2014

How Partial Transparency Influences the Processing of Compound Words

Alexandria Stathis
University of Windsor

How Partial Transparency Influences the Processing of Compound Words

By

Alexandria Stathis

A Thesis

Submitted to the Faculty of Graduate Studies

through Clinical Psychology

in Partial Fulfillment of the Requirements for

the Degree of Master of Arts at the

University of Windsor

Windsor, Ontario, Canada

How Partial Transparency Influences the Processing of Compound Words

by

Alexandria Stathis

APPROVED BY:

_____

N. McNevin
Human Kinetics

_____

A. Baird
Psychology

_____

L. Buchanan, Advisor
Psychology

January 17, 2014

Author's Declaration of Originality

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

Abstract

Compound words are words with multiple constituents that individually have their own meaning and that combine to make another meaning (e.g. *doghouse*). When these constituents help us infer the meaning of the whole compound word, they are known as transparent constituents (e.g. either constituent of *blueberry*). In contrast, opaque constituents do not help us infer the meaning of the whole compound word (e.g. *moonshine*). The current study sought to describe the processing of partially transparent words, in which one constituent relates to the total meaning, whereas the other does not (like *strawberry*, which is a berry, but not made of straw). Thirty-seven participants were asked to complete a lexical decision task in which they indicated whether an item was a word or a nonword. The fully transparent compound words enjoyed a processing advantage when compared to the monomorphemic words. This processing advantage disappeared when any constituent (or both) became opaque.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

How Partial Transparency Influences the Processing of Compound Words

The primary goal of psycholinguistics research is to learn how people process and understand language. One important question is how meaning is derived from single words. In psycholinguistic terms, a word's meaning reflects its morphemic content. A morpheme is the smallest meaningful unit of a word, such as the word *giraffe* (Ji, Gagne, & Spalding, 2011). This word cannot be divided into any smaller, meaningful units. A monomorphemic word is just that: a word comprised of a single morpheme (Ji, Gagne, & Spalding, 2011). In contrast, a multimorphemic word is comprised of multiple morphemes, for example *giraffes* has two meaningful units (*giraffe* and '*s*' that designates more than one giraffe). A special case of multimorphemic words is the compound word (Ji, Gagne, & Spalding, 2011). A compound word traditionally contains two components, known as constituents, for example *blueberry* with its constituents *blue* and *berry*.

Compound words are found in a multitude of languages. Each language contains its own rules about the development and usage of these compound words. In Finnish, Dutch, and German novel compound words are common and are developed spontaneously (Hittmair-Delazer, Andree, Semenza, Blesser, & Benke, 1994; Juhasz, Starr, Inhoff, & Placke, 2003). English is not such a spontaneous language and there are no systematic rules concerning how free morphemes ultimately become paired to form a compound word (Juhasz et al, 2003).

In English compound words, the second constituent is generally the "head" of the word (Libben & Jarema, 2006; McGregor, Rost, Guo, & Sheng, 2008). Essentially, the head of the compound indicates something about the semantic (i.e., meaning) category of the word, and the other constituent modifies it (McGregor et al, 2008). For example, *houseboat* describes a type of boat, whereas *boathouse* describes a type of house.

Research by Juhasz et al (2003) supports the idea that the second constituent is used to determine the semantic categorization of a compound word. In their study, the authors examined the processing of compound words by studying eye fixation, naming, and lexical decisions. They reported increased eye fixations, as well as faster lexical decisions and naming for the head, or second constituent (Juhasz et al, 2003). Evidently, the meaning conveyed in the second constituent holds comparably more weight than the first constituent.

Compound words offer psycholinguists a unique window into morphemic influences in single word reading, and are consequently important for theoretical reasons. At a more applied level, an understanding of normal compound word reading may assist educators charged with helping students with specific language impairments (SLI). Previous research indicates that children with SLIs have difficulty ordering noun-noun compound words (like *doghouse*). This had been attributed to a lexical deficit (Grela, Snyder, & Hiramatsu, 2005), but McGregor et al's (2008) expansion of this research suggests that the deficit may be one of compounding rather than accessing.

This consideration of processing deficits as either compounding or accessing maps onto one of the central theoretical debates in compound word research. There are two main theories of compound word comprehension: whole word processing and decomposition (Libben & Jarema, 2006). In the whole word processing model, a compound word is represented and accessed as a whole unit, just as monomorphemic words are (Butterworth, 1983). In the decomposition model, however, the compound words are accessed and assembled via their two constituents (Andrews, 1986; Andrews, Miller, & Rayner, 2004; Libben, 1998; Pollastek, Hyona, & Bertram, 2000; Taft & Forster, 1975; Zwisterlood, 1994).

Each theory entails a different combination of costs and benefits for the person processing the compound word. The decomposition model has a processing burden, whereas the whole word model has a storage burden. For the former, only the constituents are stored in the lexicon, but the morphological parsing is associated with a time cost (Bertram, Laine, & Karvinen, 1999). For the latter, the whole word representation must be stored in addition to its constituents, but that storage eliminates the need for additional parsing (Butterworth, 1983; Libben & Jarema, 2006). An argument in favour of the whole word model is that decomposition is thought to be more error-prone than whole word access (Lehtonen et al, 2007). Despite this, there have been a number of studies that have found that morphological decomposition occurs regularly in compound words (Andrews, 1986; Andrews, Miller, & Rayner, 2004; Libben, 1998; Pollastek, Hyona, & Bertram, 2000; Taft & Forster, 1975; Zwisterlood, 1994).

It may not be that decomposition and whole word access are opposing processes. A number of researchers have proposed a horse race model (see Laine, Vainio, & Hyona, 1999), whereby decomposition occurs at the same time as whole word processing, with the faster method being determined by theoretically important variables, such as word presentation, word frequency, and transparency as described below (Ji, Gagne, and Spalding, 2011; Libben et al., 2003; Baayen, Dijkstra, & Shreuder, 1997; Bertram et al, 1999; Frauenfelder & Shreuder, 1991).

The transparency of a compound describes the semantic content of a constituent as it relates to meaning of the whole word (Libben & Jarema, 2006). If a constituent is transparent, then it is related to the whole meaning of the compound word. In the word *blueberry*, both constituents are transparent because a blueberry is, indeed, a berry that is blue. A constituent can also be opaque, which means that it does not explicitly relate to the whole compound word (Libben & Jarema, 2006). This is clearly displayed with the word *moonshine*. Moonshine is a

form of alcohol, but it has nothing to do with either the moon or its shine. The matter is further complicated by the fact that partial transparency exists. That is, one of the constituents is transparent, while the other remains opaque (Libben & Jarema, 2006). *Strawberry* is a common example, in which the head is transparent (because a strawberry is a berry), while the modifier is opaque (because a strawberry is not a berry made of straw). *Jailbird* is another example where the modifier is transparent and the head is opaque.

Laudana and Burani (1995) have suggested that semantic transparency determines whether a compound word will be processed via whole word recognition or through decomposition. Similarly, Schreuder and Baayen's (1995) meta-model of morphological processing postulates that semantic transparency determines whether a compound word has its own representation in the mental lexicon, or whether it is represented by its constituents. In this model, totally opaque compound words must be stored as a whole word, because decomposing the word into its constituents gives no information about the meaning of the word (Libben, Gibson, Yoon, & Sandra, 2003).

Libben et al (2003) examined how transparency influences the processing of compound words in a lexical decision task. Libben et al (2003) presented the participants with four types of compound words based on constituent transparency: transparent-transparent (TT), opaque-transparent (OT), transparent-opaque (TO), and opaque-opaque (OO). There were processing differences across the word types, with responses to words with transparent heads (TT and OT) being faster than to words with opaque heads (TO and OO). Importantly, the advantage disappeared for the OT words in a second condition that presented the compound words with spaces between the constituents. The spaces introduced between the constituents made decomposition more likely, so only those words that could maximally benefit from the

decomposition showed the advantage. Therefore, when decomposition was encouraged via the addition of spaces, TO, OT, and OO words showed similar processing patterns.

Physiological support for the influences of transparency comes from research by Koester et al. (2007) who studied compound words by tracking event-related brain potentials (ERPs). They found that the ERP signals (N400) differed as a function of constituent transparency. Fully transparent compound words elicited an increased negativity relative to opaque and partially transparent compound words. This seems to indicate that the transparency of the constituents is important in the accessing and understanding of compound words.

There is both behavioural and physiological support for a horse race model of compound word processing that is influenced by transparency of the constituents. However, to the extent that compound words have two pathways to recognition, they should show a processing advantage over monomorphemic words. Ji, Gagne, and Spalding (2011) tested just this in their lexical decision study contrasting transparent and opaque compound words with monomorphemic words. It is important to note that the Ji, Gagne, and Spalding (2011) study included both partially transparent and fully opaque words within their opaque group. Their results indicated faster reaction times for both groups of compound words when compared to monomorphemic words. Like Libben et al (2003), they also encouraged decomposition processing in a second condition in which spaces were added. In that second condition, the processing advantage for opaque words disappeared. They suggested that this might have something to do with the conflict between the lexical representation of the whole word and the meaning of the constituents. Of course, in the case of the fully transparent words this conflict does not exist, so they still retained the processing advantage. These findings (Ji, Gagne and Spalding, 2011; Libben, 2003) suggest that if the horse race model is true, the fastest route for

compounds that are not fully transparent is likely to be, in most cases, the whole word route. For fully transparent words, both routes are equally useful.

By contrasting responses for compound words to responses for monomorphemic words, this research (Ji, Gagne, & Spalding, 2011) complements Libben's (2003) findings and adds to our knowledge regarding how compound words are processed in English. However, a consideration of the two studies reveals that a gap in the literature exists. The processing of partially transparent compound words has never been compared to the processing of monomorphemic words. While Libben's (2003) research did not reveal differences between any of the groups with opaque constituents when the decomposition paradigm was used, they were not compared to monomorphemic words. We are still faced with the query of whether words with partial transparency are able to derive a processing advantage relative to monomorphemic words. The question remains: compared to monomorphemic words, do compound words with transparency of head or modifier alone give rise to advantages derived from having two pathways to recognition? The purpose of this study is to answer this question.

To summarize, we know that there is a processing advantage associated with compound words compared to monomorphemic words (Ji, Gagne, & Spalding, 2011). We also know that this processing advantage disappears for opaque words when decomposition is forced by incorporating spaces between words (Ji, Gagne, & Spalding, 2011). What we do not know is whether partial transparency is sufficient to maintain this processing advantage over monomorphemes in the forced decomposition paradigm.

This study has one main hypothesis: fully or partially transparent compound words that are presented with a space between the constituents will be recognized with greater rapidity and

accuracy than will monomorphemic words. In contrast, fully opaque words presented in the same manner will not be.

Methods

Participants

Thirty-seven University of Windsor undergraduates were drawn from the Participant Pool. The participants in this sample were predominantly female (83%) and predominantly in their 20s (range = 18 to 43). Participants were granted bonus points towards the Participant Pool in exchange for their participation. These points could be allotted towards eligible psychology courses. As this study took approximately 30 minutes to complete, the participants were all compensated with 0.5 bonus points, as is consistent with Participant Pool guidelines.

There was a single exclusionary criterion used for this study: participants were required to be fluent in the English language. The purpose of this study is to look at how words are perceived and understood. If a participant is not familiar with English words, then it will be more difficult for them to decide whether an item a word or a non-word.

Materials

This study used a modified version of the word lists generated by Ji, Gagne, and Spalding (2011). Expanding on this study the current compound word lists were divided into the four transparency groups: transparent-transparent (TT), opaque-transparent (OT), transparent-opaque (TO), and opaque-opaque (OO). Using similar word selection criteria (e.g. by matching each item's letters and syllables) as Ji, Gagne, and Spalding (2011), additional items were added to the partially transparent lists so that there were an equal number of words in each category. The monomorphemic words of Ji, Gagne and Spalding (2011) were used along with additional monomorphemic words that were selected in such a way that, when broken up into components

(consistent with the broken presentation pattern), there was a mixture of component types. For instance, there were an equal proportion of monomorphemic words that had two real component (i.e. p*lank ton*), words that had a real first component (i.e. c*on clude*), words that had a real second component (i.e. p*ron ounce*), and words where neither of the components were real (i.e *chole sterol*). These additional words were added using WordMine (Durda & Buchanan, 2006) and were selected according to the number of letters (range: 6 - 11) and syllables (range: 2 - 4), as well as orthographic neighbourhood frequencies (ONF; whole ONF range: 0 – 2, but usually 0).

An orthographic neighbour is defined as a word of the same length as the target word that differs by only one letter (Perea & Rosa, 2000). For instance, *cat* is an orthographic neighbour to *bat.* The orthographic neighbourhood frequency describes the average frequency at which the orthographic neighbours are encountered in presentations of text (Perea & Rosa, 2000). This is important because research suggests that the number of orthographical neighbours a word has may affect lexical access, which in turn may influence reaction times (Perea & Rosa, 2000).

Finally, an equal amount of nonwords were added to the word list. There were two types of nonwords: those that had characteristics of monomorphemic words (i.e., *flouch*), and those that had characteristic of compound words (i.e. *topdrug*). Compound nonwords were taken from Taft and Forster (1976). These were constructed in such a way that nonword constituent pairs used letter combinations that would very rarely be seen together in monomorphemic words (such as *k* and *b*, which are rare consonant pairs). There were four types of compound nonwords. The **first** is a word-word pair, in which both constituents are real words, but the total compound word is not found in the English language (i.e. *brief tax*). The **second** is a word-nonword pair, in which the first constituent is a real word, but the second constituent is a pronounceable nonword. The

third is a nonword-word pair and the **fourth** is a nonword-nonword pair. Monomorphemic nonwords were taken from the Ji, Gagne, and Spalding (2011) study, and additional ones were also creating using Wuggy, the multilingual pseudoword generator (Keuleers and Brysbaert, 2010).

In summation, there were 40 compound words (10 each TT, OO, TO and OT), 80 monomorphemic words, 40 compound-like nonwords, and 80 monomorpheme-like nonwords for a total of 240 words. There were twice as many monomorphemic words as compound words and twice as many monomorpheme-like nonwords as compound-like nonword because this strategy made controlling for compound words less difficult. With more monomorphemic words, participants would probably remain unaware that this was a study about compound words and would presumably be less inclined to quickly identify constituent nonwords and then reject the word based on the first constituent.

The items used in this study can be seen in the Appendices, as well as their respective number of letters, syllables, and orthographic neighbourhood frequencies.

Procedures

The study began by having each participant fill out demographic information (age and sex) on the computer. The participants then began the lexical decision task in which they were asked to indicate as quickly and accurately as possible whether a presented letter string was a real English word or a nonword (ignoring the spaces in the items). The lexical decision task was run using DirectRT (Jarvis, 2012), which displayed the instructions, practice trials, and the experimental trials on a computer screen. Each word was presented in the centre of the screen, with size 24 Times New Roman font, in bright aqua on a black background. The keyboard responses that each participant made were also recorded using DirectRT. The responses were

recorded with a timing resolution of 1 millisecond, which allows for an accurate capture of lexical decision reaction times. Accuracy was also recorded.

The items were presented one at a time on the computer display in randomized order. Each item followed a 300 ms presentation of a fixation cross at the center of the display and the item remained on the display until a response was made. Participants made the lexical decision by pressing the designated key on the button bar with their index finger (they were instructed to press 1 for real words and 3 for nonwords). They were also instructed that after each decision, they were to return their finger to the 2 key (located between buttons 1 and 3) in preparation for the next trial. This was to reduce the potential influence of predictive hovering. Before the test trials began, the participants were asked to complete a series of practice trials to ensure that they understood the instructions. They received feedback as to whether their lexical decision was correct immediately after each practice trial. Upon completion of the practice trials, the participants were then able to continue onto the test trials.

Given the length of this lexical decision task, participants were allotted breaks every 50 words to prevent fatigue. When the participants were ready to proceed, they were able to press a key to continue with the trials.

Results

To examine how partially transparent compound words behave relative to monomorphemic words, reaction times and accuracy rates were evaluated using separate Repeated Measures ANOVAs. Both reaction time and accuracy are thought to be estimates of a word's ease of processing. Indeed, words that are easy to process should be identified more rapidly and more accurately than words that are more difficult to process.

To help analyze reaction time and accuracy rate together, the inverse efficiency score (IES; Townsend & Ashby, 1978) was calculated for each word type per participant. Inverse efficiency scores are calculated by dividing each participant's average reaction time by their average accuracy rate, per word type. As such, the IES is basically an estimate of reaction time that has been adjusted according to accuracy. The lower the accuracy the more units it adds on to the adjusted time, whereas a perfect score means that the reaction time remains the same. It also has the advantage of being a measure that accounts for both reaction time and accuracy in a single quantity. A study by Bruyer and Brysbaert (2011) describes that the unit of measurement attributed to IES is milliseconds.

The IES can be thought of as a measurement that gauges the average effort used by the system over time (Bruyer & Brysbaert, 2011; Townsend & Ashby, 1978). A higher IES value can represent the idea that more effort was expended in the processing of a word, and consequently is processed less efficiently than items with lower IES values.

Correct reaction times were analyzed using a Repeated Measures ANOVA, followed with a planned contrast (in which TT, OT, TO, and OO words were compared to the monomorphemic words). Before conducting any analytical procedures, the data was examined for normality and outliers. Three participants were excluded from data analysis due to low accuracy (less than 70%). Any reaction times that were greater than 2.5 standard deviations away from the mean per condition were removed. One reaction time was removed for being less than 300ms. Altogether, this led to the removal of 2.5% of the observations.

To assess for departures of normality, skewness, kurtosis, and the Shapiro-Wilk's statistic were all considered. Shapiro-Wilk's was significant, indicating that normality cannot be assumed. This remained to be the case following the removal of the outliers. When considering

skewness and kurtosis, each word type showed a slight positively skewed distribution, while

kurtosis was roughly within normal limits (or slightly platykurtic). Homogeneity of variance was

evaluated by examining Levene's test, which was not significant, indicating that the assumption

of homogeneity of variance was met.

Sphericity is the final assumption that must be considered when conducting a repeated

measures ANOVA. To determine whether any departures from sphericity were present,

Mauchley's test ($W = .235$, $p < .05$) was used, and results indicated that the assumption of

sphericity was violated. Given that estimates of sphericity fell below 0.75, the Greenhouse-

Geisser correction was used when assessing within-subject effects. Cohen's d was calculated

using Morris and DeShon's (2002) equation, which corrects for dependence in within-subject

studies.

The results reveal a main effect of word type in the reaction time data, $F1(2.57, 84.65) =$

4.650, $p < .01$. In particular, TT words were responded to faster than monomorphemic words,

$F1(1, 33) = 18.954$, $p < .01$, $d = .712$. In contrast, OT ($F1(1,33) = .959$, $p = .335$, $d = .161$), TO

($F1(1,33) = .225$, $p = .638$, $d = .078$), and OO words ($F1(1,33) = 1.743$, $p = .196$, $d = .218$) were

not responded to at a different rate than monomorphemic words. Post-hoc analysis using

Bonferroni correction found no differences between the reaction times of TO and OT words, $p >$

0.05. A visual representation of the reaction time for each word type can be seen in Figure 1

below.

Figure 1

Graph of Mean Reaction Times by Word Type
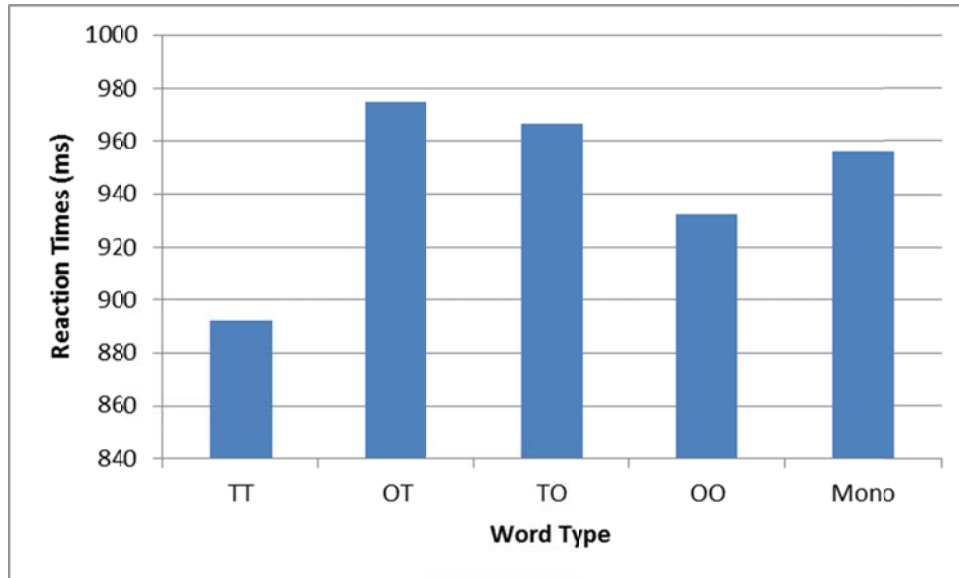
Accuracy rates were analyzed using a Repeated Measures ANOVA, followed with a planned contrast (in which TT, OT, TO, and OO words were all compared to monomorphemic words). Shapiro-Wilk's was significant, indicating that normality cannot be assumed. Homogeneity of variance was evaluated by examining Levene's test, which was not significant, indicating that the assumption of homogeneity of variance was met. To determine whether any departures from sphericity were present, Mauchley's test ($W = .460$, $p < .05$) was used, and results indicated that the assumption of sphericity was violated. Given that estimates of sphericity fell below 0.75, the Greenhouse-Geisser correction was used when assessing within-subject effects.

Results indicate that word type had a main effect on accuracy rate ($F1(2.83, 98.38) = 28.881$, $p < .01$). Monomorphemic words were processed less accurately than TT ($F1(1, 33) = 7.507$, $p < .05$, $d = .486$), but more accurately than OO ($F1(1, 33) = 34.813$, $p < .01$, $d = .880$) and OT words ($F1(1, 33) = 21.707$, $p < .01$, $d = .879$). No significant differences were found between monomorphemic words and TO words ($F1(1, 33) = .712$, $p = .402$, $d = .189$). Post hoc analysis using the Bonferroni correction revealed that TO words were processed more accurately

than OT words, $p < 0.05$. A visual of representation of the accuracy rates by word type can be seen in Figure 2.

Figure 2

Graph of Accuracy Rate by Word Type
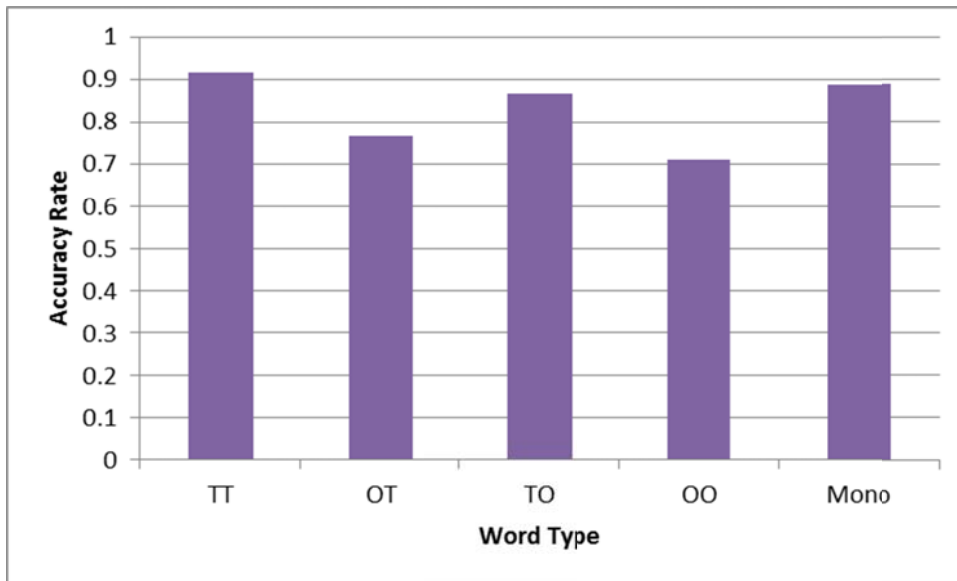
To analyze the inverse efficiency scores a Repeated Measures ANOVA was conducted, followed by a planned contrast (in which TT, OT, TO, and OO words were all compared to monomorphemic words). Assessments of sphericity (W = .125, p<.05) indicated that the assumption of sphericity was not met. Since estimates of sphericity fell below 0.75, the Greenhouse-Geisser correction was used when assessing within-subject effects.

Results indicate that word type had a main effect on the inverse efficiency score, $F1(2.486, 82.051) = 14.346$, $p < .001$. Specifically, TT words resulted in lower inverse efficiency scores than monomorphemic words, $F(1,33) = 20.130$, $p < .001$, $d = 0.774$. In contrast, OT ($F1(1,33) = 10.118$, $p < .01$, $d = 0.719$) and OO words ($F1(1,33) = 17.545$, $p < .001$, $d = 0.906$) revealed higher inverse efficiency scores than monomorphemic words. Finally, the inverse efficiency scores of the TO words were not significantly different from monomorphemic words,

$F(1,33) = 2.335$, $p = .136$, $d = 0.294$. Post hoc analysis using the Bonferroni correction revealed that OT words were processed less efficiently than TO words, $p < 0.05$. A visual representation of each group's inverse efficiency scores can be seen in figure 3 below.

Figure 3

Graph of Inverse Efficiency Scores by Word Type
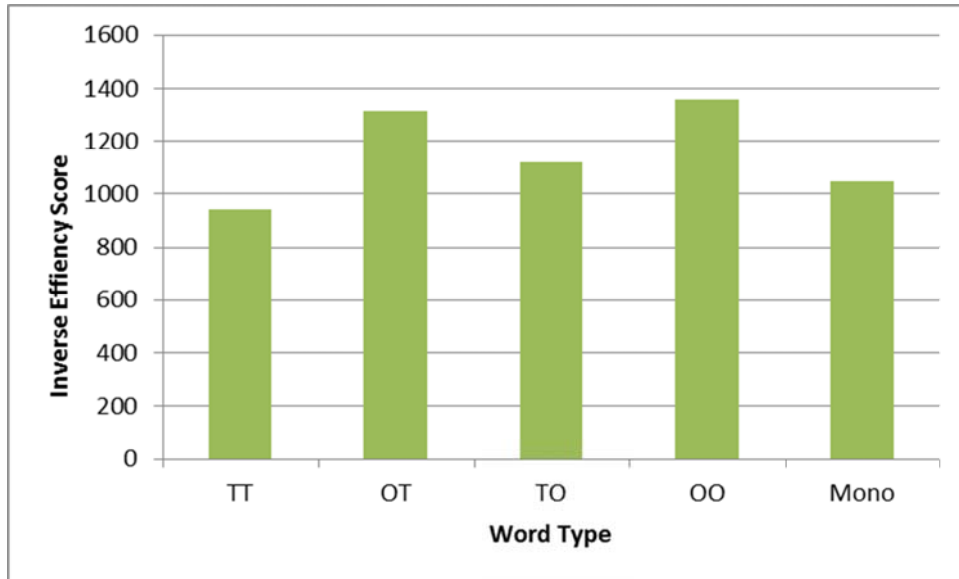
Means and standard deviations for reaction times, accuracy rates, and IES units for each word group can be seen in table 1 below.

Table 1

Means and Standard Deviations of Calculations by Word Type

		Word Type				
		TT	OT	TO	OO	Mono
RT (ms)	**Mean**	**892.113**	**974.960**	**966.820**	**932.379**	**956.040**
	SD	186.892	222.928	223.451	235.277	208.827
Accuracy[1]	**Mean**	**.935**	**.786**	**.879**	**.723**	**.896**
	SD	.101	.175	.151	.194	.074
IES	**Mean**	**970.710**	**1351.964**	**1151.618**	**1398.892**	**1078.885**

	SD	258.856	632.753	408.509	572.0136	273.419

1. Accuracy is represented as percentage correct.

Discussion

The aim of the current study was to evaluate how partially transparent compound words are processed relative to monomorphemic words. Previous research (Ji, Gagne, & Spalding, 2011; Libben et al, 2003; Libben, 1998) has shown that compound words, regardless of transparency, are processed with more ease than monomorphemic words in lexical decision tasks. When decomposition processing is encouraged via the addition of spaces between the constituents, Ji, Gagne, and Spalding (2011) found that opaque compound words lost their processing advantage relative to monomorphemic words. The current study sought to examine what would happen in the same paradigm under conditions where the number and position of transparent constituents in the compound word was varied.

Consistent with previous literature (Ji, Gagne, & Spalding, 2011; Libben, 1998), semantic transparency seems to facilitate compound processing relative to monomorphemic processing, but only when both constituents are transparent. In the current study, fully transparent compound words were processed more rapidly, more accurately, and consequently yielded a lower IES than monomorphemic words (indicating comparably more processing efficiency).

Also as seen in Ji, Gagne, and Spalding (2011), fully opaque compound words did not show a processing advantage relative to monomorphemic words. Opaque compound word reaction times were not significantly different from monomorphemic words, while opaque compound words were processed less accurately than monomorphemic words. Opaque compound words showed a higher IES than monomorphemic words, thereby indicating comparably less processing efficiency.

These results make sense in light of what is already known about how compound words are processed. In the current study, decomposition was encouraged by adding spaces between the constituents/components of both compound and monomorphemic words. In such a situation it appears that participants read each of the constituents separately, combine them in their mental lexicon, and then decide whether the item is a real word. When both constituents are transparent, both are able to provide consistent semantic information about the meaning of the word (Libben & Jarema, 2006). As such, the word should be identified more rapidly because the reader can take advantage of the decomposition route. With fully opaque compound words, however, the reader cannot readily take advantage of the decomposition route because the information in the constituents indicate nothing about the total meaning of the word (Libben & Jarema, 2006). Therefore, the reader needs additional time in order to process fully opaque compound words using the whole word route. Consequently there is no advantage afforded to the processing of fully opaque compound words relative to monomorphemic words

Having one transparent constituent was not sufficient to produce a processing advantage, and, in fact, having a transparent head following an opaque modifier seemed to hinder processing accuracy. This is an interesting finding, given that the head of the compound word describes the primary semantic categorization about the whole compound (Libben & Jarema, 2006). Given that the information contained in a transparent head is relevant to the total meaning of the word, it should follow that participants would be able to use this information to help identify the word as being real. OT and TO showed similar reaction times to monomorphemic words. Additionally, OT words were processed less accurately than monomorphemic words, whereas TO words maintained roughly the same accuracy as monomorphemic words. OT words showed a higher IES than monomorphemic words, whereas TO words showed a similar IES to monomorphemic

words. To summarize, there was no situation using the decomposition paradigm in which partial transparency provided any advantage to the processing of compound words as compared to monomorphemic words.

The reason for this is not totally clear; however, it may be reasonable to speculate that the processing affiliated with partial transparency follows the same processing pattern that it shows with monomorphemic and opaque compound words. Being provided semantically relevant information from only one transparent constituent is not enough to take advantage of the decomposition route, and therefore the whole word route may still be necessary. In any case, a single opaque constituent dissolves any potential processing advantage.

Another notable finding was that OT and TO words were processed differently; OT words were processed less accurately and with less efficiency than TO words. This is surprising given that, as previously mentioned, there should be a processing advantage afforded to words with transparent heads. The current study seems to indicate that the addition of an opaque modifier results in some sort of deficit when a participant tries to make a lexical decision with the decomposition paradigm in place.

This can be considered using the horse race model of compound word processing (Laine, Vainio, & Hyona, 1999). This model proposes that both whole word processing and decomposition processing occur simultaneously, and the faster method is determined by a number of variables (i.e., word frequency and transparency; Ji, Gagne, and Spalding, 2011; Libben et al., 2003; Baayen, Dijkstra, & Shreuder, 1997; Bertram et al, 1999; Frauenfelder & Shreuder, 1991). If decomposition is the processing method that is encouraged by the transparent head, and it cannot proceed because the modifier is opaque, then the participant may be inclined to reject the word based on that first judgement, particularly if whole word processing had not

yet identified the word as real. This behaviour may have been encouraged by directions to participants to respond as quickly as possible. Words with opaque heads would not be encouraged to undergo decomposition processing, because there is no semantic information to derive from them. As such, the participants may be displaying a processing error when OT words are presented that would not necessarily be seen when presenting TT, OO, or TO words.

The current study hypothesized that a processing advantage would be apparent in words that are either fully or partially transparent relative to monomorphemic words. The results were partially in line with the hypothesis. No processing advantage was seen for fully opaque words, and a distinct processing advantage was apparent for totally transparent words, as expected. However, the hypothesis was not consistent with the results in terms of partially transparent words. It was hypothesized that having at least one transparent constituent would be adequate to produce a processing advantage over monomorphemic words, but that was not the case. Having one transparent constituent was not advantageous when the decomposition processing route was encouraged.

Future research would benefit from examining partial transparency effects using eye tracking. It would be interesting to determine whether transparency influences how the eye scans a word. Perhaps under conditions of opacity, one might initially scan each constituent separately, and then re-scan the word holistically when meaning cannot be derived from the constituents alone. It would also be beneficial for future compound word transparency research to compile a database of transparency values, so that research across studies can use consistent values for matching words and evaluating results.

In conclusion, the goal of the current study was to evaluate how partial transparency influences the processing of compound words. The results indicated that fully transparent

compound words enjoyed a processing advantage when compared to the monomorphemic words. This processing advantage was no longer apparent when even one constituent was opaque. This study also revealed processing differences between the two types of partial transparencies (OT vs TO). While the study's main hypothesis was not entirely supported, the results help to advance our understanding of transparency as a processing facilitator.

Running Head: PARTIAL TRANSPARENCY IN COMPOUND WORDS

References

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language, 25*, 726–740.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*, 285–311.

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

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of Psycholinguistic Research, 28*, 213–226.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica, 51*(1), 5 – 13.

Burani, C., & Thornton, A. M. (2003). The interplay of root, suffix and whole-word frequency in processing derived words. In H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 157–208). Berlin: Mouton de Gruyter.

Butterworth, B. (1983). Lexical representation. *Language production*, *2*, 257-294.

Durda, K., Buchanan, L. (2006). WordMine2 [Online] Available: http://web2.uwindsor.ca/wordmine

Frauenfelder, U. H., & Schreuder, R. (1991). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. Booji & J. van Marle (Eds.), *Yearbook of morphology* (pp. 165–183). Kluwer: Dordrecht.

Jarvis, B. G. (2012). DirectRT (Version 2012.4.164) [Computer Software]. New York, NY:

Empirisoft Corporation.

Ji, H., Gagne, C. L., & Spalding, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language, 65*, 406 – 430.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming lexical decision and eye fixations. *British Journal of Psychology, 94*, 223-244.

Hittmair-Delazer, M., Andree, B., Semenza, C., De Blesser, R. & Benke, T. (1994) Naming by german compounds. *Journal of Neurolinguistics*, *8*(1), 27-41.

Hyönä, J., Vainio, S., & Laine, M. (2002). A morphological effect obtains for isolated words but not for words in sentence context. *European Journal of Cognitive Psychology*, *14*, 417–433.

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

Koester, D., Gunter, T. C., & Wagner, S. (2007).The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPS. *Brain and Language*, 102(1), 64-79.

Laine, M., Vainio, S., & Hyönä, J. (1999). Lexical access routes to nouns in a morphologically rich language. Journal of Memory and Language, 40, 109–135.

Laudanna, A., & Burani, C. (1995). Distributed properties of derivational affixes: Implications for processing. In Feldman, L. B. (ed), *Morphological Aspects of Language Processing*. Lawrence Erlbaum, Hillsdale, New Jersey.

Lehtonen, M., Cunillera, T., Rodriguez-Fornells, A., Hulten, A., Tuomainen, J., & Laine, M. (2007). Recognition of morphologically complex words in Finnish: Evidence from event-

related potentials. *Brain Research*, *1148*, 123–137

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for

representation, processing, and impairment. *Brain and Language, 61*, 30–44.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of

semantic transparency and morphological headedness. *Brain and Language, 84*, 50 - 64

Libben, G., & Jerema, G. (2006). *The Representation and Processing of Compound Words*.

Oxford University Press Inc: New York.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with

repeated measures and independent-groups designs. Psychological Methods, 7, 105-125.

Perea, M. & Rosa, E. (2000) The effects of orthographic neighbourhood in reading and

laboratory word identification tasks: A review. *Psicologica, 21*, 327-340.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading

Finnish compound words. *Journal of Experimental Psychology: Human Perception and*

Performance, 26, 820–833.

Raaijmakers, J. G. W., Schrijnemakers, J.M.C., & Gremmen, F.(1999). How to deal with the

"language-as-fixed-effect fallacy": Common misconceptions and alternative solutions.

Journal of Memory & Language, 41, 416-426.

Schreuder, R. and Baayen, R. (1995) Modeling morphological processing. In Feldman, L. B.

(ed), *Morphological Aspects of Language Processing*. Lawrence Erlbaum, Hillsdale,

New Jersey, 131–154.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. Journal of

Verbal Learning and Verbal Behavior, 14, 638–647.

Townsend, J.T., & Ashby, F.G. (1978). Methods of modeling capacity in simple

processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory*. Vol. 3.

Running Head: PARTIAL TRANSPARENCY IN COMPOUND WORDS

(pp. 200-239). Hillsdale, N.J.: Erlbaum

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of

Dutch compounds. *Language and Cognitive Processes, 9*, 341–368.

Running Head: PARTIAL TRANSPARENCY IN COMPOUND WORDS

Appendix A

Table A1

Words Used with their Lengths, Syllables, and Orthographic Neighbourhood Frequencies

				Orthographic Neighbourhood Frequencies		
Condition	Word	Letters	Syllables	Whole Word	First Constituent	Second Constituent
Compound TT	SUN RISE	7	2	0	17	12
Compound TT	TEA CUP	6	2	0	9	10
Compound TT	HOME TOWN	8	2	0	11	6
Compound TT	ANT EATER	8	3	0	7	8
Compound TT	HOVER CRAFT	10	3	0	7	3
Compound TT	KEY HOLE	7	2	0	6	15
Compound TT	HAIR PIN	7	2	0	5	18
Compound TT	NOTE PAPER	9	3	0	9	7
Compound TT	BARN YARD	8	2	0	14	8
Compound TT	PAY DAY	6	2	1	24	17
Compound TO	LAW SUIT	7	2	0	19	6
Compound TO	WED LOCK	7	2	0	12	15
Compound TO	NAME SAKE	8	2	0	9	17
Compound TO	HEAT WAVE	8	2	0	11	16
Compound TO	TEN FOLD	7	2	0	21	12
Compound TO	NEWS CASTER	9	3	0	9	10
Compound TO	WAR FARE	7	2	0	16	19
Compound TO	HONEY COMB	9	3	0	6	6
Compound TO	JAIL BIRD	8	2	0	11	4
Compound TO	STALE MATE	9	2	0	5	25

Running Head: PARTIAL TRANSPARENCY IN COMPOUND WORDS

Condition	Word	Letters	Syllables	Orthographic Neighbourhood Frequencies		
				Whole Word	First Constituent	Second Constituent
Compound OT	FORK LIFT	8	2	0	9	9
Compound OT	GOD SON	6	2	0	13	21
Compound OT	FOLK LORE	8	2	0	3	22
Compound OT	PEPPER MINT	10	3	0	2	13
Compound OT	QUICK SAND	9	2	0	2	11
Compound OT	TRAP DOOR	8	2	0	6	6
Compound OT	SKY LARK	7	2	0	7	10
Compound OT	TURTLE NECK	10	3	0	1	5
Compound OT	MAY FLY	6	2	0	22	6
Compound OT	BACK LOG	7	2	0	16	17
Compound OO	BILL FOLD	8	2	0	17	12
Compound OO	MOON SHINE	9	2	0	13	11
Compound OO	HEAD WAY	7	2	0	12	21
Compound OO	RAG TIME	7	2	0	22	10
Compound OO	SMALL POX	8	2	0	3	12
Compound OO	PAN HANDLE	9	3	1	22	2
Compound OO	HOLLY HOCK	9	3	0	9	13
Compound OO	HOG WASH	7	2	0	16	12
Compound OO	DUMB BELL	8	2	0	2	13
Compound OO	SHOW DOWN	8	2	0	13	5
Monomorpheme	PLANK TON	8	2	0	-	-
Monomorpheme	HIB ISCUS	8	3	0	-	-
Monomorpheme	ECLI PSE	7	2	1	-	-
Monomorpheme	BIL LIARD	7	2	0	-	-

Running Head: PARTIAL TRANSPARENCY IN COMPOUND WORDS

					Orthographic Neighbourhood Frequencies	
Condition	Word	Letters	Syllables	Whole Word	First Constituent	Second Constituent
Monomorpheme	CRI PPLE	7	2	0	-	-
Monomorpheme	BUC KLE	6	2	2	-	-
Monomorpheme	SOPHO MORE	9	2	0	-	-
Monomorpheme	SHRAP NEL	8	2	0	-	-
Monomorpheme	TOURNI QUET	10	3	0	-	-
Monomorpheme	TUNG STEN	8	2	0	-	-
Monomorpheme	MAN DOLIN	8	3	0	-	-
Monomorpheme	MANNE QUIN	9	3	0	-	-
Monomorpheme	LIME RICK	8	2	0	-	-
Monomorpheme	PAN CREAS	8	3	0	-	-
Monomorpheme	LA RYNX	6	2	0	-	-
Monomorpheme	SCAR LET	7	2	1	-	-
Monomorpheme	TEM PEST	7	2	0	-	-
Monomorpheme	PHOE NIX	7	2	0	-	-
Monomorpheme	GIRA FFE	7	2	0	-	-
Monomorpheme	SCA LLOP	7	2	1	-	-
Monomorpheme	JAS MINE	7	2	0	-	-
Monomorpheme	MON GOOSE	8	2	0	-	-
Monomorpheme	TROM BONE	8	2	0	-	-
Monomorpheme	CASH MERE	8	2	0	-	-
Monomorpheme	BRO CCOLI	8	3	0	-	-
Monomorpheme	PORCU PINE	9	3	0	-	-
Monomorpheme	GUILLO TINE	10	3	0	-	-
Monomorpheme	CHOLE STEROL	11	4	0	-	-

Running Head: PARTIAL TRANSPARENCY IN COMPOUND WORDS

| | | | | | Orthographic Neighbourhood Frequencies | |
Condition	Word	Letters	Syllables	Whole Word	Condition	Word
Monomorpheme	DAN DRUFF	8	2	0	-	-
Monomorpheme	RAC COON	7	2	0	-	-
Monomorpheme	CHEE TAH	7	2	0	-	-
Monomorpheme	CHIP MUNK	8	2	0	-	-
Monomorpheme	THERMO STAT	10	3	0	-	-
Monomorpheme	SEN TINEL	8	3	0	-	-
Monomorpheme	THROT TLE	8	2	0	-	-
Monomorpheme	MALI CIOUS	9	3	0	-	-
Monomorpheme	CON SOLE	7	2	0	-	-
Monomorpheme	SCOUN DREL	9	2	0	-	-
Monomorpheme	SYR INGE	7	2	0	-	-
Monomorpheme	MACK EREL	8	3	0	-	-
Monomorpheme	CAROU SEL	8	3	0	-	-
Monomorpheme	HOWL ING	7	2	0	-	-
Monomorpheme	GAZE LLE	7	2	0	-	-
Monomorpheme	ANT HEM	6	2	0	-	-
Monomorpheme	COM POST	7	2	0	-	-
Monomorpheme	VIC TOR	6	2	1	-	-
Monomorpheme	TUR QUOISE	9	2	0	-	-
Monomorpheme	TUR KEY	6	2	0	-	-
Monomorpheme	LAR VAE	6	2	0	-	-
Monomorpheme	AF RAID	6	2	0	-	-
Monomorpheme	SUP POSE	7	2	0	-	-
Monomorpheme	SYNO NYM	7	3	0	-	-

| | | | | | Orthographic Neighbourhood Frequencies | |
Condition	Word	Letters	Syllables	Whole Word	First Constituent	Second Constituent
Monomorpheme	TAT TOO	6	2	0	-	-
Monomorpheme	AD VERSE	7	2	0	-	-
Monomorpheme	SEW AGE	6	2	0	-	-
Monomorpheme	NURT URE	7	2	0	-	-
Monomorpheme	SYM BOL	6	2	0	-	-
Monomorpheme	SEV ERAL	7	3	0	-	-
Monomorpheme	CHUC KLE	7	2	0	-	-
Monomorpheme	MINI MUM	7	3	0	-	-
Monomorpheme	SNUG GLE	7	2	0	-	-
Monomorpheme	SABO TAGE	8	3	0	-	-
Monomorpheme	GOR GEOUS	8	2	0	-	-
Monomorpheme	TROU BLE	7	2	0	-	-
Monomorpheme	ILL NESS	7	2	0	-	-
Monomorpheme	PRON OUNCE	9	2	0	-	-
Monomorpheme	CON CLUDE	8	2	0	-	-
Monomorpheme	LEAK AGE	7	2	0	-	-
Monomorpheme	CON SENT	7	2	0	-	-
Monomorpheme	ASH AMED	7	2	0	-	-
Monomorpheme	JEA LOUS	7	2	0	-	-
Monomorpheme	SUM MARIZE	10	3	0	-	-
Monomorpheme	MAN AGE	6	2	0	-	-
Monomorpheme	SUFF ICIENT	10	3	0	-	-
Monomorpheme	WHE THER	7	2	0	-	-
Monomorpheme	OB TUSE	5	2	0	-	-

| | | | | Orthographic Neighbourhood Frequencies | | |
Condition	Word	Letters	Syllables	Whole Word	First Constituent	Second Constituent
Monomorpheme	ELE VEN	6	3	0	-	-
Monomorpheme	SOL DIER	6	2	0	-	-
Monomorpheme	BROW SER	7	2	0	-	-
Monomorpheme	FREE DOM	7	2	0	-	-

Appendix B

Table B1

Nonwords Used with their Lengths and Syllables

Condition	Word	Letters	Syllables
Compound	BRIEF TAX	8	2
Compound	CLEAN MIP	8	2
Compound	THERN LOW	8	2
Compound	SPILK WUT	8	2
Compound	TOP DRUG	8	2
Compound	RED BLIN	7	2
Compound	HOM RANK	7	2
Compound	RAD MOSH	7	2
Compound	HEAR FEW	7	2
Compound	HALL WUB	7	2
Compound	GURM DAY	7	2
Compound	VASH PON	7	2
Compound	FLOW GUN	7	2
Compound	FLAT BEW	7	2
Compound	BELF HIT	7	2
Compound	HALD NEG	7	2
Compound	LOT COOL	7	2
Compound	ASK TARP	7	2
Compound	NAS FUND	7	2
Compound	ERK FAND	7	2
Compound	ODD HARD	7	2
Compound	SUN WOLL	7	2
Compound	ORK TYPE	7	2
Compound	BIX MOOK	7	2
Compound	TOAST PULL	9	2
Compound	SPELL CUNG	9	2
Compound	FLURB PAIR	9	2
Compound	THRIM NADE	9	2
Compound	FORM MIND	8	2
Compound	BEST PILT	8	2
Compound	TOOP CASE	8	2
Compound	GIND TREM	8	2
Compound	DUST WORTH	9	2

Condition	Word	Letters	Syllables
Compound	FOOT MILGE	9	2
Compound	TROW BREAK	9	2
Compound	MOWD FLISK	9	2
Compound	GAS BAY	6	2
Compound	OIL RAD	6	2
Compound	LIS FAT	6	2
Compound	ROG CHY	6	2
Monomorpheme	FOS TEN	6	2
Monomorpheme	BAN GUL	6	2
Monomorpheme	TEA REN	6	2
Monomorpheme	ROF FLE	6	2
Monomorpheme	GRE NCH	6	2
Monomorpheme	HEL PRE	6	2
Monomorpheme	DEB UINE	7	2
Monomorpheme	MID MER	6	2
Monomorpheme	CON COVER	8	3
Monomorpheme	MAS TION	7	3
Monomorpheme	IMPEC ULATE	10	4
Monomorpheme	AUTHE BRIATE	11	4
Monomorpheme	ATHO RATE	8	3
Monomorpheme	NAN TOCK	7	2
Monomorpheme	RUN COWL	7	2
Monomorpheme	PAR LEEN	7	2
Monomorpheme	TEM BLE	6	2
Monomorpheme	TOT TLE	6	2
Monomorpheme	WAN BASH	7	2
Monomorpheme	HET TOM	6	2
Monomorpheme	LIS TOP	6	2
Monomorpheme	SON TARE	7	2
Monomorpheme	DEB AND	6	2
Monomorpheme	MES MIER	7	2
Monomorpheme	RUT SINIC	8	3
Monomorpheme	ANP IRE	6	2
Monomorpheme	MEN NISE	7	2
Monomorpheme	PAN COR	6	2
Monomorpheme	REDU OUS	7	3
Monomorpheme	HASS URE	7	2

Condition	Word	Letters	Syllables
Monomorpheme	STR UCE	6	2
Monomorpheme	ARB AND	6	2
Monomorpheme	ERU SANT	7	3
Monomorpheme	COS EER	6	2
Monomorpheme	WOU GHT	6	2
Monomorpheme	STR OPE	6	2
Monomorpheme	PEN FRAND	8	2
Monomorpheme	BIS HION	7	2
Monomorpheme	EVER POL	7	3
Monomorpheme	TOR ROW	6	2
Monomorpheme	BRO MUS	6	2
Monomorpheme	LAR BET	6	2
Monomorpheme	COR AND	6	2
Monomorpheme	BEN ALK	6	2
Monomorpheme	SUMB TION	8	2
Monomorpheme	DRAT SICAL	9	3
Monomorpheme	TUNI LOUS	8	3
Monomorpheme	LOM PLE	6	2
Monomorpheme	DEB ORN	6	2
Monomorpheme	SUN CHEL	7	2
Monomorpheme	MOT USE	6	2
Monomorpheme	NOM ETY	6	3
Monomorpheme	RET OFT	6	2
Monomorpheme	FOL LAP	6	2
Monomorpheme	TEL LOW	6	2
Monomorpheme	ORN EEL	6	2
Monomorpheme	AIR UST	6	2
Monomorpheme	FOM AND	6	2
Monomorpheme	HOS ENT	6	2
Monomorpheme	TUR POD	6	2
Monomorpheme	TOR WAD	6	2
Monomorpheme	TOSS ERATE	9	3
Monomorpheme	FALIN RAN	8	3
Monomorpheme	ANG EST	6	2
Monomorpheme	ARB IST	6	2
Monomorpheme	SPES TIC	7	2
Monomorpheme	FLO UCH	6	2

Condition	Word	Letters	Syllables
Monomorpheme	ALS AVE	6	2
Monomorpheme	INS ANCE	7	2
Monomorpheme	RUS SAGE	7	2
Monomorpheme	EXP LOM	6	2
Monomorpheme	WAS TLE	6	2
Monomorpheme	HOU PER	6	2
Monomorpheme	TON ALL	6	2
Monomorpheme	SPER ANE	6	2
Monomorpheme	BOD DLE	6	2
Monomorpheme	RUN STER	7	2
Monomorpheme	SWAG MITE	8	2
Monomorpheme	HEIS NER	7	2

VITA AUCTORIS

NAME: Alexandria Stathis

PLACE OF BIRTH: Windsor, ON

YEAR OF BIRTH: 1989

EDUCATION: University of Windsor