-Onset preference with low-frequency words only. This result is similar to that obtained in a standard naming paradigm whereby syllable length effects are only observed for low-frequency words (Jared & Seidenberg, 1990). One explanation for this association is that high-frequency words tend to be recognized quickly, primarily on the basis of orthographic information, with minimal phonological involvement. In contrast, the recognition of low-frequency words involves the exploitation of various sources of information to supplement the information provided by the relatively weak orthographic representations with which they are associated in memory. This additional information may be semantic or phonological (for theoretical discussion, see Harm & Seidenberg, 2004). To the extent that phonological information influences processing, an advantage for the phonological Max Onset syllable may be observed.

Syllable Preference & Feedback Phonology

Other sources of evidence converge on a similar idea: Max Onset preference emerges as the result of disproportionate reliance on phonological information in making visual lexical decisions. One of the central findings reported by Thompson & Desrochers (2009b) was that sensitivity to feedback phonology, as indexed by the magnitude of the homophone effect, was a reliable correlate of Max Onset syllable preference. Participants who showed the largest homophone interference effects showed correspondingly larger Max Onset division advantages, while those who showed little or no interference effects tended to show no preference either way (Experiments 1 & 2), or a Max Coda advantage (Experiment 3). The former finding is consistent with the idea that the Max Onset syllable is a phonological unit, and that manifest Max Onset preference in visual word recognition is due to feedback from the phonological to the orthographic domain. On the other side of the coin, the fact that lack of sensitivity to homophony was associated with Max Coda preference under some conditions supports the idea that Max Coda preference is a correlate of strong orthographic processing, as suggested repeatedly by Taft (1979, 1987, and 2001). Again, however, the Max Coda effect was specific to low-frequency words.

Reader Skill & Syllable Preference

The specificity of the Max Coda effect to low-frequency words confirms Chen & Vaid's (2007) conclusions, but it is at least somewhat at odds with Taft (1979)'s initial proposal, which was that the Max Coda unit (i.e., BOSS) is an efficient unit of access to the orthographic lexicon. If this is the case, it is unclear why it should cease to be relevant with high-frequency words, which should be the most efficiently recognized of all words. One way of addressing this issue is to verify whether Max Coda preference is specific to orthographic readers, who would ostensibly engage in more efficient word recognition strategies.

A correlate of efficient orthographic reading is superior ability to comprehend passages of text. At least for the above average readers in the Chen & Vaid (2007) study, there was not sufficient individual differences in reading ability (i.e., verbal SAT scores) to allow the detection of a skill effect on syllable preference. Unlike Chen & Vaid, Taft

(2001, 2002) has observed a positive correlation between reading comprehension skill and Max Coda preference, and a negative correlation with Max Onset preference (i.e., low-skill participants recognized words more rapidly when they were divided at the Max Onset boundary). Given that even the poor readers in the Chen & Vaid sample were above average on the SAT, the positive results reported by Taft have to be weighted more heavily. The Chen & Vaid (2007) result could simply be an artifact of restricted range on the comprehension skill variable.

However, as Thompson & Desrochers (2009a, 2009b) pointed out, Taft's results are difficult to interpret given that reader skill, defined as performance on a comprehension test, has no clear reference in computational models of reading. Inter alia, this means that there is no way for computational models to simulate Taft's results. Moreover, good reading comprehension can be achieved via compensatory strategies at various levels of analysis (Stanovich, 1980). At the extreme, some good comprehenders might actually be below average at recognizing words visually, or at decoding. The multiple strategies one might employ to comprehend text make direct extensions to the more specific task of word recognition ambiguous. It is perhaps more informative then to operationalize individual differences at the level of single-word processing so as to remove this potential source of ambiguity. Examining word recognition variables directly has the added advantage of admitting model-relevant variables into the operationalization of individual differences. Model-relevant variables are any word-recognition behaviors that are simulated by computational models of reading and word recognition (for discussion, see Thompson & Desrochers, 2009a). Further, if a good reader is supposed to be an 'orthographic' reader, then an estimate of print exposure may be a more pertinent 'non-model' indicator when investigating syllable effects than simple skill (e.g., Stanovich & West, 1989).

In an implicit recognition of the fact that 'skill' is essentially an atheoretical construct, Taft (2001) suggested more concrete, process-based interpretations for the dependency he observed between skill and syllable preference. He proposed that low-skill participants tend to rely disproportionately heavily on phonology when recognizing words, which would explain their preference for the phonological 'Max Onset' syllable unit, and their greater sensitivity to phonological difficulty factors (Chateau & Jared, 2000; Jared, Levy, & Rayner, 1999; Jared & Seidenberg, 1991; Sears, Siakaluk, Chow, & Buchanan, 2008; Strain & Herdman, 1999; Unsworth & Pexman, 2003). It is important to note that low-skill readers are not better at using phonology. It is an impoverish pool of orthographic information that forces reliance on other sources of information which are themselves impoverished (for reviews, evidence, and discussion, see Briggs & Underwood, 1987; Harm & Seidenberg, 2004; Strain & Herdman, 1999). Thus, the term sensitivity is not used here to denote a positive attribute. Rather, it is employed here to capture the idea that poor readers allow inconsistent phonological information to influence their behavior. Good orthographic readers have access to the same information, but are considered insensitive here because they direct their attention to more reliable sources of information during the word recognition process.

Indeed, with respect to the high-skill readers, Taft (2001) proposed that they are efficient word recognizers because they can base their decisions on orthographic information alone. To the extent that the Max Coda syllable is an orthographic unit, this would explain the positive association with skill. The more consistently a reader relies on orthographic information, the more likely a preference is to be demonstrated for an orthographic unit. In addition to Max Coda preference, Taft's view implies that orthographic readers will demonstrate additional characteristics. For instance, readers with strong orthographic representations of the words that they know should be relatively insensitive to a variety of difficulty factors associated with the three principal domains of knowledge: orthographic, phonological and semantic attributes. Again, the term sensitivity here is used here in a narrow sense. It reflects the extent to which readers allow a given lexical attribute to influence their observable behavior. Again, this view does not preclude the possibility that skilled readers have access to these sources of information. To the contrary, skilled readers have access to a rich database of linguistic information, which means they may have the option of ignoring inconsistent information to focus on more reliable sources. For instance, if a given word's phonology is inconsistent (e.g., PINT), then it may be more useful to focus on orthography or semantics. If this view is correct, then insensitivity to lexical attributes that make words more difficult to recognize should be characteristic of readers who prefer the Max Coda syllable. Further, this pattern should be especially marked if the sensitivity being considered is an associate of Max Onset preference. This would be the case, for example, with the phonological attributes of words.

Syllable Preference and Sensitivity to Difficulty Factors

This general logic was employed by Thompson & Desrochers (2009b) in evaluating Taft's hypothesis that syllable preference was determined in part by the degree to which participants rely on phonology during word recognition. In examining this question, two indices of participant sensitivity to phonological difficulty factors were obtained: a) regularity (irregularly pronounced vs. regularly pronounced words) and b) homophony (homophones vs. control words). Participants demonstrating relatively large disadvantages for homophones and irregular words are considered sensitive to these difficulty factors.

These estimates of participant sensitivity were subsequently used to predict performance on a word-splitting task. Two purposes were served by this analysis. First, it addressed directly the issue of whether reliance on phonology was associated with syllable preference. Second, it did so in a way that would be informative for computational theorists trying to determine where syllable units will fit in the next generation of models (i.e., both regularity and homophony effects are simulated by extant models). In sum, the analysis allowed the empirical relationship between these variables to be estimated as well as, indirectly, the association between syllable preference and the model parameters that govern simulated homophony and regularity effects. In principle, this type of information could be used to inform the simulations that guide model construction and parameter specification (e.g., Coltheart et al., 2001).

The predictive value of phonology was found to be specific to feedback phonology: the homophone effect. Specifically, Desrochers & Thompson (2009b) found that sensitivity to homophony was positively associated with Max Onset preference under certain conditions (Experiments 1, 2, 3). Conversely, insensitivity to homophony was associated with Max Coda preference under a more limited set of conditions (Experiment 3). These findings are broadly consistent with the idea that Max Onset and Max Coda preference have different sources. The former unit appears to be controlled by phonological knowledge, while the latter appears to have a different basis, ostensibly orthographic knowledge given the 'insensitivity' demonstrated by those who demonstrate such preference. Theoretical accounts of syllable preference would have to account for this distinction as both effects are unlikely to result from the same mechanism.

As useful as these findings are, they come with some important limitations. For instance, a very limited number of 'process-based' or 'sensitivity' variables were considered, which makes it difficult to draw firm conclusions. It is perhaps the case that sensitivity to all types of feedback (e.g., semantic and phonological) is positively associated with Max Onset preference. Further, it may be possible to more reliably predict Max Coda preference if a broader set of indicators was collected. The best possible indicator of strong orthographic processing would be insensitivity to a variety of difficulty factors in word recognition (e.g., Orthographic N, Imageability, and Homophony). Thus, it is important to replicate Thompson & Desrochers (2009b) with a broader set of empirical indicators to verify the generality of their results. In addition, it would be useful to include print exposure as an indicator so as to a) map its relationship to the process-based variables considered here, and b) verify the idea that reader skill, specifically orthographic processing, is predictive of syllable preference (Taft, 2001).

Purpose

The present work is an attempt to replicate and extend previous work investigating individual differences in syllable preference using the word-splitting paradigm (Chen & Vaid, 2007; Taft, 2001; Thompson & Desrochers, 2009b). It was expected that lexical frequency, print exposure, and sensitivity to homophony would moderate syllable preference in replication of previous findings. If the account of the relationships developed above is sound, then a more general pattern of results should also be observed with an extended pool of word-attribute sensitivity indicators. First, high print exposure should be associated with insensitivity to a variety of attributes that make words harder to recognize (e.g., Chateau & Jared, 2000; Sears et al., 2008). Second, joint high print exposure and general insensitivity should be positive correlates of Max Coda preference. If this pattern is not observed, then the status of the Max Coda syllable as an orthographic unit will have been weakened (see also the low-frequency specificity of Max Coda preference).

The process of generating the participant sensitivity estimates involves testing the effect of a variety of item-attributes on lexical decision performance. A secondary purpose of the present study is therefore to replicate previous work studying the importance of these variables, a complete list of which is available in the method.

Method

Participants

The sample consisted of 122 undergraduate students at the University of Ottawa, 29 men, 93 women, mean age= 19.58, SD = 4.61. All participants cited English as their maternal language, but 80 percent of the sample was bilingual. All participants had normal or corrected-to-normal vision. On average, the sample had 1.31 years of university study, SD=0.79.

Materials

Reading skill questionnaire. An estimate of Print Exposure was obtained in lieu of a measure of comprehension skill. The two measures are strongly correlated in any case (Sears et al., 2008; Stanovich & West, 1989), but print exposure is a concept that is more directly relevant to the problem at hand (i.e., skill at orthographic processing). Conceptually, it is a direct estimate of reading practice, and thus the strength of a participant's orthographic representations. This a priori supposition is borne out by empirical data demonstrating that high print exposure participants are more efficient orthographic readers (Chateau & Jared, 2000; Sears et al., 2008; Stanovich & West, 1989).

Print exposure was estimated using an updated version of the Author Recognition questionnaire (Stanovich & West, 1989). For the author recognition questionnaire, participants are presented with a list of real authors mixed in with decoy names (60 percent real, 40 percent decoys). They are asked to indicate with a mark which names they recognize as professional writers. The total number of authors recognized is the Author Recognition score, minus false positives (No participant in the present sample generated more than two false positives). Note that Print Exposure is an estimate of a distal explanatory variable, whereas those to follow might be considered proximal, reflecting how the reading system of participants performs currently.

Linguistic materials. Two general classes of stimuli were used. The first class was designed to estimate the sensitivity of participants to various item attributes. Again, the term sensitivity is used here to capture the idea that participants vary in the extent to which their performance is affected by the difficulty factors associated with letter-strings (e.g., length, lexical frequency). Participants whose performance is affected by a difficulty factor more than average is said to be 'sensitive'. The second class of stimuli was intended to be presented within the context of the word splitting manipulation. These stimuli were presented with an artificially introduced boundaries designed to evaluate potential syllable-preference effects. The latter class of stimuli comprised exclusively multisyllable words and were the same 480 items (240 words, 240 pseudowords) used by Thompson & Desrochers (2009b). These stimuli contrasted items in terms of frequency (continuous variable) and their syllable structure: medial consonant (e.g., pedal, tulip) versus medial consonant cluster (e.g., thunder, bluster), which are hereafter referred to as simple and complex words respectively. In contrast, the first class of stimuli was composed exclusively of monosyllable words which were assembled as follows.

The lexical attributes measured were imageability (20 high vs. 20 low items; mean imageability values of 5.55, SD=0.86, versus 1.87, SD=.028), homophony (20 homophones vs. 20 controls) and regularity (10 irregular words vs. 10 regular controls). In each case, the comparison conditions were matched on objective frequency, length, and orthographic neighborhood. The homophone and regularity items were taken from the stimulus list reported by Pexman, Lupker, and Reggin (2002). The imageability items were selected based on the imageability norms reported by Cortese and Fugett (2004) and other lexical characteristics extracted from the English Lexicon Project database (Balota et al., 2007). The mean log HAL frequency for the high- and low-imageability words was 6.16 (SD=0.73) and 6.17 (SD=0.96) respectively. The imageability items were selected so as to keep their length (average=3.5 letters), orthographic neighborhood, and bigram frequency of the imageability items were kept constant as well.

The pseudowords were constructed so as to resemble the words in terms of their average length and structure. That said, there was sufficient variation in length, bigram frequency, and orthographic neighborhood (N; Coltheart, Davelaar, Jonasson, & Besner, 1977) for the influence of these factors to be evaluated within the context of a regression analysis. In addition, 10 pseudohomophones taken from the Unsworth & Pexman (2003) study were included in the list of nonwords so as to evaluate sensitivity to lexical phonology in the sample of participants, controlling statistically for the other possible confounding factors.

Procedure

The experiment involved a lexical decision task which consisted of three blocks of stimuli (1 practice, 2 experimental). The practice block (20 items) and experimental blocks (200 and 480 items respectively) each contained an equal number of words and pseudowords. Participants were asked to discriminate between the words and pseudowords by button press. They were asked to make the decision as quickly as possible while keeping their error rate under ten percent.

Item-sensitivity block. The first experimental block contained the items intended to estimate participant sensitivity to various item characteristics. The items were presented in a single list whose order was randomized for each participant. All such items were presented in lowercase, black against a white background, including the 20 practice items.

Word-splitting block. The materials, counterbalancing, randomization, and trial-level procedures employed were identical to those of Thompson & Desrochers (2009b, Experiment 2). Specifically, three word lists were created that were matched as close as possible on objective frequency, length, and orthographic N. Each list was assigned to one of three word-splitting conditions (Max-Onset, Max-Coda, Max-Coda+1) according to a counterbalancing scheme across participants. This precaution provided further insurance that list characteristics were orthogonal to the word-splitting manipulation, since all were presented in the three conditions an equal number of times across participants. A similar type of control was achieved with pseudowords by drawing these items randomly without replacement from a single list. The pseudoword items were assigned to one of the three conditions with equal probability under the constraint that an equal number of nonwords appeared in each condition, within each presentation block.

During the experiment, word-splitting items were presented in two sub-blocks, which were created by dividing each of the three lists above in two. Including words and nonwords, the two blocks of word-splitting stimuli comprised 240 items each. The order of these sub-blocks was counterbalanced across participants, and randomized within, so as to attenuate bias due to order effects.

In order to produce the word-splitting manipulation, items were presented in colour against a white background (Rouibah & Taft, 2001; Thompson & Desrochers, 2009b). The word-initial segment was colored blue, while the word-final segment was colored red. As noted earlier, the segmentation confirmed to one of three rules: Max-Onset, Max-Coda, & Max-Coda+1 (For a more detailed explanation, see Rouibah & Taft, 2001).

Trial-level procedure. For the practice block and the first experimental block, the trial-level procedure conformed to the standard lexical decision task. Each trial began with the presentation of a fixation point (+), which remained visible for 1 second. The fixation point was then replaced by the target item which remained visible until a response was collected by button press. The procedure was the same for the word-splitting task except target items were replaced by a blank screen after 500 ms in this condition. A 1 s inter-trial interval was observed. In all cases, words were identified using the dominant hand.

Results and Discussion

The results are reported in three sections according to their logical sequence of execution. First, participant sensitivity to various item difficulty factors is estimated and reported. The tests of group sensitivity constitute attempts at replicating previously reported effects in the literature. Second, the correlation among the participant attributes are reported along with the results of a principal component analysis that was intended to a) reduce the number of variables available for predicting performance on the subsequent analysis of word-splitting performance and b) elucidate the dimensions underlying these indicators. Finally, the results of the word-splitting manipulation are reported. Two sets of results are of interest: a) tests of the word-splitting effect at the item level (and word-attribute moderators like frequency) and b) tests of the moderating role of participant attributes. The results of each analysis are reported in turn.

Sensitivity to Item Attributes

The attributes of both words and pseudowords are known to affect response latencies in the lexical decision task. Accordingly, the attributes evaluated here comprise lexical characteristics (e.g., homophony, regularity, imageability) and nonwords characteristics (pseudohomophony, bigram frequency, orthographic N). The lexical characteristics were evaluated independently of each other with dedicated sets of items for which item-selection controls were implemented (see Method). In contrast, the effects of nonword characteristics were evaluated within the context of a multiple regression analysis, which allowed statistical control of ambiguous explanatory variance shared amongst the letter-string properties considered here. As a result, the unstandardized beta coefficients evaluated in the pseudoword regression models are relatively free of confounding factors.

Both word and pseudoword effects were evaluated using the mixed-effects procedures (lme4) of R version 2.8.1 for both continuous dependent variables (i.e., reaction time, RT; see Baayen, Davidson, & Bates, 2008) and binary outcomes (i.e., error rate, ER; see Jaeger, 2008). Wherever possible, statistical significance was determined via the Markov Chain Monte Carlo (MCMC) procedure. RT observations were screened prior to analysis using a lower-bound filter of 250 and an upper-bound filter of 3.5 standard deviations above the participant mean. This screening procedure was performed separately for words and pseudowords, resulting in a loss of less than two percent of observations.

The results of the mixed-model analyses performed on the resulting dataset are reported in Table 1.

Word attributes

Inspection of Table 1 reveals three statistically significant effects for the word stimuli. A homophone interference effect was observed on RT, but not ER. This replicates previous work (Pexman & Lupker, 1999; Pexman, Lupker, & Jared, 2001; Pexman et al., 2002; Thompson & Desrochers, 2009b), and is consistent with the idea that the present sample is not particularly low-skilled overall (otherwise a corresponding ER effect would have been detected). The two remaining effects indicated statistically significant facilitation on RT and ER with high- relatively to low-imageability words (see also, Balota et al., 2004; James, 1975; Strain & Herdman, 1999; Strain, Patterson, & Seidenberg, 1995, 2002).

Pseudoword attributes

Analysis of the pseudoword attributes revealed a number of statistically significant effects. In the first place, pseudohomophones took longer to correctly reject as nonwords (RT) and were more likely to be falsely identified as being words (ER) than non-pseudohomophone pseudowords. This effect was significant even after controlling for the other item characteristics considered here (see also, Briesemeister, Hofmann, Tamm, Kuchinke, Braun, & Jacobs, 2009; McCann, Besner, Davelaar, 1988; Seidenberg, Peterson, MacDonald, & Plaut, 1996; Yates, Locker, & Simpson, 2003; Ziegler, Jacobs, Kluppel, 2001). In addition, orthographic N was found to significantly influence both RT and ER. Specifically, pseudowords with a larger orthographic neighborhood were rejected relatively less consistently as words and after a longer delay (see also, Andrews, 1992, 1997;

Coltheart et al., 1977; Siakaluk, Sears, & Lupker, 2002). Finally, both RT and ER increased significantly as a function of pseudoword length in letters (see also, New, Ferrand, Pallier, & Brysbaert, 2006).

Conclusion

A number of common item-attribute effects were replicated with the present sample of participants. These findings provide a sort of double cross-validation in the sense that a) previous results were replicated; supporting the validity of previous findings, and more importantly b) the present sample has been shown to behave like participants in previously published studies, supporting the contention that the present sample responded as expected to the presented stimuli.

Participant Attributes: Inter-correlations and Dimensions

The mixed-models of item-attribute effects reported in the preceding section were re-estimated as random coefficient models (Lorch & Myers, 1990; Thompson, 2008). The purpose of this re-estimation was to provide custom estimates of the item-attribute effects for each participant which could then serve as participant-level predictors in the subsequent mixed-model of performance on the lexical decision word-splitting task. In addition to these variables, Print Exposure was considered as an important source of individual differences (i.e., Author Recognition questionnaire). The correlations among these measures are reported in Table 2.

Data Reduction Strategy: Principal Components Analysis

The number of predictor variables was prohibitively large relative to the sample size and a desire to manage the Type-I error rate. For this reason, a variable reduction strategy was implemented that would extract all the important sources of variance in the dataset regardless of whether this variance was shared or unique to a particular indicator: principal component analysis. First, analyses were conducted to determine whether the correlation matrix was suitable for factorization. Bartlett's test of sphericity indicated that the observed matrix differed significantly from identity, χ^2 (102)= 288.79, $p < .001$. In other words, the test indicates the presence of common variance in the dataset which is available to be extracted. Similarly, the Kaiser-Meyer-Olkin (KMO) index of sampling adequacy was .54, which exceeds the minimum value of .50 necessary for factorization (Hutcheson & Sofroniou, 1999; Kaiser, 1974). The KMO index confirms and extends the results obtained with Bartlett's tests in that it considers jointly the role of sample size and communality among the variables in determining the reliability of results that can be obtained by principal components analysis. Taken together, these tests suggest that reasonably stable results can be obtained from the dataset in the analyses to follow.

To determine the number of dimensions to extract, a revised version of the parallel analysis macro reported by O'Connor (2000) was employed. The parallel analysis was based on a thousand bootstrap iterations of the raw dataset (15 variables, 122 participants). This analysis indicated that six variance components were worth extracting from the dataset. Accordingly, a principal components analysis was performed that requested the extraction of six orthogonal dimensions (allowing them to be correlated had no substantive impact on the results). The 6 components together accounted for roughly 63 percent of variance in the variables. Note that because the explained variance was relatively low, a separate series of analyses was conducted on the residual variance in each variable to verify whether individual differences relevant to syllable preference escaped the 6 extracted variance components (i.e., variance in the observed indicators that was uncorrelated with the 6 components). This residual variance was found to be unrelated to syllable preference, which is a finding that tends to validate the decision to only retain the 6 components. The full set of results is not reported here for the sake of brevity. In any case, the loadings of the observed variables that were obtained from the rotated solution principal components analysis are reported in Table 2.

Substantive Interpretation of the Variance Components

Participants were attributed a score on each of the six dimensions. These component scores were then employed as participant-level predictors of performance on the word-splitting task. In order to facilitate the interpretation of these components in the analyses to follow, the loadings associated with each component are considered in turn. At this point, it is worth calling attention to the obvious: no single component captured the print-exposure/sensitivity characteristic that was developed in the introduction. The data suggest that this one-dimensional account was overly simplistic. Instead, elements of the skill-sensitivity dimension are distributed across multiple dimensions.

Variance component 1: Lexical Phonology ↑ & Nonword Sensitivity ↓ (RT). The first dimension is associated with the homophone RT effect and the nonword variables bigram frequency, orthographic N, and length in letters. The direction of these relationships suggests that higher scores on the first dimension indicate stronger lexical processing, specifically phonological lexical processing and a corresponding insensitivity to some of the nonword attributes. The first dimension is positively correlated with the magnitude of the homophone effect on RT, and even more strongly associated with a decrease in sensitivity to bigram frequency, orthographic N, and length in letters on RT. Thus, the higher scores on the first dimension represent primarily the tendency to be insensitive to the orthographic attributes of nonwords, and heightened sensitivity to (lexical) feedback phonology. The homophone RT component of this dimension leads to the hypothesis that it should yield a pattern of results similar to that reported by Thompson & Desrochers (2009b): robust Max Onset effects as scores increase (with low-frequency words). Note that this dimension is essentially orthogonal to print exposure (see also, Unsworth & Pexman, 2003). If this dimension is related to syllable preference in the same way as homophone sensitivity in Thompson & Desrochers (2009b), then this suggests that the effect that was observed is unrelated to general comprehension skill, or at least print exposure.

Variance component 2: ↑Sensitivity to nonword difficulty factors (accuracy, ↓Print Exposure). The second dimension is associated with the bigram frequency effect, the N effect, and the length effect on ER. It is also related to print exposure. The common variance being captured here seems to be specifically sensitivity to the attributes of nonwords. The underlying dimension might be the extent to which participants can quickly dismiss nonwords as plausible lexical candidates. Those who generally have difficulty rejecting nonwords would show more sensitivity to the attributes that make them more or less similar to real words. The second dimension was most strongly related to increased sensitivity to bigram frequency, N, and length on ER. Thus, higher scores on this dimension represent more liberal use of the distributional properties of pseudowords in making judgments. In this respect, it is the mirror image of Component 1, which indicated stronger lexical processing with increases values. Note also that print exposure loads negatively on this dimension, suggesting that this type of sensitivity to nonwords attributes is at least in part the result of deficient orthographic mastery due to limited exposure to print (see also, Chateau & Jared, 2000; Sears et al., 2008). This dimension could potentially discriminate between Max Onset and Max Coda preference as an indicator of strong orthographic processing.

Variance component 3: ↑ Orthographic Reading. The third dimension was related to print exposure, the imageability effect, and the pseudohomophony effect. High print exposure, ease of rejecting pseudohomophones, and insensitivity to imageability are correlates of strong word recognition processing (e.g., Strain & Herdman, 1999; Unsworth & Pexman, 2003). The common dimension represented here then would seem to be orthographic processing. This third dimension is strongly and positively associated with print exposure, insensitivity to Imageability, and insensitivity to Pseudohomophony on ER. This set of loadings is consistent with a dimension where higher scores represent strong reliance on orthographic information in making lexical decisions due to high print exposure. Thus, if the logic developed in the introduction holds, we would expect this dimension to be positively associated with Max Coda preference.

Variance component 4: ↑ Sensitivity to Word and Nonword Difficulty Factors (RT). The fourth variance component was associated with imageability, pseudohomophony, and orthographic N on RT. The common dimension represented here appears to be sensitivity, as defined earlier. Specifically, higher values on the fourth dimension represent greater sensitivity to imageability, pseudohomophony, and nonword orthographic N on RT. As with component 3, this pattern suggests that higher values on the dimension indicate relatively inefficient use of orthographic information and reliance on compensatory sources of information like phonology (pseudohomophony) and semantics (imageability). If this component is associated with syllable preference, then it should be positively associated with Max Onset division. What distinguishes this component from Component 3 is the lack of association with Print Exposure and the fact that it is the RT measures specifically that are involved. As RT is typically the more sensitive/subtle measure, Component 3 may capture more extreme individual differences.

Variance component 5: ↑ Speed accuracy tradeoff. In contrast to the previous variance components, the fifth appears to reflect a less interesting source of variability: a speed-accuracy tradeoff involving the homophone effect. The RT and ER homophone sensitivity variables both load on this dimension, but with different sign. This dissociation is the symptom of speed-accuracy tradeoff. It is unclear what value this dimension would have as a predictor of word-splitting task performance, unless it is taken as a general indicator of a participant's propensity to trade performance on one indicator for another.

Variance component 6: ↑ Sensitivity to Word Difficulty Factors. Finally, the sixth dimension was associated with regularity on RT and ER, and Imageability on RT. This dimension captures varying degrees of sensitivity to semantic and phonological difficulty factors. Specifically, higher values on the sixth dimension represent primarily increased sensitivity to regularity on RT and ER, which has been attributed to inefficient or messy use of phonological information during word recognition (Unsworth & Pexman, 2003). Note that higher values on the sixth dimension are also associated with increased sensitivity to imageability on RT. Similar findings in the past have been interpreted in terms of sensitivity to regularity and imageability having a common cause: poor phonological coding leads to increased compensatory reliance on semantics (Strain & Herdman, 1999). When imageability, phonology, and orthography are deficient, performance suffers. Skilled word recognizers are more flexible in that they use all kinds of information efficiently, including orthographic information. Thompson & Desrochers (2009b) failed to observe a relationship between regularity sensitivity and syllable preference, but given the phonology component and the fact such sensitivity is associated with low skill, they had predicted a positive relationship with Max Onset preference. However, if the Max Coda unit is implicated in the retrieval of semantics as Taft (2001) has claimed, it is conceivable that a positive relationship with Max Coda preference will be observed instead. This dimension is essentially independent of print exposure, which means that a response strategy rather than word recognition skill may be the operative factor. Participants that tend to rely on semantics may prefer the Max Coda syllable, not because they are skillful per se, but because they place greater weight on semantics in making their decision (e.g., via the indirect phonology-semantics route; Coltheart et al., 2001; Harm & Seidenberg, 2004).

Word-Splitting Task: Mixed Model of Syllable Preference

Chi-square statistical tests indicated that random effects for both participants and items were justified. Differences among participants explained 22.8 percent of the total estimated variance in RT (S^2=46, 299.46), while random item-level variance explained 11.5 percent. Similar random effect estimates are unavailable for the ER data. Subsequently, fixed effects were fitted to the word-splitting RT data for at the item-level and then the participant-level in attempt to explain the systematic random variance that was just described. For the sake of economy, results for the ER data are only reported at the item-level because the effects at the participant-level were not of interest.

Fitting the item-level model

The following fixed-factor variables were entered as item-level predictors of word-splitting task RT and ER: Word-splitting condition, Lexical Frequency, Syllabic Complexity, and the complete set of interactions among them. The results of the RT and ER analysis are reported in Table 3.

The analysis revealed that increased Lexical Frequency significantly improved response speed and accuracy. No other variable made a significant unique contribution to explaining word-splitting performance. For the RT analysis, the variables in the item-level model explained 39.4 percent of the random effect variance associated with items, and 4.5 percent of the total variance. Obviously, the item-level variables did not explain any participant-level variance.

Fitting the Participant-level Model: Mean RT

Following Thompson & Desrochers (2009b), we though it important to statistically control for overall response speed so as to distinguish between this factor and the more substantive variables discussed above. As a preliminary step, the mean RT from the preceding item-sensitivity analysis was entered as participant-level predictor variable in a mixed-model analysis along with the item-level factors. The direct effect of mean RT (henceforth Participant Speed) was entered along with the associated interaction terms. The results of the mixed-model of word-splitting RT are reported in Table 4.

Unsurprisingly, Participant Speed was a strong predictor of individual differences in RT on the word-splitting task. By itself, it accounted for 45.9 percent of the random effect variance associated with participants or 10.4 percent of the total variance. The direct effect was confounded by two statistically significant interactions: a two-way interaction with lexical frequency and a three-way interaction involving word-splitting and syllable complexity. Combined, the interaction effects explained an additional 1 percent of the item random-effect variance.

The two-way interaction between mean RT and Lexical Frequency replicates a common finding in the literature (Wagenmakers, Ratcliff, Gomez, & McKoon, 2008; Taft & Russell, 1992), which is that the magnitude of the frequency effect is larger when participants respond slowly (+1 SD), b=-33.69, 95 percent CI -38.45 to -28.94, than when they tend to respond quickly (-1 SD), b=-19.89, 95 percent CI -24.68 to 15.26. At present, it is impossible to distinguish between attenuation of the frequency effect due to a faster response criterion (same participants, different instructions; Wagenmakers et al., 2008) and attenuation due to differences in skill or print exposure per se (Chateau & Jared, 2000; Sears et al, 2008), though the latter effect is typically only observed when pseudohomophonic pseudowords are employed instead of typical pseudowords.

More interesting for present purposes is the three-way interaction, which analysis revealed to be caused by slow responders who manifested a significant Max-Coda advantage with Complex words, b=-12.33, 95 percent CI -23.29 to -0.73, but not with simple words, b=4.86, 95 percent CI -7.02 to 16.81. The word-splitting effect was null with fast responders. All other things being equal, slow responding tends to be characteristic of lower-skill/print exposure (Chateau & Jared, 2000; Lewellen, Goldinger, Pisoni, & Greene, 1993; Sears et al., 2008), which means that the present results would seem to contradict previous work indicating that only high-skill readers prefer Max-Coda segmentation (Taft, 2001, 2002). This point is taken up again in light of the analyses to follow which consider the joint contribution of Response Speed and the other explanatory variables considered here. The additional information provided by this analysis may clarify the picture.

Fitting the Participant-level Model: Mean RT + Principal Components

To address questions relating to the syllable effect, RT and ER on the word-splitting task were regressed against the item-level and participant-level predictors. In addition to Participant Speed (i.e., Mean RT), the six variance-component scores generated by the principal component analysis described above were included as predictors, one at a time, in six independently estimated models of the word-splitting data. The results of these tests are reported in turn.

Component 1: Lexical Phonology & Nonword Sensitivity (RT). The first component and its interactions with mean RT and the item-level variables were entered as predictors of word-splitting RT. The specification of the model and the associated results are reported in Table 5. The results revealed a statistically marginal three-way interaction among Component 1, Word-splitting and Syllable complexity. The pattern of the interaction indicates that participants who score high on Component 1 tend to prefer Max Onset segmentation with complex words, b=15.46, 95 percent CI -11.28 to 45.49, while those who score low on Component 1 prefer Max Coda segmentation with the same items, b=-12.92, 95 percent CI -31.18 to 7.65. While the interaction is not significant, the pattern of results mirrors that observed by Thompson & Desrochers (2009b, Experiments 1 & 2) who linked greater sensitivity to homophony with preference for the Max Onset syllable. Recall that Component 1 is associated with heightened sensitivity to the homophone effect. In the present research, the homophone items were presented in a list along with many additional items, including pseudohomophones. This difference in experimental context may account for the weaker association between homophone sensitivity variance and syllable preference observed here. It should be noted that Component 1 was not directly related to word-splitting RT even when it was entered in a model where it acted as the sole participant-level predictor. Participant Speed did not eliminate the predictive power of homophone sensitivity in the Thompson & Desrochers (2009b) analysis either.

Component 2: Sensitivity to nonword difficulty factors (accuracy). A new regression model was then estimated, this time with Component 2 entered as a participant-level variable along with mean participant RT. The specification of this model and the associated results are reported in Table 6. As with Component 1, no direct relationship was discerned between Component 2 and word-splitting task performance either before or after statistically controlling for mean RT and the interaction terms. Component 2 was, however, implicated in three interactions. First, a two-way interaction with Participant mean RT reflected an antagonistic relationship between these two factors: the effect of both variables is greatest when the other is low. Among other things, this means that the component scores of slow responders were uninformative in predicting word-splitting RTs.

Second, this relationship was apparently qualified by two higher-order three-way interactions: one involving lexical frequency and the other the word-splitting manipulation. The interaction involving frequency was due to the fact that the normal deflation of the frequency effect associated with fast responding was magnified by low scores on Component 2, frequency effect of $b=-18.22$, 95 Percent CI -22.99 to -13.44, when compared to high scorers, frequency effect of $b=-25.15$, 95 Percent CI -29.95 to -20.26. Given that this component is negatively associated with print exposure, the result supports the claim that some aspect of greater print exposure (i.e., that which is associated with insensitivity to the distribution properties of pseudowords) tends to attenuate the frequency effect over and above attenuation due to participant response speed, at least for those who respond quickly. Previous work examining the relationship between print exposure and moderation of the frequency effect did not consider Participant Speed as a potential mediator (Chateau & Jared, 2000; Sears et al., 2008).

Finally, the three-way interaction involving the Word-Splitting factor was due to a statistically significant Max Onset advantage for participants who are both fast and low on Component 2, b=12.88, 95 percent CI 1.25 to 24.46. Inter alia, Component 2 is indicative of sensitivity to the distributional properties of pseudowords in the lexical decision task, and to a lesser extent lack of print exposure. Low scores on this component and fast responding are associated with a Max Onset advantage here, which is a result that runs counter to the expectations given that insensitivity was supposed to entail a Max Coda advantage. A potential explanation may lay in the fact that skilled/high-exposure readers access phonology rapidly (Chateau & Jared, 2003). For such readers, fast responding could reasonably yield preference for a phonological unit like the Max Onset syllable.

Component 3: Orthographic Reading (accuracy sensitivity). The same basic model was then re-estimated using Component 3 as a participant-level predictor (see Table 7). Component 3 was directly and significantly related to word-splitting RT, but only before entering mean RT and interactions into the model. Analysis of the complete model nevertheless indicated three statistically significant interaction terms. The most relevant effect for present purposes is the four-way interaction among Component 3, RT, Frequency, and Word-Splitting. Decomposition of the four-way interaction revealed that conditional effects were observed specifically for the low-frequency words. With low-frequency words, fast responders who were high in Component 3 showed a trend towards a Max Coda advantage, b=-61.32, SE=36.73, 95 Percent CI -131.79 to 11.04. This result is consistent with Taft (2001, 2002). However, a Max Coda advantage was observed also for slow participants who were low in Component 3, b=-71.09, SE=35.69, 95 Percent CI -141.85 to -4.01. In summary, the analysis of Component 3 fails to support the idea that there is a consistent relationship between Print Exposure (a strong, positive determinant of Component 3) and Max Coda preference. Rather, it seems that this relationship depends on response speed.

And so, high-exposure fast responders showed a Max Coda preference, but so did relatively slow low-exposure participants. The first result was expected, but the latter was not. It is unclear what could account for such a result. One could posit that the two effects have entirely different bases. High-exposure fast responders are clearly a 'Max Coda' demographic, characterized by strong orthographic processing. There is little ambiguity there. It is not quite as clear what the low-exposure slow responders were doing. Perhaps the result is due to response criterion setting. High-exposure participants only show Max Coda preference when they respond quickly, because this is when such units are relevant to them. Low-exposure participants only show Max Coda preference when they respond slowly, because it takes longer for the source of Max Coda preference to become active in memory for them. The present experiment used non-standard instructions. Participants were asked to respond as quickly as possible, while keeping their error rate under ten percent. Standard instructions require participants to respond as quickly and as accurately as possible. If the instructions used here resulted in greater freedom in the response rates participants selected, this might account for the conflict with Taft (2001, 2002)'s study.

An interesting question is whether response speed can be manipulated (e.g., via instructions) to induce Max Coda preference or make it disappear. For example, if we assume that low-exposure participants are relatively slow to access certain types of information (e.g., semantics), then, to the extent that Max Coda preference is due to such retrieval processes, we should be able to eliminate this pattern of preference by inducing them to respond quickly. Indeed, instructions emphasizing speed may have improved the odds of detecting a Max Onset preference with the low-exposure fast responders. Conversely, inducing high-exposure participants to respond more slowly (e.g., by emphasizing accuracy) may have eliminated the Max Onset preference. Taft (1979) began with the assumption that a single syllable unit might be general to all. He abandoned this idea later to propose that some readers prefer the Max Onset syllable, while others prefer the Max Coda. The truth may be that everyone can demonstrate a Max Coda preference under the right circumstance. Of course, this account of the results is quite speculative at this point.

With Component 3, the only property that consistently emerges as an associate of Max Coda advantage is for words to be of low frequency. This observation is broadly consistent with previous work (Chen & Vaid, 2007). It is interesting to note that Chen & Vaid failed to observe any skill dependency in Max Coda preference. This result may have been due to their sample being uniformly high-skill (i.e., above average on the SAT verbal) or it may have reflected a more general relationship. It is possible that both high and low skill participants can demonstrate Max Coda preference with low-frequency words given the right circumstances (e.g., the right response-timing criteria; for a theoretical model of timing processes, see Ratcliff, Gomez, & McKoon, 2004; Wagenmakers et al., 2008).

Components 4 & 5: Sensitivity to Word and Nonword Difficulty Factors (RT) & Speed accuracy tradeoff. The analysis was performed again with Components 4 and 5, the results of which are reported in Tables 8 and 9 respectively. These results are reported here for the sake of completeness, but are not discussed further given that none of the effects involving the word-splitting manipulation are significant, and they did not suggest interesting patterns.

Component 6: Sensitivity to Word Difficulty Factors (Imageability, Regularity). Finally, the parameters of the base-model specification used throughout this section were re-estimated with the sixth variance component as a participant-level predictor (see Table 10). A scan of the results indicates that the four-way interaction involving Component 6, Participant Speed, Complexity, and the Word-splitting manipulation was marginally significant. Post-hoc exploration of the effect indicated that it was being driven by a statistically significant Max-Coda advantage on complex words for slow participants who are high on Component 6, $b=-13.83$, $SE=6.30$, 95 Percent CI -26.09 to -1.79. As discussed earlier, this component was unrelated to Print Exposure, reflecting instead an exposure-independent sensitivity to regularity and imageability (an index of semantic richness, for a review, see Desrochers & Thompson, 2009). This result is inconsistent with the idea that Max Coda preference is driven by efficient orthographic representation (i.e., independence from auxiliary sources of information like phonology and semantics). Rather, it seems that a variance component associated with inefficient phonological representation and dependence on semantic information (Strain & Herdman, 1999) biases responses towards a Max Coda preference with low-frequency words. This result is inconsistent with the 'sensitivity' idea that was developed in the introduction, but is consistent with the idea that those who place weight on semantics in making their decisions will tend to favor a Max Coda syllable. Recall that this dimension is unrelated to Print Exposure, and so may merely reflect an implicit decision to rely on semantics more heavily (a decision which may involve the recruitment of phonological information, increasing sensitivity to regularity).

General Discussion

The central purpose of this research was to verify whether Max Coda syllable preference was demonstrated by participants who are high in print exposure (a proxy for skill) and generally insensitive to various factors that make words more or less difficult to recognize. This hypothesis was not confirmed. Previously reported effects of many item-attributes were obtained, including the effect of pseudohomophony, orthographic N, length, homophony and that of imageability with low-frequency words. The expected relationship between Print Exposure and sensitivity to difficulty factors was also observed: Print exposure was positively associated with insensitivity on two distinct principal component dimensions. Syllable preference, however, was not related to these factors in the expected manner. The pattern of results was relatively complex, which is an aspect to the data that could only be captured using the multivariate approach adopted here. We now consider this matter in more detail.

The Dimensionality of Syllable Preference

As was just noted, the observed pattern of results is complex, and points to the fact that participants vary along multiple sensitivity dimensions, many of which contribute to apparent patterns of syllable preference in the word-splitting task. Participant sensitivity to difficulty factors, participant speed, and lexical frequency are each implicated in one case or another. These seeming codeterminants of syllable preference are discussed in what follows.

Participant sensitivity. The results reported here clearly indicate that a single dimension is incapable of capturing the observed individual differences in participant sensitivity. Further, the multiple dimensions identified here each have their own distinct relationship to word-splitting performance. In some cases no relationship was discerned, but in others the pattern of results was relatively complicated when compared to the straight-forward predictions that were developed in the introduction based on previous work.

Indeed, adopting a unidimensional approach, Taft (2001, 2002) reported that high-skill participants prefer the Max Coda syllable. As has been demonstrated, individual differences in word recognition performance are multi-faceted, only one of which yielded a results that was partially consistent with Taft (2001, 2002). Counter-examples were numerous. For instance, participants who were situated in the low print exposure and\or high-sensitivity end of some dimensions showed Max Coda preference under some conditions (e.g., Components 3 & 6: Orthographic Reader, Sensitivity to Word Attributes (Regularity)). If Max Coda preference is characteristic of orthographic readers (i.e., high print exposure and low-sensitivity), then such an association should not be observed. Equally inconsistent was the observation that low-sensitivity variance was positively associated with Max Onset preference (i.e., Component 2: Nonword sensitivity). If low-sensitivity is an indicator of efficient word recognition, then this means that Max Onset preference is characteristic of efficient word recognition. A clear Max Coda preference for strong 'orthographic reader' participants was only observed with Component 3 (Orthographic Reader) with fast responders, and even in that case weak orthographic slow responders also showed a Max Coda preference. It is obvious however that there is more to predicting syllable preference than that simple sensitivity to difficulty factors.

Participant sensitivity: independent of Print Exposure. We can distinguish, for instance, between sensitivity that is correlated with Print Exposure and sensitivity that is not. Sensitivity that is uncorroborated by low print-exposure may reflect a response strategy rather than a fundamental difference between high and low skill word recognizers. Two sensitivity components that were related to syllable preference yet unrelated to print exposure were Component 1 and Component 6. With Component 1, a non-significant trend provided some corroboration (i.e., in the meta-analysis sense) of the finding that participants who are insensitive to homophony tend prefer the Max Coda syllable, while those who are homophone sensitive prefer the Max Onset syllable (Component 1 of this paper, a lexical phonology indicator; Thompson & Desrochers, 2009b). With Component 6, sensitivity to semantics (and regularity) was positively associated with Max Coda preference with low frequency words. This finding is inconsistent with Max Coda preference being an attribute of efficient 'orthographic' readers. It is however consistent with the notion that Max Coda preference can emerge as the result of a tendency to rely on semantics, specifically semantic reliance via the indirect route (orthography-phonology-semantics). Alternatively, it is also consistent with the idea that Max Coda syllable preference is related to resolving inconsistencies (such as regularity) in the relationship between print and sound (Chateau & Jared, 2003). The former explanation seems more likely given the clear null effects that were obtained when regularity was considered as a predictor in isolation (Thompson & Desrochers, 2009b). The distinguishing feature of the present analysis is the variance shared between Imageability and Regularity sensitivity. Whatever the source of Max Coda preference, it is not clear that orthographic processing is the primary source.

Participant speed. Two dimensions interacted with Participant speed in determining syllable preference: Components 2 and 3. Both dimensions were correlated with Print Exposure, a known associate of participant response speed. In general, faster responders tend to perform better on reading skill tests and to be less sensitive to word recognition difficulty factors. Yet, slow-responders here preferred the Max Coda syllable, which was supposed to be a unit preferred by low-sensitivity, high-skill, orthographic readers. This unexpected result manifested itself in the way participant speed interacted with the observed sensitivity dimensions.

The strongest test of the idea that skill/sensitivity determines syllable preference was provided by Component 3, a dimension that was positively associated with Print Exposure and inversely related to sensitivity to nonwords difficulty factors. As noted above, the predicted results were obtained with the strong orthographic readers (as with high-skill readers, Taft, 2001, 2002) and low-frequency words (as in Chen & Vaid, 2007). The weak-orthographic slow responders however yielded a contradictory result: they too showed Max Coda preference with low-frequency words. It was suggested in the results that perhaps strong orthographic participants gain access to the information that generates Max Coda preference quickly, while weak orthographic participants only access this information after a delay. Such a mechanism would capture the observed response-speed dissociation.

One way of better understanding this response-time dynamic empirically would be to manipulate response criterion via participant instructions or factorial manipulation of the emphasis of speed and accuracy using a payoff matrix to bias responding predictably for speed, accuracy, or both (for discussion of how computational decision making theory relates to response biases induced by a payoff matrix, see Voss, Rothermund, & Voss, 2004). If apparent syllable preference is malleable and time-course dependent, then this should be reflected in the observed pattern of results. A number of interesting possibilities present themselves. For instance, could it be that all readers represent both Max Onset and Max Coda syllables, but only manifest this preference under particular circumstances (i.e., a particular response criterion)? The present results suggest that this may be the case, but additional experimentation is needed to confirm the reported pattern of results and map the boundary conditions on response speed for obtaining a syllable effect.

A weaker associate of print exposure was Component 2. The quick responders who were high in print exposure/low in sensitivity to nonword properties (as per this dimension) actually preferred the Max Onset syllable, which again is contrary to expectations. This result tends to support the idea that the nature of syllable effects is less a matter of personal preference than it is one of time course in the activation of linguistic knowledge. Unlike Component 3, this result was independent of frequency, perhaps because skilled readers are extremely efficient at accessing phonological information from print (Chateau & Jared, 2000). If the lexical decision task were made more difficult (e.g., by making it harder to discriminate between words and nonwords) or if instructions to participants emphasized accuracy, then fast responding would be discouraged this relationship might disappear, and perhaps be replaced by a Max Coda preference.

Frequency. In general, lexical frequency was an inconsistent moderator of the above relationships, only emerging in the analysis of Component 3 (Skill, Word Insensitivity) and Component 6 (Inefficient phonology: Regularity sensitivity, Sensitivity to Imageability). Chen & Vaid (2007) repeatedly found a Max Coda preference with low-frequency words with an unselected group of participants. The present study failed to detect such a general effect. Rather, the low-frequency specificity of syllable preference was conditional on participant characteristics. It is perhaps the case that the Chen & Vaid (2007) results were due to the combined contributions of the frequency-sensitive dimensions identified here and their interaction with word-splitting. The answer to such questions requires that more detailed descriptions of participant samples be collected and reported when syllable effects are investigated in the future.

Other issues: Modeling Sensitivity, syllable preference, and response time. Thompson & Desrochers (2009a) suggested that process-based variables were superior to comprehension skill as predictors of syllable preference. They argued that conceptualizing individual differences in terms of behavioral variables that are simulated by current models provides a surer way of understanding the sources of apparent syllable preference effects in lexical decision. The six variance components reported here provide a first approximation of how such process-based variables covary. Models that purport to simulate the effect of these letter-string difficulty factors would also be expected to account for these patterns of covariance. Developmental models like that proposed by Harm & Seidenberg (2004) or models of the static reading system like that proposed by Coltheart et al (2001) have parameters that control (directly or indirectly) the magnitude of such effects. These same parameters allow individual differences to be modeled. If a given parameter set is adequate, then it should be able to accommodate the heterogeneous covariance structure of the 'sensitivity' indicators considered here. An adequate model should be able to explain why certain sensitivities covary and others do not.

Modeling the relationship of such sensitivity to syllable preference requires more careful attention to additional factors. Existing models of reading assume that linguistic information is accessed gradually over time and have parameters than govern this spread of activation (Coltheart et al., 2001; Harm & Seidenberg, 2004), a property that allows such models to capture the individual differences in speed of access to information. Such models of the word recognition system could be combined with models of the decision mechanism (e.g., response timing, evidence evaluation; Ratcliff et al., 2004) to provide a complete account of how response speed affects the involvement of various sources of information, and in turn how this relates to syllable effects. The work reported here provides clues regarding where to begin situating the representation of syllables.

Conclusion

The present research attempted to verify a very simple proposition: do participants who are insensitive to various linguistic difficulty factors in lexical decision tend to prefer the Max Coda syllable. The answer to this question is not simple. When a broad set of participant characteristics are considered together, it is clear that several dimensions contribute to the apparent syllable preference of participants on the word-splitting task. In some cases, sensitivity to difficulty factors (taken as an indicator of weak orthography) is positively related to Max Coda preference, and in others Max Onset. Future work examining how such dimensions interact with explicitly manipulated respondent speed and non-manipulated participant-characteristics like Print Exposure may help to clarify the theoretical relevance of some of these dimensions.

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*, 234–254.

Andrews, S. (1997). The effect of orthographic similarity on lexical re- trieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*, 439–461.

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*, 390-412.

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

Briesemeister, B.B., Hofmann, M.J., Tamm, S., Kuchinke, L., Braun, M., & Jacobs, A.M. (2009). The pseudohomophone effect: Evidence for an orthography-phonology-conflict. *Neuroscience Letters, 455*, 124-128.

Briggs, P., & Underwood, G. (1987). The nature of reader ability differences in lexical access. *Journal of Research in Reading, 10*, 57-74.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory &* Cognition, 28, 143-153.

Chen, H.-C. & Vaid, J. (2007). Word frequency interacts with the BOSS in orthographic processing of English polysyllable words. *Language and Cognitive Processes, 22,* 52-82.

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Cortese, M.J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*, 384-387.

Desrochers, A., & Thompson, G.L. (2009). Subjective frequency and imageability ratings for 3,600 French nouns. *Behavior Research Methods*, *41*, 546-557.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin*, *123*, 71-99.

Harm, M. & Seidenberg, M. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Hutcheson, G. & Sofroniou, N. (1999). *The multivariate social scientist.* London: Sage.

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 130-136.

Jared, D., & Seidenberg, M. (1990). Naming multisyllabic words. *Journal of Experimental* Psychology: Human Perception and Performance, 16, 92-105.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? Journal of Experimental Psychology: General, 120, 358-394.

Jared, D., Levy, B., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General, 128*, 219-264.

Jordan, T. (1986). Testing the BOSS hypothesis: Evidence for position-insensitive orthographic priming in the lexical decision task. *Memory & Cognition, 14*, 523-532.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36.

Katz, L., & Baldasare, J. (1983). Syllable coding in printed-word recognition by children and adults. Journal of Educational Psychology, 75, 245-256.

Lewellen, M.J., Goldinger, S.D., Pisoni, D.E., & Greene, B.G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General, 122*, 316-330.

Lima, S. D., & Pollatsek, A. (1983). Lexical access via orthographic code? The Basic Orthographic Syllabic Structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior, 22*, 310-332.

Lorch, R. & Myers, J. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149-157.

McCann, R. S. , Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14*, 693-706.

New, B., Ferrand, L., Pallier, C., Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13*, 45-52.

O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*, 396-402. Retrieved November 12, 2008: http://people.ok.ubc.ca/brioconn/nfactors/nfactors.html

Pexman, P., & Lupker, S., (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects. *Canadian Journal of Experimental Psychology, 53*, 323-334.

Pexman, P., Lupker, S., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 139-156.

Pexman, P., Lupker, S., & Reggin, L. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28, 572-584.

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

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

Rouibahh, A., & Taft, M. (2001). The role of syllabic structure in French visual word recognition. *Memory & Cognition, 29*, 373-381.

Sears, C.R., Siakaluk, P.D., Chow,V.C., & Buchanan, L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research, 37,* 269-291.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention & Performance XII* (pp.245-263). Hillsdale, NJ: Erlbaum.

Seidenberg, M. S., Petersen, A., MacDonald, M. C., & Plaut, D. C. (1996). Pseudohomophone effects and models of word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 48-62.

Siakaluk, P.D., Sears, C.R., & Lupker, S.J. (2002). Orthogrpahic neighborhood effects in lexical decision: the Effects of nonwords orthographic neighborhood size. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 661-681.

Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16,* 32-71.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 402-433.

Strain, E., & Herdman, C.M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology*, *53*, 347-359.

Strain, E., Patterson, K., Seidenberg, M. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140-1154.

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 207-214.

Taft, M. (1979). Lexical Access via an orthographic Code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behaviour, 18*, 21-39.

Taft, M. (1987). Morphographic processing: The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and performance, XII: Reading* (pp. 265-299). Hillsdale, NJ: Erlbaum.

Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language & Speech, 44*, 351-376.

Taft, M. (2002). Orthographic processing of polysyllabic words by native and nonnative English speakers. *Brain & Language, 81*, 532-544.

Taft, M. & Russel, B. (1992). Pseudohomophone naming and the word frequency effect. *The Quarterly Journal of Experimental Psychology, 45*A, 51-71.

Thompson, G.L. (2008). Eliminating aggregation bias in experimental research: Random coefficient analysis as an alternative to performing a 'by-subjects' and/or 'by-items' ANOVA. *Tutorials in Quantitative Methods for Psychology, 4*, 21-34.

Thompson, G.L., & Desrochers, A. (2009a). Towards a Process-Based Characterization of Syllable Effects in Visual Word Recognition. Unpublished chapter of a doctoral dissertation, Cognitive Psychology Laboratory, University of Ottawa.

Thompson, G.L., & Desrochers, A. (2009b). Sensitivity to Homophone Interference as Correlate of Syllable Preference in Visual Word Recognition. Unpublished chapter of a doctoral dissertation, Cognitive Psychology Laboratory, University of Ottawa.

Unsworth, S.J. & Pexman, P.M. (2003). The impact of reader skill on phonological processing in visual word recognition. *The Quarterly Journal of Experimental Psychology, 56A,* 63-81.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation, *Memory & Cognition*, *32*, 1206-1220.

Wagenmakers, E-J, Ratcliff, R., Gomez, P. & McKoon, G. (2008). A diffusion model account of criterion shifts in lexical decision tasks. *Journal of Memory and Language*, *58,* 140-159.

Yates, M., Locker, L. JR., Simpson, G.B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones, *Memory and Cognition*, *31*, 856-866.

Ziegler, J., Jacobs, A., M., Kluppel, D. (2001). Pseudohomophone effects in lexical decision : Still a challenge for current word recognition models, *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 547-559.

Authors Note

Glenn L. Thompson and Alain Desrochers, Psychology Department, University of Ottawa, Ottawa, Ontario, Canada.

This research was funded in parts by graduate scholarships awarded to the first author by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Correspondence concerning this manuscript should be addressed to: Glenn L. Thompson, School of Psychology, University of Ottawa, 145 Jean-Jacques Lussier, P. O. Box 450, Station A, Ottawa, Ontario, Canada, KIN 6N6, or e-mail (GlennLThompson@gmail.com).

Table 1. Mixed-model tests of the sample's sensitivity to target item attributes on RT and ER (N=122).

Word Analysis (Effects evaluated separately, Experimental control)				HPD 95 Percent CI	
Attribute	Intercept	Coefficient	ER Z-value	Lower	Upper
Homophony-RT	665.87	44.97*	-	1.46	90.66
(Items=40, obs=4297)	(16.06)	(23.21)			
Homophony-ER	-3.16	0.53	1.07	-	-
(Items=40, Obs=4781)	(0.26)	(0.50)			
Regularity-RT	622.82	28.19	-	-14.77	69.04
(Items=20, Obs=2319)	(14.36)	(20.30)			
Regularity-ER	-3.86	0.22	0.63	-	-
(Items=20, Obs=2430)	(0.22)	(0.35)			
Imageability-RT	752.79	-114.34†	-	-159.3	-70.84
(Items=40, Obs=3140)	(17.60)	(22.67)			
Imageability-ER	-1.01	-2.36†	-7.55†	-	-
(Items=40, Obs=4819)	(0.18)	(0,31)			
Pseudoword Analysis (Evaluated Together, Statistical control)					
	Intercept	Coefficient	ER Z-value	Lower	Upper
Pseudohomophone –RT	794.26	51.92*		18.81	84.88
(Items=60, Obs=6214)	(20.89)	(16.90)			
Pseudohomophone –ER	-2.62	0.81	2.82*		
(Items=60, Obs=7220)	(0.17)	(0.29)			
Bigram Frequency- RT		-11.92		-28.33	5.16
		(8.74)			
Bigram Frequency - ER		-0.16	-1.08		
		(0.15)			
Orthographic N - RT		23.48*		7.01	39.87
		(8.48)			
Orthographic N - ER		0.32	2.16*		
		(0.15)			
Length –RT		50.94†		33.52	69.24
		(9.22)			
Length - ER		0.55	3.41†		
		(0.16)			

Note. For the columns Intercept and Coefficient, values within parentheses denote standard errors. ER (error rate) coefficients are logit values (i.e., log odds). Bigram Frequency, Orthographic N, and Length were standardized prior to analysis to achieve convergence. Statistical significance of RT coefficients determined by Markov Chain Monte Carlo. * $p < .05$, † $p < .001$

Table 2. Inter-correlations among participant-level predictors (N=122)

Variable Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Author Recognition																		
2. Subject RT [a] – W	-0.25																	
3. Subject ER [a] – W	-0.52	-0.04																
4. Subject RT [a] –PW	-0.08	0.79	-0.33															
5. Subject ER [a] – PW	-0.27	0.07	0.00	0.29														
6. IM Effect – RT [b]	0.24	-0.53	-0.16	-0.39	0.04													
7. IM Effect – ER [b]	0.45	0.06	-0.44	0.16	0.10	0.07												
8. Hom Effect -RT [b]	-0.14	0.47	0.05	0.37	0.00	-0.22	-0.09											
9. Hom Effect –ER [b]	-0.11	-0.02	0.23	-0.11	0.07	-0.07	-0.11	-0.26										
10. Reg Effect –RT [b]	-0.01	0.19	-0.01	0.20	-0.05	-0.06	0.04	0.03	0.08									
11. Reg Effect - ER [b]	-0.01	0.20	-0.01	0.14	0.03	-0.05	0.04	0.01	0.07	0.07								
12. PHom Effect -RT [b]	-0.15	0.17	0.00	0.00	0.12	-0.22	-0.10	0.03	0.06	-0.12	-0.01							
13. PHom Effect - ER [b]	-0.15	-0.01	0.13	-0.03	0.02	-0.04	-0.13	-0.13	0.01	-0.03	0.03	0.03						
14. BF Effect –RT [b]	-0.01	0.16	0.01	0.04	0.07	-0.01	0.19	0.23	0.09	0.06	0.05	-0.20	0.08					
15. BF Effect – ER [b]	0.13	-0.03	0.00	-0.16	0.04	0.11	0.08	-0.14	-0.02	-0.08	0.04	0.11	0.07	-0.04				
16. N Effect –RT [b]	0.03	0.04	-0.10	-0.03	-0.22	-0.20	-0.03	-0.16	0.02	-0.16	0.07	0.27	-0.11	-0.48	0.16			
17. N Effect - ER [b]	-0.08	0.03	-0.07	0.16	0.30	-0.05	0.05	-0.01	-0.06	0.12	-0.04	0.01	-0.08	-0.11	-0.32	-0.13		
18. L Effect - RT [b]	-0.09	0.14	-0.13	0.22	-0.07	-0.10	-0.08	-0.14	-0.03	0.04	0.02	-0.07	0.05	-0.49	-0.05	0.39	-0.01	
19. L Effect - ER [b]	-0.17	-0.08	0.18	0.02	0.46	0.00	0.00	0.03	0.09	0.06	-0.20	-0.07	-0.09	0.09	-0.32	-0.22	0.51	-0.04

Note. IM =Imageability, Hom = Homophony, Reg = Regularity, PHom = Pseudohomophony, BF = Bigram Frequency on nonwords, N = Orthographic Neighborhood on pseudowords, W= Words, PW=Pseudowords, RT = Reaction time, and ER = Error Rate. RT effects based on correct responses. ER effects expressed as logit values. Correlation values of 1 for nonwords errors reflect the fact that the model was over-parameterized. Correlations of .18 and greater are $p<.05$, Correlations of .28 and greater are $p<.001$

[a] Random Intercepts

[b] Random Slopes

Table 2. Participant variables: Loadings on the four dimensions extracted from the Principal Components analysis (N=122).

				Principal Component Dimensions					
Variable	Mean	SD	Communality	1	2	3	4	5	6
1. Author Recognition	10.37	5.83	0.68		-0.20	0.75	-0.26		
2. IM Effect – RT [b]	-114.33	139.15	0.61				-0.64	0.20	-0.34
3. IM Effect – ER [b]	-0.37	0.20	0.64			0.77			
4. Hom Effect -RT [b]	40.78	61.11	0.77	0.28			0.24	-0.78	
5. Hom Effect –ER [b]	0.04	0.09	0.72				0.23	0.75	0.25
6. Reg Effect –RT [b]	27.72	70.76	0.49		0.21				0.65
7. Reg Effect - ER [b]	0.01	0.09	0.47		-0.24				0.62
8. PHom Effect -RT [b]	51.14	86.50	0.67				0.75		-0.28
9. PHom Effect - ER [b]	0.08	0.11	0.43		-0.23	-0.54	-0.22		
10. BF Effect –RT [b]	-0.28	1.25	0.75	0.84					
11. BF Effect – ER [b]	0.00	0.00	0.48		-0.63				
12. N Effect –RT [b]	5.34	11.35	0.73	-0.68	-0.24		0.42		
13. N Effect - ER [b]	0.01	0.01	0.64		0.79				
14. L Effect - RT [b]	84.77	71.81	0.73	-0.81					0.20
15. L Effect - ER [b]	0.08	0.09	0.70		0.81				

Note. RT and ER denote reaction time and error rate respectively. Error rate is expressed here as in the form of logit coefficients, where increasingly negatively intercepts (i.e., means in the table) are associated with a corresponding decrease in the probability of observing a mistake and where for slopes, as usual, the value zero denotes the perfect absence of a relationship. Overall, the six dimensions explained 14.97 (after rotation: 13.21), 13.08 (13.00), 10.76 (10.60), 8.92 (9.67), 7.98 (8.78), and 7.64 (8.10) percent of variance in the indicators respectively for a total of 63.35 percent of variance explained. Loadings inferior to .20 were omitted.

Table 3. Mixed-Effects model of word-splitting performance on RT and ER containing only the item-level fixed-effects predictors (Level-1).

				HPD 95% CI[b]		Random Effects
	b	SE	t[a]	Lower	Upper	
RT- Intercept	644.03	11.45		624.97	661.85	
Word-Level						**S^2=2788**
Contrast A (1A)	0.74	2.96	0.25	-5.24	6.32	
Contrast B (1B)	3.99	3.42	1.17	-2.77	10.65	
Log F (F)	-29.03	2.26	-12.82*	-33.10	-24.83	
Complexity (C)	-3.07	7.83	-0.39	-17.17	11.71	
1A x F	2.35	1.77	1.33	-1.11	5.75	
1A x C	-4.14	5.91	-0.70	-15.97	7.28	
1B x F	-3.23	2.02	-1.60	-7.02	0.88	
1B x C	4.02	6.82	0.59	-9.32	17.49	
F x C	0.65	4.53	0.14	-7.54	8.94	
1A x F x C	0.13	3.51	0.04	-6.70	6.93	
1B x F x C	-1.92	4.01	-0.48	-9.62	6.13	
Participant-Level						**S^2=9278**
Residual						**S^2=26455**
	logit	SE	Z	Lower	Upper	
ER- Intercept	-1.69	0.10	-17.59*	-	-	
Word-Level						**S^2=1.32**
Contrast A (1A)	-0.03	0.04	-0.63	-	-	
Contrast B (1B)	-0.06	0.05	-1.23	-	-	
Log F (F)	-0.53	0.05	-11.62*	-	-	
Complexity (C)	-0.22	0.15	-1.44	-	-	
1A x F	-0.01	0.03	-0.52	-	-	
1A x C	-0.05	0.08	-0.58	-	-	
1B x F	-0.02	0.03	-0.64	-	-	
1B x C	-0.04	0.10	-0.41	-	-	
F x C	0.09	0.09	0.96	-	-	
1A x F x C	-0.04	0.05	-0.70	-	-	
1B x F x C	0.006	0.06	0.10	-	-	
Participant-Level						**S^2=0.39**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on 29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster). * $p < .05$, [†]$p < .001$ [a] Statistical significance determined via MCMC simulation.

Table 4. Mixed-Effects model of word-splitting performance on RT as a function of Item-Level (Level-1) and Mean RT as a participant-level predictor.

	b	SE	t[a]	HPD 95% CI[b] Lower	Upper	Random Effects
Intercept	650.68	8.15		635.68	664.76	
Word-Level						**S^2=2785**
Contrast A (1A)	0.85	2.95	0.29	-4.80	6.68	
Contrast B (1B)	3.87	3.41	1.14	-2.83	10.44	
Log F (F)	-29.01	2.26	-12.81	-33.25	-24.90	
Complexity (C)	-3.11	7.83	-0.40	-17.24	11.39	
1A x F	2.13	1.76	1.21	-1.34	5.70	
1A x C	-4.74	5.89	-0.80	-15.96	6.94	
1B x F	-2.99	2.01	-1.49	-6.85	0.93	
1B x C	4.32	6.80	0.64	-9.09	17.45	
F x C	0.67	4.53	0.15	-7.39	9.08	
1A x F x C	0.20	3.50	0.06	-6.32	7.34	
1B x F x C	-1.39	4.00	-0.35	-9.23	6.64	
Participant-Level						**S^2=5003**
Overall RT (RT)	0.68	0.05	12.31*	0.58	0.77	
RT x 1A	-0.03	0.02	-1.52	-0.08	0.01	
RT x F	-0.05	0.01	-10.30*	-0.04	0.06	
RT x C	0.02	0.02	1.07	-0.02	0.05	
RT x 1A x F	-0.01	0.01	-0.42	-0.03	0.02	
RT x 1A x C	-0.09	0.04	-2.08*	-0.18	-0.01	
RT x F x C	-0.01	0.01	-0.40	-0.02	0.02	
RT x 1A x F x C	0.01	0.02	0.38	-0.04	0.06	
Residual						**S^2=26247**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on 29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds. * p <.05, †p <.001

[a] Statistical significance determined via Markov Chain Monte Carlo simulation.

Table 5. Mixed-model of word-splitting RT with Item-Level predictors and Mean RT and Component 1 as Participant-level predictors.

	b	SE	t^{a}	$Lower^{b}$	$Upper^{b}$	Random Effects
Intercept	650.30	8.18		635.78	664.72	
Word-Level						**S^2=2784**
Contrast A (1A)	0.97	2.97	0.33	-4.70	7.06	
Contrast B (1B)	4.06	3.42	1.19	-2.52	7.06	
Log F (F)	-29.05	2.26	-12.84	-33.12	-24.66	
Complexity (C)	3.08	7.82	-0.39	-17.03	11.80	
1A x F	2.40	1.77	1.36	-1.06	5.83	
1A x C	-4.03	5.93	-0.68	-16.43	7.24	
1B x F	-3.20	2.02	-1.59	-7.17	0.72	
1B x C	4.29	6.82	0.63	-9.36	17.45	
F x C	0.69	4.52	0.15	-7.68	8.91	
1A x F x C	0.67	3.52	0.19	-5.87	7.79	
1B x F x C	-2.02	4.01	-0.50	-9.67	6.27	
Participant-Level						**S^2=5028**
Overall RT (RT)	0.66	0.06	11.29*	0.56	0.76	
RT x 1A	-0.03	0.02	-1.17	-0.07	0.02	
RT x F	-0.05	0.01	-9.50	-0.14	3.07	
RT x C	0.02	0.02	1.00	-6.78	3.76	
RT x F x C	-0.008	0.01	-0.74	-1.49	6.22	
RT x 1A x F	-0.0004	0.01	0.03	-0.03	0.03	
RT x 1A x C	-0.08	0.05	-1.69	-0.17	0.01	
RT x 1A x F x C	-0.02	0.03	0.62	-0.04	0.08	
Component 1(Compl)	-9.11	8.30	-1.10	-23.64	5.55	
Compl x 1A	-0.64	3.35	-0.19	-7.35	5.88	
Compl x F	0.72	0.71	1.01	-0.69	2.12	
Compl x C	-1.12	2.40	-0.46	-5.84	3.73	
Compl x 1A x F	2.69	1.99	1.35	-1.56	6.34	
Compl x 1A x C	12.76	6.45	1.91^{m}	-0.09	25.23	
Compl x F x C	-1.29	1.43	-0.90	-5.11	1.19	
Compl x 1A x F x C	1.39	3.97	0.35	-6.18	9.21	
Compl x RT	0.04	0.06	0.79	-0.05	0.14	
Compl x RT x 1A	-0.03	0.02	-1.51	-0.08	0.01	
Compl x RT x F	-0.009	0.01	-0.66	-0.03	-0.01	
Compl x RT x C	0.03	0.04	-0.78	-0.02	0.04	
Compl x RT x 1A x F	2.64	1.99	1.33	-0.04	0.02	
Compl x RT x 1A x C	12.8	6.69	1.91	-0.12	0.04	
Compl x RT x F x C	-0.009	0.03	-0.37	-0.01	0.03	
Compl x RT x 1A x F x C	-0.01	0.03	-0.39	-0.06	0.04	
Residual						**S^2=26247**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on

29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds, Comp1 =Component 1. $^{m}p < .10$, $^{*}p < .05$, $^{\dagger}p < .001$

[a] Statistical significance and [b] 95 percent confidence intervals determined via MCMC simulation.

Table 6. Mixed-model of word-splitting RT with Item-Level predictors and Mean RT and Component 2 as Participant-level predictors.

	b	SE	t[a]	Lower[b]	Upper[b]	Random Effects
Intercept	650.54	8.08		635.31	664.03	
Word-Level						**S^2=2779**
Contrast A (1A)	0.69	2.96	0.23	-5.22	6.44	
Contrast B (1B)	3.93	3.42	1.15	-2.65	10.82	
Log F (F)	-28.95	2.27	-12.78	-32.94	-24.60	
Complexity (C)	-3.14	7.83	-0.40	-17.51	11.31	
1A x F	2.30	1.77	1.30	-1.20	5.76	
1A x C	-4.28	5.91	-0.72	-16.07	7.24	
1B x F	-3.17	2.01	-1.57	-7.15	0.60	
1B x C	4.12	6.82	0.60	-8.73	17.58	
F x C	0.56	4.53	0.12	-7.78	8.90	
1A x F x C	0.39	3.51	0.11	-6.54	7.29	
1B x F x C	-1.67	4.01	-0.42	-9.45	6.34	
Participant-Level						**S^2=4967**
Overall RT (RT)	0.67	0.05	12.44*	0.57	0.77	
RT x 1A	-0.04	0.02	-1.82	-0.08	0.003	
RT x F	-0.05	0.005	-10.23*	-0.06	-0.04	
RT x C	0.02	0.02	1.08	-0.02	0.05	
RT x F x C	-0.004	0.01	-0.41	-0.03	0.02	
RT x 1A x F	-0.002	0.01	-0.14	-0.03	0.02	
RT x 1A x C	-0.09	0.04	-2.00*	-0.17	-0.003	
RT x 1A x F x C	0.02	0.03	0.58	-0.03	0.07	
Component 1(Comp1)	2.07	7.22	0.29	-10.97	14.43	
Comp2 x 1A	0.54	2.97	0.18	-5.17	6.39	
Comp2 x F	-0.97	0.71	-1.35	-2.39	0.46	
Comp2 x C	4.01	2.42	1.66	-0.68	8.98	
Comp2 x 1A x F	-3.25	1.76	-1.85	-6.59	0.23	
Comp2 x 1A x C	-10.22	5.94	-1.72	-21.93	1.17	
Comp2 x F x C	-0.57	1.43	-0.40	-3.35	2.18	
Comp2 x 1A x F x C	1.46	3.50	0.42	-5.27	8.35	
Comp2 x RT	-0.10	0.05	-2.16*	-0.19	-0.02	
Comp2 x RT x 1A	0.05	0.02	-2.30*	-0.09	-0.006	
Comp2 x RT x F	0.02	0.005	3.92*	0.01	0.03	
Comp2 x RT x C	-0.02	0.02	-1.24	-0.05	0.01	
Comp2 x RT x 1A x F	-0.007	0.01	-0.56	-0.03	0.02	
Comp2 x RT x 1A x C	-0.007	0.04	-0.18	-0.09	0.07	
Comp2 x RT x F x C	0.002	0.01	0.23	-0.02	0.02	
Comp2 x RT x 1A x F x C	0.002	0.02	0.09	-0.05	0.05	
Residual						**S^2=26214**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on

29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds, Comp2 = Component 2. $^{m}p < .10$, $^{*}p < .05$, $^{\dagger}p < .001$

[a] Statistical significance and [b] 95 percent confidence intervals determined via MCMC simulation.

Table 7. Mixed-model of word-splitting RT with Item-Level predictors and Mean RT and Component 3 as Participant-level predictors.

	b	SE	t[a]	Lower[b]	Upper[b]	Random
Intercept	649.52	8.27		633.95	663.66	Effects
Word-Level						**S^2=2798**
Contrast A (1A)	1.02	2.99	0.34	-5.10	6.72	
Contrast B (1B)	3.79	3.41	1.11	-2.69	10.81	
Log F (F)	-28.75	2.27	-12.67	-32.93	-24.56	
Complexity (C)	-2.39	7.84	-0.30	-16.93	12.02	
1A x F	1.38	1.80	0.77	-2.28	4.79	
1A x C	-3.91	5.99	-0.65	-15.03	8.15	
1B x F	-3.24	2.01	-1.61	-7.23	0.66	
1B x C	3.97	6.80	0.58	-9.61	17.31	
F x C	0.60	4.54	0.13	-7.51	8.93	
1A x F x C	0.01	3.57	<0.01	-6.45	7.47	
1B x F x C	-1.45	4.01	-0.36	-9.36	6.63	
Participant-Level						**S^2=4986**
Overall RT (RT)	0.66	0.06	11.18*	0.55	0.76	
RT x 1A	-0.03	0.02	-1.41	-0.08	0.01	
RT x F	-0.05	0.005	-8.72*	-0.06	-0.04	
RT x C	0.02	0.02	0.94	-0.02	0.05	
RT x F x C	0.003	0.01	0.24	-0.02	0.02	
RT x 1A x F	-0.02	0.01	-1.10	-0.04	0.01	
RT x 1A x C	-0.07	0.05	-1.45	-0.16	0.03	
RT x 1A x F x C	0.001	0.03	0.04	-0.05	0.06	
Component 1(Comp1)	-13.75	7.81	-1.76	-27.55	0.01	
Comp3 x 1A	-0.91	3.13	-0.29	-6.99	5.23	
Comp3 x F	0.89	0.74	1.19	-0.52	2.39	
Comp3 x C	-6.66	2.53	-2.63*	-11.51	-1.74	
Comp3 x 1A x F	-1.77	1.85	-0.96	-5.62	1.76	
Comp3 x 1A x C	6.09	6.24	0.98	-6.36	18.08	
Comp3 x F x C	2.64	1.49	1.77	-0.30	5.52	
Comp3 x 1A x F x C	-4.33	3.67	-1.18	-11.72	2.84	
Comp3 x RT	-0.02	0.07	-0.34	-0.14	0.09	
Comp3 x RT x 1A	0.008	0.03	0.31	-0.05	0.06	
Comp3 x RT x F	0.02	0.006	2.36*	0.002	0.02	
Comp3 x RT x C	0.02	0.02	0.99	-0.02	0.06	
Comp3 x RT x 1A x F	-0.03	0.02	-1.98*	-0.07	-0.001	
Comp3 x RT x 1A x C	0.06	0.05	1.07	-0.05	0.16	
Comp3 x RT x F x C	0.01	0.01	0.94	-0.01	0.04	
Comp3 x RT x 1A x F x C	-0.01	0.03	-0.42	-0.07	0.05	
Residual						**S^2=26215**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on

29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds, Comp3 = Component 3. [m] $p < .10$, * $p < .05$, † $p < .001$

[a] Statistical significance and [b] 95 percent confidence intervals determined via MCMC simulation.

Table 8. Mixed-model of word-splitting RT with Item-Level predictors and Mean RT and Component 4 as Participant-level predictors.

	b	SE	t[a]	Lower[b]	Upper[b]	Random
Intercept	651.95	8.27		637.51	666.60	Effects
Word-Level						**S^2=2791**
Contrast A (1A)	1.12	3.02	0.37	-4.77	7.13	
Contrast B (1B)	3.88	3.41	1.14	-2.69	10.48	
Log F (F)	-29.33	2.27	-12.93*	-33.39	-25.09	
Complexity (C)	-2.60	7.84	-0.33	-16.66	11.89	
1A x F	1.40	1.80	0.78	-2.07	4.96	
1A x C	-4.37	6.04	-0.72	-15.19	8.21	
1B x F	-2.80	2.01	-1.39	-6.62	1.24	
1B x C	4.20	6.80	0.62	-9.16	17.51	
F x C	0.47	4.54	0.10	7.72	9.10	
1A x F x C	-0.22	3.58	-0.06	-7.16	6.81	
1B x F x C	-1.18	4.00	-0.29	-9.12	6.71	
Participant-Level						**S^2= 4879**
Overall RT (RT)	0.71	0.06	12.54*	0.61	0.81	
RT x 1A	-0.03	0.02	-1.42	-0.08	0.02	
RT x F	-0.06	0.005	-10.88*	-0.07	-0.05	
RT x C	0.02	0.02	1.23	-0.01	0.06	
RT x F x C	-0.004	0.01	-0.38	-0.19	-0.008	
RT x 1A x F	-0.01	0.01	-0.98	-0.04	0.01	
RT x 1A x C	-0.10	0.05	-2.21*	-0.19	-0.008	
RT x 1A x F x C	0.006	0.03	0.22	-0.05	0.06	
Component 4 (Comp4)	-13.25	9.52	-1.39	-30.28	3.19	
Comp4 x 1A	0.02	3.89	0.01	-7.43	7.89	
Comp4 x F	1.66	0.93	1.78[m]	-0.18	3.48	
Comp4 x C	-0.29	3.17	-0.09	-0.19	-0.008	
Comp4 x 1A x F	1.67	2.33	0.72	-0.04	0.01	
Comp4 x 1A x C	7.48	7.77	0.96	-0.19	-0.008	
Comp4 x F x C	-0.78	1.86	-0.42	-4.29	3.00	
Comp4 x 1A x F x C	-0.53	4.63	-0.11	-9.19	8.96	
Comp4 x RT	-0.03	0.04	-0.62	-0.10	0.05	
Comp4 x RT x 1A	0.006	0.01	0.70	-0.04	0.03	
Comp4 x RT x F	0.009	0.004	2.09*	0.0004	0.02	
Comp4 x RT x C	-0.01	0.01	-0.95	-0.04	0.01	
Comp4 x RT x 1A x F	0.02	0.01	1.51	-0.005	0.04	
Comp4 x RT x 1A x C	-0.01	0.04	-0.26	-0.08	0.06	
Comp4 x RT x F x C	0.006	0.01	0.70	-0.01	0.02	
Comp4 x RT x 1A x F x C	0.007	0.02	0.33	-0.04	0.05	
Residual						**S^2=26218**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on 29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A=

Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds, Comp4 = Component 4. [m] $p < .10$, *$p < .05$, [†] $p < .001$

[a] Statistical significance and [b] 95 percent confidence intervals determined via MCMC simulation.

Table 9. Mixed-model of word-splitting RT with Item-Level predictors and Mean RT and Component 5 as Participant-level predictors.

	b	SE	t[a]	Lower[b]	Upper[b]	Random Effects
Intercept	651.20	8.48		636.67	666.75	
Word-Level						**S^2= 2801**
Contrast A (1A)	1.45	3.06	0.47	-5.01	7.10	
Contrast B (1B)	3.61	3.41	1.06	-2.59	10.86	
Log F (F)	-29.74	2.28	-13.06*	-33.92	-25.39	
Complexity (C)	-3.44	7.87	-0.44	-18.28	10.62	
1A x F	1.74	1.83	0.95	-1.80	5.38	
1A x C	-3.91	6.12	-0.64	-15.70	7.70	
1B x F	-2.99	2.01	-1.49	-6.99	0.84	
1B x C	4.57	6.80	0.67	-8.83	17.64	
F x C	1.03	4.55	0.23	-7.73	9.13	
1A x F x C	1.41	3.65	0.39	-5.67	8.54	
1B x F x C	-1.49	4.00	-0.37	-9.08	6.58	
Participant-Level						**S^2=5091**
Overall RT (RT)	0.66	0.06	11.12*	0.56	0.76	
RT x 1A	-0.04	0.02	-1.58	-0.08	0.008	
RT x F	-0.06	0.006	-10.42*	-0.07	-0.05	
RT x C	0.02	0.02	0.88	-0.02	0.06	
RT x F x C	-0.008	0.01	-0.72	-0.03	0.02	
RT x 1A x F	-0.005	0.01	-0.39	-0.03	0.02	
RT x 1A x C	-0.09	0.05	-1.85	-0.03	0.02	
RT x 1A x F x C	0.003	0.03	0.11	-0.05	0.06	
Component 5(Comp5)	-7.44	8.08	-0.92	-21.61	6.77	
Comp5 x 1A	-3.24	3.29	-0.98	-9.63	3.18	
Comp5 x F	-0.27	0.79	-0.34	-1.88	1.22	
Comp5 x C	0.09	2.67	0.04	-5.29	5.19	
Comp5 x 1A x F	0.68	1.95	0.35	-3.15	4.47	
Comp5 x 1A x C	0.01	6.57	0.002	-12.32	13.19	
Comp5 x F x C	-2.32	1.57	-1.47	-5.39	0.73	
Comp5 x 1A x F x C	-6.24	3.87	-1.61	-14.30	1.06	
Comp5 x RT	0.01	0.04	0.20	-0.07	0.09	
Comp5 x RT x 1A	0.01	0.02	0.60	-0.03	0.04	
Comp5 x RT x F	-0.02	0.004	-3.58*	-0.02	-0.007	
Comp5 x RT x C	-0.007	0.01	-0.46	-0.04	0.02	
Comp5 x RT x 1A x F	-0.01	0.01	-1.00	-0.03	0.01	
Comp5 x RT x 1A x C	0.02	0.04	0.45	-0.05	0.09	
Comp5 x RT x F x C	0.005	0.008	0.58	-0.01	0.02	
Comp5 x RT x 1A x F x C	0.02	0.02	0.94	-0.02	0.06	
Residual						**S^2=26228**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on

29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds, Comp5 = Component 5. [m] $p < .10$, *$p < .05$, [†]$p < .001$

[a] Statistical significance and [b] 95 percent confidence intervals determined via MCMC simulation.

Table 10. Mixed-model of word-splitting RT with Item-Level predictors and Mean RT and Component 6 as Participant-level predictors.

	b	SE	t^a	$Lower^b$	$Upper^b$	Random Effects
Intercept	649.70	8.71		634.53	665.69	
Word-Level						**S^2=2808**
Contrast A (1A)	0.98	3.16	0.31	-4.96	7.37	
Contrast B (1B)	3.76	3.41	1.10	-2.86	10.51	
Log F (F)	-28.67	2.28	-12.57*	-32.99	-24.51	
Complexity (C)	-3.01	7.88	-0.38	-17.31	11.80	
1A x F	2.05	1.89	1.09	-1.70	5.72	
1A x C	-0.13	6.31	-0.02	-12.41	12.37	
1B x F	-2.91	2.01	-1.45	-6.76	1.23	
1B x C	4.41	6.80	0.65	-8.63	17.87	
F x C	0.84	4.56	0.18	-7.55	9.01	
1A x F x C	0.22	3.76	0.06	-7.14	7.57	
1B x F x C	-1.48	4.00	-0.37	-9.26	6.53	
Participant-Level						**S^2=5087**
Overall RT (RT)	0.68	0.06	10.68*	0.56	0.79	
RT x 1A	-0.03	-0.03	-1.30	-0.08	0.02	
RT x F	-0.04	0.006	-6.58*	-0.05	-0.03	
RT x C	0.02	0.02	1.11	-0.02	0.06	
RT x F x C	-0.003	0.01	-0.28	-0.03	0.02	
RT x 1A x F	-0.003	0.02	-0.20	-0.03	0.03	
RT x 1A x C	-0.05	0.05	-0.90	-0.03	0.02	
RT x 1A x F x C	0.007	0.03	0.22	-0.06	0.06	
Component 6 (Comp6)	-2.43	8.22	-0.30	-16.31	12.48	
Comp6 x 1A	-0.20	3.38	-0.06	-6.52	6.57	
Comp6 x F	-3.36	0.82	-4.10*	-4.89	-1.72	
Comp6 x C	-1.10	2.75	-0.40	-6.38	4.26	
Comp6 x 1A x F	-0.63	2.04	-0.31	-4.57	3.45	
Comp6 x 1A x C	-4.72	6.76	-0.70	-17.77	8.54	
Comp6 x F x C	0.23	1.63	0.14	-3.00	3.38	
Comp6 x 1A x F x C	0.41	4.06	0.10	-7.81	8.19	
Comp6 x RT	0.02	0.05	0.33	-0.06	0.09	
Comp6 x RT x 1A	0.0006	0.02	0.03	-0.04	0.04	
Comp6 x RT x F	-0.005	0.004	-1.24	-0.01	0.003	
Comp6 x RT x C	-0.0004	0.01	-0.03	-0.03	0.03	
Comp6 x RT x 1A x F	0.0002	0.01	0.02	-0.02	0.02	
Comp6 x RT x 1A x C	-0.07	0.04	-1.94^m	-0.14	0.004	
Comp6 x RT x F x C	-0.002	0.009	-0.27	-0.02	0.02	
Comp6 x RT x 1A x F x C	0.0009	0.02	0.04	-0.04	0.04	
Residual						**S^2=26223**

Note. RT analysis based on 18, 387 observations (Correct responses, after outlier screening) and ER based on

29, 280 observations, 24.1 percent of which were errors (240 stimuli, 122 Participants). F=Frequency, A= Max Onset vs Max Coda, B= Syllable Heuristic versus Comparison condition (Max Coda plus 1), C =Syllable Complexity (single consonant versus consonant cluster), RT= Reaction Time in milliseconds, Comp6 = Component 6. [m] $p < .10$, * $p < .05$, † $p < .001$

[a] Statistical significance and [b] 95 percent confidence intervals determined via MCMC simulation.

Chapter 6

General Discussion: Characterizing the Syllable Effect in Visual Word Recognition: A Novel Methodology & New Evidence

Abstract

Thompson & Desrochers (2009a) argued that extending single-syllable models to the case of multi-syllable words requires additional empirical information. For instance, what syllable units are relevant (e.g., Max Onset, Max Coda), when and why are they relevant, and for whom? A common paradigm for addressing this issue is the word-splitting manipulation in visual lexical decision, where a visual boundary is introduced within words that is either consistent or inconsistent with a syllable boundary. If an advantage is observed for the consistent condition, then a 'syllable effect' has been obtained. Recent work has focused on identifying variables that moderate the effect of such word-splitting manipulations in an attempt to explain the empirical inconsistency in this literature (Chen & Vaid, 2007; Taft, 2001, 2002). The results of a research program designed to contribute to this area (see Thompson & Desrochers, 2009b) are reviewed. In the first place, this research serves to illustrate how random coefficient analysis and multi-level regression can be used in a flexible manner to efficiently address subtle psycholinguistic issues. In the second place, a number of moderators of the word-splitting effect were confirmed or discovered: lexical frequency, syllabic complexity, print exposure and various indices of participant sensitivity to word and nonword difficulty factors. The results have both practical and theoretical implications for the syllable effect literature in that they further specify the conditions under which syllable effects are likely to be observed and they link syllable preference to model parameters via the participant sensitivity indicators.

Introduction

Thompson & Desrochers (2009a, 2009b) outlined a research program that proposed to clarify the role of syllable-like units in visual word recognition, thereby removing an obstacle to the extension of computational models of monosyllabic word processing to the more general case of multi-syllable words. This research program required the use of a novel combination of random coefficient analysis and multi-level modeling (Thompson, 2008; Thompson & Desrochers, 2009c, 2009d) to explore the contingencies among item-level and participant-level variables that could potentially reveal stable and theoretically interpretable conditions under which syllable effects emerge. The term syllable effect here refers to the finding in the word-splitting literature that words are recognized more rapidly when they are divided in a manner consistent with a syllable boundary than when divided elsewhere in a letter string. Such findings are interpreted as evidence that the syllable boundary is psychologically real or functionally relevant.

The operational definition of such word-splitting manipulations can take many forms, including the introduction of a mid-word foreign character (THUN/DER) or change of case (THUNder) or letter color (e.g., Rouibahh & Taft, 2001; Thompson & Desrochers, 2009c, 2009d). Following previous work (for reviews, see Chen & Vaid, 2007; Taft, 2001), Thompson & Desrochers (2009b) proposed to contrast two syllabification heuristics: the Max Onset syllable (THUN/DER) and the Max Coda syllable (THUND/ER) using such word-splitting techniques in the visual lexical decision task. Within the context of such a comparison, syllable effects are termed instead 'syllable preference' effects, where division according to one or the other heuristic yields a relative advantage. Two general conclusions can be extracted from their research: a) participants who are sensitive to homophone interference tend to prefer the more 'phonological' type of syllable (Thompson & Desrochers, 2009c, 2009d) and b) when a broad set of participant-level indicators is considered, sources of individual differences appear to be heterogeneous, each free to vary independently with the syllable preference tendencies of participants (Thompson & Desrochers, 2009d). Interestingly, some variance components interacted with print exposure and/or frequency and/or participant response speed in predicting syllable preference, while others did not. Some dimensions were correlated with print exposure, while others were not. Taken together, the findings serve to emphasize that syllable effects are determined by multiple parallel factors and subtle interactions.

In what follows, the results of this research program are discussed in more detail. First, the broad lines of the research are sketched. Second, the methodology used to investigate questions related to the syllable effect is discussed in terms of their technical contribution. The preceding sections serve to set the ground for a discussion of the variables that act in concert to produce apparent patterns of syllable preference. Some of these variables are attributes of words and others are attributes of participants. At this point, the practical and theoretical implications of these findings are discussed. The paper concludes by highlighting some of the limitations of the present research program and by offering some concluding remarks.

A Research Program in Review

Thompson & Desrochers (2009a) argued that ambiguity with respect to the role of the syllable in visual word recognition is an important obstacle for most computational models of word recognition. Such models are currently only equipped to handle single-syllable words, and extensions to handle multi-syllable words are a logical next step (e.g., see Harm & Seidenberg, 2004). To take this step, theorists need to understand how and where such units are represented within the word recognition system. With such information, the most basic modeling decisions can be taken so that a first approximation of an integrated theoretical account of single-syllable and multi-syllable word processing may be undertaken.[1] An important limit to current understanding of how this might be achieved is the fact that the empirical literature comprises a number of inconsistent findings (for reviews, see Chen & Vaid, 2007; Taft, 2001; Thompson & Desrochers, 2009a). In an attempt to address this issue, recent work has focused on identifying the characteristics of participants (e.g., comprehension skill, Taft, 2001) and words (e.g., lexical frequency; Chen & Vaid, 2007) that could reveal the contingencies at the root of this inconsistency, allowing theoretical work to move forward.

Thompson & Desrochers (2009b) set about pursuing this line of inquiry by way of a research program designed to indentify item- (e.g., frequency) and participant-level (e.g., print exposure) moderators of syllable preference. As noted earlier, syllable preference here refers to the finding in the word-splitting literature that words are recognized more rapidly when they are divided in a manner consistent with a syllable boundary. Following Chen & Vaid (2007) and Taft (2001), Thompson & Desrochers (2009b) placed the focus on contrasting two syllabification heuristics: the Max Onset syllable (THUN/DER) and the Max Coda syllable (THUND/ER). They proposed to take up investigation of this problem using a novel combination of random coefficient analysis and multi-level modeling. Following this suggestion Thompson & Desrochers (2009c, 2009d) set about mapping an extended set of conditions under which preference for one or the other type of syllable would be demonstrated.

Item- and participant-level predictors. Across two studies, item- and participant-level associates of specific types of syllable preference were identified. The first study reported the results of three experiments (Thompson & Desrochers, 2009c; 3 x N=48) while the second (Thompson & Desrochers, 2009d) reported the results of a larger-sample correlation analysis (N=122). In both cases, item-attributes that were considered relevant in previous research were examined (for discussion, see Thompson & Desrochers, 2009c), including syllabic complexity (a single inter-vocalic consonant versus multiple between-vowel consonants, i.e., TU<u>L</u>IP versus THU<u>ND</u>ER) and lexical frequency (e.g., Chen & Vaid, 2007). Both complexity and frequency were found to be consistent moderators of apparent syllable preference (see more detailed discussion below; Thompson & Desrochers, 2009c,2009d). Lack of systematic control of these factors in previous research may have contributed to the empirical instability in the literature (see also, Chen & Vaid, 2007). Previous work examining the role of participant-level moderators had focused on distal predictors of syllable preference like reading comprehension skill (Taft, 2001). Thompson & Desrochers (2009a, 2009b, 2009c, 2009d) argued instead that individual differences should be operationalized at the level of single word online processing (i.e., at the level of current models of word recognition), reasoning that the behavior measured should be a behavior that is simulated directly by the models of reading that are of current interest (e.g., Coltheart et al., 2001; Harm & Seidenberg, 2004), which is to say a 'process-based' behavioral variable. Process-based variables, whose values are controlled by some reasonably well-defined element of the cognitive machinery that presides over word recognition, are termed indices of participant sensitivity and used as predictors of syllable preference in what follows. The advantage of such a strategy is that any association that is observed between a process-based variable and syllable preference is directly interpretable in terms of existing models. This interpretability arises from the fact that certain model parameters are known to control the magnitude of process-based effects. An association with an empirically estimated process-based variable that varies across participants is akin to an association with changing values of the theoretical parameter than controls it. Note that the goal of this enterprise was not to explain the skill-syllable preference relationship reported by Taft (2001) per se, but primarily to operationalize individual differences in a different 'online' way and map these differences to syllable preference.

Initially, the focus was on a specific type of process-based variable, specifically the extent to which participants were susceptible to interference from phonological information when recognizing words visually (Thompson & Desrochers, 2009c). The idea here was to directly test a hypothesis for the differential patterns of syllable preference reported by Taft (2001). If Taft's low-skill comprehenders preferred the Max Onset syllable because they relied preferentially on phonology, then surely some direct evidence of this tendency could be picked up using process-based estimates of phonological reliance.

Here a key idea was introduced: sensitivity. Sensitivity to a letter-string attribute refers to a participant's susceptibility to having their performance influenced by an attribute that makes word recognition more difficult, a difficulty factor. In other words, the term sensitivity is used here to denote sub-optimal performance (i.e., focusing attention on information that tends to hinder rather than help performance). At least two accounts of such sensitivity might be envisaged. The first would posit that the ability to suppress irrelevant information varies across participants. According to this view, all sources of information are equally available for each participant, but control of this information varied. The second view would attribute the source of sensitivity to a lack of a choice, whether implicit or explicit, in selecting information for the decision-making process. Here certain participants are sensitive to phonology for example because this is they lack other sources of information to fall back on when such information becomes unreliable. Thompson & Desrochers (2009d) argued for the second view, but both accounts are potentially useful in this context. Some combination of the two accounts could be at work. In both cases, a lack of flexibility or control would be the root of the larger effect of some phonological difficulty factors with poor readers. In contrast, efficient word recognizers have the option of focusing on more reliable sources of information and/or a greater degree of control. Individual differences in the weighting of such information in the decision making process can be modeled either by assuming that the value of certain model parameters can be varied strategically (for an example of how emphasis of different types of information can be modeled within models of word recognition, see Kello & Plaut, 2003) or by assigning the role of weighting various sources of information and response timing to an independent decision-making process (see Norris, 2006; Ratcliff, Gomez, & McCoon, 2004; Wagenmakers, Steyvers, Raaijmakers, Shiffrin, van Rijn, & Zeelenberg, 2004). The obtained results partially supported the initial hypothesis, but the link with syllable preference was not observed with the index of phonological sensitivity that was most strongly associated with skill in previous work (i.e., regularity; Unsworth & Pexman, 2003). Instead, the link with syllable preference was observed with sensitivity to the homophone interference effect. The significance of this finding is taken up in more detail below.

Once the link between sensitivity to phonology and syllable preference was examined, the idea of sensitivity was subsequently applied more generally to a variety of characteristics that can vary so as to make words and nonwords more difficult to discriminate in the lexical decision task. This list of characteristics touched upon each of the three principle domains of knowledge: phonology, orthography, and semantics. Multiple indices of participant sensitivity were estimated, which allowed their covariance to be evaluated (i.e., to gain a better understanding of their covariance structure) and to then be entered as a predictor of syllable preference. Thus, the goal was to examine the idea of participant 'sensitivity' more closely as a construct in addition to studying its relationship to syllable preference. A distal 'skill' measure was also taken (i.e., print exposure) to potentially link these observed individual differences in participant sensitivity with a developmental process that is indirectly related to word-recognition performance, namely reading experience (for an example of how experience can cause fundamental differences in the way information is accessed in the word recognition system, see Harm & Seidenberg, 2004). This issue is discussed in more detail below.

Individual differences: a participant sensitivity approach. Initially, Thompson & Desrochers (2009c) failed to observe a strong relationship between homophone sensitivity and the magnitude of the regularity effect. This observation is consistent with the partial dissociation of these effects reported elsewhere (Unsworth & Pexman, 2003), and it does not support the idea of a general sensitivity factor that would capture a common source of individual differences in sensitivity to a variety of difficulty factors. This idea was examined more formally by Thompson & Desrochers (2009d) with a broad set of sensitivity indicators. In this case, when the basic idea of 'sensitivity to difficulty factors' was extended beyond the homophone interference and regularity effects to include a wide range of predictors, including phonological (e.g., pseudohomophone), semantic (e.g., imageability), and orthographic (e.g., orthographic N) difficulty factors, the pattern of covariance was found to be heterogeneous in a principal components analysis (Thompson & Desrochers, 2009d). What distinguishes a principal components analysis from a factor analysis is that the algorithm attempts to maximize explanation of variance, both that which shared among predictors and unique to individual predictors. Thus, the goal was to summarize an important percentage of the observed individual differences with a reduced set of variables rather than to focus strictly on latent variables which are based exclusively on shared variance.

With this in mind, six noteworthy variance components were identified (arrows denote sign of loading): 1) Lexical Phonology ↑ & Nonword Sensitivity ↓ (RT), 2) Sensitivity to nonword difficulty factors ↑ (accuracy) and print exposure ↓, 3) ↑ Orthographic Reading (accuracy), 4) ↑ Sensitivity to various Word and Nonword difficulty factors (RT), 5) ↑ Speed-Accuracy Tradeoff on the homophone effect, and 6) ↑ Sensitivity to word difficulty factors (Imageability, Regularity). These variance components summarized approximately 64 percent of the variance in the Thompson & Desrochers (2009d) participant attribute indicators. This amount of explained variance was relatively modest, but the residual variance in the observed variables was found to be unrelated to syllable preference in any case. In sum, the main goal of summarizing important sources of individual differences in this set of indicators was achieved.

As an exercise in understanding the covariance of sensitivity indicators, the principal components analysis is a good first step. It is of course an open question whether this structure would generalize to other samples, and to indices of sensitivity estimated under different experimental conditions (e.g., emphasizing speed over accuracy; different list composition). Nevertheless, these findings strongly suggest a single dimension does not underlie participant sensitivity. If this were the case, stronger inter-correlations would have been observed among all indicators. The absence of any semblance of a general factor calls the idea of an 'insensitive' orthographic reader into question. The real story appears to be more nuanced, requiring that time-course of activation and response timing mechanism be invoked (see below).

Because some of the components were associated with Print Exposure and others were not, it is tempting to speculate that the components that are associated with Print Exposure are the result of a more stable difference between participants arising from varying degrees of experience with print, while other variance components may reflect a potentially less stable strategic choice, such as modifying a response-time criterion (Lupker, Brown, Colombo, 1997) or emphasizing a particular source of information in the decision-making process (Kello & Plaut, 2003; Ratliff et al., 2004). Even given the possibility of flexibility (e.g., word recognition skill), some readers may choose to emphasize particular sources of information. Regardless of whether such individual differences are the result of skill or strategic choice, the relationship of these components to syllable preference provides a source of contingency that may serve to pin down their theoretical locus. Regardless of the conditions under which such an association is observed, an association between a sensitivity (process-based variable) and syllable preference remains interpretable.

Participant sensitivity and syllable preference. The main focus of the research was not an examination of the factor structure of participant sensitivity to difficulty factors.

Rather, its main purpose was to map out the conditions under which syllable effects occur, which an emphasis on participant sensitivity as a potential moderator. Indeed, the idea that participant 'sensitivity' to difficulty factors of various configurations could predict syllable preference was tested repeatedly with the variance components discussed above, yielding a large set of results. The most consistent of these was the finding that feedback phonology (i.e., sensitivity to the homophone interference effect) tends to be positively associated with Max Onset preference with low-frequency words (especially syllabically complex words) across experiments (Thompson & Desrochers, 2009c, 2009d). The relationship was not always found to be statistically significant (i.e., Thompson & Desrochers, 2009c, Experiment 1, Thompson & Desrochers, 2009d), but the pattern is consistent enough across studies to be considered reliable.

As Thompson & Desrochers (2009a) argued, such an empirical fact, obtained with so-called 'process-based' variables, is more theoretically useful than a general measure of reading comprehension skill in that the level of analysis is at the same level as the theoretical models (i.e., isolated words) and the variable considered (i.e., sensitivity to homophone interference) is a behavior that is explicitly simulated by current computational models (Coltheart, Rastle, Langdon, Perry, & Ziegler, 2001; Harm & Seidenberg, 2004). In this case, it would appear that apparent Max Onset syllable effects are controlled by lexical phonology. The parameter(s) that control the magnitude of lexical phonology effects in visual lexical decision would, by extension, be expected to control the magnitude of Max Onset syllable preference effects. A first attempt at modeling Max Onset syllable effects might involve a processing mechanism implicated with lexical phonology. For example, this Max Onset syllable structure could be explicitly coded in terms of consonant-vowel structure frames for phonological units (Plaut et al., 1996). The relative weighting given to activation arising from lexical phonology could then be varied to see whether the observed effects in lexical decision can be simulated. If a given model cannot simulate the observe pattern of results, then one might consider whether the architecture of the model is flawed in some way.

The larger regression study involving the six variance components mentioned above did not yield a neat pattern with respect to the relationship between sensitivity and syllable preference. It was expected that low-sensitivity participants would prefer the Max Coda syllable because such participants would be high-skill. It was further expected that high-sensitivity participants would prefer the Max Onset syllable because such participants would be relatively low skill at recognizing words. Thompson & Desrochers (2009d) found limited support for this idea, but more general support for the idea that the relationship between some types of sensitivity and syllable preference is moderated by overall response speed and/or frequency in unanticipated ways.

To accommodate this finding, they proposed that apparent syllable preference may vary as a function of time-course of the activation of information. Participants who tend to respond early will demonstrate syllable preference consistent with the information that is most active at the time the response was made. Similarly, participants who tend to respond late will demonstrate syllable preference consistent with the information that is dominant at the time they made the response. The time-course of access to various sorts of information may be supposed to vary as a function of reader skill (Chateau & Jared, 2000). If we assume that the information that gives rise to Max Coda syllable preference is available early to high-skill word recognizers and becomes available later to low-skill word recognizers, then we might expect the pattern of results that was obtained by Thompson and Desrochers (2009d; Component 3 with low-frequency words). High-skill recognizers who respond quickly show the effect and so do low-skill slow responders, but others tend to respond at a point when competing tendencies towards Max Coda and Max Onset preference cancel out, or when the source of Max Coda preference is not sufficiently strong to generate an observable effect. The appealing aspect of this account is that it proposes a single theoretical source of Max Coda preference regardless of skill level. A similar mechanism could be invoked to explain the counter-intuitive results that Thompson & Desrochers (2009d) obtained with Components 2 & 6. In both cases, we could assumed that all readers are capable of demonstrating either Max Onset or Max Coda preference if they tend to respond at the 'right' time.

Determining the validity of this explanation would require follow-up experimentation where the 'natural' response speed of participants is manipulated either through instructions (e.g., emphasizing accuracy), by making the lexical decision task more difficult, or by influencing behavior systematically via reward (e.g., a payoff matrix). The payoff matrix is particularly interesting in that it should reduce variability in the adherence to instructions. For example, rewarding both speed and accuracy would achieve the goal of speeding up responding as much as possible while minimizing problematic behavior like a speed-accuracy tradeoff. In this case, heterogeneity in the setting of internal response criterion will be reduced. If the same participants can be induced to alter their syllable preference under fast and slow conditions, this would tend to support the interpretation that we have given to the results reported here. Modeling such effects would require a computational mechanism to account for individual differences in time-course of activation (see models such as Coltheart et al., 2001; Harm & Seidenberg, 2004; and related models based on the principles of activation) and a mechanism for controlling response criterion.

Random Coefficients and Multi-level analysis

In addition to its empirical contribution (reviewed in more detail below), the present research program acts as a demonstration in principle of how random coefficient analysis and multi-level analysis can be used in conjunction to investigate subtle psycholinguistic issues. Random coefficient analysis provides a natural way of modeling individual differences in sensitivity to item attributes (Thompson, 2008; Thompson & Desrochers, 2009c, 2009d). The regression framework used here is flexible enough to accommodate both categorical and continuous predictors, which means that naturally continuous variables like lexical frequency do not have to be scaled down by factorization (for a discussion of factorization, see Desrochers, Thompson, & Fréchette, in press). In addition to beta coefficients, random intercepts can be modeled. In the research discussed here, participant response speed was treated as a continuous variable resulting from a random intercept analysis (Thompson & Desrochers, 2009d). The possibility of modeling custom intercepts and beta coefficients for each participant is very powerful because it opens up a new class of indicators that can be entered as predictors in regression analyses, of which multi-level models are an example. The sensitivity analysis discussed above depended on the possibility of modeling slopes differentiated for each participant.

Multi-level models themselves are an improvement on traditional methods of analyzing psycholinguistic data (e.g., Thompson, 2008). The advantages of this approach are multifold and hinge on its ability to simultaneously model variance and covariance within and across levels of analysis. For psycholinguistic experiments this is useful because it means that participant and item variability can be modeled and related hypotheses explored, within a single analytical framework. As demonstrated by Thompson & Desrochers (2009c, 2009d), the technique is flexible enough to be applied to both reaction time and accuracy data. An added advantage is that traditional problems that sometimes emerge with reaction time data, such as speed accuracy tradeoff, can be examined in great detail by exploiting the ability to model the probability of making a mistake for each participant-item combination (Thompson & Desrochers, 2009c, Experiment 2). The estimated probability of making an error on each item can be entered as a predictor of item-level variability in reaction time. A necessary condition for arguing that the observed effect on RT is not reducible to a speed-accuracy tradeoff is that the effect should persist even after controlling statistically for estimated item-level error probability. Indeed, the effect reported by Thompson & Desrochers (2009c) was robust to the inclusion of this error covariance. More generally, the inclusion of error probabilities as a covariate in multi-level models of RT is an interesting way to verify to what extent the RT and Error effects are redundant. Completely redundant effects of RT and Error rate would constitute evidence that both effects arise from a common source. Currently, such inferences are based on far weaker evidence such as parallel effects in the analysis of mean differences (i.e., if the RT and Error rate effects go in the same direction, they are assumed to arise from the same factor), which are intrinsically ambiguous because the effects could be driven by distinct sets of items. We propose that the method employed by Thompson & Desrochers (2009c) is more rigorous.

Moderators of Syllable Preference

Studies investigating the role of the syllable in visual word recognition have yielded inconsistent findings. Overall, the research program described here confirmed this inconsistency in that a direct effect of word-splitting manipulation was only observed in a single experiment (Thompson & Desrochers, 2009c, Experiment 1). Otherwise, the main effect of word splitting was null. These observations validate the decision made by Thompson & Desrochers (2009d) to adopt a more exploratory approach. An approach with sufficient flexibility to reveal the heretofore unidentified contingencies that preside over the manifestation of syllable preference effects would shed more light on the empirical inconsistency in the discipline than more narrowly focused experiments. The proposed strategy constitutes a relaxation of experimental constraint in the spirit of Balota et al (2004) which is to be subsequently followed up with systematic experimental investigation. Having broadly sketched the general pattern of results above, we now turn to a more detailed examination of the findings reported by Thompson & Desrochers (2009c, 2009d).

Lexical Frequency. Taft (1979, 1987, 2001, and 2002) never identified lexical frequency as a relevant control variable, preferring instead to sample stimulus materials liberally from a wide range of frequencies. This practice was called into question by Chen & Vaid (2007) who observed that studies which employed lower frequency stimulus materials tended to show the largest Max Coda syllable effect advantages. Follow-up experimental work confirmed that Max Coda syllable effects tend to be specific to low-frequency words (Chen & Vaid, 2007). This result was confirmed by Thompson & Desrochers (2009c, Experiment 3) when syllable preference was conditioned on (in)sensitivity to homophone interference. Thompson & Desrochers (2009c) extended this low-frequency specificity to include manifest Max Onset preference as well (Experiments 1, 2, & 3), this time conditional on heightened sensitivity to homophone interference. The practical consequence of this finding is that frequency should not be left uncontrolled when assembling stimulus lists. Drawing more or less randomly from a wide range of frequency could lead to unstable results. The theoretical significance of this finding is that the syllable, whether it is the Max Onset or Max Coda syllable, is unlikely to be a unit of orthographic access. If the syllable where indeed an efficient unit of orthographic access, then it should be as relevant for high-frequency words as for low-frequency words (i.e., as per the sometimes implicit assumption made by Taft, 1979, 1987, 2001). To the contrary, the observed frequency dependency suggests that syllable preference, like other frequency sensitive effects (see Coltheart et al., 2001), can be understood as an emergent property of the online processing of letter-strings and the accrual of activation.

Syllable Complexity. Taft (2002) acknowledged that the strength of syllable effects could depend on the structural properties of the syllables involved, such as the length its nucleus (e.g., short or long vowel) or the number of consonants populating the boundary of two syllable units (e.g., the presence of a consonant cluster between the two vowels in the word THUNDER). Thompson & Desrochers (2009c) observed Max Onset syllable effects with both single consonant words (e.g., tulip) and words with complex consonant clusters (e.g., thunder), but in distinct experimental contexts. The effect with simple words only emerged when the experimental conditions were made exceedingly difficult. Specifically, the simple word effect emerged and the complex word effect disappeared when the word-initial segment was made more difficult to discriminate against the background. In contrast, the syllable effect with complex words emerged specifically when the word-initial segment was easy to discriminate. Syllabic complexity is a known difficulty factor in the acquisition of literacy (Seymour, Aro, & Erskine, 2003). When the visual familiarity of a word is disrupted, as is the case with the word-splitting paradigm, words with the more complex structure may be the first to show a syllable advantage. It may take a more extreme manipulation to induce a syllable effect with words with a simple structure, at least when letter colour is the method of splitting words. The argument here is that the complexity of a word's structure may make it easier to influence normal processing. A complete account of the results requires the additional assumption that too much disruption of normal processing wipes out syllable preference effect. This explanation is speculative because it is unclear how much is too much disruption or what a shift away from 'normal' processing would mean exactly. Yet, other methods of word splitting are arguably more extreme (e.g., insertion of a foreign character), which may explain why larger effects seem to be observed with items that have a simpler structure under such conditions (for discussion, see Taft, 2002). It must be acknowledged that this explanation for the weaker effect observed with complex syllable structure items reported by Taft is at odds with the phonetics-based account he proposed.

Distal Predictors: Skill and Print Exposure. Taft (2001, 2002) found that reader skill moderated the effect of word splitting. High-skill participants tended to show preference of the Max Coda unit, while low-skill participants tended to prefer the Max Onset unit. Thompson & Desrochers (2009d) used Print Exposure rather that comprehension skill as their distal predictor (which in any case they argued is a strong associate of comprehension skill). They confirmed that participants who score high on a component dominated by print-exposure prefer Max Coda division (when they respond quickly). Paradoxically, as noted earlier, they also found that slow responders with low-print exposure also preferred Max Coda division. The findings indicate that the story developed by Taft (2001) to account for the relationship between skill and syllable preference may not be so straightforward. Under the right conditions, both high and low skill word recognizers may show preference for the Max Coda syllable (and possibly Max Onset preference). This possibility imposes a different theoretical view of syllable preference from that advanced by Taft. This view would link particular syllable types to specific bodies of knowledge regardless of skill. The degree to which a unit is relevant to word recognition would depend on the involvement of that body of knowledge in the decision making process. This involvement could be determined by a participant's proficiency as a word recognizer and contextual response strategies, or other attributes. The importance of participant characteristics would in this case be secondary, mere contingencies that would serve to improve our ability to focus on the treatment of a particular type of information. The disagreement in the results obtained by Thompson & Desrochers (2009d) and Taft (2001) might be due to the instructions used. Different instructions could induce different response strategies for accommodating the competing goals of response speed and accuracy (see Thompson, 2009d). This sort of explanation can be contrasted with Taft (2001)'s proposal that the difference between good and poor readers is qualitative, structural. Taft's account might be maintained if we accept that with various word-recognition skill levels there exist sub-types of participants who share qualitatively distinct syllabification strategies. A priori, the latter explanation seems less parsimonious.

Sensitivity to Phonology. The initial problem examined by Thompson & Desrochers (2009c) was an investigation of the relationship between reliance on phonology and syllable preference. As noted earlier, both sensitivity to regularity and homophony were examined, but only homophony was found to be a reliable associated of syllable preference. This idea was pursued by Thompson & Desrochers (2009d) with an additional index of phonological dominance in lexical decision: sensitivity to pseudohomophony. It is worth noting that in both cases, the relationships between these variables and syllable preference were independent of print exposure. To the extent that print exposure can be taken as a proxy for reading skill (for discussion, see Thompson & Desrochers, 2009a, 2009d), this finding would tend to infirm Taft's explanation for low-skill participants preferring the phonological syllable. Of course, given that the correlation between these two measures is far from perfect, this result cannot be considered a refutation of the idea.

Sensitivity to non-lexical phonology did not relate to syllable preference in the expected way. Thompson & Desrochers (2009d) found that the part of sensitivity to regularity that is associated with sensitivity to imageability is positively related to Max Coda syllable preference for slow participants (with complex items). Similarly, again with slow responders, sensitivity to pseudohomophony was positively associated with Max Coda syllable preference as a part of the third variance component they estimated (associated with print exposure). Contrary to expectations then, reliance on non-lexical phonology seems to be positively associated with Max Coda preference. This is not inconsistent with the observation that the Max Coda syllable could be useful for resolving print-to-sound correspondences that are inconsistent when linked using other units of analysis (Chateau & Jared, 2003), but is inconsistent with the idea that the Max Coda syllable is an efficient unit of orthographic access. It should be noted that the Chateau & Jared (2003) effect seemed to be limited to particular words, where the finding reported here implies that Max Coda preference might be indicative of a more general strategy. For these variance components at least, the Max Coda syllable appeared to be a unit favored by those who weigh non-lexical phonology heavily in their decision-making. As Thompson & Desrochers (2009d) suggested, this weighting may be associated with indirect access to semantic information. Indirect (phonologically) mediated access to semantics is slower than direct access to semantics via orthography (see Harm & Seidenberg, 2004). Assuming that this particular Max Coda effect arises from processes native to the indirect route would explain why it is only observed with slow responders, for whom information obtained via the indirect route might be especially salient at response time. If this is the case, then the source of this Max Coda preference may lie in the interface between non-lexical phonology and semantics.

The opposite pattern was observed with lexical feedback phonology across studies (Thompson & Desrochers, 2009c, and 2009d). Participants who were especially sensitive to homophone interference in lexical decision tended to show preference of the Max Onset syllable with low frequency words. This result suggests that Max Onset syllable preference arises from feedback from phonology to orthography. An alternative possibility is that the Max Onset syllable is a unit of orthographic access (see Taft, 2001) and those who employ it extract phonological information rapidly. The latter possibility seems less likely given that readers who are skilled at reading for comprehension appear to actually access phonological information faster (Chateau & Jared, 2003) and prefer the Max Coda syllable (Taft, 2001). Taken together, the evidence would seem to support an implementation of the Max Onset syllable that is seated within units that control lexical phonology. In the Coltheart et al. (2001) model lexical phonology means dedicated representations of word identity in the phonological domain. In contrast, parallel distributed models like that proposed by Harm & Seidenberg (2004) must invoke semantics to attribute the label 'lexical' an activation pattern. The common thread in both models would be that the effect seems to arise due to strength of feedback from phonology to orthography.

Theoretically, the findings discussed in the previous paragraph provide clues as to where the phonological locus of these effects may lie. The Max Onset syllable appears to be associated with the (lexical) phonology-to-orthography interface while the Max Coda syllable appears to be associated with the (arguably non-lexical) interface between semantics and phonology. This information provides useful clues for modelers in developing an initial account of syllable representation. If the explanations suggested here are correct, then inserting Max Onset and Max Coda syllable structure in the lexical and non-lexical phonology processing streams respectively should allow a computational model to simulate most of the effects reported here. All of this does not discount the possibility that manifest Max Coda preference may have additional sources, such as segmentation at the orthographic level (e.g., see Max Coda preference with Homophone insensitivity, Thompson & Desrochers, 2009c, Experiment 3).

Response Speed. Taft (2001) and Thompson & Desrochers (2009c) failed to identify overall reaction time as a relevant predictor of syllable preference in the lexical decision word-splitting paradigm. Thompson & Desrochers (2009c) further observed that statistically controlling for the direct effect of response time did not eliminate the effect of the moderators considered here. Estimation of this redundancy was achieved via the inclusion of a term representing the two way interaction of overall RT and the word-splitting manipulation in a model predicting lexical decision performance. Thompson & Desrochers (2009d) took the idea a step further and considered reaction time as a potential moderator of the influence of other predictors. Estimation of this effect took the form of three-way interactions involving overall participant RT, word-splitting, and a third predictor variable. Indeed, they discovered that reaction time interacted with other predictor variables in determining syllable preference (i.e., participant sensitivity variance components).

Three of the six variance components estimated by Thompson & Desrochers (2009d) interacted with overall reaction time in predicting syllable preference. The analysis of the second component indicated that fast responding/high-exposure/low nonword sensitivity participants tended to prefer Max Onset division. Analysis of the third component indicated that both high-exposure and low-exposure participants both show Max Coda preference, the former when responding quickly and the latter when responding slowly. Finally, the sixth component revealed that slow responders who were sensitive to imageability and regularity demonstrate an apparent Max Coda syllable preference. Clearly, these results are inconsistent with the idea that insensitive participants would prefer the Max Coda syllable. The broad theoretical significance of these findings remains to be determined, but one way of reconciling them with the logic that was developed for the predicted relationship between sensitivity and syllable preference is to assume that the proximal variable governing observed syllable preference is the type of information dominating processing when the response is made. According to this view, the two major determinants of observed syllable preference should be a) response initiation criterion (fast versus slow within participant) and b) the time-course of activation transmission. The latter source of variability would be expected to vary as a function of skill, whereas the former would be subject to strategic influence. Manipulating the response criterion may be sufficient to cause a change in syllable preference while the participant attribute remains constant. If this is the case, then primacy of reliance on a particular kind of information as opposed to a qualitative difference in the structure of the reading system across individuals (Taft, 1991, 2001) will have been demonstrated. As noted earlier, the idea that the functional units of reading should be determined by the constraints imposed by the three main domains of knowledge and their inter-relationships as opposed to unspecified developmental processes that would give rise to qualitative structural differences in the reading system is appealing for its parsimony.

Experimentally manipulating the response speed of participants is a relatively straightforward matter. A payoff matrix could be introduced that rewards accuracy or response speed or both (e.g., Voss, Rothermund, & Voss, 2004). Such manipulations should allow the role of response speed in the production of apparent syllable effects to be evaluated. Of particular interest is the idea that a key factor in determining syllable preference would be the participant's response criterion. Rewarding both speed and accuracy should give a solid estimate of syllable effects when participants are responding as quickly as they can. Rewarding accuracy should provide a good estimate of syllable preference when responses are more considered. The idea here is to reduce heterogeneity in response strategy by instituting some degree of control. If this control has the predicted effect, then we can have greater confidence in the response-timing account developed here. What is more, if the same participants show Max Onset or Max Coda preference depending on the experimental context, then this would provide support for the idea that the two units have a different locus in the reading system, and that either can be more or less relevant depending on which source of information is dominant when the response is initiated. An alternative possibility is that a given participant only ever prefers a particular type of syllable and that manipulation of response only tends to accentuate this fact. From a practical standpoint, the status of response speed as a moderator means that this variable should be controlled or included in the design of experiments examining syllable effects in the word-splitting paradigm.

Theoretical and Practical Implications

The results discussed here identify many associates of apparent syllable preference. Practically, these results are useful in they suggest contingencies for the observation of particular patterns of syllable preference that researchers may exploit for the purpose of better controlling the behavior of participants. Participant response speed, lexical frequency, the complexity of syllable structure, print exposure, and participant sensitivity to various difficulty factors are confirmed associates of syllable preference. The results demonstrate that it is not sufficient to consider these variables in isolation. Rather, it is in interaction that these variables account for a significant portion of the variance in syllable preference. The covariance structure of the participant sensitivity measures suggests that multiple independent dimensions contribute the apparent syllable preference. Each component provides a potential source of experimental control and an avenue for follow-up empirical investigation. The combination of random coefficient analysis and multi-level modeling that was used here provides a model of how this can be done.

Theoretically, the results indicate that modeling syllable preference requires a mechanism for representing individual differences in the time course of activation through the word recognition system. Existing activation-based models satisfy this criterion (e.g., Coltheart et al., 2001; Harm & Seidenberg, 2004). Unlike robust effects like lexical frequency, extant computational model may need to be supplemented with a response criterion mechanism if they are to simulate the effects reported here. If all readers represent both the Max Onset and Max Coda syllable, then both time-course and response criterion may be needed to isolate the moments in the decision making process when a responses will yield a manifest preference for one syllabification heuristic over another. The decision making process may be responsible for some of the individual differences in participant sensitivity, especially those that are unrelated to print exposure. An account of this type of individual differences in sensitivity may require that the decision-making mechanism have the ability to control attention, emphasizing specific types of information (for examples of models that do just that, see above discussion).

Limitations

The most obvious limitation of the present research is the correlational nature of the results. Stronger interpretation of the participant attribute effects would be possible if these were manipulated experimentally. If, for example, participant sensitivity was manipulated by inducing participants to strategically emphasize particular types of information (e.g., Kello & Plaut, 2003), then the case for a causal relationship between these characteristics and syllable preference might be strengthened. The same applies to the role of participant speed in determining syllable preference.

Another limitation concerns the selection of measures that were used to operationalize individual differences. Print exposure for example was selected instead of an estimate of comprehension skill because it was deemed to be more directly interpretable in terms of word recognition processing. That said, this choice could be criticized on the grounds that previous work (Chen & Vaid, 2007; Taft, 2001) examined the role of comprehension skill specifically. The inclusion of comprehension skill would have come at the cost of increasing participant testing time, but would have substantially enriched the empirical database upon which the studies discussed here are based, and allowed an explanation advanced by Taft (2001) for his results to be definitively ruled out or confirmed. As it stands, all that can be said is that there is some evidence that Max Onset preference is associated with a specific type of phonological reliance, which would tend to support his phonology-skill idea (but see Unsworth & Pexman, 2003). Another contentious decision was the choice to avoid tasks other than lexical decision. Different forced-choice tasks have been used to assess the strength of orthographic processing directly (e.g., homophone-pseudohomophone discrimination; Sears, Siakaluk, Chow, & Buchanan, 2008). The decision to keep the task consistent was in accord with the stated objective of focusing on predictors that extant models of word recognition are designed to simulate, but came at the cost of discarding a potentially more direct measure of orthographic processing.

Conclusion

The goal of the present document was to review the results of the research program outlined by Thompson & Desrochers (2009b) and pursued by Thompson & Desrochers (2009c, 2009d). The overarching goal of this research program was to indentify the conditions for observing particular patterns of syllable preference effects in lexical decision. Participant- and item-level characteristics were examined on their own and in concert. Consistent with previous work, lexical frequency was identified as a moderator of syllable preference (Chen & Vaid, 2007). The distinction between lexical and non-lexical phonology proved relevant in predicting syllable preference in that each seems to relate independently to syllable preference. The association of lexical phonology with Max Onset syllable preference provided partial support for the idea that such syllable preference is the result of preferential reliance on phonological information. The association of high-skill and fast responding with Max Coda preference tended to support the idea that this effect arises from efficient orthographic processing. Other results were inconsistent with these ideas and it was proposed that a combination of response criterion and individual differences in speed of responding to information might in some measure account for these inconsistencies. Interestingly, the putatively ‘orthographic’ syllable was associated with non-lexical phonology (and possibly use of the indirect semantics route) in some analyses, which would tend to indicate that another source Max Coda preference is situated at the interface of phonology and semantics. An obvious avenue for future research would be the systematic examination of response criterion and experimental manipulations thereof with the intention of estimating the range of intra-participant syllable preference.

References

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105*, 678-723.

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*, 390-412.

Balota, D.A., Cortese, M.J., Sergent-Marshall, S.D., Spieler, D.H., & Yap, M.J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283-315.

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Desrochers, A., Thompson, G. L., & Fréchette, S. (in press). The Impact of Information Technology on the Selection of Lexical Items for Research in Human Experimental Psychology. In F. Columbus (Ed.), *Advances in Psychological Research.* Hauppauge, NY: Nova Science Publishers, Inc.

Harm, M. & Seidenberg, M. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Kello, C.T., & Plaut, D.C. (2003). Strategic control of rate of processing in word reading: a computational investigation. *Journal of Memory and Language*, *48*, 207-232.

Locker, L.Jr., Hoffman, L., & Bovaird, J.A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, *39*, 723-730.

Lupker, S.J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, band Cognition*, 23, 570-590.

Norris, D. (2006). The Bayesian Reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*, 327-357.

Sears, C.R., Siakaluk, P.D., Chow,V.C., & Buchanan, L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research, 37,* 269-291.

Seymour, P.H.K., Aro, M., & Erskine, J.M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143-174.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 402-433.

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

Rouibahh, A., & Taft, M. (2001). The role of syllabic structure in French visual word recognition. *Memory & Cognition, 29*, 373-381.

Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language & Speech*, *44*, 351-376.

Taft, M. (2002). Orthographic processing of polysyllabic words by native and nonnative English speakers. *Brain & Language, 81*, 532-544.

Thompson, G.L. (2008). Eliminating aggregation bias in experimental research: Random coefficient analysis as an alternative to performing a 'by-subjects' and/or 'by-items' ANOVA. *Tutorials in Quantitative Methods for Psychology*, *4*, 21-34.

Thompson, G.L., & Desrochers, A. (2009a). Towards a Process-Based Characterization of Syllable Effects in Visual Word Recognition. Manuscript submitted for publication.

Thompson, G.L., & Desrochers, A. (2009b). Characterizing the Syllable Effect in Visual Word Recognition: A Novel Methodology & New Evidence

Thompson, G.L., & Desrochers, A. (2009c). Sensitivity to Homophone Interference as Correlate of Syllable Preference in Visual Word Recognition. Manuscript submitted for publication.

Thompson, G.L., & Desrochers, A. (2009d). Individual Differences and Syllable Preference: Distal and Proximal Correlates. Manuscript submitted for publication.

Unsworth, S.J., & Pexman, P. (2003). The impact of reader skill on phonological processing in visual word recognition. *Quarterly Journal of Experimental Psychology*. *56A*, 63-81.

Wagenmakers, E-J, Steyvers, M., Raaijmakers, J.G.W., Shiffrin, R.M., van Rijn, H., & Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology*, *48*, 332-367.

Authors Note

Glenn L. Thompson & Alain Desrochers, School of Psychology, University of Ottawa, Ottawa, Ontario, Canada.

The research reported in this article was supported by the National Sciences and Engineering Council of Canada (NSERC) in the form of Master- and Doctoral-level scholarships awarded to the first author, as well as NSERC grants awarded to the second author.

Correspondence should be addressed to Glenn Thompson at the School of Psychology, University of Ottawa, 145 Jean-Jacques Lussier, P.O. Box 450, Station A, Ottawa, Canada, K1N 6N5, or e-mail: GlennLThompson@gmail.com.

Notes

[1] Ans, Carbonnel, and Valdois (1998), for example, modeled syllable-length effects in visual word recognition effects. Such modeling has its place, but we prefer to focus on the cumulative nature of the model building enterprise (Jacobs & Grainger, 1994). A model of polysyllabic word reading should account for both syllable effects in addition to the stable effects generated by the rich literature on the recognition of single-syllable words. It is in this sense that an adequate explanatory model of multi-syllable word reading is lacking.