University of Alberta

The Influence of Morphological Complexity on Word Processing

by

Hongbo Ji

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

Edmonton, Alberta
Spring 2008

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-45538-8
Our file Notre référence
ISBN: 978-0-494-45538-8

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Seven lexical decision experiments were conducted to examine the influence of complex structure on the processing speed of English compound words. The first two experiments revealed that semantically transparent compounds (e.g., *rosebud*) were processed more quickly than matched monomorphemic words (e.g., *giraffe*). Experiment 3 investigated the influence of the constituents on processing speed of transparent compounds by manipulating constituent frequencies while controlling overall compound frequencies. Compounds with high-frequency first constituents were responded to more quickly than compounds with low-frequency first constituents. No such effect was found for the second constituent. In Experiment 4, opaque compounds (e.g., *jailbird* or *hogwash*) were processed more quickly than monomorphemic words. When the decomposition route was reinforced in Experiments 5-7, however, the advantage for opaque compound processing disappeared. In addition, there was even evidence of inhibition due to constituent frequency in opaque compound processing in that high-frequency constituents were associated with slower responses. This research suggests that morphological decomposition initiated by complex structure aids rather than hinders English transparent compound processing because this access route activates consistent information with the direct retrieval of whole word representations. On the other hand, morphological decomposition does not necessarily aid opaque compound processing because this access route can compute a meaning that conflicts with the meaning retrieved by the direct access. For example, the decomposition route would yield the meaning for *jailbird* as "a bird that lives in jail". This interpretation and the retrieved meaning — a prisoner would interfere with each other.

DEDICATION

To my parents, my sister, my daughter, and my husband.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Christina Gagné, who had opened her lab and office to me even before I was admitted to the graduate program. In the past five years, she has offered invaluable training at every aspect of my research. She has been a role model for quality research and professionalism.

I would like to thank my co-supervisor, Thomas Spalding, for his intriguing ideas and perspectives throughout my graduate research and for his wisdom in helping me make my career decisions.

I would like to thank my committee member, Peter Dixon, for his solid advice and support. Peter was the first professor I met in the department. I knew this was the place I wanted to be after our meeting.

Special thanks also to Gary Libben who always opens his lab and resource to me, his vision shaped my research; to Connie Varnhagen for her mentorship of my teaching and a role model for dedication and service, to Rachèl Kemps, Chris Westbury, Shelley Josey, Elena Nicoladis, Allison Mullaly, Lauren Figueredo and Mike Hoff. They all helped me in different ways during my graduate education.

Finally, I would like to thank my loving husband Jinhui, for his love, care and encouragement. If it were not for him, I would have never pursued my dream in Psychology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

The Influence of Morphological Complexity on Word Processing

Overview

The central question for research of representation and processing of complex words such as *snowball* (snow + ball), *editor* (edit + or) and *reunion* (re + union) has been whether morphological decomposition occurs in complex word processing. That is, whether morphological components of complex words (e.g., *snow* and *ball* for *snowball*, *edit* and *or* for *editor*) become available in word processing. Recent studies reveal that morphological decomposition is involved in the processing of complex words (e.g., Andrews, 1986; Andrews, Miller, & Rayner, 2004; Baayen, Dijkstra, & Schreuder, 1997; Bergman, Hudson, & Eling, 1988; Burani & Laudanna, 1992; Caramazza, Laudanna, & Romani, 1988; Hyönä & Pollatsek, 1998; Juhasz, Starr, Pollatsek, Hyönä, & Bertram, 2003; Laudanna, Burani, & Cermele, 1994; Libben, 1998; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Pollatsek, Hyönä, & Bertram, 2000; Pollatsek & Hyönä, 2005; Sereno & Jongman, 1997; Taft & Forster, 1975; 1976; Taft, 1979; Taft, 1994; Zwitserlood, 1994). These findings of morphological decomposition suggest that complex words can be recognized via access to their components in addition to direct retrieval of whole word representations.

The aim of the current dissertation is to expand previous research by: 1) exploring the influence of morphological decomposition on the processing time of English compound words; 2) investigating whether semantic transparency influences the processing time of compound words by influencing the likelihood of morphological decomposition; 3) investigating whether the relationship of morphological decomposition and the direct retrieval of whole word representations is mutually facilitative or inhibitory and how this relationship influences the speed of compound word processing; 4) investigating whether the relationship of morphological decomposition and direct retrieval is flexible by examining whether morphological decomposition can be accelerated by experimental manipulation and whether this acceleration mediates the interplay of morphological decomposition and direct retrieval.

Investigating these questions is important for the following reasons; first, no current model of complex word processing explicitly predicts the influence of morphological decomposition on the processing speed of compound words although it is directly relevant to the big picture question of the balance of computation and storage in word recognition and processing; morphological decomposition involves meaning computation and it is important to investigate how morphological decomposition interacts with direct retrieval of the stored whole word representations of compound words in memory. Second, the answer to the question of the influence of morphological decomposition on the processing speed of compound words might not be a simple "advantage" or "disadvantage" because previous research suggests that factors such as semantic transparency, influence the extent to which a compound word is decomposed (e.g., Sandra, 1990; Zwitserlood, 1994). In specific, the processing of semantically transparent compounds, whose meanings can be derived from their constituents (e.g., *rosebud*, or *teacup*), involve morphological decomposition at both the lexical level and semantic level, whereas the processing of semantically opaque compounds, whose meanings cannot be derived from their constituents (e.g., *hogwash*, or *jailbird*), might involve morphological decomposition only at the lexical level. Thus, examining the influence of semantic transparency on the processing speed of compound words can help understand the balance of computation and storage more exclusively at both the lexical level and semantic level and possibly the interplay of the balance at these two levels.

Finally, although most current models of complex word processing assume that both morphological decomposition and the direct retrieval of whole word representations are involved in complex word processing, the nature of the relationship of these two access routes has not been resolved. Some models, such as Taft's (and Forster, 1975; 1976) automatic decomposition model, are stage models that assume serial processing where obligatory decomposition precedes the direct retrieval of whole word representations. Other models, such as the dual-route race model (Baayen, Dijkstra, & Schreuder, 1997; Frauenfelder & Schreuder, 1992; Schreuder & Baayen, 1995), assume simultaneous but independent involvement of the decomposition and the direct retrieval routes. Yet, the third kind of model, such as Libben's (1998) model of compound word processing, can be interpreted as assuming cooperation of these two access routes at the lexical level and inhibition of these two access routes at the semantic level. Thus, it is necessary to closely examine the relationship of these two access routes in various types of situations to test the validity and generaliability of these models of complex word processing. Furthermore, it is of theoretical interest to investigate whether encouraging morphological decomposition alters the relationship and the balance of the two access routes. If so, it would suggest that the relationship of the two access routes is flexible. In the following eight chapters, I will propose that it is plausible in principle that the two access routes are involved simultaneously and facilitate each other at both the lexical and semantic levels in compound word processing. In addition, a more abstract conceptual level might also be involved in that the activation of semantic representations of the constituents could be extended to the conceptual level. As a result, the decomposition route might compose a meaning for the whole based on the constituents that would be either consistent or inconsistent with the retrieved meaning depending on whether the compound is semantically transparent or opaque. Thus, a more appropriate label for the "decomposition" route discussed in the context of the proposed theoretical framework in this dissertation might be "composition" route due to its emphasis on meaning composition.

I will also provide empirical evidence to show that these two access routes are mutually facilitative in the processing of transparent compound words because the two access routes activate consistent information and boost each other in this kind of situation. For example, for the transparent compound word *rosebud*, the decomposition route activates *rose* and *bud* at the lexical level, which should increase the activation of *rosebud* at the direct access through facilitative activation. At the semantic level, similar processes occur. The semantic activation of *rose* and *bud* should increase the activation of *rosebud* as a whole because the meaning of *rosebud* is related to both *rose* and *bud* — rosebud is the bud of a rose. Moreover, there might be facilitation from the conceptual level; the activated representations for the concepts *rose* and *bud* can be composed to compute a meaning *the bud of a rose* which is consistent with the retrieved legitimate meaning of the whole word. The computed meaning and the retrieved meaning should increase the activation of each other. In contrast, the two access routes are either mutually facilitative or inhibitory in the processing of opaque compound words because the two access routes activate consistent information at the lexical level but conflicting information across semantic and conceptual levels. For example, for the opaque compound word *jailbird*, the decomposition route activates *jail* and *bird* at the lexical level, which should increase the activation of *jailbird* via the direct access through facilitative activation. However, things are different at the semantic level. Although there might be a facilitative link connecting *jail* and *jailbird* at this level, no such link should connect *bird* and *jailbird* because *bird* does not contribute to the meaning of *jailbird*. At the conceptual level, the decomposition route could compute a meaning, e.g., "a bird that lives in jail", which is not compatible with the retrieved whole word meaning, "a prisoner". The computed meaning and the retrieved legitimate meaning might conflict with each other and result in interference. Because the advantage due to facilitation between activated units of the two access routes at the lexical level and the disadvantage due to conflict of activated units of the two access routes across semantic level and conceptual level have to trade off, whether the ultimate influence is advantage or disadvantage depends on whether morphological decomposition is encouraged so that lexical and semantic units are more strongly activated at this route. If morphological decomposition is not encouraged, then there should be an advantage for opaque compound processing because in this kind of situation, the decomposition route is less likely to compute a meaning that would conflict with the retrieved meaning. If the decomposition route is encouraged, then there should be no advantage or even disadvantage for opaque compound processing because in this kind of situation, the decomposition route is more likely to compute a meaning that conflicts with the retrieved legitimate meaning at the direct access. This should introduce interference between the two access routes and slow down word recognition which should be eventually determined by the direct access.

Introduction

Major questions addressed by previous research of compound word processing

How do people recognize compound words such as *snowball* and *rosebud*? One possibility is that compound words are represented in the mental lexicon as whole word units like monomorphemic words such as *giraffe* and *chimpanzee* (e.g., Butterworth, 1983). However, in the past three decades, ample evidence has shown that compound words also can be accessed via their constituents (e.g., Andrews, 1986; Andrews, Miller, & Rayner, 2004; Hyönä & Pollatsek, 1998; Juhasz, Starr, Pollatsek, Hyönä, & Bertram, 2003; Libben, 1998; Pollatsek, Hyönä, & Bertram, 2000; Pollatsek & Hyönä, 2005; Taft & Forster, 1975; 1976; Taft, 1979; Taft, 1994; Zwitserlood, 1994).

The access to compound constituents is called morphological decomposition. Evidence of morphological decomposition in compound word processing mainly came from two lines of research. Researchers have manipulated the constituent frequency of compound words and demonstrated that constituent frequency influenced the ease of compound word processing in either lexical decision times (e.g., Andrews, 1986; Juhasz, Starr, Pollatsek, Hyönä, & Bertram, 2003; Taft & Forster, 1976) or eye movement measurements (e.g., Andrews, Miller, & Rayner, 2004; Hyönä & Pollatsek, 1998; Juhasz, Starr, Pollatsek, Hyönä, & Bertram, 2003; Pollatsek, Hyönä, & Bertram, 2000; Pollatsek & Hyönä, 2005). Researchers had also used a priming procedure to investigate the presence of morphological decomposition in compound word processing. In this procedure, pairs of words were presented to participants sequentially. Here, either compound words were first presented to participants, and then followed by their constituents (e.g., Zwitserlood, 1994), or constituents were first presented and followed by the original compounds (e.g., Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999; Libben, Gibson, Yoon, & Sandra, 2003; Monsell, 1985). Either way, evidence of morphological decomposition was identified in that compound words and their constituents speeded the recognition of each other.

Besides the evidence of morphological decomposition, there was also evidence of whole word access in compound word processing. Andrews et al. (2004) and Pollatsek et al. (2000) found that whole word frequency influenced the ease of English compound and Finnish compound processing respectively in that the higher the whole word frequency, the faster the word recognition. Taken together the evidence of morphological decomposition and the evidence of the retrieval of whole word representations, it appears that two access routes are involved in compound word processing. One is the decomposition route and the other is the direct access.

In addition, the issue of semantic transparency was also investigated, together with the issue of morphological decomposition in compound word processing. In general, compound words can be further divided into semantically transparent compounds whose meanings can be derived from their constituents (e.g., *rosebud*) and semantically opaque compounds whose meanings cannot be derived from their constituents (e.g., *hogwash* or *jailbird*). The question that has been addressed in this line of research is whether semantic transparency mediates morphological decomposition of compound words (e.g., Libben, 1998; Pollatsek & Hyönä, 2005; Sandra, 1990; Zwitserlood, 1994). In specific, Zwitserlood found priming effect not only for transparent Dutch compounds and their constituents (e.g., *kerkorgel,* "church organ" vs. *orgel*, "organ") but also for opaque Dutch compounds and their constituents (e.g., *klokhuis* , "clockhouse", meaning, "core of an apple" vs. *huis*, "house"). However, Sandra only found priming effect for transparent Dutch compounds (e.g., *brood*, "bread" vs. *boterpot*, "butter-dish") but not for opaque Dutch compounds (e.g., *brood*, "bread" vs. *boterbloem*, "buttercup") when the primes were semantically related to the constituents of target compounds. These two findings together were usually interpreted as evidence to suggest that constituents are accessed at both the lexical level and semantic level in transparent compound processing but only at the lexical level in opaque compound processing (e.g., Libben, 1998). However, this interpretation does not explain what prevents the activation spreading from the lexical level to semantic level.

The purpose of the current studies

The purpose of the current studies is to expand previous research by exploring the influence of morphological decomposition on the processing time of compound words. Unlike previous research that has been focused on whether morphological decomposition occurs in compound word processing, the current studies investigate how morphological decomposition influences the processing speed of compound words by investigating the relationship of the decomposition route and direct access. In other words, I explore whether the relationship of the two access routes is mutually facilitative or inhibitory and how this relationship influences the speed of compound word processing. I also investigate whether the relationship of the two access routes can be altered by experimental manipulation, such as the use of filler items, intra-word spaces and color contrast of constituents of compound words, to further explore the relationship of the two access routes in these kinds of situations. In the meantime, the current studies also investigate whether semantic transparency influences the processing time of compound words by influencing the likelihood of morphological decomposition and the relationship of the two access routes.

The motivation for the current research is that current models of complex word processing mainly focus on the explanation for the presence and locus of morphological decomposition, which has been successful in interpreting the abundant findings of morphological decomposition in complex word processing. These models, however, do not explicitly address how morphological decomposition influences the processing time of complex words. This question is of particular theoretical interest because it touches the fundamental question of the balance of storage and computation for the function of the mental lexicon. Based on the metaphor of how computers work, researchers used to assume that computation and storage have to trade off in word processing and storage (e.g., Baayen, Dijkstra, & Schreuder, 1997; Sandra, 1994; Schreuder & Baayen, 1995; Taft & Forster, 1975; 1976). That is, if the mental lexicon is designed for processing efficiency, it would probably have to experience a storage burden in that it needs to store all complex forms. On the other hand, if the mental lexicon is designed for storage efficiency, it would probably have to experience a computational cost in that it needs to compute meanings for complex items online based on their constituents. In other words, if the mental lexicon takes the advantage of saving storage, it has to endure a computational cost; if the mental lexicon takes the advantage of saving computation, it has to endure a storage burden. In complex word processing, the decomposition route involves computation because this route should first decompose complex words into their constituents and then combine the constituents to generate the meaning of the whole. Here, computation means nothing more than cognitive efforts involved in the process of decomposition and composition. In the meantime, direct retrieval might also be involved in complex word processing. This process directly maps complex words to their whole word representation stored in memory. Thus, direct access in complex word processing is relevant to the storage of complex forms in the mental lexicon.

The issue of the balance of computation and storage in word processing is important because it can help understand the organization and structure of the mental lexicon. For example, although most current models of word processing assume the existence of decomposed representations of morphologically complex words and allow for decomposition/computation in word processing, it is still not clear what the benefit is of having morphological decomposition/computation and how this morphological decomposition/computation interplays with the direct retrieval of whole word representations, if any. If the benefit is saving storage, then why is there evidence of whole word representations for complex words (e.g., Andrews et al., 2004; Baayen, Dijkstra, & Schreuder, 1997; Pollatsek et al., 2000)? Thus, it is possible that having both decomposed representations and whole word representations is more beneficial than having either alone. If so, how is this benefit realized and what are the processes involved? Answers to these questions can certainly shed light on more detailed description of the structure and organization of the mental lexicon.

Current models of complex word processing

Among current models of complex word processing, the automatic decomposition model (Taft & Forster, 1975; 1976) posits that morphological decomposition is obligatory in complex word processing because the mental lexicon is organized and centered on stems and components of complex words. According to this model, morphological decomposition occurs from left to right in complex word processing with an attempt to identify an access code. When this access code is identified, it is used as an index to search for the whole word representation for the target word according to the frequency of words that share the same access code. For example, in compound word processing, the first constituent might be identified as an access code. Then, the system would use the first constituent as an index to search a list of words that share the same first constituent according to the frequency of their occurrence. Thus, this model assumes that morphological decomposition occurs before direct retrieval of whole word representation. Later on, Taft (1994) integrated this model into an interactive activation framework, where decomposition becomes an automatic process driven by the interactive connections among lexical entries. This later version of the model differs from the early version in that it implies simultaneous instead of sequential involvement of the two access routes. Also, this version takes the issue of semantic transparency into consideration in that it depicts the processing of semantically transparent and opaque complex words differently which I will provide more details when discuss predictions of current models in the section to follow. In sum, the focus of the automatic decomposition model, as its name implies, is to emphasize the dominant role of morphological decomposition in complex word recognition. The motivation behind such a proposal is that decomposed representation and processing can save storage and make the organization of the mental lexicon more efficient or structured.

However, some researchers suggest that saving storage implies introducing a computational cost in morphological decomposition because in the process of decomposition, people have to first decompose complex words into their constituents and then combine the constituents to generate the meaning of the whole. Thus, complex word representation and processing might be organized in a more balanced manner. Based on this kind of philosophy, researchers proposed a dual-route race model (Baayen, Dijkstra, & Schreuder, 1997; Frauenfelder & Schreuder, 1992; Schreuder & Baayen, 1995). According to this model, a direct access route and a decomposition route are simultaneously involved in complex word processing. The route that wins the race determines the word recognition time. For high-frequency complex words with low-frequency constituents, the direct access route is more likely to win because the threshold of the whole word representation is relatively low. For low-frequency complex words with high-frequency constituents, the chance for the decomposition route to win is increased because the activation threshold of the whole word representation is relatively high. Thus, the dual-route race model aims to bring a balanced view of storage and computation by introducing two access routes working simultaneously.

A third model that needs to be discussed is the automatic progressive parsing and lexical excitation (APPLE) model by Libben (1998). Unlike the first two models that more or less aim to encompass the issues of complex word processing in general, the APPLE model focuses on compound words with the aim to address the issue of semantic transparency. This model assumes three levels of representations for compound words, a stimulus level, a lexical level and a conceptual level. Transparent and opaque compounds are represented in a similar way up to the lexical level where within-level facilitative links connect constituents and whole word representations. For example, *blue* and *berry* are both linked to *blueberry* and *straw* and *berry* are both linked to *strawberry* at the lexical level. Semantic transparency is captured by cross-level links between the lexical level and semantic level where whole word representations at the lexical level are connected to transparent constituents but not opaque constituents at the conceptual level. For example, *strawberry* as a whole word at the lexical level is linked to *strawberry* and *berry* but not *straw* at the conceptual level because the meaning of *berry* is transparent but the meaning of *straw* is opaque here. In addition, the whole and the constituents inhibit each other at the conceptual level regardless of semantic transparency. The APPLE model is also dual-route in nature and the assumption of the facilitative links between compounds and their constituents at the lexical level distinguishes this model from the dual-route race model because it essentially assumes cooperation instead of race of the two access routes at the lexical level. Thus, this model does not necessarily assume that computation and storage have to trade off as assumed by the dual-route race model.

It is apparent that all three models touch the issue of the balance or the relationship of computation and storage. Yet, due to their different focuses, none of them directly address the influence of morphological decomposition on the processing time of complex words relative to monomorphemic words. Comparing the processing time of these two types of words is important because the processing of monomorphemic words only involves direct access, which provides a baseline to investigate the possible influence of morphological decomposition on the speed of complex word processing.

As of the relationship of the two access routes, the early version of the automatic decomposition model (Taft & Forster, 1975; 1976) is a serial stage model where morphological decomposition precedes whole word access although this position is somewhat weakened in the later version (Taft, 1994). In contrast, the dual-route race model assumes simultaneous and independent involvement of both routes. In principle, however, as implied by the APPLE model to some degree but not completely, it is also possible that the two access routes are involved simultaneously and communicate with each other via facilitative activation. Importantly, such mutual facilitation might not only occur at the lexical level as assumed by the APPLE model, but also at the semantic level because if one assumes lexical access to compound constituents, activation at the lexical level should automatically spread to the semantic level.

Such a position makes interesting predictions for the influence of morphological decomposition on the processing time of compound words. That is, the communication of the two access routes should aid the processing of transparent compound words through mutual facilitation between the units activated at the two access routes throughout the process. For example, for the transparent compound word *rosebud*, the decomposition route activates *rose* and *bud* at the lexical level, which should increase the activation of *rosebud* via the direct access through facilitative activation. At the semantic level, similar processes occur. The semantic activation of *rose* and *bud* should increase the activation of *rosebud* because the meaning of *rosebud* is related to both *rose* and *bud* — rosebud is the bud of a rose. Moreover, there might be facilitation from the conceptual level;[1] the activated representations for the concepts *rose* and *bud* can be composed to compute a

---

[1] The conceptual level here is different from the conceptual level proposed within the structure of the APPLE model. In the APPLE model, the "conceptual level" is more inclusive and should correspond to both the semantic level which represents word meanings and the conceptual level which represents more general conceptual knowledge as proposed in this dissertation.

meaning *the bud of a rose* which is consistent with the retrieved legitimate meaning of the whole word. The computed meaning and the retrieved meaning should increase the activation of each other. On the other hand, the communication of the two access routes might be more complicated in the processing of opaque compound words because the two access routes should be mutually facilitative at the lexical level but mutually inhibiting across semantic and conceptual levels. For example, for the opaque compound word *jailbird*, the decomposition route activates *jail* and *bird* at the lexical level, which should increase the activation of *jailbird* via the direct access through facilitative activation. However, things are different at the semantic level. Although there might be a facilitative link connecting *jail* and *jailbird* at this level, no such link should connect *bird* and *jailbird* because *bird* does not contribute to the meaning of *jailbird*. At the conceptual level, the decomposition route could compute a meaning, e.g., "a bird that lives in jail", which is inconsistent with the retrieved whole word meaning, "a prisoner". The computed meaning might inhibit the retrieved legitimate meaning. Whether the ultimate influence is beneficial or hindrance depends on whether the decomposition route is encouraged so that the computed non-target meaning is more likely to introduce inhibition between the computed meaning and the retrieved meaning. Thus, there might be either evidence of advantage or disadvantage in opaque compound processing depending on whether the decomposition route is boosted.

Predictions of current models on processing speed of compound words

What kinds of predictions can current models make for the processing time of transparent compounds and opaque compounds relative to monomorphemic words which only involve direct access?

The predictions of the automatic decomposition model. In general, both the early version (Taft & Forster, 1975; 1976) and the later version (Taft, 1994) of the automatic decomposition model emphasize the involvement of morphological decomposition in complex word processing. However, they differ in terms of how they depict the relationship of morphological decomposition and direct retrieval in complex word processing and whether they address the issue of semantic transparency. Thus, I discuss the predictions of the two versions of this model respectively. First, based on the early version of the automatic decomposition model which is a serial stage model, and the general view that computation is time consuming, the most natural prediction of the early version of this model on the processing speed of compound words seems to be that compound words should be processed more slowly than monomorphemic words. This is because morphological decomposition occurs prior to direct retrieval which is the searching and verification stage in compound word processing. Thus, the processing of compound words should be slower than the processing of monomorphemic words which only involves direct retrieval; the former involves more stages and computation than the latter. In addition, unlike the later version of this model, this early version predicts no difference on the relative processing speed of transparent and opaque compounds because it does not distinguish semantically transparent and opaque complex words in its framework.

Second, based on the later version of this model (Taft, 1994) where the serial searching nature is tempered, morphological decomposition in compound word processing may occur simultaneously with the direct access at the lexical level but not at the semantic level. According to this version of the model, for transparent complex word processing, only decomposition/composition but not direct retrieval is available at the semantic level, whereas for opaque complex word processing, only direct retrieval but not decomposition/composition is available at the semantic level. The rationale for assuming only computation at the semantic level for transparent complex words is saving storage. For example, the meaning of *reheat* "make hot again" can be derived from the two morphemes *re*, meaning "again", and *heat*, meaning, "make hot". So, there is no need to store the whole word meaning which is redundant. In contrast, the reason for assuming only direct retrieval at the semantic level for opaque complex word processing is that the meaning of opaque complex words cannot be computed based on the meanings of their parts. For example, the meaning of *relate* is "pertain to" instead of "not on time again". If there is any attempt to compute the meanings of opaque complex words via the decomposition route (e.g., compute the meaning for *relate* as "not on time again"), it should be suppressed by the direct retrieval of whole word meanings (e.g, retrieve the meaning for *relate* as "pertain to"). Note, however, that assuming such a suppression mechanism implies that people somehow know *a priori* whether a complex word is semantically opaque or transparent before the access reaches the semantic level. Based on such a framework, the most straightforward predictions for the processing speed of compound words versus monomorphemic words and for the difference between transparent and opaque compounds would be as the following.

Among the three types of words, the processing of opaque compounds should be the fastest, and there should be no difference between transparent compound processing and monomorphemic word processing. In specific, if the lexical activation of compound constituents and the whole can increase the activation of each other via their links (e.g., *rose* and *bud* should be both linked to *rosebud*, *jail* and *bird* should be both linked to *jailbird*) as assumed by this model, compound word processing in general should have an advantage over monomorphemic word processing at the lexical level because monomorphemic word processing only involves direct retrieval and cannot benefit from facilitative activation between the parts and the whole. At the semantic level, there should be no difference between the processing of opaque compounds and monomorphemic words because the meanings of both types of words are accessed by direct retrieval. There should be a computational cost, however, for transparent compounds at the semantic level — the whole word meanings of transparent compounds are not represented or stored, so they must be computed based on the meanings of their constituents. This computational cost at the semantic level should trade off the advantage gained at the lexical level due to the involvement of both access routes. Thus, overall, there should be no difference in processing time for transparent compounds and monomorphemic words.

The predictions of the dual-route race model. In contrast with the automatic decomposition model, the dual-route race model (Baayen et al., 1997; Schreuder & Baayen, 1995) provides different predictions of the processing speed of the two types of compound words relative to monomorphemic words and the processing speed of transparent compounds relative to opaque compounds. For the comparison between transparent compounds and monomorphemic words, it is possible for this model to predict an advantage for transparent compounds. This prediction is based on the assumption of statistical facilitation made within the framework of this model. According to the dual-route race model, a direct access route and a decomposition route are simultaneously involved in complex word processing. The route that wins the race determines the word recognition time. This model assumes that responses for a certain item following each access route form a distinct frequency distribution. When the distributions of the two routes overlap, mean response time to a certain item should be less than following either access route alone because the faster end of the distribution of the slow route wins over and replaces the slower end of the distribution of the faster route. This facilitation in response times due to the existence of two independent routes is called statistical facilitation (Raab, 1962). Thus, if one assumes that the distributions of the two access routes overlap in transparent compound processing, it should be predicted that transparent compounds are processed more quickly than monomorphemic words which can only be recognized by a single direct access.

For the processing time of opaque compounds, however, the dual-route race model predicts no statistical facilitation because it assumes that opaque compounds can only be recognized via the direct access in that their whole word meanings cannot be computed based on their constituents (Schreuder & Baayen, 1995). In other words, this model assumes that opaque compounds are recognized in a similar way as monomorphemic words. Consequently, the dual-route race model should predict no difference in processing time for opaque compounds and monomorphemic words. Following the same logic, transparent compounds should be processed more quickly than opaque compounds because the former can benefit from statistical facilitation while the latter cannot.

The predictions of the APPLE model. A third different set of predictions can be made by the APPLE model (Libben, 1998) for the relative processing speed of transparent compounds, opaque compounds and monomorphemic words. As mentioned earlier, this model assumes three levels of representations, a stimulus level, a lexical level and a conceptual level. At the lexical level, within-level facilitative links connect compound words and their constituents regardless of semantic transparency. In addition, for transparent compounds, cross-level facilitative links connect whole word representations of compounds at the lexical level to both whole word representations and constituent representations at the conceptual level. For opaque compounds, cross-level facilitative links only connect whole word representations at the lexical level to whole word representations and the transparent constituents, if any, at the conceptual level. Furthermore, this model also assumes inhibition between representation units (within-level cross-level alike) where facilitative links are absent to prevent the activation of non-target units. For example, because no link connects *strawberry* at the lexical level and *straw* at the conceptual level, the meaning of *straw* should be suppressed to prevent it from interfering with the activation of the whole word meaning of *strawberry*. Importantly, no links between the whole and the parts are assumed at the conceptual level regardless of semantic transparency. For example, *blueberry* is not linked to *blue* or *berry*, *strawberry* is not linked to *straw* or *berry* at the conceptual level. This structure of the model implies that the meanings of compound words are accessed by direct retrieval instead of computation regardless of semantic transparency. It should be noted though that assuming absence of facilitative links might not be the best way to implement inhibition because if lexical and semantic units are not connected by links, it is difficult to appreciate how these units inhibit each other. That is, if *strawberry* at the lexical level is not linked to *straw* at the conceptual level, then *straw* as a concept should not be activated at all, and therefore there seems to be no need to further assume inhibition between the two. In any event, by assuming this inhibition mechanism, the bottom line of the APPLE model is to stress that meaning composition does not occur in compound word processing regardless of semantic transparency.

The APPLE model essentially depicts cooperation of the two access routes at the lexical level regardless of semantic transparency because it posits facilitative links between compound words and their constituents at this level, transparent and opaque compounds alike. Thus, this model should predict an advantage for both transparent and opaque compound processing relative to monomorphemic word processing because the latter cannot benefit from morphological decomposition like the former.

As for the relative processing speed of transparent and opaque compounds, according to this model, advantage for compound processing should occur at the lexical level but not at the semantic level because meanings of compounds are retrieved rather than computed regardless of semantic transparency. Thus, this model should predict no difference in processing speed for transparent and opaque compounds. Alternatively, a slightly different prediction could be made if one takes into consideration the inhibition mechanism assumed by this model. In specific, the model depicts within-level inhibition between compound words and their constituents at the conceptual level regardless of semantic transparency (e.g., *blueberry* should inhibit *blue* and *berry* and *strawberry* should inhibit *straw* and *berry* at the conceptual level). In terms of cross-level inhibition, the model assumes inhibition from lexical representations of whole words to opaque constituents at the conceptual level but not to transparent constituents (e.g., *strawberry* at the lexical level inhibits *straw* at the conceptual level whereas *blueberry* at the lexical level does not inhibit *blue* at the conceptual level). Thus, in total, opaque compound processing should receive more inhibition from the direct access than transparent compound processing. In other words, the decomposition route should be less interfering in opaque compound processing than in transparent compound processing at the conceptual level. As a result, one could also predict that the processing of opaque compounds should be quicker than transparent compounds by further assuming that the cross-level inhibition between whole word representations at the lexical level and constituent representations at the conceptual level influences processing time significantly.

To summarize, the predictions of the three models of the processing time of transparent compounds, opaque compounds and monomorphemic words are as the following. The early version of the automatic decomposition model predicts that compound words should be processed more slowly than monomorphemic words regardless of semantic transparency. In addition, there should be no difference in the processing of transparent and opaque compounds. The later version of the automatic decomposition model predicts that opaque compounds should be processed more quickly than transparent compounds and monomorphemic words. In addition, there should be no difference in the processing of transparent compounds and monomorphemic words. In contrast, the dual-route race model predicts that transparent compounds should be processed more quickly than opaque compounds and monomorphemic words, and there should be no difference in the processing of opaque compounds and monomorphemic words. Finally, the APPLE model has two possible predictions. The first possible prediction is that both transparent and opaque compounds should be processed more quickly than monomorphemic words, and there should be no difference in processing the two types of compound words. The second possible prediction is that opaque compounds should be processed more quickly than transparent compounds, and transparent compounds should be processed more quickly than monomorphemic words. In the section to follow, I will turn to previous studies of the processing speed of complex words and how they are related to the current studies.

Previous findings of the processing speed of complex words and their implications

Previous studies comparing the processing speed of complex words and monomorphemic words have provided mixed results with regard to the issue of the influence of morphological decomposition on the processing time of complex words. These results and the interpretation provided for these results are insightful for the current studies that exclusively examine the processing time of compound words. For example, two studies showed that Finnish case inflectional forms (e.g., elative case form *auto+sta* "car+from") were processed more slowly than monomorphemic words in lexical decision tasks (Bertram, Laine, & Karvinen, 1999; Laine, Vainio, & Hyönä, 1999). In contrast, a study comparing Dutch derivational complex words with monomorphemic controls yielded the opposite result; derivational forms were processed more quickly than monomorphemes in both lexical decision and naming tasks (Hudson & Buijs, 1995). Similarly, Bertram and colleagues found that Finnish derivational forms with a productive denominal suffix *stO* (e.g., *kirja+sto* [book+collective noun] "library") were processed more quickly than monomorphemic controls.

Bertram and colleagues suggested that the assumption of statistical facilitation within the framework of the dual-route race model (Baayen, Dijkstra, & Schreuder, 1997;

Frauenfelder & Schreuder, 1992; Schreuder & Baayen, 1995) could explain why Finnish derivational forms with a productive suffix are processed more quickly than monomorphemic words. As mentioned earlier, according to the dual-route race model, a direct access route and a decomposition route are simultaneously involved in complex word processing. The route that wins the race determines the word recognition time. The direct access route is relatively fast because it directly maps the complex word to the stored whole unit representation in the system. The decomposition route is relatively slow because computation is involved in the process. When the two response distributions corresponding to the direct access route and the decomposition route overlap, word recognition decisions should get statistical facilitation, and thus be faster than decisions made using a single route alone. However, if statistical facilitation occurs in dual-route access, it is then difficult for the same model to explain the cost or disadvantage associated with the processing of some complex words such as Finnish case inflectional forms (Bertram et al., 1999; Laine et al., 1999). To make the assumption of statistical facilitation consistent with the structure of the model, one has to assume that dual-route race does not occur in the processing of all kinds of complex words. For some complex words, such as Finnish case inflectional forms, the decomposition route might be the only route at work.

In contrast, based on their findings, Laine and colleagues have suggested that the relatively slow speed in processing Finnish case inflectional forms is due to the fact that these complex words are processed via the decomposition route, which is time consuming. Laine and colleagues further suggested that the relationship between the decomposition route and the direct access route is inhibitory because they also found that morphologically ambiguous words, that can be either interpreted as case inflectional forms or monomorphemic words, were processed more quickly than case inflectional forms but more slowly than monomorphemic words. Based on this finding, they proposed that both access routes are activated in the processing of morphologically ambiguous words for the two possible interpretations respectively, then, inhibit the activation of each other and slow down both processes. Morphologically ambiguous words are processed more quickly than case inflectional forms because these words are recognized via the direct access route most of the time. Morphologically ambiguous words are processed more slowly than monomorphemic words because the inhibited direct access in the processing of morphologically ambiguous words is slower than the non-inhibited direct access in the processing of monomorphemic words. Although the assumption about the dominant role of the decomposition route in the processing of case inflectional forms and the assumption of the inhibition mechanism in the processing of morphologically ambiguous forms can explain the relatively slow speed in processing Finnish inflectional forms and morphologically ambiguous forms, these mechanisms cannot be extended to explain the relatively fast processing of other complex words such as Dutch and Finnish derivational forms (Bertram et al., 1999; Hudson & Buijs, 1995; Laine et al., 1999).

Based on these findings of other types of complex words, it appears that both a facilitative mechanism and an inhibition mechanism might be a special case of a neutral structure where each mechanism corresponds to a specific situation. In other words, no universal principle governs the processing of different types of complex words in different languages. This observation is consistent with the early proposal for compound word processing where the two access routes might be mutually facilitative in transparent compound processing and mutually inhibitory in opaque compound processing.

The motivation of the current studies

Proposals assuming either statistical facilitation or inhibition of the two access routes are based on studies using inflectional or derivational complex words. The influence of morphological decomposition on the processing time of compound words, however, has not been systematically investigated. Unlike inflections and derivations, which consist of one independent meaning unit plus an affix, compounds consist of two or more stems that have independent meanings. Whether the independent meaning units demand more computation in word recognition and hence slow down the speed of compound processing, or rather, the involvement of two or more real stems can make the processing benefit more from facilitative activation between the constituents and the whole is an important theoretical question that can help understand the role of semantics and the balance of computation and storage in complex word processing. In the meantime, this question is also relevant to the applicability of the assumption of statistical facilitation and the mechanism of inhibition in compound processing.

Another reason to use compound words in the current research is that as mentioned earlier, the issue of semantic transparency has been systematically investigated in compound word processing and that semantic transparency might be an important factor that influences the speed of compound word processing. As mentioned earlier, previous studies found repetition priming for both transparent and opaque compounds but semantic priming only for transparent compounds (e.g., Sandra, 1990; Zwitserlood, 1994). These findings suggest that meaning composition might play different roles in processing transparent and opaque compounds and this should influence the relative processing speed of transparent and opaque compounds. In particular, because the computed meaning is compatible with the retrieved meaning in transparent compound processing, this might speed the recognition of transparent compounds. In contrast, because the computed meaning is inconsistent with the retrieved meaning in opaque compound processing, this conflict might introduce inhibition between the two and slow down word recognition.

Two recent studies compared the processing speed of compound words and monomorphemic words. For example, Fiorentino and Poeppel (2007) found that compound words were processed more quickly than monomorphemic words using lexical decision task. This result, however, was interpreted as evidence of morphological decomposition. In addition, the authors suggested that the advantage of compound word processing is due to access to compound constituents that increases the ease of word recognition. They argued that this finding supports dual-route models that assume early decomposition in compound word processing because a late decomposition model predicts no advantage of compound processing relative to monomorphemic word processing. The authors did not define whether the access to compound constituents was lexical or semantic in nature because they did not manipulate semantic transparency in the experiment. Similarly, Juhasz (2006) compared the processing of compound words with monomorphemic words when they were presented in sentence context in two experiments. She found that compound words with high constituent frequencies were recognized more quickly than compound words with low constituent frequencies and monomorphemic words using eye movement measurements. Again, the difference between compound words and monomorphemic words were interpreted as evidence of morphological decomposition in compound word processing. Also, the author suggested that the advantage of compound word processing is due to lexical access to compound constituents that facilitates whole word processing as proposed by the APPLE model of Libben (1998).

The current studies differ from the two previous studies in the following ways. First, neither of the previous two studies further investigated the issue of how semantic transparency mediates the degree of morphological decomposition and influences the speed of compound word processing. In contrast, the current research emphasizes the issue of semantic transparency and how it influences the degree of morphological decomposition and the relative speed in processing transparent and opaque compounds. This is of particular theoretical interest because as discussed earlier, current models of complex word processing make different predictions of how semantic transparency influences the degree of morphological decomposition and the speed of compound word processing.

It should be acknowledged that Fiorentino and Poeppel (2007) provided insightful speculations that further examining the influence of semantic transparency as an extension of their studies can help identify the locus of the advantage of compound processing found in their studies that used semantically transparent compounds. The results of the current studies suggest that the processing of transparent compounds takes advantage from both lexical and semantic access to their constituents whereas the processing of opaque compounds takes advantage mainly from lexical access to their constituents. This pattern is in general consistent with Fiorentino and Poeppel's speculations based on findings of semantic transparency effects in the literature. Namely, as just mentioned earlier in this Introduction, there was no semantic transparency effect in repetition priming between compounds and their constituents (e.g., Zwitserlood, 1994) and there was semantic transparency effect in semantic priming in that primes that were semantically associated with compound constituents speeded the recognition of transparent compounds but not opaque compounds (e.g., Sandra, 1990). Importantly, the results of the current studies further suggest that meaning composition is an important component in compound processing. This component might benefit transparent compound processing but cause interference in opaque compound processing because in the former case, it generates a meaning that is consistent with the retrieved meaning at the direct access, in the latter case, it generates a meaning that conflicts with the retrieved meaning at the direct access.

Second, neither of the previous two studies aimed to focus on the relationship of the two access routes and how this relationship influences the processing speed of compound words. In contrast, the current research aims to depict the relationship of the two access routes and how this relationship influences processing speed in light of the issue of semantic transparency. Furthermore, the current research also aims to explore the flexible nature of the relationship of the two access routes by accelerating the decomposition route. The findings point to a theoretical perspective that emphasizes the flexibility of the relationship of the two access routes. Some times, the interpretations activated at the two routes increase the activation of each other because they are compatible information. Other times, the interpretations activated at the two routes hinder the activation of each other because they conflict with each other. Such a theoretical position has never been reported in the literature before.

Predictions of the current studies

In the current studies, I compare the processing of English two-constituent compounds with monomorphemic words that are matched by whole word frequency and word length. In addition, I compare the processing of transparent compounds and opaque compounds respectively with the processing of matched monomorphemic words. Transparent compound processing should be quicker than monomorphemic word processing because the decomposition route activates consistent information as direct access throughout the process, and thus the two routes can boost the activation of each other. For example, during the processing of transparent compounds such as *rosebud*, the lexical representations of *rose* and *bud* will become activated and increase the activation of the whole word representation *rosebud* at this level through facilitative activation. Likewise, the semantic representations of *rose* and *bud* will also become activated and increase the activation of the whole word representation at this level through facilitative activation because the meaning of the whole is directly related with both parts (*rosebud* is the bud of a rose). Moreover, there might be facilitation from the conceptual level; the activated representations for the concepts *rose* and *bud* can be composed to compute a meaning *the bud of a rose* which is consistent with the retrieved legitimate meaning of the whole word. The computed meaning and the retrieved meaning should increase the activation of each other. The influence of morphological decomposition should be more complicated for opaque compounds because in opaque compound processing, the decomposition route activates consistent information with the direct access at the lexical level but inconsistent information across the semantic and conceptual level. For example, during the processing of opaque compounds such as *jailbird*, the lexical representations of *jail* and *bird* will become activated and increase the activation of the whole word representation *jailbird* at this level through facilitative activation. In contrast, although the semantic representations of *jail* and *bird* will also become activated automatically via their lexical activation, such activation cannot increase the activation of the whole word representation for *jailbird* at this level as much as in the case for *rosebud* because both constituents are not semantically related to the whole. Thus, when the decomposition route is not boosted (e.g., the fillers are composed of a real word and a nonword *rostpepper* [rost+pepper], so that the decomposition route would easily fade out after the lexical level because for this route to continue, there should be two activated lexical units at the same time), opaque compounds should be processed more quickly than monomorphemic words because the decomposition route can still boost the direct access up to the lexical level. In the meantime, the computed meaning at the decomposition route (e.g., interpret *jailbird* as *a bird that lives in jail*) is not activated strong enough, and thus is less likely to inhibit the legitimate meaning retrieved via the direct access (a prisoner). On the other hand, opaque compounds should not be processed more quickly than monomorphemic words when the decomposition route is boosted by using fillers composed of two real words *restpepper* (rest+pepper), by inserting spaces between the two constituents of compounds and by presenting the two constituents of compounds in contrasting colors. In this kind of situation, although the activation of *jail* and *bird* increases the activation of *jailbird* at the lexical level, the activation of these two constituents at the semantic level and then the conceptual level should compute a meaning (e.g., a bird that lives in jail) that inhibits the retrieved whole word meaning (a prisoner). Suppose that word recognition is eventually determined by the direct access, the inhibition from the computed meaning should slow down the process at the direct access and make word recognition more difficult. This disadvantage of computation at the conceptual level should trade off the advantage of decomposition at the lexical level and might result in no overall difference in the processing time of opaque compounds and monomorphemic words. Finally, the influence of semantic transparency is also expected in that transparent compounds should be processed more quickly than opaque compounds because the decomposition route facilitates transparent compound processing throughout the process whereas it mainly facilitates opaque compound processing at the lexical level and may cause inhibition between the computed meaning and the retrieved meaning across the semantic and conceptual levels.

In Experiments 1 and 2, I compare English transparent compounds with matched monomorphemic words. In Experiment 3, I further test the nature of the advantage for transparent compound processing by manipulating the frequency of the two constituents. Experiment 4 compares opaque compounds with matched transparent compounds and monomorphemic words. Experiments 5-7 further investigate the nature of the advantage/disadvantage of opaque compound processing by accelerating the decomposition route.

Chapter II

Experiment 1

In the current experiment, I examine whether responses to transparent compounds (e.g., *rosebud*) are faster than to matched monomorphemic words (e.g., *giraffe*). If transparent compounds are easier to recognize than monomorphemic words, then this finding would suggest that transparent compounds are not recognized via a single direct access. Rather, decomposition is involved in transparent compound processing and increases the ease of word recognition. If transparent compounds are more difficult to recognize than monomorphemic words, then this finding would suggest that transparent compounds are processed mainly via a decomposition route and the computational cost associated with decomposition slows down the recognition of transparent compounds. I expect an advantage in the processing of transparent compounds over monomorphemic words because the decomposition route and direct access activate consistent information throughout the process and should boost the activation of each other. In terms of the predictions of the three models discussed in the Introduction, both the dual-route model and the APPLE model predict an advantage of transparent compounds over monomorphemic words. In contrast, the early version of the automatic decomposition model predicts a disadvantage of transparent compounds relative to monomorphemic words, whereas the later version of this model predicts no difference in the processing of these two types of words.

Method

Materials. Sixty pairs of English transparent compounds and monomorphemic words were selected from the CELEX corpus (Baayen, Piepenbrock, & Gulikers, 1995), where frequencies are calculated based on 17.9 million word tokens. Each pair of items was matched on both lemma frequency and surface frequency.[2] Lemma frequency is the sum of the frequencies of all inflectional forms of a word. Surface frequency is the frequency of a particular word form. Each pair was also matched on word length (by number of letters and number of syllables). The mean surface frequencies (per million count, from now on all reported frequencies are per million counts) in the compound and monomorphemic word conditions were 1.59 and 1.58 respectively. The mean lemma frequencies in the two conditions were 1.92 and 1.91 respectively. The mean number of letters in both conditions was 7.85 and the mean number of syllables in both conditions was 2.30. Transparency ratings for compound words were collected from 10 judges after the experiment had been run. None of the judges participated in the experiment. The rating scale was a 7-point scale (1: totally opaque, 7: totally transparent). The mean transparency rating was 4.68.

A set of 120 nonword filler items was also constructed so that the number of *yes* and *no* responses in the experiment was balanced. Sixty nonwords used compound formats and were constructed by randomly pairing a real word and a nonword (e.g., *rostpepper* or *chivesonse*). Thirty of the nonword compounds had their first constituents as nonwords and 30 had their second constituents as nonwords. The remaining 60 nonwords used monomorphemic word formats and were constructed by changing one to

---

[2] Bigram frequencies for the two conditions in the current experiment and for conditions in the remaining experiments reported in this dissertation were checked using online database MCWord by Medler & Binder (2005). No difference was found.

two letters in real words. For example, *arithmutia* was constructed from *arithmetic*. No phonological or orthographic constraint was explicitly applied in constructing nonword fillers, though fillers were made as pronounceable as possible, and as orthographically legal as possible. None of the words used to construct filler items appeared in the experimental items.

The design was a within-subject design. The 120 experimental compound and monomorphemic items were presented in a randomized order along with the 120 filler items. Each participant saw a different random order of list.

Procedure. A lexical decision task was used. Participants sat in front of a computer screen and placed the index finger of their left hands on the F key (labeled *no*) of the keyboard and the index finger of their right hands on the J key (labeled *yes*). Trial presentation was self-paced; each trial began with the message "Ready?" on the computer screen and participants initiated the trial by pressing the space bar. The stimulus item then appeared and participants indicated whether the item was a word by pressing the appropriate key. There were eight practice trials before the start of the experiment.

Participants. Twenty-five undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from one participant were excluded due to low accuracy (less than 60%) and the data from another participant were excluded due to overall slow responses (mean response time more than 1500 ms). Thus, the data from 23 participants were used in the analyses.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition means were regarded as outliers and removed. In addition, six trials with response times more than 5000 ms were deleted and one trial was deleted due to a computer error. In total, 3.1% of correct responses were removed. Mean reaction times for correct responses, and accuracy rates of both conditions for the experimental items are listed in Table 1. The accuracy rates for the compound nonwords and monomorphemic nonwords were 93% and 88% respectively. Mean reaction times for correct responses, and accuracy rates of filler items are listed in Table 2. The two types of words were compared to investigate the effect of complex structure on the time to make a lexical decision by fitting linear mixed-effect (multi-level) regression models using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects (for details about this procedure see Baayen, 2008; Baayen, 2004; Bates, 2005; Dixon, in press; Pinheiro & Bates, 2000). The model was fitted such that adjustments to the intercept were made on the basis of both subjects and items. Analyses of accuracy hereafter were conducted by fitting models using quasibinomial distribution. This distribution allows for skewness and is more appropriate for human response times. The advantage of fitting the linear mixed-effect model is that this model takes subject and item as random factors simultaneously, reduces the error term, and consequently increases statistical power to detect true effects.[3]

---

[3] Originally, separate ANOVAs using subject and item as random variables were conducted for all the analyses in this dissertation. The results were consistent with the results reported here. More specifically, analyses using linear mixed-effect model were significant only when both subject and item analyses were significant in ANOVA, but not significant if only one of the two analyses was significant.

Transparent compounds (*M* = 6.57, raw RT = 742 ms) were responded to more quickly than were monomorphemic words (*M* = 6.75, raw RT = 908 ms), *t*(2269) = 8.94, *SE* = .02, *p* < .0001. The accuracy data also showed a difference, *t*(2751) = 8.20, *SE* = .32, *p* < .0001.

In addition, correlation analyses revealed association of whole word frequencies of compound words and log response times; both log lemma frequency (r = -.08, *p* < .005), and log surface frequency (r = -.07, *p* < .01) were correlated with log response time. These correlations were calculated over the full data set using linear mixed-effect regression model. The evidence of the association of whole word frequency of compound words and response times suggests that direct access is involved in compound word processing.

One might argue that the advantage for transparent compound processing is restricted to low frequency words. If so, there should be an interaction between whole word frequency and word type. However, for log response times, neither log lemma frequency of the whole word, *F* (1, 2267) = 2.25, *MSE* = .06, *p* > .13, nor the log surface frequency of the whole word, *F* (1, 2267) = 2.96, *MSE* = .06, *p* > .09, interact with word type. Similarly, for accuracy, neither log lemma frequency of the whole word, *F* < 1, nor the log surface frequency of the whole word, *F* (1, 2749) = 1.06, *MSE* = .66, *p* > .30, interact with word type.

I also re-analyzed the data without 13 pairs of items because the monomorphemic words of these pairs had low accuracy (less than 50%). This was to rule out the alternative interpretation of the findings being due to these low accuracy items. Mean reaction times for correct responses, and accuracy rates for the experimental items of the re-analysis are listed in Table 3.

This re-analysis yielded consistent results with the original analysis, transparent compounds ($M = 6.58$, raw RT = 747 ms) were responded to more quickly than were monomorphemic words ($M = 6.73$, raw RT = 880 ms), $t(1892) = 6.57$, $SE = .02$, $p < .0001$. The accuracy data also showed a difference, $t(2156) = 5.22$, $SE = .28$, $p < .0001$. The results of the re-analysis demonstrated that the advantage of transparent compounds was not due to the low accuracy items in the data.

The findings of the current experiment suggest that morphological decomposition is involved in transparent compound processing because, if not, the processing speed of transparent compounds would have not differed from that of frequency-matched monomorphemic words. More importantly, the advantage for transparent compound processing suggests that the availability of morphological decomposition does not necessarily slow down word recognition. Rather, it might accelerate word processing via facilitative links between the constituents and the whole throughout the process.

Table 1

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 1

Condition	LogRT(SE)	RT (SE)	Accuracy (SE)
Compound	6.57 (.02)	742 (19)	93 (1)
Monomorpheme	6.75 (.03)	908 (32)	72 (2)

Note. Descriptive statistics reported in all Tables hereafter were calculated by averaging data over subjects, standard errors of means were calculated for each condition.

Table 2

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 1

Condition	RT (SE)	Accuracy (SE)
Nonword Compound Left	972 (46)	93 (1)
Nonword Compound Right	967 (46)	92 (2)
Nonword Monomorpheme	993 (45)	88 (2)

Note. Nonword Compound Left = nonword compounds whose first constituents were nonwords and second constituents were real words (e.g., *rostpepper*); Nonword Compound Right = nonword compounds whose second constituents were nonwords and first constituents were real words (e.g., *pepperrost*).

Table 3

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for re-analysis of Experiment 1

Condition	LogRT (SE)	RT (SE)	Accuracy (SE)
Compound	6.58 (.02)	747 (18)	93 (1)
Monomorpheme	6.73 (.03)	880 (30)	83 (2)

Chapter III

Experiment 2

An alternative interpretation for the results of Experiment 1 is based on the participants engaging in strategic processing. Because the nonword compound fillers were constructed by pairing a nonword and a real word, recognition of individual constituents might be sufficient for participants to distinguish real compounds from nonword compounds. Thus, constituents of compounds might have been accessed separately without being combined into one whole word. In other words, the compound might never have been accessed. This would have led to faster response times. The current experiment uses a different set of filler items, which eliminates the use of this strategy.

Method

Materials. The materials were the same as used in Experiment 1, except that the 60 nonword compounds were constructed by randomly pairing two real words (e.g., *restpepper* [rest+pepper]).

Procedure. The procedure was identical to the procedure used in Experiment 1.

Participants. Twenty-five undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition means were regarded as outliers and removed. In addition, three trials with response times more than 4000 ms were deleted. In total, 3.2% of correct responses were removed. Mean reaction times for correct responses, and accuracy rates of both conditions for the experimental items are listed in Table 4. The accuracy rates for the compound nonwords and monomorphemic nonwords were 90% and 85% respectively. Mean reaction times for correct responses, and accuracy rates for filler items are listed in Table 5. The two types of words were compared to investigate the effect of complex structure on the time to make a lexical decision by fitting linear mixed-effect (multi-level) regression models using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects.

As for Experiment 1, compounds ($M = 6.65$, raw RT = 805 ms) were responded to more quickly than were monomorphemic words ($M = 6.76$, raw RT = 902 ms), $t(2449) = 5.25$, $SE = .02$, $p < .0001$. The accuracy data also showed a difference, $t(2988) = 4.33$, $SE = .32$, $p < .0001$.

In addition, both log lemma frequency ($r = -.07$, $p < .008$) and log surface frequency ($r = -.09$, $p < .001$) were correlated with log response time to compound words. These correlations were calculated over the full data set using linear mixed-effect regression model. The evidence of the influence of whole word frequency on compound word processing suggests that direct access route is involved in compound word processing.

Furthermore, for log response time, neither log lemma frequency of the whole word, $F(1, 2447) = 1.02$, $MSE = .06$, $p > .31$, nor the log surface frequency of the whole word, $F(1, 2447) = 2.12$, $MSE = .06$, $p > .15$, interacts with word type. Similarly, for accuracy, neither log lemma frequency of the whole word, $F < 1$, nor the log surface frequency of the whole word, $F < 1$, interacts with word type. This finding suggests that the advantage of transparent compound processing is not restricted to low frequency words.

I also re-analyzed the data without 12 pairs of items. Among the deleted 12 pairs, 10 were deleted due to the low accuracy of the monomorphemic words (less than 50%) and two were deleted due to the low accuracy of the compound words (less than 50%). This re-analysis was to rule out the alternative interpretation that the advantage for transparent compound processing was due to these low accuracy items. Mean reaction times for correct responses, and accuracy rates for experimental items of the re-analysis are listed in Table 6.

Transparent compounds ($M = 6.64$, raw RT = 794 ms) were responded to more quickly than were monomorphemic words ($M = 6.74$, raw RT = 885 ms), $t(2101) = 4.95$, $SE = .02$, $p < .0001$. The accuracy data also showed a difference, $t(2392) = 3.65$, $SE = .27$, $p < .0003$. The results of the re-analysis demonstrated that the advantage for transparent compounds was not due to the low accuracy items in the data.

In sum, the findings again demonstrate that transparent compounds are processed more quickly than frequency-matched monomorphemic words regardless of the type of fillers being used, and suggest that the advantage for transparent compound processing observed in Experiment 1 was not due to strategic processing.

The advantage of transparent compound processing over frequency-matched monomorphemic word processing suggests that morphological decomposition is involved in transparent compound processing. The results of the correlation analyses suggest that the direct access is also involved in transparent compound processing because higher whole word frequencies were associated with faster response times. The decomposition route and the direct access, however, do not necessarily inhibit each other and slow down compound processing as would have been suggested by Laine and colleagues (1999). Rather, the availability of two access routes increases the ease for the processing of transparent compounds.

One possible mechanism that can explain the advantage of transparent compound processing is that transparent compounds can benefit from morphological decomposition because constituents that are activated via this access route are connected to the whole word representations that are activated via the direct access at the lexical level. Consequently, lexical units that are activated at the two access routes boost the activation of each other and increase the ease of transparent compound processing relative to monomorphemic word processing. Such an explanation is equivalent to the framework of the APPLE model by Libben (1998) where facilitative links are assumed to connect compounds and their constituents at the lexical level.

Alternatively, the advantage of transparent compound processing is due to statistical facilitation assumed by the dual-route race model (Baayen et al., 1997). Presumably, both the decomposition route and the direct access are involved in transparent compound processing. Given that word recognition times are always determined by the faster route, overall, transparent compounds should be processed more quickly than monomorphemic words because word recognition times of the latter are only determined by a single direct access.

The advantage of transparent compound processing is not compatible with the predictions of the automatic decomposition model (Taft & Forster, 1975; 1976; Taft, 1994). The early version of this model (Taft & Forster, 1975; 1976) predicts a disadvantage of transparent compound processing based on the assumptions of serial search and computational cost in transparent compound processing. The latter version of this model (Taft, 1994) predicts no advantage of transparent compound processing because the benefit gained at the lexical level with the involvement of both access routes should be traded off by the computational cost at the semantic level with the involvement of only the decomposition route.

Table 4

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 2

Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Compound	6.65 (.02)	805 (19)	88 (2)
Monomorpheme	6.76 (.03)	902 (25)	75 (1)

Table 5

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 2

Condition	RT (SE)	Accuracy (SE)
Nonword Compound	1074 (41)	90 (2)
Nonword Monomorpheme	1114 (58)	85 (2)

Table 6

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for re-analysis of Experiment 2

Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Compound	6.64 (.02)	794 (20)	91 (2)
Monomorpheme	6.74 (.02)	885 (22)	84 (2)

Chapter IV

Experiment 3

The findings of Experiments 1 and 2 suggest that the availability of the decomposition route increases the ease of transparent compound processing. One possible explanation for the effects is that transparent compounds and their constituents are connected by facilitative links at the lexical level. Consequently, lexical units that are activated at the two access routes can boost the activation of each other and speed up responses. Based on such an explanation, one should further predict that factors that influence the ease of lexical access, such as word frequency, should influence the ease of the activation of compound constituents, and consequently the speed of transparent compound processing. In specific, one should predict that transparent compounds with high-frequency constituents should be responded to more quickly than transparent compounds with low-frequency constituents.

In the current experiment, I manipulate the frequency of the first and second constituents to examine whether they influence the ease of transparent compound processing. Although previous studies have investigated the influence of constituent frequency on the ease of compound processing (e.g., Andrews et al. 2004; Hyönä & Pollatsek, 1998; Juhasz et al., 2003; Pollatsek et al., 2000), the purpose of those studies was to investigate whether morphological decomposition is involved in compound processing, which was different from the purpose of the current experiment that aims to provide evidence that lexical access to compound constituents contributes to the advantage for transparent compound processing. That is, I include monomorphemic words as a control condition to examine the advantage for transparent compounds with different constituent frequencies. Thus, the current experiment is also a further replication of Experiments 1 and 2, and is necessary to serve the purpose of the current project.

Another reason to run the current experiment is that previous studies have yielded different results in terms of the influence of constituent frequency in compound processing. In particular, two previous studies about English compounds had manipulated constituent frequencies of compounds as in the current experiment. Andrews et al. (2004) measured participants' eye movements when they read compounds that were embedded in sentences. In their first experiment they manipulated the frequencies of both constituents while controlled for whole word frequencies. They found that both constituent frequencies influenced the time spent on compound processing. Juhasz et al. (2003) also manipulated the frequencies of both compound constituents and measured participants' eye movements in a sentence reading task (Experiment 3). In addition, they used a lexical decision task (Experiment 1) and a naming task on the same set of stimuli (Experiment 2). Overall, their studies revealed more robust influence of the frequency of the second constituent than that of the first constituent; for item analyses, the effect of the first constituent did not reach statistical significance in any of the three experiments.

My explanation for the advantage of transparent compound processing predicts that the frequency of both constituents should influence the speed of compound word processing. Given that the previous two studies yielded different results, it is necessary to conduct the current experiment to investigate the issue further.

Method

Materials. There were five sets of stimuli in this experiment, four for semantically transparent compounds and one for monomorphemic words. In total, 120 transparent compounds and 30 monomorphemic words were selected from the CELEX corpus (Baayen et al., 1995). The compounds were further divided into four sets by two factors, frequency of the first constituent (high vs. low) and frequency of the second constituent (high vs. low). Thus, there was one set for compounds with both high-frequency first constituents and high-frequency second constituents (HH); one set for compounds with high-frequency first constituents and low-frequency second constituents (HL); one set for compounds with low-frequency first constituents and high-frequency second constituents (LH) and one set for compounds with both low-frequency first constituents and low-frequency second constituents (LL). All five sets of stimuli were matched on whole word frequency (lemma frequency, surface frequency) and word length (by number of letters and number of syllables) in the group means. The mean surface frequencies in the HH, HL, LH, LL and monomorphemic word sets were .56, .60, .50, .55 and .62 respectively. The mean lemma frequencies in the five sets were .73, .81, .72, .72 and .69 respectively. The mean numbers of letters in the five sets were 8.60, 8.63, 8.50, 8.77 and 8.30 respectively. The mean numbers of syllables in the five sets were 2.27, 2.20, 2.33, 2.43 and 2.40 respectively.

Transparency ratings for compound words were collected from 26 judges after the experiment had been run. None of the judges participated in the experiment. The rating scale was a 7-point scale (1: totally opaque, 7: totally transparent). In order, mean transparency ratings for the HH, HL, LH, and LL word sets were, 4.92, 5.23, 4.90 and 4.44 respectively.

A set of 150 nonword filler items was also constructed so that the number of *yes* and *no* responses was balanced in the experiment. A hundred and twenty nonwords used compound formats and were constructed by randomly pairing a real word and a nonword (e.g., *rostpepper*). Half of the compound nonwords had their first constituents as nonwords and the other half had their second constituents as nonwords. The remaining 30 nonwords used monomorphemic word formats and were constructed by changing one to two letters in real words.

The design was a within-subject design. The 150 experimental compound and monomorphemic items were presented in a randomized order along with the 150 filler items. Each participant saw a different random order of list.

Procedure. The procedure was identical to the procedure used in Experiment 2.

Participants. Forty-three undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from fifteen participants were excluded due to low accuracy rates (less than 60%) and the data from one participant were excluded due to slow responses (mean response times more than 1700 ms). Thus, the data from 27 participants were used in the analyses.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition mean were regarded as outliers and trimmed. In addition, two trials were deleted due to a computer error. In total, 3% of correct responses were trimmed. Reaction times and standard errors for correct responses and accuracy rates for experimental items are listed in Table 7. The accuracy rates for the compound nonwords and monomorphemic nonwords were 93% and 84% respectively. Reaction times and standard errors for correct responses and accuracy rates for filler items are listed in Table 8. Linear mixed-effect (multi-level) regression models were fitted to investigate the effect of complex structure and constituent frequency of compounds on the time to make a lexical decision. I fitted models using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects.

Evidence of advantage of transparent compounds. I conducted one planned comparison to evaluate the influence of complex structure on lexical decision time. For this comparison, the log response time of the four compound sets together was compared with the log response time in the monomorpheme condition. As has been found in Experiments 1 and 2, responses to compounds ($M = 6.61$, raw RT = 778 ms) were faster than responses to monomorphemic words ($M = 6.78$, raw RT = 915 ms), $t(3337) = 8.29$, $SE = .02, p < .0001$. The accuracy analysis mirrored the result with response time, $t(4048) = 6.03$, $SE = .26, p < .0001$.

Evidence of effect of the first constituent. I conducted a 2 (constituent position: first constituent or second constituent) by 2 (word frequency of constituents: high frequency or low frequency) factorial analysis to evaluate the effect of constituent frequencies on log response times. Log reaction times to compounds with high-frequency first constituents were faster than to compounds with low-frequency first constituents, $F\,(1, 2778) = 4.70$, $MSE = .06, p < .03$. The accuracy analysis mirror the result with response time, $F\,(1, 3236) = 7.73$, $MSE = .67, p < .005$.

In contrast, no main effect of the frequency of the second constituent was found for reaction times, $F < 1$or for accuracy, $F < 1$. In addition, no interaction of the frequencies of the two constituents was found for reaction times, $F < 1$or for accuracy, $F < 1$.

Correlation analyses. Both log lemma frequency ($r = -.10$, $p < .0001$) and log surface frequency ($r = -.08$, $p < .0001$) were correlated with log response time to compound words. These correlations were calculated over the full data set using linear mixed-effect regression model. The evidence of the influence of whole word frequency on compound word processing suggests that direct access is involved in compound word processing.

To summarize, the results again demonstrate that transparent compounds are processed more quickly than monomorphemic words. In addition, the finding that the frequency of the first constituent influenced response times of transparent compound processing provides further evidence that lexical access to compound constituents contributes to the advantage for transparent compound processing. However, I did not find evidence for the influence of the frequency of the second constituent on response times.

A close observation of the transparency ratings in the four compound conditions revealed that compounds with high-frequency first constituents were rated more transparent than compounds with low-frequency first constituents, $F(1, 3116) = 3.77$, $MSE = 3.11$, $p < .05$. No such difference was found between compounds with high-frequency second constituents and compounds with low-frequency second constituents, $F < 1$. It is possible then that the finding of the influence of the first constituent was due to the confounding of transparency ratings.

Thus, I conducted a follow-up experiment for the current experiment to better control the transparency ratings. A subset of 15 words from each of the five conditions was selected to better balance the semantic transparency ratings for the two compound factors and match the conditions on other dimensions as in the current experiment. The design of the follow-up experiment was identical to the current experiment except that 45 monomorphemic words were added in so that the total number of compounds was equal to the total number of monomorphemic words. The filler items were different from those in the current experiment in that fillers that took the format of compounds were constructed by randomly pairing two real words (e.g., *restpepper*). The mean transparency ratings in the HH, HL, LH, and LL compound word sets were, 5.30, 5.64, 5.66 and 5.39 respectively. The mean surface frequencies in the HH, HL, LH, and LL compound word sets were .60, .67, .48 and .55 respectively. The mean lemma frequencies in the four sets were .82, .96, .72 and .73 respectively. The mean numbers of letters in the four sets were 8.53, 8.40, 8.60 and 8.60 respectively. The mean numbers of syllables in the five sets were 2.20, 2.20, 2.33 and 2.47 respectively. In order, for the overall match of compound words and monomorphemic words, the mean surface frequencies in the two sets were, .58 and .69 respectively. The mean lemma frequencies in the two sets were both .81. The mean numbers of letters in the two sets were 8.53 and 8.17 respectively. The mean numbers of syllables in the two sets were not matched, and the numbers were 2.30 and 2.57 respectively.

Participants. Thirty-two undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from nine participants were excluded due to low accuracy rates (less than 60%). Thus, the data from 23 participants were used in the analyses.

The results showed a similar pattern to the current experiment. Reaction times and standard errors for correct responses and accuracy rates for experimental items of the follow-up experiment are listed in Table 9. For log response times, there was evidence of advantage of compounds over monomorphemic words; responses to compounds (*M* = 6.63, raw RT = 792 ms) were faster than responses to monomorphemic words (*M* = 6.70, raw RT = 851 ms), *t*(2279) = 16.83, *SE* = .02, *p* < .0001. The accuracy rates mirrored the pattern of response times, *t*(2752) = 5.32, *SE* = .24, *p* < .0001.

In terms of the influence of constituent frequencies, for log response times, there was neither an effect of the first constituent, *F* (1, 1222) = 1.45, *MSE* = .06, *p* > .23, nor an effect of the second constituent, *F* < 1. Besides, there was no interaction of the frequencies of the two constituents, *F* < 1. The non-effect of the first constituent in log response time might not be surprising because the statistical power of this follow-up experiment was not as strong as the original experiment due to the reduced number of items included in this experiment.

The accuracy analyses, however, showed a main effect of the first constituent where responses to high-frequency first constituents were more accurate than responses to low-frequency first constituents, *F* (1, 1374) = 7.46, *MSE* = .60, *p* < .01. In contrast, no main effect of the frequency of the second constituent was found, *F* < 1. Also, no interaction of the frequencies of the two constituents was found, *F* < 1.

In addition to the follow-up experiment, I also analyzed the data of 13 participants with low accuracy from the data of the16 participants that have been excluded from the original analysis. The data of three participants were not used because two of them had extremely low accuracy (less than 10%) probably due to the switch of the "yes" and "no" response keys and one of them had slow responses (mean response times greater than 1700 ms). Reaction times and standard errors for correct responses and accuracy rates for the 13 participants with low accuracy are listed in Table 10.

I conducted one planned comparison to evaluate the influence of complex structure on lexical decision time. Responses to compounds ($M = 6.58$, raw RT = 750 ms) were faster than responses to monomorphemic words ($M = 6.75$, raw RT = 899 ms), $t(1446) = 6.77$, $SE = .03$, $p < .0001$. The analysis of accuracy mirrored the result of response times, $t(1948) = 9.26$, $SE = .28$, $p < .0001$.

The effect of constituent frequencies on log response times was also evaluated. Responses to compounds with high-frequency first constituents were faster than to compounds with low-frequency first constituents, $F(1, 1275) = 6.80$, $MSE = .05$, $p < .009$. The analysis of accuracy mirrored the result of log response time, $F(1, 1556) = 5.32$, $MSE = .67$, $p < .02$.

In contrast, no main effect of the frequency of the second constituent was found for log response times, $F < 1$ or for accuracy, $F < 1$. In addition, no interaction of the frequencies of the two constituents was found for log response times, $F < 1$ or for accuracy, $F(1, 1556) = 2.17$, $MSE = .67$, $p > .12$.

The results of the analysis of the data of the 13 participants with low accuracy rates again showed a similar pattern to the original analysis that had excluded these participants. This finding indicated that the data pattern revealed by the original analysis was not due to exclusion of participants with low accuracy.

In sum, the current experiment replicated the advantage of transparent compound processing in Experiments 1 and 2. The results of the current experiment also demonstrate that constituent frequencies of transparent compounds influence the ease of transparent compound processing in that the higher the frequency of the first constituent, the faster the responses. Such a finding lends supports to the proposal that lexical activation of the constituents increases the ease of transparent compound processing. However, the current experiment and its follow-up experiment failed to find the influence of the second constituent which should be expected.

The reason that I failed to find an influence of the frequency of the second constituent might be due to the stimuli properties of the current experiment. As suggested by Juhasz and colleagues (2003), who found stronger evidence for the influence of the second constituent than for the influence of the first constituent in their Experiment 1, the degree of lexicalization of compound stimuli might be the reason that some experiments found stronger evidence for the influence of the first constituent, whereas some other experiments found stronger evidence for the influence of the second constituent.
In Juhasz and colleagues' Experiment 1, an experiment with similar task and design as the current experiment, they explicitly selected compounds with frequencies of at least 1 per million from the Kucera and Francis (1967) corpus when they prepared their stimuli. In contrast, the mean surface frequency of all compounds in the current experiment was 0.6 per million.[4] Thus; stimuli in the current experiment appeared to be more novel than stimuli in Juhasz and colleges' Experiment 1 in terms of frequency count. Juhasz and colleagues suggested that the effect of the first constituent was more robust when novel compounds were used as stimuli. In contrast, the effect of the second constituent was more robust when lexicalized compounds were used as stimuli. For example, van

---

[4] The mean surface frequencies of all compound conditions in all of the remaining Experiments were more than 1 per million.

Jaarsveld and Rattink (1988) found greater effect of the first constituent in three experiments, where novel Dutch compounds were used as stimuli. However, when both novel compounds and lexicalized compounds were used as stimuli in their Experiment 4, novel compounds yielded an effect of the first constituent but not the second constituent, and lexicalized compounds yielded an effect of the second constituent but not the first constituent. Thus, the relatively high degree of novelty of stimuli in terms of frequency count in the current experiment might be the reason that I only found the influence of the first constituent.

Table 7

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 3

Word Type	Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Compound	HH	6.60 (.02)	761 (19)	89 (2)
	HL	6.60 (.02)	764 (20)	89 (2)
	LH	6.64 (.03)	800 (27)	81 (3)
	LL	6.63 (.03)	788 (23)	85 (2)
Monomorpheme		6.78 (.03)	915 (25)	69 (2)

Table 8

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 3

Condition	RT (SE)	Accuracy (SE)
Nonword Compound Left	872 (35)	95 (1)
Nonword Compound Right	898 (36)	92 (2)
Nonword Monomorpheme	922 (34)	84 (3)

Note. Nonword Compound Left = nonword compounds whose first constituents were nonwords and second constituents were real words (e.g., *rostpepper*); Nonword Compound Right = nonword compounds whose second constituents were nonwords and first constituents were real words (e.g., *pepperrost*).

Table 9

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for the follow-up Experiment

Word Type	Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Compound	HH	6.62 (.04)	791 (32)	91 (2)
	HL	6.60 (.03)	767 (23)	92 (2)
	LH	6.65 (.03)	812 (25)	87 (2)
	LL	6.63 (.03)	796 (28)	86 (2)
Monomorpheme		6.70 (.03)	851 (26)	76 (2)

Table 10

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for 13 participants with low accuracy of Experiment 3

Word Type	Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Compound	HH	6.55 (.04)	729 (29)	88 (3)
	HL	6.55 (.04)	728 (29)	83 (2)
	LH	6.60 (.05)	774 (42)	75 (5)
	LL	6.60 (.04)	767 (36)	82 (3)
Monomorpheme		6.75 (.04)	899 (36)	43 (2)

Chapter V

Experiment 4

The results of Experiment 3 demonstrate that constituent frequency, a factor that influences the ease of lexical access, influences the speed of transparent compound processing. This finding supports my proposal that lexical access to compound constituents facilitates the processing of transparent compounds. However, the results of Experiments 1- 3 do not rule out the possible influence of other factors such as semantic factors in compound processing. Given that I used semantically transparent compounds as stimuli, the results of Experiments 1-3 could be either due to pure lexical facilitation from compound constituents or due to a combination of lexical facilitation and semantic facilitation from compound constituents. In particular, because the meanings of the constituents and the whole of transparent compounds are directly related, there should be facilitative activation between the parts and the whole at the semantic level, which would speed response times.

To investigate the influence of semantic transparency on the speed of compound processing, I compare the processing of opaque compounds with that of transparent compounds and monomorphemic words. If the advantage for transparent compound processing is solely due to lexical access to compound constituents as suggested by the APPLE model (Libben, 1998), one would expect that opaque compound processing exhibits a similar advantage as transparent compound processing assuming that the constituents of opaque compounds can be accessed at the lexical level. In contrast, if other factors, such as semantic transparency, also contribute to the advantage for transparent compound processing, opaque compound processing should still exhibit an advantage over monomorphemic word processing, but with a lesser degree than the advantage for transparent compound processing. Of course, if the constituents of opaque compounds cannot be accessed at the lexical level or no decomposition occurs in opaque compound processing at all as suggested by the dual-route race model (Schreuder & Baayen, 1995), opaque compounds should be processed in a similar speed as matched monomorphemic words.

Method

Materials. Thirty triples of items were selected from the CELEX database (Baayen et al., 1995). Ratings of semantic transparency of the stimuli were collected from nine individuals after the experiment was completed. None of the judges participated in the experiment. The rating scale was a 7-point scale (1: totally opaque, 7: totally transparent). The mean rating for the transparent compound condition was 5.97 and the mean rating for the opaque compound condition was 3.73. The ratings for the two groups were different, t (538) = 8.66, $p < .0001$. In addition, each triple was matched on whole word frequency (lemma frequency, surface frequency) and word length (number of letters and number of syllables). The mean surface frequencies in the transparent, opaque, and monomorphemic word conditions were 1.04, 1.03 and 1.06 respectively. The mean lemma frequencies in the three conditions were 1.19, 1.18 and 1.20 respectively. The mean number of letters in all three conditions was 7.90. The mean numbers of syllables in the three conditions were 2.20, 2.33 and 2.33 respectively.

A set of 90 nonword filler items was also constructed so that the number of *yes* and *no* responses was balanced in the experiment. Sixty nonwords used compound formats and were constructed by randomly pairing a real word and a nonword (e.g., *rostpepper*). Half of the compound nonwords had their first constituents as nonwords and the other half had their second constituents as nonwords. The remaining 30 nonwords used monomorphemic word formats and were constructed by changing one or two letters in real words.

The design was a within-subject design. The 90 experimental compound and monomorphemic items were presented in a randomized order along with the 90 filler items. Each participant saw a different random order of list.

Procedure. The procedure was identical to the procedure used in Experiment 3.

Participants. Thirty-one undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from four participants were excluded because, for two participants, their accuracy rates were low (less than 60%), for two participants, their responses were slow with their average response times both greater than 2 seconds. Thus, the data from 27 participants were used in the analyses.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition mean were trimmed. In addition, 14 trials with reaction times more than 3000 ms were removed. In total, 3.7% of correct responses were trimmed as outliers. Reaction times and standard errors for correct responses, and accuracy rates are listed in Table 11. The accuracy rates for the compound nonwords and monomorphemic nonwords were 94% and 87% respectively. Reaction times and standard errors for correct responses and accuracy rates for filler items are listed in Table 12. Linear mixed-effect (multi-level) regression models were fitted to investigate the effect of complex structure and semantic transparency on the time to make a lexical decision. Models were fitted by using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects.

I conducted two planned comparisons to evaluate the advantage for compound processing. There was evidence for an advantage in compound processing regardless of semantic transparency, transparent compounds ($M = 6.61$, raw RT = 774 ms) were responded to more quickly than were monomorphemic words ($M = 6.71$, raw RT = 869 ms), $t(2057) = 3.30$, $SE = .03$, $p < .001$, 95% confidence interval (-.17, -.04). Opaque compounds ($M = 6.66$, raw RT = 812 ms) were responded to more quickly than were monomorphemic words, $t(2057) = 1.94$, $SE = .03$, $p > .05$, 95% confidence interval (-.004, .13).

For the most part, the accuracy analyses showed a similar pattern of results. There was a difference for the comparison between transparent compounds and monomorphemic words where responses to transparent compounds were more accurate than responses to monomorphemic words, $t(2395) = 2.53$, $SE = .46$, $p < .01$. For the comparison between opaque compounds and monomorphemic words, no difference was found, $t(2395) = .37$, $SE = .44$, $p > .71$.

There was no evidence for the effect of semantic transparency in terms of log response time,[5] $t(2057) = 1.36$, $SE = .03$, $p > .17$, 95% confidence interval (-.11, .02).

The accuracy analysis, however, showed an effect of semantic transparency where

---

[5] There was no interaction between semantic transparency and whole word frequency in the current experiment and Experiments 5-7 to follow.

responses to transparent compounds were more accurate than responses to opaque compounds, $t(2395) = 2.18$, $SE = .46$, $p < .03$.

In addition, correlation analyses revealed association of whole word frequencies and log response times. Collapsing transparent and opaque compounds, both log lemma frequency ($r = 0.10$, $p < .0001$) and log surface frequency ($r = -.08$, $p < .003$) of whole word were correlated with log response time. This finding suggests that direct access is involved in compound word processing. Furthermore, log lemma frequency ($r = -.12$, $p < .0001$) of the first constituent of compound words was correlated with log response time. This negative correlation indicates that the higher the frequency of the first constituent, the faster the response. All correlations were calculated over the full data set using linear mixed-effect regression model.

In sum, the current experiment demonstrates that both transparent compounds and opaque compounds are easier to process than frequency-matched monomorphemic words. This finding suggests that decomposition not only occurs in the processing of transparent compounds but also occurs to a certain extent in the processing of opaque compounds. The effect of semantic transparency in the accuracy data indicates that the advantage for transparent compound processing is greater than the advantage for opaque compound processing. This implies that the advantage for transparent compound processing is not solely due to lexical access to compound constituents and that semantic transparency is also part of the story. A possible explanation for this finding is that the facilitative activation between compounds and their constituents at the semantic level benefits the recognition of transparent compounds more than opaque compounds because the meanings of the constituents are more directly related with their wholes in the former than in the latter. Specifically, in transparent compound processing, one can assume facilitative activation between the parts and the whole at the semantic level. For example, the facilitative activation from *rose* and *bud* at the semantic level might accelerate the activation of *rosebud* at the semantic level because both constituents are semantically related with the whole. In contrast, opaque compounds cannot benefit from facilitative activation between the parts and the whole at the semantic level as much as transparent compounds because both constituents are not semantically related to the whole (e.g., *hog*, *wash* versus *hogwash*, or *jail*, *bird* versus *jailbird*). As a result, the differential facilitative activation at the semantic level leads to greater advantage in processing transparent compounds than that of opaque compounds.

Alternatively, one can argue that if both transparent compounds and opaque compounds are decomposed at the lexical level, one should expect that frequencies of the two constituents influence the ease of processing, or at least the frequency of the first constituent influences the ease of processing as was found in Experiment 3. Thus, the advantage in transparent compound processing over opaque compound processing might be due to the possibility that the transparent compound condition had higher constituent frequency than the opaque compound condition in the current experiment. Paired-samples t-tests were conducted to compare the constituent frequency of the two compound conditions. No difference was found. For lemma frequency of the first constituent, $t(29) = 1.17, p > .25$; for surface frequency of the first constituent, $t(29) = 1.16, p > .26$. For lemma frequency of the second constituent, $t(29) = .02, p > .98$; for surface frequency of the second constituent, $t(29) = .11, p > .91$. Thus, the finding that transparent compounds were processed more quickly than opaque compounds cannot be interpreted as due to the difference of constituent frequencies in the two compound conditions.

In terms of the predictions of the three models that have been discussed in the Introduction, none of them is fully compatible with the findings of the current experiment. According to the early version of the automatic decomposition model (Taft & Forster, 1975; 1976), it should be more difficult to process compound words than monomorphemic words regardless of semantic transparency. This prediction is neither consistent with the advantage of compound words in general, nor is it consistent with the finding of semantic transparency in the accuracy analysis. The later version of the automatic decomposition model (Taft, 1994) predicts that it should be easier to process opaque compounds than transparent compounds, and there should be no difference in processing transparent compounds and monomorphemic words. This prediction is not consistent with the advantage of transparent compounds in the accuracy data and it is not consistent with the advantage of transparent compounds over monomorphemic words either. In contrast with the automatic decomposition model, the dual-route race model's (Baayen et al., 1997; Schreuder & Baayen, 1995) prediction of the advantage of transparent compounds over monomorphemic words is consistent with the findings of the current experiment and previous experiments on this matter. This model, however, cannot accommodate the advantage of opaque compounds over monomorphemic words found in the current experiment because it predicts no such difference. As for the APPLE model (Libben, 1998), it predicts the advantage of both types of compounds over monomorphemic words, which is consistent with the findings of the current experiment. This model, however, cannot accommodate the advantage of transparent compounds over opaque compounds in the accuracy data because it predicts no such difference.

Table 11

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 4

Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Transparent Compound	6.61 (.02)	774 (18)	90 (2)
Opaque Compound	6.66 (.02)	812 (21)	85 (2)
Monomorpheme	6.71 (.03)	869 (27)	82 (1)

Table 12

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 4

Condition	RT (SE)	Accuracy (SE)
Nonword Compound Left	989 (55)	93 (2)
Nonword Compound Right	991 (52)	94 (2)
Nonword Monomorpheme	1071 (59)	87 (2)

Note. Nonword Compound Left = nonword compounds whose first constituents were nonwords and second constituents were real words (e.g., *rostpepper*); Nonword Compound Right = nonword compounds whose second constituents were nonwords and first constituents were real words (e.g., *pepperrost*).

Chapter VI

Experiment 5

The results of Experiment 4 suggest that the availability of the decomposition route not only speeds up transparent compound processing but also opaque compound processing. Based on the evidence, I have proposed that opaque compound processing can benefit from the mutual facilitation of the two access routes and the locus of the benefit is mainly at the lexical level. The purpose of the current experiment is to investigate whether the advantage for opaque compound processing can be removed through experimental manipulation. If this advantage can be removed, that means that the facilitative mechanism is not *a priori* or a fixed setting of the two access routes. Rather, it might depend on properties of complex word stimuli and the particular context in which compound words are processed.

The removal of the advantage for opaque compounds is possible because the meanings of the constituents are not directly related to the meaning of the whole. Consequently, if one computes a meaning for the whole based on the meanings of the constituents, it would not be consistent with the legitimate meaning of that particular opaque compound. For example, when people see an opaque compound like *jailbird*, the initial decomposition at the lexical level, then the decomposition at the semantic level and eventually the composition at a conceptual level that corresponds to abstract and conceptual knowledge might introduce a different meaning from the stored semantic representation of the whole word. That is, people might attempt to interpret *jailbird* as "a bird that lives in jail". This attempted interpretation is different from the stored meaning which refers to "a prisoner". How strongly this non-target interpretation would be activated in the same time course as the retrieved meaning to cause interference by inhibiting each other might depend on how much the decomposition route is reinforced. In Experiment 4 in which evidence was found for the advantage for opaque compound processing over monomorphemic word processing, attempts to compute meanings for opaque compounds were less likely to cause a problem because the semantic representations of the constituents might not have been activated strong enough to compute a meaning for the whole that would introduce inhibition between the computed meaning and the retrieved meaning. In that study, nonword compound fillers were constructed by a nonword and a real word. To make a correct decision about a nonword compound, decomposition at the lexical level might be sufficient because the lexical representation of the nonword constituent could not be identified. Consequently, decomposition/composition might not extend to the semantic level and conceptual level because for the decomposition route to continue, it should need two activated lexical units. The characteristics of these filler items might have influenced how real word stimuli were processed. That is, the computed meaning at the decomposition/composition route (e.g., interpret *jailbird* as a bird that lives in jail) might be less likely to interfere with the legitimate meaning retrieved via the direct access (a prisoner).

An obvious way to boost the activation of compound constituents in attempting to introduce conflict between the decomposition route and the direct access in opaque compound processing is to use nonword compound fillers constructed by two real words. In this situation, the activation of real-word constituents of nonword compounds at the decomposition route is more likely to extend to the semantic level and conceptual level because both constituents have lexical representations and the lexical activation of the constituents should automatically spread to and beyond the semantic level. The properties of these nonword compounds should influence the processing of real-word compounds in that semantic representations of compound constituents should be activated stronger to compute a meaning at the conceptual level so that it introduces inhibition between the computed meaning and the retrieved meaning in the case of opaque compound processing.

In the current experiment, I compare the processing of opaque compounds with matched transparent compounds and monomorphemic words. Nonword compound fillers are constructed by two real words to introduce inhibition between the two access routes in opaque compound processing. Such a manipulation should make morphological decomposition easier because both constituents of nonword compound fillers have lexical entries. Consequently, computation of non-target meanings for opaque compounds should be more likely to introduce inhibition between the computed meaning and the retrieved meaning, and this should delay responses. In the meantime, such a manipulation should not introduce inhibition in the processing of transparent compounds and monomorphemic words. In the former case, the two access routes generally activate compatible information throughout the process because the computed meaning does not conflict with the retrieved meaning; in the latter case, only the direct access is available and no inhibition should occur. If interference due to inhibition can be introduced in opaque compound processing, the advantage of opaque compounds over monomorphemic words observed in Experiment 4 should disappear or even turn into a disadvantage. On the other hand, if interference cannot be introduced in opaque compound processing, the same pattern of results as Experiment 4 should be observed.

The design of the current experiment allows for a further exploration of the potential role of morphological decomposition at the semantic and conceptual levels in opaque compound processing. This aspect has not been explicitly pursued by the three models of complex word processing discussed earlier because all of the three assume no activation of constituents at the semantic level in opaque compound processing and thus, predict no disadvantage for opaque compounds due to morphological decomposition.

Method

Materials. Like Experiment 4, there were three types of stimuli, one for semantically transparent compounds, one for semantically opaque compounds, and one for monomorphemic words. One hundred and thirty-five compounds were pre-selected from the CELEX database. Ratings of semantic transparency of these compound items were then collected from 37 individuals. None of these judges participated in the experiment. The rating scale was a 7-point scale (1: totally opaque, 7: totally transparent). Based on the ratings, 36 pairs of transparent and opaque compounds were selected. The mean rating for the transparent compound condition was 5.57 and the mean rating for the opaque compound condition was 3.09. The ratings for the two groups were different, $t(2662) = 14.34, p < .0001$. In addition, constituent frequencies of the two compound conditions were controlled. For lemma frequency of the first constituent, $t(35) = .08, p > .93$; for surface frequency of the first constituent, $t(35) = .39, p > .70$. For lemma frequency of the second constituent, $t(35) = .05, p > .96$; for surface frequency of the second constituent, $t(35) = .15, p > .88$. A set of 36 monomorphemic words were also selected from the CELEX database. Each triple was matched on whole word frequency (lemma frequency, surface frequency) and word length (number of letters and number of syllables). The mean surface frequencies in the transparent, opaque, and monomorphemic word conditions were 1.25, 1.19 and 1.18 per million respectively. The mean lemma frequencies in the three conditions were 1.43, 1.35 and 1.34 per million respectively. The mean number of letters in all three conditions was 7.92. The mean numbers of syllables in the three conditions were 2.25, 2.31 and 2.33 respectively.

A set of 108 nonword filler items was also constructed so that the number of *yes* and *no* responses was balanced in the experiment. Seventy-two nonwords used compound formats and were constructed by randomly pairing two real words (e.g., *restpepper*). The remaining 36 nonwords used monomorphemic word formats and were constructed by changing one or two letters in real words.

The design was a within-subject design. The 108 experimental compound and monomorphemic items were presented in a randomized order along with the 108 filler items. Each participant saw a different random order of list.

Procedure. The procedure was identical to the procedure used in Experiment 4.

Participants. Thirty-two undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from one participant were excluded due to overall slow responses (mean response time more than 1500 ms). Thus, the data from 31 participants were used in the analyses.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition mean were regarded as outliers and trimmed. In addition, three trials with response times more than 5000 ms were deleted. In total, 3% of correct responses were trimmed. Reaction times and standard errors for correct responses to word stimuli, and accuracy rates are listed in Table 13. The accuracy rates for the compound nonwords and monomorphemic nonwords were 90% and 89% respectively. Reaction times and standard errors for correct responses to nonword fillers, and accuracy rates are listed in Table 14. Linear mixed-effect (multi-level) regression models were fitted to investigate the effect of complex structure on the time to make a lexical decision when nonword compound fillers were constructed by two real words. The models were fitted by using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects

I conducted two planned comparisons to evaluate the advantage for compound processing. There was evidence for an advantage in transparent compound processing, transparent compounds ($M = 6.63$, raw RT = 782 ms) were responded to more quickly than were monomorphemic words ($M = 6.68$, raw RT = 825 ms), $t(2898) = 2.21$, $SE = .03$, $p < .03$, 95% confidence interval (-.11, -.01). In contrast, there was no evidence for an advantage in opaque compound processing ($M = 6.67$, raw RT = 818 ms), $t(2898) = .10$, $SE = .03$, $p > .92$, 95% confidence interval (-.05, -.06).

For the most part, the accuracy analyses showed a similar pattern of results. For the comparison between transparent compounds and monomorphemic words, there was a difference in that responses to transparent compounds were more accurate than responses to monomorphemic words, $t(3341) = 2.13$, $SE = .37$, $p < .03$. Importantly, for the comparison between opaque compounds and monomorphemic words, there was indication of disadvantage in opaque compound processing, responses to opaque compounds were less accurate than responses to monomorphemic words, $t(3341) = 2.02$, $SE = .34$, $p < .04$.

There was also evidence for the effect of semantic transparency such that transparent compounds were responded to faster than were opaque compounds, $t(2898) = 2.08$, $SE = .03$, $p < .04$, 95% confidence interval (-.105, -.003). The accuracy analysis showed a similar pattern, $t(3341) = 4.11$, $SE = .36$, $p < .0001$.

In addition, evidence of the facilitation of constituent frequencies of transparent compounds and of the interference of those of opaque compounds was revealed by correlation analyses. In specific, log lemma frequencies of both constituents of transparent compounds were negatively correlated with log response times. For the first constituent, $r = -.11$, $p < .0007$, for the second constituent, $r = -.09$, $p < .003$. In contrast, log lemma frequency of the first constituent of opaque compounds was positively correlated with log response time, $r = .08$, $p < .02$. There was no correlation between the log lemma frequency of the second constituent and log response time, $r = -.04$, $p < .18$. Furthermore, there was evidence of the influence of whole word frequencies. Collapsing transparent and opaque compounds, log response time was negatively correlated with log lemma frequency, $r = -.10$, $p < .0001$, and log surface frequency of compounds, $r = -.08$, $p < .0006$. All correlations were calculated over the full data set using linear mixed-effect regression model.

The results of the current experiment indicated that when the decomposition route was encouraged by using nonword compounds that were constructed by two real words, the advantage of opaque compounds over frequency-matched monomorphemic words observed in Experiment 4 disappeared. Moreover, there was even evidence of disadvantage in opaque compound processing because opaque compounds were responded to less accurately than were monomorphemic words. This suggests that the two access routes eventually inhibit each other in opaque compound processing and this inhibition makes it more difficult to recognize these words. Correlation analyses lent further support for inhibition of the decomposition route in opaque compound processing. The positive correlation between log lemma frequency of the first constituent and log response time indicated that opaque compounds with relatively high-frequency first constituents were responded to more slowly than were opaque compounds with relatively low-frequency first constituents. This suggests that in the case of opaque compound processing, the meaning computed at the decomposition route inhibits the meaning retrieved at the direct access, cancels out the benefit gained at the lexical level or even makes the processing more difficult. In addition, the disadvantage of opaque compounds relative to monomorphemic words in terms of accuracy and the positive correlation between the frequency of the first constituent and log response time to opaque compounds also suggests that although the current experiment revealed no difference for the processing speed of opaque compounds and monomorphemic words, one cannot conclude that opaque compounds are processed via the direct access in a similar way as monomorphemic words as suggested by the dual-route race model.

The findings of the current experiment again demonstrated the effect of semantic transparency in that transparent compounds were responded to more quickly than opaque compounds. The semantic transparency effect was more solid in the current experiment than in Experiment 4. This was probably due to the inhibition between the two access routes in opaque compound processing introduced by reinforcing the decomposition route in the current experiment. The negative correlations between log response time to transparent compounds and log lemma frequencies of both constituents indicated that the higher the constituent frequency, the faster the responses. This evidence of facilitation of constituent frequencies in transparent compound processing was the opposite to the evidence of inhibition of constituent frequencies in opaque compound processing where higher first constituent frequencies were associated with slower responses. In sum, the correlation analyses suggest that the decomposition route is facilitative in transparent compound processing but inhibiting in opaque compound processing.

The results of the current experiment, together with the findings in Experiment 4, indicate that the relationship of the two access routes is flexible and context dependent. When the fillers are not constructed by pairing a nonword and a real word, the two access routes are mutually facilitative in the processing of both transparent and opaque compounds. In contrast, when the fillers are constructed by pairing two real words, the two access routes are still mutually facilitative in transparent compound processing but mutually inhibitory in opaque compound processing. The evidence of inhibition from the decomposition route in opaque compound processing suggests that lexical activation of opaque constituents can automatically spread further to the semantic level and conceptual level to compute non-target meanings. In other words, composition might be an important component and process in compound word processing. The three major models of complex word processing have difficulties to accommodate such a finding because none of them assumes the activation of constituents at the semantic level in the processing of opaque compounds and none of them predicts meaning computation in opaque compound processing. I will provide a fuller discussion on this matter in the General Discussion.

Table 13

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 5

Condition	LogRT (SE)	RT (SE)	Accuracy (SE)
Transparent Compound	6.63 (.02)	782 (16)	92 (2)
Opaque Compound	6.67 (.02)	818 (15)	81 (2)
Monomorpheme	6.68 (.02)	825 (16)	87 (1)

Table 14

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 5

Condition	RT (SE)	Accuracy (SE)
Nonword Compound	989 (25)	90 (1)
Nonword Monomorpheme	1020 (33)	89 (1)

Chapter VII

Experiment 6

The findings of Experiment 5, together with the findings of Experiment 4, suggest that the advantage for opaque compound processing can be removed by encouraging the decomposition route. This means that the relationship of the two access routes is flexible; the two access routes do not always facilitate each other. Rather, they can inhibit each other when they activate conflicting information across the semantic level and conceptual level as in the case of opaque compound processing in Experiment 5.

Another way to promote the decomposition route is inserting spaces into compounds. Juhasz, Inhoff, and Rayner (2005) found (Experiment 1) that inserting a space into normally concatenated English compounds increases the ease of word recognition using lexical decision task. They suggested that the advantage was due to increased ease for decomposition. In Juhasz et al. (2005), they did not indicate the semantic transparency status of their stimuli, nor did they provide stimuli in the Appendix. I suspect that the majority of the compound word stimuli used in their study were semantically transparent because increased ease for decomposition should decrease the ease of opaque compound processing due to the fact that the decomposition route would be more likely to compute a different meaning than the retrieved semantic representation of the whole word (e.g., compute the meaning for *jailbird* as “a bird that lives in jail” versus the legitimate meaning “a prisoner”).

In the current experiment, I insert a space into transparent compounds, opaque compounds and monomorphemic words respectively and compare the ease of processing of these three types of words. I expect that this manipulation should decrease the ease of recognition of opaque compounds due to the fact that the decomposition route would be encouraged and more likely to compute a non-target meaning in the same time course as the retrieved target meaning at the direct access. These two interpretations should inhibit each other and introduce interference in opaque compound processing. As for the processing of transparent compounds, the advantage should remain because the meaning computed via the decomposition route is compatible with the meaning retrieved via the direct access. None of the three major models of complex word processing outlined in Introduction would predict inhibition of the two access routes in opaque compound processing because these models do not assume the activation of opaque constituents at the semantic level.

Method

Materials. The word stimuli were identical to those used in Experiment 5 except that a space was inserted into the two types of compounds to separate their two constituents. A space was also inserted into monomorphemic word stimuli (e.g., *ecli pse*, *sopho more*, *plank ton*) to avoid strategic responses based on the appearance of spaces.

A set of 108 nonword filler items were constructed so that the number of *yes* and *no* responses was balanced in the experiment. The 72 nonwords that used compound formats were identical to those used in Experiment 5 except that a space was inserted to separate the two constituents. The remaining 36 nonwords were constructed based on the lexical status of the two halves after inserting a space into monomorphemic word stimuli. Eighteen nonwords were constructed by a nonword and a real word (e.g., *rost pepper*, or *pepper rost*), thirteen were constructed by two nonwords (e.g., *kan fidole*) and the remaining 5 were constructed by two real words. A space was inserted into the two constituents of this group of nonword fillers as well.

As in Experiment 5, the design of the current experiment was a within-subject design. The 108 experimental compounds and monomorphemic items were presented in a randomized order along with the 108 filler items. Each participant saw a different random order of list.

Procedure. The procedure was identical to the procedure used in Experiment 5.

Participants. Twenty-eight undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from nine participants were excluded due to overall slow responses (mean response time more than 1500 ms). Thus, the data from 19 participants were used in the analyses.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition mean were regarded as outliers and trimmed. In total, 2.9% of correct responses were trimmed. Reaction times and standard errors for correct responses to real word stimuli, and accuracy rates are listed in Table 15. The accuracy rates for the compound nonwords and monomorphemic nonwords were 91% and 93% respectively. Reaction times and standard errors for correct responses to nonword fillers, and accuracy rates are listed in Table 16. Linear mixed-effect (multi-level) regression models were fitted to investigate the effect of complex structure on the time to make a lexical decision when a space is inserted into compound words. Models were fitted by using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects.

I conducted two planned comparisons to evaluate the advantage for compound processing. There was evidence for an advantage in transparent compound processing, transparent compounds ($M = 6.66$, raw RT = 810 ms) were responded to more quickly than were monomorphemic words ($M = 6.75$, raw RT = 882 ms), $t(1796) = 3.41$, $SE = .03$, $p < .0007$, 95% confidence interval (-.138, -.039). In contrast, there was no evidence for an advantage in opaque compound processing ($M = 6.72$, raw RT = 867 ms), $t(1796) = 1.02$, $SE = .03$, $p > .31$, 95% confidence interval (-.03, .08).

The accuracy analyses showed a similar pattern of results. For the comparison between transparent compounds and monomorphemic words, responses to transparent compounds were more accurate than to monomorphemic words, $t(2049) = 3.13$, $SE = .35$, $p < .002$. In contrast, there was no evidence for an advantage in opaque compound processing, $t(2049) = .17$, $SE = .33$, $p > .86$.

There was also evidence for the effect of semantic transparency such that transparent compounds were responded to more quickly than were opaque compounds, $t(1796) = 2.35$, $SE = .03$, $p < .02$, 95% confidence interval (-.12, -.01). The accuracy analysis showed a similar pattern, $t(2049) = 3.30$, $SE = .35$, $p < .001$.

In addition, evidence of the facilitation of constituent frequencies of transparent compounds and of the inhibition of those of opaque compounds was revealed by correlation analyses. For transparent compounds, log response time was negatively correlated with log lemma frequency of the first constituent, $r = -.14$, $p < .0003$. This negative correlation suggests that high-frequency constituents are facilitative in transparent compound processing. In contrast, for opaque compounds, log response time was positively correlated with log lemma frequency of the first constituent, $r = .14$, p

$< .0009$. This positive correlation suggests that high-frequency constituents are interfering in opaque compound processing. Furthermore, there was evidence of the influence of whole word frequencies. Collapsing transparent and opaque compounds, log response time was negatively correlated with log lemma frequency, $r = -.11, p < .0002$, and log surface frequency of compounds, $r = -.07, p < .02$. All correlations were calculated over the full data set using linear mixed-effect regression model.

Like Experiment 5, the results of the current experiment indicated that reinforcing the decomposition route by inserting spaces into compound words decreased the ease of opaque compound processing because the advantage of opaque compounds over frequency-matched monomorphemic words observed in Experiment 4 disappeared. Moreover, there was evidence of inhibition of the two access routes in opaque compound processing because opaque compounds with high-frequency first constituents were associated with slower responses. This suggests that in the case of opaque compound processing, the meaning computed at the decomposition route inhibits the meaning retrieved at the direct access, cancels out the facilitation at the lexical level or even makes the processing more difficult.

The current experiment replicated the effect of semantic transparency found in previous experiments in that there was an advantage in transparent compound processing over opaque compound processing. The semantic transparency effect again was more solid in the current experiment than in Experiment 4, which might be due to the interference in opaque compound processing introduced by inserting spaces into compound words. Correlation analyses indicated that constituent frequencies played different roles in transparent and opaque compound processing because high-frequency first constituents were associated with faster responses to transparent compounds but slower responses to opaque compounds. These findings suggest that the decomposition route is facilitative in transparent compound processing but inhibiting in opaque compound processing. The decomposition is facilitative in transparent compound processing because it computes a meaning that is compatible with the retrieved whole word meaning. The decomposition is inhibiting in opaque compound processing because it might compute a meaning that is incompatible with the retrieved whole word meaning.

Table 15

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 6

Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
Transparent Compound	6.66 (.03)	810 (25)	92 (1)
Opaque Compound	6.72 (.03)	867 (25)	84 (2)
Monomorpheme	6.75 (.03)	882 (29)	87 (2)

Table 16

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 6

Condition	RT (SE)	Accuracy (SE)
Nonword Compound	1127 (46)	91 (1)
Nonword Monomorpheme	1138 (62)	93 (1)

Note. The Nonword monomorphemes were not really monomorphemes here because they might have a real word component.

Chapter VIII

Experiment 7

The findings of Experiments 5 and 6 indicate that the decomposition route can be encouraged to introduce inhibition between the two access routes in opaque compound processing by using compound nonword fillers that are formed by two real words or by inserting spaces into normally concatenated compounds. The goal of the current experiment is to further examine the consequence of enhancing the decomposition route. A third way to promote the decomposition route is using different colors (red vs. black) to present the two constituents of compounds. The color contrast of the two constituents highlights the fact that the two constituents of compounds are two individual words, and thus should encourage decomposition and make it easier to identify the constituents.

The purpose of the current experiment is to investigate whether color contrast of the two constituents decreases the ease of opaque compound processing. I expect that color contrast of the two constituents should make morphological decomposition easier and accelerate processes at the decomposition route, so that it is more likely for this route to compute a non-target meaning in the same time course as the meaning retrieved at the direct access. In this kind of situation, inhibition between the meanings activated at the two access routes should occur and slow down word recognition which should be eventually determined by the direct access. When inhibition between the two access routes occurs, there should be no advantage in opaque compound processing over monomorphemic word processing or opaque compound processing would even exhibit a disadvantage. Among the three major models of complex word processing, none would predict an inhibition between the two access routes because all three models assume that opaque constituents are not supposed to be activated at the semantic level, and thus no computation of meaning should be executed.

Method

Materials. The 108 word stimuli were identical to those used in Experiment 6 except that no space was inserted into the word. In addition, the two constituents of compounds were presented in two different colors (red-black or black-red across groups). Likewise, monomorphemic word stimuli were also divided into two parts and presented in two different colors to avoid strategic responses based on the appearance of colors. Monomorphemic words were divided into two parts in the same way as they were divided in Experiment 6 in which a space was inserted into all stimuli.

A set of 108 nonword filler items were included so that the number of *yes* and *no* responses was balanced in the experiment. The filler items were identical to those used in Experiment 6. The filler items were presented in a similar way as their corresponding experimental items in terms of constituent colors.

The design of the current experiment was different from the previous experiments. It was a mixed-design in that the factor of color contrast was a between-subject factor whereas the factor of word type was a within-subject factor. For the factor of color contrast, one group of participants saw the first constituents in red color and the second constituent in black color. The other group saw the first constituent in black color and the second constituent in red color. Color contrast was kept consistent within group to prevent participants from interpreting the purpose of this manipulation. The 108 experimental compound and monomorphemic items were presented in a randomized order along with the 108 filler items. Each participant saw a different random order of list.

Procedure. The procedure was identical to the procedure used in Experiment 6.

Participants. Forty undergraduates from the University of Alberta participated in the study for partial course credit. All participants spoke English as a first language. The data from four participants were excluded due to low accuracy (less than 60%) in one or more experimental conditions or filler conditions. In addition, the data from two participants were excluded due to overall slow responses (mean responses were more than 1400 ms). Thus, the data from 34 participants were used in the analyses.

Results and Discussion

Reaction times more than 2.5 standard deviations away from the condition mean were regarded as outliers and trimmed. In addition, two trials were deleted due to a computer error, one trial with a response less than 200 ms and three trials with responses more than 4000 ms were removed. In total, 2.9% of correct responses were trimmed. Reaction times and standard errors for correct responses to real word stimuli, and accuracy rates are listed in Table 17. The accuracy rates for the compound nonwords and monomorphemic nonwords were 90% and 95% respectively when the first constituent was presented in black color, and were 92% and 94% respectively when the first constituent was presented in red color. Reaction times and standard errors for correct responses to nonword fillers, and accuracy rates are listed in Table 18. Linear mixed-effect (multi-level) regression models were fitted to investigate the effect of complex structure on the time to make a lexical decision when the two constituents of compound words were presented in different colors. Models were fitted using log response time and accuracy as dependent variables, word type as predictor variable, and subject and item as random effects.

I conducted two planned comparisons to evaluate the advantage for compound processing. There was evidence for an advantage in transparent compound processing, transparent compounds ($M = 6.61$, raw RT = 768 ms) were responded to more quickly than were monomorphemic words ($M = 6.66$, raw RT = 799 ms), $t(3116) = 2.09$, $SE = .03$, $p < .04$, 95% confidence interval (-.11, -.002). In contrast, although responses to opaque compounds ($M = 6.68$, raw RT = 825 ms) were slower than responses to monomorphemic words, this difference was not significant, $t(3116) = .21$, $SE = .03$, $p > .83$, 95% confidence interval (-.06, .05).

The accuracy analyses showed no difference in either comparison. For the comparison between transparent compounds and monomorphemic words, although responses to transparent compounds were more accurate than to monomorphemic words, this difference did not reach significance, $t(3658) = 1.47$, $SE = .45$, $p > .14$. Likewise, although responses to opaque compounds were less accurate than to monomorphemic words, this difference did not reach significance, $t(3658) = 1.49$, $SE = .43$, $p > .14$.

There was also evidence for the effect of semantic transparency such that transparent compounds were responded to more quickly than were opaque compounds, $t(3116) = 2.29$, $SE = .03$, $p < .02$, 95% confidence interval (-.12, -.01). The accuracy analysis showed a similar pattern, $t(3658) = 2.95$, $SE = .44$, $p < .003$.

In addition, there was no main effect of the factor of color contrast; for response times, $F < 1$, and for accuracy, $F < 1$. This finding suggests that word processing in general did not differ in the two color presentation conditions. Also, there was no evidence of interaction between the two levels of color contrast and the three types of words. For log response times, $F(1, 3116) = 2.14$, $MSE = .05$, $p > .12$, and for accuracy,

$F < 1$. The absence of interaction suggests that the two ways of color contrast influenced different types of word processing the same regardless of whether the first constituent or the second constituent was presented in red color as long as the two constituents were presented in different colors.

Evidence of the facilitation of constituent frequencies of transparent compounds and of the inhibition of those of opaque compounds was revealed by correlation analyses. For transparent compounds, log response time was negatively correlated with log lemma frequency of the first constituent, $r = -.12, p < .0001$. This negative correlation suggests that high-frequency constituents are facilitative in transparent compound processing. In contrast, for opaque compounds, log response time was positively correlated with log lemma frequency of the first constituent, $r = .12, p < .0001$. This positive correlation suggests that high-frequency constituents are inhibiting in opaque compound processing. Furthermore, there was evidence of the influence of whole word frequencies. Collapsing transparent and opaque compounds, log response time was negatively correlated with log lemma frequency, $r = -.15, p < .0001$, and log surface frequency of compounds, $r = -.13$, $p < .0001$. All correlations were calculated over the full data set using linear mixed-effect regression model.

In sum, as in Experiments 5 and 6, the current experiment again provides evidence of inhibition of the two access routes in opaque compound processing and further suggests that the relationship of the two access routes in compound word processing is flexible and not fixed. In the case of transparent compound processing, the two access routes can boost each other because they activate compatible information throughout the process. In the case of opaque compound processing, the two access routes can inhibit each other across the semantic level and conceptual level because they activate conflicting information. The inhibition of the two access routes should slow down the process and cancel out the benefit due to mutual facilitation of the two routes at the lexical level.

It should be noted that in both Experiment 6 and the current experiment, the decomposition route might have been accelerated more than in the case of Experiment 5 where no spaces were inserted into compound words or no color contrast had been applied to the two constituents of compound words. One might have expected to see more overt evidence of inhibition to be reflected in the comparison of opaque compounds and monomorphemic words. However, both experiments failed to show such an overt effect of inhibition. One possible reason is that space inserting and color contrast might have slowed down the direct access. This would slow down the processing of monomorphemic words and reduce the chance to reveal an overt disadvantage of opaque compound processing. Another possible reason is that as mentioned before, although the computed illegitimate meaning for opaque compound should interfere with the retrieved legitimate meaning at the direct access, the mutual facilitation of the two access routes at the lexical level might have compensated the disadvantage due to mutual inhibition of the two access routes across semantic level and conceptual level to some extent and made it less likely to detect an overt disadvantage of opaque compound processing.

Table 17

Mean log response times, response times (in ms) and accuracy rates (%) with standard errors for word stimuli of Experiment 7

Color	Condition	Log RT (SE)	RT (SE)	Accuracy (SE)
First Constituent Black	Transparent Compound	6.61 (.02)	769 (16)	90 (1)
Second Constituent Red	Opaque Compound	6.67 (.02)	817 (21)	80 (2)
	Monomorpheme	6.67 (.02)	809 (16)	85 (2)
First Constituent Red	Transparent Compound	6.61 (.02)	766 (16)	91 (1)
Second Constituent Black	Opaque Compound	6.68 (.03)	833 (21)	79 (2)
	Monomorpheme	6.64 (.02)	790 (16)	86 (2)

Table 18

Mean response times (in ms) and accuracy rates (%) with standard errors for filler items of Experiment 7

Color	Condition	RT (SE)	Accuracy (SE)
First Constituent Black	Nonword Compound	1047 (39)	90 (2)
Second Constituent Red	Nonword Monomorpheme	962 (46)	95 (2)
First Constituent Red	Nonword Compound	942 (39)	92 (2)
Second Constituent Black	Nonword Monomorpheme	877 (46)	94 (2)

Note. The Nonword monomorphemes were not really monomorphemes here because they might have a real word component.

Chapter VIIII

General Discussion

In this section, I summarize the findings of Experiments 1-7 and interpret these results in terms of their theoretical implications for compound word processing and their compatibility with previous findings. I then discuss explanations for these results provided by current models of complex word processing. I end by proposing a framework that aims to accommodate the findings of the current research and by discussing future research.

Lexical access and the advantage for compound processing

Experiments 1-3 demonstrated that semantically transparent compounds were processed more quickly than frequency-matched monomorphemic words. This finding suggests that the availability of morphological decomposition initiated by complex structure aids rather than hinders transparent compound processing. Experiment 3 further indicated that lexical access to the constituents influenced the speed of transparent compound processing; transparent compounds with high-frequency first constituents were processed more quickly than compounds with low-frequency first constituents. No evidence was found for the influence of the frequency of the second constituents. As mentioned previously, the reason might be that the compound stimuli in Experiment 3 were relatively novel in terms of whole word frequency count comparing to compound stimuli being used in other similar studies that found evidence for the influence of frequency of the second constituent (e.g., Juhasz et al., 2003; van Jaarsveld & Rattink, 1988). Experiment 4 provided additional evidence for the role of lexical access to compound constituents in that semantically opaque compounds were also processed more quickly than frequency-matched monomorphemic words. Because the only difference between transparent and opaque compounds is semantic transparency, if the processing of these two types of compound words shows some commonality like the processing advantage over monomorphemic words, one can infer that the cause of the common advantage must not be due to semantic transparency. Instead, the advantage of compound processing in general is likely due to lexical access to compound constituents.

Based on these findings, I propose that the advantage of compound processing in general is due to lexical access to compound constituents. More specifically, the lexical entries of the constituents of both transparent and opaque compounds are accessed during compound processing. The activation of the constituents increases the activation of the whole via facilitative links that connect them. Consequently, the recognition of compound words is easier than the recognition of monomorphemic words because the latter only involves direct access to whole word representations and cannot benefit from facilitative activation from constituents.

The current findings and proposal about the activation of lexical entries of compound constituents in compound processing are consistent with previous findings of repetition priming in compound word processing regardless of semantic transparency in that both transparent and opaque compound primes facilitated the recognition of their constituents, or constituent primes facilitated the recognition of the compounds (see for example, Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999; Libben, Gibson, Yoon, & Sandra, 2003; Monsell, 1985, Zwitserlood, 1994). These previous findings suggested that lexical entries of compound constituents are accessed in compound processing regardless of semantic transparency. The findings in the current studies not only suggest that compound constituents are accessed but also suggest that the access to compound constituents facilitates compound processing relative to monomorphemic word processing.

Semantic access and the advantage/disadvantage for compound processing

Experiment 4 also indicated that although both transparent and opaque compounds were processed more quickly than frequency-matched monomorphemic words, it was easier to recognize transparent compounds than opaque compounds because responses to transparent compounds were more accurate than responses to opaque compounds. Experiments 5-7 provided further evidence of semantic transparency effects in that transparent compounds were processed more quickly than opaque compounds in all three experiments. These findings suggest that in addition to lexical factors, semantic transparency also influences the ease of compound processing. I propose that the most plausible explanation for this finding is that constituents of compounds are being accessed at the semantic level and influence the ease of compound processing. In transparent compound processing, I assume facilitative activation between the constituents and the whole at the semantic level. For example, the activation from *rose* and *bud* at the semantic level should facilitate the activation of *rosebud* because they are all semantically related. In contrast, opaque compounds cannot benefit much from facilitative activation at the semantic level because both constituents are not semantically related to the whole. For example, *jail*, *bird* versus *jailbird*, where "*bird*" does not contribute much meaning to "*jailbird*" and thus does not facilitate the activation of "*jailbird*" at the semantic level. As a result, the differential facilitative activation from the semantic level leads to the different degrees of advantage in processing transparent and opaque compounds in Experiment 4. This differential facilitative activation is also partially responsible for the semantic transparency effects in Experiments 5-7. In addition, the semantic transparency effects in these three experiments are partially due to inhibition between the computed meaning and the retrieved meaning at the two access routes in opaque compound processing. This inhibition should slow down the processing of opaque compounds and contribute to semantic transparency effects.

My interpretation of the influence of semantic transparency on compound processing implies that the constituents of both transparent and opaque compounds are being accessed at the semantic level. In other words, I suggest that not only transparent compounds but also opaque compounds are being decomposed at the semantic level during word recognition. Alternatively, one can propose that opaque compounds are not being decomposed at the semantic level and that semantic representations of opaque compounds are activated via the direct mapping route. Based on this assumption, there should be no facilitative activation between the constituents and the whole at the semantic level because the constituents are not activated at all. Consequently, opaque compound processing should be more difficult than transparent compound processing due to the lack of facilitative activation from the constituents at the semantic level.

However, the evidence of negative influence of constituent frequency of opaque compound processing in Experiments 5-7 suggests that such an explanation is implausible. Under this explanation, constituent frequency should not have negative influences in opaque compound processing because constituents are not supposed to be activated at the semantic level. In addition to the difficulty to accommodate the data of Experiments 5-7, such an explanation also has a theoretical barrier. One should explain how compound constituents being activated at the lexical level are not being further activated at the semantic level. If one introduces an inhibition mechanism to stop the activation of the constituents spreading from the lexical level to the semantic level, one has to assume that people some how know *a priori* that a compound word is semantically opaque even before their processing reaches the semantic level. Such an assumption is not logically solid and a more plausible explanation is that the constituents of opaque compounds are automatically activated at the semantic level although they might not be as helpful as the constituents of transparent compounds or could even be harmful if attempts are made to compute an illegitimate meaning for the whole.

The evidence of semantic transparency effects on processing speed found in Experiments 5-7 is compatible with Sandra's (1990) finding of the effect of semantic transparency for Dutch compound words. His studies indicated that an item that was semantically related to the constituents of a compound word could not prime the compound if the compound was semantically opaque. Priming effects were observed, however, if the compound was semantically transparent. There are at least three possible explanations for the effect of semantic transparency found in Sandra's studies. First, the effect can be explained by assuming no access to the constituents in opaque compound processing. However, such an explanation cannot accommodate findings of repetition priming in opaque compound processing where opaque compounds and their constituents speeded the recognition of each other (e.g., Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999; Libben, Gibson, Yoon, & Sandra, 2003; Zwitserlood, 1994). Second, taken together these previous findings of semantic transparency effect in semantic priming studies and the absence of semantic transparency effect in repetition priming studies, the effect can be explained by assuming lexical access but not semantic access to the constituents in opaque compound processing. However, such an explanation has the difficulty to explain why the activation of the constituents at the lexical level does not spread to the semantic level in opaque compound processing. Thus, a third way to explain the effect is to assume both lexical and semantic access to the constituents and an inhibition mechanism across the semantic and conceptual levels in opaque compound processing. That is, the semantic representations of opaque constituents are initially activated at the decomposition route. The activation at the semantic level further spreads to the conceptual level to compute meanings for the whole based on the conceptual knowledge of the constituents. However, this computed meaning is inconsistent with and eventually suppressed by the retrieved meaning at the direct access so that the legitimate meaning is warranted.

Among the three possible interpretations for the effect of semantic transparency, the last proposal receives the most support from the current data. Experiments 5-7 provided consistent evidence of an inhibition mechanism because the higher the constituent frequency, the slower the processing of opaque compounds. In addition, the accuracy data of Experiment 5 showed a disadvantage of opaque compounds versus monomorphemic words. The findings of these three experiments suggest that automatic activation of compound constituents spreads from the lexical level to the semantic level in word recognition. In addition, the decomposition route might attempt to compute a meaning for the compound word based on the constituents depending on whether the decomposition route is encouraged. In the case of opaque compound processing because the meaning computed at the decomposition route is not compatible with the meaning retrieved at the direct access (e.g., *jailbird* is not "a bird that lives in jail" but "a prisoner"), it introduces inhibition between the two possible interpretations activated at the two access routes. The meaning computed via the decomposition route should be eventually inhibited by the direct access so that the legitimate interpretation is warranted. Thus, the current research contributes to the literature of semantic transparency by identifying the potential inhibition mechanism in opaque compound processing. This contribution is unique because to my knowledge, no such data for normal population have been reported in the literature. Libben (1998) has reported data from an aphasia patient that implied an inhibition mechanism in opaque compound processing. This patient tended to interpret opaque compounds as if they were transparent. Libben suggested that this might be due to the failure of inhibition of the meanings of the constituents in opaque compound processing. In the text to follow, I will propose a framework that assumes inhibition between the decomposition route and the direct access across the semantic and conceptual level in opaque compound processing. This inhibition mechanism, however, does not apply to the processing of transparent compounds as assumed in Libben's APPLE model.

Current models and the advantage/disadvantage for compound processing

The automatic decomposition model. What explanations do some current models of complex word processing provide for the advantage for transparent compound processing and the effect of semantic transparency? The early version of the automatic decomposition model (Taft & Forster, 1975; 1976) is not readily compatible with the findings of the advantage for transparent compound processing because although this model emphasizes the involvement of decomposition in complex word processing, it does not make explicit predictions about whether decomposition should aid or hinder complex word processing. If one takes into consideration the general view that computation is time consuming, this version of the model should predict that transparent compounds are processed more slowly than monomorphemic words. Similarly, the later version of this model (Taft, 1994) does not predict a clear advantage for transparent compounds either because meanings of transparent compounds are computed rather than retrieved. The cost of computation might trade off the advantage gained at the lexical level where activation of constituents increases the activation of the whole.

As for the effect of semantic transparency in Experiments 4-7, the early version of the automatic decomposition model (Taft & Forster, 1975; 1976) was not designed to address the issue of semantic transparency in complex word processing and thus predicts no difference in the processing of transparent and opaque compounds. The later version of the model (Taft, 1994) proposed that although both transparent complex words and opaque complex words are automatically decomposed at the access level, the meaning of the former is computed by combining the meanings of the constituents, whereas the meaning of the latter is directly retrieved from the representation of the whole word meaning, and this direct retrieval dominates any competitions from attempts to compute the meaning of the whole based on constituents. According to this proposal, the later version of the automatic decomposition model should predict an advantage for opaque compounds over transparent compounds because there should be a computational cost associated with transparent compound processing. Although the author mentioned the "suppression" mechanism when outlining how the model works, this mechanism was not explicitly implemented in the structure of the model, where no assumption was made for the activation of the transparent meaning of opaque constituents in the processing of opaque complex words. Therefore, the "suppression" mechanism might work in vain because it has nothing to suppress. Thus, this "suppression" mechanism cannot accommodate the findings of semantic transparency in the current research. For the same reason, it is difficult to appreciate how this later version of the automatic decomposition model would accommodate the negative influence of constituent frequency in opaque compound processing found in Experiments 5-7.

The dual-route race model. In contrast with the automatic decomposition model, the dual-route race model (Baayen et al., 1997; Schreuder & Baayen, 1995) can provide an alternative explanation for the data about the advantage for transparent compound processing with the assumption of statistical facilitation. According to this model, the direct access route should be faster than the decomposition route because the former directly maps the complex word to the stored whole unit representation in the system whereas the latter involves meaning computation that is assumed to be time consuming. When the two response distributions corresponding to the direct access route and the decomposition route overlap such as in the case of transparent compound processing, word recognition decisions should get statistical facilitation, and thus be faster than decisions made using a single route alone.

Although this model can provide an explanation for the findings about transparent compound processing, as mentioned in the Introduction, this model might have difficulties explaining why statistical facilitation occurs for some complex words, such as compounds and some derivational forms, but not for some other complex words, such as Finnish case inflectional forms. This difficulty arises because one of the key assumptions of this model is that both the direct access route and the decomposition route are involved in complex word processing simultaneously. Consequently, statistical facilitation should be a general prediction for all complex words instead of a specific prediction for some complex words. In addition, I also want to point out that one important assumption of statistical facilitation is that the two access routes are independent from each other (see Raab, 1962). However, this assumption is not readily consistent with a dual-route race model that is based on an interactive activation framework (e.g., Baayen et al., 1997; Schreuder & Baayen, 1995). In such a framework, transparent compounds and their constituents should be connected by facilitative links. Hence, the decomposition route and the direct access cannot work independently in compound processing.

Although the dual-route race model can provide a potential explanation for the advantage for transparent compound processing with the assumption of statistical facilitation, it is not readily compatible with the advantage for opaque compound processing over frequency-matched monomorphemic word processing found in Experiment 4 and the negative influence of constituent frequency (positive correlation between log response times and constituent frequency) for opaque compounds found in Experiments 5-7. According to this model, no statistical facilitation should occur for opaque compound processing because it assumes that opaque compounds can only be recognized via direct access in that their meanings cannot be computed based on their constituents (Schreuder & Baayen, 1995). In other words, this model assumes that opaque compounds are recognized in a similar way to monomorphemic words. This assumption, however, is incompatible with the findings of the advantage for opaque compound processing in Experiment 4 and the negative influence of constituent frequency in opaque compound processing in Experiments 5-7 because the constituents of opaque compounds are not supposed to be activated.

The APPLE model. Comparing to the two general models of complex word processing, the APPLE model by Libben (1998) makes more explicit assumptions about how compound words are represented and recognized, and thus is more compatible with the findings in the current research. As mentioned earlier in the Introduction, the APPLE model is also dual-route in nature. This model essentially implies the cooperation of the two access routes at the lexical level because it assumes facilitative links between compound words and their constituents. Thus, this part of the model is readily compatible with the findings of advantage for both transparent and opaque compound processing in Experiment 4. This model, however, cannot explain the negative influence of constituent frequency for opaque compound processing in Experiments 5-7. According to this model, opaque constituents should not be accessed at the conceptual level. This means, it is impossible to compute a meaning of the whole for opaque compounds and no inhibition between the "computed meaning" and the retrieved meaning should be expected. Thus, no negative influence of constituents should occur in opaque compound processing. For example, according to this model, the meaning of the opaque constituent *straw* of *strawberry* is not supposed to be activated at the conceptual level and therefore, no attempt should be made to compute a meaning for *strawberry* based on *straw* and *berry*. This model also has difficulties to explain the semantic transparency effect in Experiments 4-7. According to this model, advantage for compound processing should occur at the lexical level but not at the conceptual level because meanings of compounds are retrieved rather than computed regardless of semantic transparency. In specific, the model depicts within-level inhibition between compounds and their constituents at the conceptual level regardless of semantic transparency (e.g., *blueberry* should inhibit *blue* and *berry* and *strawberry* should inhibit *straw* and *berry* at the conceptual level) so that attempts of computation of meanings are prevented. Thus, it should predict no difference in processing the two types of compounds. Alternatively, this model could predict an advantage of opaque compounds over transparent compounds if one considers the assumption of cross-level inhibition in compound processing. Based on this assumption, there is inhibition from the lexical representation of the whole word to opaque constituent but not transparent constituent (e.g., *strawberry* at the lexical level inhibits *straw* at the conceptual level whereas *blueberry* at the lexical level does not inhibit *blue* at the conceptual level). Thus, in total, opaque compound processing should receive more inhibition from the direct access than transparent compound processing. In other words, the decomposition route should be less interfering in opaque compound processing than transparent compound processing. Consequently, when the decomposition route is encouraged like in the case of Experiments 5-7, the direct access should be more likely to win in the case for opaque compound processing than in the case for transparent compound processing. This alternative prediction is inconsistent with the semantic transparency effects found in Experiments 5-7 where transparent compounds were processed more quickly than opaque compounds.

A flexible framework for the recognition of English compound words

Given that current models do not fully explain the findings of the current research, in this section, I outline a framework for word recognition that aims to accommodate the advantage for transparent compound processing, the effect of semantic transparency, and the flexibility of the relationship of the two access routes. This framework is heavily indebted to previous models of complex word processing, especially the APPLE model by Libben (1998).

As showed in Figure 1, this framework assumes three levels of representations, a lexical level, a semantic level and a conceptual level.

The function of the lexical level is to capture non-semantic formal representations of morphological structure. The reason for assuming such a structure is that as argued by Libben (1998), previous studies found repetition priming for both semantically transparent and opaque compounds (e.g., Jarema, Busson, Nikolova, Tsapkini, & Libben, 1999; Libben, Gibson, Yoon, & Sandra, 2003; Zwitserlood, 1994) while only semantic priming for semantically transparent compounds (e.g., Sandra, 1990), and these findings suggest the representation of morphological structure at a level that is separate from a semantic level. At the lexical level, within-level facilitative links connect compound words and their constituents regardless of semantic transparency (e.g., *blue* and *berry* are both linked to *blueberry* and *straw* and *berry* are both linked to *strawberry*). In addition, cross-level facilitative links connect the whole word representation of compounds at the lexical level and the whole word representation at the semantic level. This part of the framework directly inherits the APPLE model. It is different from the APPLE model in that cross-level facilitative links connect the constituent representations at the lexical level and constituent representations at the semantic level regardless of semantic transparency. This is to assume that activation of constituents at the lexical level automatically spreads to the semantic level regardless of semantic transparency. For example, *blue* and *berry* as constituents of *blueberry* at the lexical level are linked to *blue* and *berry* as constituents of *blueberry* at the semantic level respectively; *straw* and *berry* as constituents of *strawberry* at the lexical level are linked to *straw* and *berry* at the semantic level respectively.

The major function of the semantic level is to represent word meanings and to capture semantic transparency. Word meanings of both compounds and their constituents are represented at this level. Semantic transparency is depicted by assuming within-level facilitative links between transparent constituents and the whole so that the parts and the whole can increase the activation of each other at this level (e.g., *blue* and *berry* are both linked to *blueberry* at the semantic level). In contrast, no such links are assumed between opaque constituents and the whole at the semantic level, and, therefore, opaque parts and the whole cannot increase the activation of each other (e.g., the opaque constituent *straw* is not linked to *strawberry* at this level because *straw* does not contribute to the meaning of *strawberry*). It should be noted that this framework does not assume inhibition where facilitative links are absent. That is, no inhibition is assumed between *straw* and *strawberry* at the semantic level. This is to distinguish from the APPLE model that assumes inhibition where links are absent.

The function of the conceptual level is to emphasize the involvement of computation of word meaning in compound word processing. This emphasis is implemented by assuming a composition/computation process where constituents of compound words at the semantic level are linked to conceptual representations of these constituents. In turn, the conceptual representations of the two constituents are linked to the computed meanings for compounds. As showed in Figure 1, these links are depicted by dotted lines to distinguish them from links that do not directly involve computation. In addition, the conceptual unit generated by computation is linked to the whole word representation at the semantic level. For transparent compounds, this link is facilitative and for opaque compounds, this link is inhibitory. This distinction is to depict the fact that the computed meaning for transparent compounds is consistent with the retrieved meaning whereas the computed meaning for opaque compounds is inconsistent with the retrieved meaning. In transparent compound processing, word meaning can be accessed when the activation accumulated from the two access routes reaches some threshold because meanings generated in the two access routes are consistent. In opaque compound processing, the computed meaning interferes with the retrieved meaning. Word meaning should be eventually determined based on the direct access. Thus, this framework assumes that when conflicts arise, the direct access has the priority. A possible mechanism behind this process is that when conflicts arise, the system might have to consult previous experiences of the usage of the compound word and these previous experiences should support the meaning retrieved at the direct access because the system should retrieve no record of interpreting *jailbird* as "a bird that lives in jail".

This proposed framework essentially assumes the communication of the two access routes at both the lexical level and semantic level for transparent compound processing because it assumes facilitative links between transparent compound words and their constituents at both levels. Thus, this framework can explain the findings of advantage for transparent compound processing relative to monomorphemic word processing because the latter cannot benefit from the facilitative activation from the constituents as the former. This framework can also explain semantic transparency effects in Experiments 5-7. According to this framework, transparent compounds should be processed more quickly than opaque compounds because in transparent compound processing, the activation of the parts and the whole can boost each other via the facilitative links at the both the lexical and semantic level, whereas in opaque compound processing, the parts and the whole mainly boost each other via the facilitative links at the lexical level. If there is boost between the parts and the whole at the semantic level, it is only between the transparent constituent and the whole but not between the opaque constituent and the whole. For example, *berry* and *strawberry* can facilitate the activation of each other but *straw* and *strawberry* cannot facilitate the activation of each other at the semantic level. In addition, transparent compound processing can potentially benefit from the conceptual level via mutual facilitation between the decomposition route and the direct access because the computed meaning and the retrieved meaning are compatible. In contrast, opaque compound processing might have a disadvantage at this aspect because the computed meaning and the retrieved meaning are not compatible and the two access routes are mutually inhibitory across the semantic and conceptual levels. Following the same logic, this framework can explain the negative influence of constituent frequency for opaque compound processing in Experiments 5-7. Due to the assumption of inhibition between the meaning computed at the decomposition/composition route and the meaning retrieved at the direct access, when the decomposition route is encouraged like in the case of Experiments 5-7, the higher the constituent frequency, the slower the responses to opaque compounds because the activation of relative high-frequency constituents should be stronger than the activation of relative low-frequency constituents and hence, the computed meaning should be more activated and introduce more inhibition. As a result, there is a negative influence of constituent frequency in opaque compound processing when the decomposition route is encouraged.

In addition, this proposed framework can also accommodate the advantage of opaque compound processing found in Experiment 4. In the situation of Experiment 4, the computed meanings for opaque compounds might not be activated strong enough to cause a problem due to the fillers being used. That is, participants did not have to compute a meaning for the filler item *rostpepper* to decide that it was not a real word because the nonword constituent *rost* does not have a semantic representation. Consequently, this might have influenced the processing of real word compounds in that the compute meanings for these compounds might not be activated strong enough to introduce inhibition. As a result, not much interference from the decomposition route was generated to trade off the benefit gained at the lexical level due to the facilitation of the two access routes.

Alternative interpretations for findings in the current studies

Although the proposed model can accommodate the findings of the current studies, there are certainly alternative interpretations for these findings. In this section, I discuss a different interpretation for the advantage for transparent compound processing and two alternative interpretations for the disadvantage for opaque compound processing found in Experiments 5-7.

A more liberal interpretation for the advantage of transparent compound processing found in the current studies is that the advantage is due to a faster decomposition route in transparent compound processing than the direct access in monomorphemic word processing. In other words, it can be assumed that in transparent compound processing, the decomposition route is the only route at work and this route is faster than the direct access in monomorphemic word processing. A question that might be raised here is that how the decomposition route that involves meaning computation and is supposed to be time consuming wins over the direct access in monomorphemic word processing that does not involve any computation. One piece of information that might support this argument is that constituent frequencies of English compounds are almost always higher than their whole word frequencies. In fact, in the entire CELEX database, I found only about five compounds whose constituent frequencies are equal or less than their whole word frequencies. This means that for the majority of English compounds, their constituent frequencies are higher than whole word frequencies. Thus, it is possible that even though morphological decomposition and meaning computation per se is time consuming in general, word recognition at the decomposition route does not have to be slow in the case of transparent compound processing. Time can be saved in activating compound constituents because these constituents are highly frequent and have low thresholds.

Another question that might be raised with respect to this interpretation is that if the decomposition route is the only route at work in transparent compound processing, why did the current studies show correlation of whole word frequency and response time to compound words? One might argue, however, that whole word frequency is not an index of direct access in transparent compound processing. Rather, it might be an index of the decomposition process. In specific, the association of whole word frequency and response time might be evidence that the more the transparent compounds are encountered and accessed via the decomposition route, the easier it becomes and the faster the process. That is, the role of the whole word frequency in transparent compound processing is not an index of whole word retrieval but an index of the ease of the decomposition process.

Based on this interpretation for the advantage of transparent compound processing, it naturally follows that whole word frequency plays a different role in opaque compound processing because the decomposition route cannot compute the legitimate meaning for opaque compounds and word recognition should be eventually determined by direct retrieval of whole word representations. Thus, the association of whole word frequency and response time might be interpreted as evidence of direct access in opaque compound processing.

In addition to the alternative interpretation for the advantage of transparent compound processing, there are at least two alternative interpretations for the disadvantage of opaque compound processing found in Experiments 5-7. First, it is possible that evidence of disadvantage of opaque compound processing in these experiments is not due to the mutual inhibition between the computed and retrieved meanings. Rather, the disadvantage is due to the fact that two possible interpretations are activated at the two access routes simultaneously and lead to ambiguity in terms of the intended meaning of the compound. To resolve this "ambiguity", participants have to consult their memory and experiences for previous usage of a particular opaque compound. Based on these previous experiences, participants eventually decide that the computed meaning is implausible because they cannot recall that they have ever interpreted *jailbird* as "a bird that lives in jail". As a result, the computed meaning has to be suppressed based on previous experiences of the compound word. Hence, the higher the constituent frequency, the slower the response to opaque compounds; the computed meaning should receive more activation when the constituents are highly frequent and should be more difficult to suppress. Importantly, this process of ambiguity resolution in consulting previous experiences might slow down the recognition of opaque compounds and contribute to the disadvantage of opaque compound processing.

Alternatively, it is also possible that evidence of disadvantage in opaque compound processing is due to divided cognitive resource for two conflicting and competing interpretations activated at the two access routes. Consequently, the accumulated activation for both the computed meaning and the retrieved meaning are reduced. This should make the processing of opaque compounds more difficult than transparent compounds because in the latter case, the two access routes activate consistent interpretation which should not introduce competition for cognitive resource. Eventually, the computed illegitimate meaning should be suppressed by the system based on previous experiences of the compound word. When the constituents have high frequencies, the computed meaning should be more strongly activated. Consequently, it should be more difficult to suppress the computed meaning and make the response to opaque compounds slower.

Future research

The time course of compound word processing. The current studies used lexical decision task to explore the influence of morphological structure on the processing speed of compound words. The findings of the involvement of constituent frequency in both transparent and opaque compound processing indicate that both the constituents of transparent and opaque compounds are activated initially although the activation might be eventually inhibited in opaque compound processing. These findings are compatible with a theoretical framework that assumes early automatic decomposition in compound processing because late decomposition models (see for example, Caramazza et al., 1988) assume that morphological decomposition only occurs after the failure of the direct access which is the default route of complex word processing. Such a model predicts that morphological decomposition should not occur in real word processing (all real words should be represented in the mental lexicon), and thus cannot accommodate the findings of morphological decomposition in the processing of both types of compounds in the current studies.

The assumption of early decomposition has recently received support from measurements that are more sensitive to the time course of word recognition such as eye movement measurements and magnetoencephalography (MEG) measurements. For example, studies using eye movement measurements found that constituent frequency of compound words influenced first fixation duration which was a measure that was sensitive to initial stage of word processing (e.g., Andrews et al., 2004; Hyönä & Pollatsek, 1998; Juhasz et al., 2003; Pollatsek et al., 2000; Pollatsek & Hyönä, 2005). In addition, Fiorentino and Poeppel (2007) found earlier peak of amplitude in compound processing than in monomorphemic word processing in a MEG component at 350 ms time window that was sensitive to early stage of word processing (e.g., Beretta, Fiorentino, & Poeppel, 2005; Pylkkänen, Llinás, & Murphy, 2006).

Future research aims to depict the time course of compound word processing might focus on the time course of the semantic transparency effect and the potential composition stage for transparent compound processing. It should be noted that I am not promoting a sequential model for compound processing here. Early decomposition does not mean that morphological decomposition precedes the direct access. Based on available data about the time course of compound word processing mentioned above, one can only rule out the possibility for late decomposition but nothing beyond that.

The two access routes in processing other types of complex words. At this point, although my proposed framework can explain the advantage/disadvantage for compound word processing, one might ask whether it can explain the disadvantage for Finnish case inflectional forms (Bertram et al., 1999; Laine et al., 1999). It seems that the sketched framework cannot readily explain the disadvantage for Finnish case inflectional forms. The reason is that one should expect an advantage for Finnish case inflectional forms by assuming facilitative links between the parts and the whole for these complex forms, which is contradictory to the findings of the disadvantage. This suggests that no facilitative links between the parts and the whole should be assumed for Finnish case inflectional forms. Thus, it appears that just as proposed by Laine et al., the decomposition route is the only route at work in the processing of Finnish case inflectional forms, and the computational cost in this process slows down word recognition. The reason that only the decomposition route is involved is probably due to the fact that these Finnish inflectional forms are extremely productive and thus it might be more parsimonious to store only the stems in the mental lexicon and to compute the meanings of the whole word forms online. This interpretation also implies that the balance of computation and storage might take different patterns for different types of complex words. The disadvantage for the processing of ambiguous Finnish inflectional forms, however, can be explained by extending the proposed inhibition mechanism for opaque compound processing here. When two possible interpretations are activated simultaneously in the processing of these ambiguous inflectional forms, they inhibit each other and slow down the process. Although most of the time the direct access wins over, the inhibited direct access should be slower than the non-inhibited direct access in monomorphemic word processing. Such an explanation is equivalent to the inhibition mechanism proposed by Laine et al. except that the proposed inhibition mechanism is not a universal mechanism in the current framework and does not apply to the processing of transparent compounds.

The proposed framework can also accommodate the advantage for derivational complex word processing over monomorphemic word processing (Bertram, et al., 1999; Hudson & Buijs, 1995). One can assume facilitative links between the parts and the whole for derivational complex words so that derivational forms can get facilitative activation from their stems as well as their suffixes in a similar way as compounds. In a recent study, Ji (unpublished raw data) found that English derivational complex words were processed more quickly than frequency-matched monomorphemic words but slower than frequency-matched transparent compounds. This finding suggests that derivational forms might not receive as much facilitation from their suffixes as from their stems.

Further discussion of the advantage/disadvantage in the processing of different types of complex words is certainly beyond the scope of this dissertation. Future research aims to study the balance and interplay of the two access routes in processing different types of complex words in different kinds of languages will help examine the generaliability and limitation of my proposal of the flexible relationship of the two access routes in compound processing.

Concluding remarks

The findings of the current studies indicate that the availability of morphological units introduced by complex structure aids rather than hinders transparent compound processing. In contrast, the availability of morphological units may or may not aid opaque compound processing depending on whether the decomposition route is encouraged to compute a meaning for the whole. The source of the advantage for transparent compound processing might be located at the lexical level as well as the semantic level, whereas the source of the advantage for opaque compound processing might be mainly located at the lexical level. This interpretation receives further support from the finding that the advantage for opaque compounds disappears when morphological decomposition is reinforced and there is even evidence of negative influence of constituent frequency. In addition, the accuracy data indicate a disadvantage for opaque compounds in Experiment 5. These findings indicate that the benefit of morphological decomposition from the lexical level can be traded off by the disadvantage due to inhibition across the semantic and conceptual levels when the decomposition route is enhanced in opaque compound processing; the decomposition route might compute a meaning that inhibits the meaning retrieved from the stored representation and slow down word recognition.

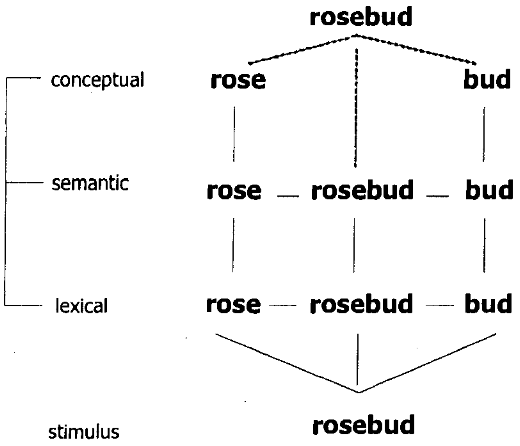
The findings in the current studies can be explained within a dual-route framework that assumes communication of morphological decomposition and whole word retrieval. This framework also assumes that the relationship of morphological decomposition and whole word retrieval is flexible. In transparent compound processing, the two access routes are mutually facilitative because they activate compatible information and boost each other throughout the process. In opaque compound processing, the relationship of the two access routes is more complicated because they are mutually facilitative within the lexical level but mutually inhibitory across the semantic and conceptual levels. Whether there is a benefit for overall processing time depends on whether the decomposition route is boosted to encourage meaning composition.

The findings of the current studies also provide a new perspective in viewing the balance of storage and computation in complex word processing. It is commonly assumed that morphological decomposition in complex word processing involves computation that is time consuming. The advantage for transparent compound processing over frequency-matched monomorphemic word processing, however, implies that even if morphological decomposition per se is time consuming, the overall processing time for complex words is not necessarily longer. Time might be saved from other sources such as the constituents of complex words that have low activation threshold (high frequency) at both the lexical level and the semantic level. Indeed, as mentioned earlier, in the entire CELEX database, I found only about five compounds whose constituent frequencies are equal or less than their whole word frequencies. This means that for the majority of English compounds, their constituent frequencies are higher than whole word frequencies. This property of English compounds might make it easier to access compound constituents and speed up the process at the decomposition route. More importantly, lexical and semantic units activated at the decomposition route and direct access increase the activation of each other through facilitative links. Thus, the balance of storage and computation in complex word processing might be more complicated than originally assumed (e.g., Schreuder & Baayen, 1995; Taft & Forster, 1975; 1976). That is, processing time can be reduced not only via the storage of whole words but also via the storage of constituents.

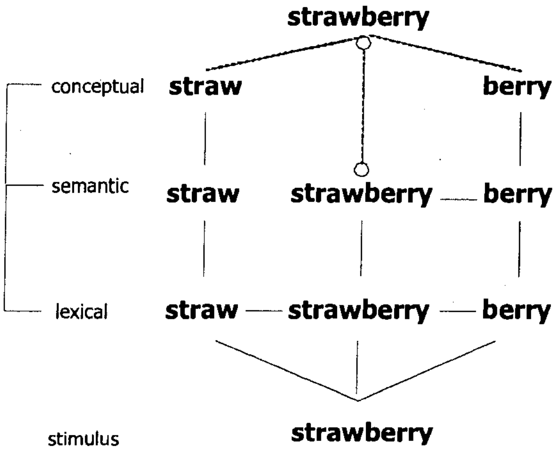
The finding of the negative influence of constituent frequency in opaque compound processing has important theoretical implications. First, this finding suggests that there is potential inhibition across the semantic and conceptual level between the two access routes in opaque compound processing, although the disadvantage due to this inhibition might be compensated by the benefit due to facilitation between the two access routes at the lexical level. Such a finding encourages models of complex word processing to depict the interplay of lexical, semantic and conceptual processing. If one only investigates and depicts compound word processing at the lexical level or semantic level, it would be impossible to examine how meaning composition interact with meaning retrieval differently in transparent and opaque compound processing and how these processes at the semantic and conceptual levels interact with the process at the lexical level. Second, this finding, together with the findings of the positive influence of constituent frequency in transparent compound processing, suggests that the relationship of the two access routes is flexible and sensitive to semantic transparency and the context in which compounds are processed. This interpretation encourages models of complex word processing to expand by considering the flexibility of the relationship of the two access routes.

Finally, the proposed framework assumes the involvement of a conceptual level which is more abstract than the semantic level and a component that implements the composition process in compound word processing. This is a theoretical exploration that aims to sketch a picture of the whole process of morphological decomposition. If compound word processing starts with morphological decomposition, it might end with morphological composition. It appears that assuming such a composition component provides more explanation power for the findings in the current research. This encourages models of compound word processing to take into consideration of the role of composition in compound word recognition and the consequences of this composition process (for research of this aspect see for example, Gagné & Shoben, 1997; Gagné & Spalding, 2004; Levi, 1978; Murphy, 1988; Sandra, 1994; Wisniewski, 1996).

Figure 1. *The proposed framework for compound recognition. Dotted lines indicate links that are connected to the constructed meaning at the conceptual level. Dark lines indicate links that are connected to the represented units. Links end with circles are inhibitory links.*



The processing of transparent compounds



The processing of opaque compounds

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. (2004). Statistics in Psycholinguistics: A critique of some current gold standards. In G. Libben & K. Nault (Eds.), *Mental Lexicon Working Papers in Psycholinguistics I*, Edmonton, (pp. 1–45).

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to Statistics*. New York: Cambridge University Press.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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

Bates, D. M. (2005) Fitting linear mixed models in R. *R News, 5*(1), 27-30.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy in lexical access: an MEG study. *Cognitive Brain Research, 24*, 57-65.

Bergman, M. W., Hudson, P. T., Eling, P. A. (1988). How simple complex words can be: Morphological processing and word representations. *The Quarterly Journal of Experimental Psychology, 40A*, 41-72.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language, 48,* 615-634.

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

Burani, C., & Laudanna, A. (1992) Units of representation for derived words in the lexicon. In R. Frost & L. Katz (Eds), *Orthography, phonology, morphology, and meaning*. (pp. 361-376). Oxford, England: North-Holland.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production, Vol. 2* (pp. 257-294). New York: Academic Press.

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

Dixon, P., (in press). Models of accuracy in repeated-measures designs. *Journal of Memory and Language.*

Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes, 22*, 1-48.

Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. Booij & J. v. Marle (Eds.), *Yearbook of morphology 1991* (pp. 165-183). Dordrecht, The Netherlands: Kluwer Academic.

Gagné, C. L., & Shoben, E. J. (1997). The influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 71-87.

Gagné, C. L., & Spalding, T. L. (2004). Effect of relation availability on the interpretation and access of familiar noun-noun compounds. *Brain and Language, 90*, 478-486.

Hudson, P. T. W., Buijs, D. (1995). Left-to-right processing of derivational morphology. In L. Feldman, (Ed.), *Morphological aspects of language processing* (pp. 383-396). Hillsdale, NJ: Erlbaum.

Hyönä, J., Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1612-1627.

Jarema, G., Busson, C., Nikolova, R.,Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain and Language, 94*, 362-369.

Ji, H. (2006). [The processing of derivational words and compound words]. Unpublished raw data.

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes, 20*, 291-316.

Juhasz, B. J. (2006). Effects of word length and sentence context on compound word recognition: An eye movement investigation. *Dissertation Abstracts International, 67*(4-B), 2252B. (UMI No. 3215917)

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 decisions and eye fixations. *British Journal of Psychology, 94*, 223-244.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English.* Providence, RI: Brown University Press.

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

Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Processes, 9*, 295-316.

Levi, J. N. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.

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.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101*, 3-33.

Medler, D.A., & Binder, J.R. (2005). MCWord: An On-Line Orthographic Database of the English Language. http://www.neuro.mcw.edu/mcword/

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language, Vol. 2* (pp. 147-195). Hillsdale, NJ: Lawrence Erlbaum

Associates, Inc.

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12*, 529-562.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. New York: Springer.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception & Performance, 26,* 820-833.

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes, 20*, 261-290.

Pylkkänen, L., Llinás, R., & Murphy, G. (2006). Representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience, 18*, 1-13.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 24*, 574-590.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 42,* 529-567.

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes, 9*, 227-269.

Sereno, J. A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory & Cognition, 25*, 425-437.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. E. Feldman (Ed.), *Morphological aspects of language processing.* (pp. 131-154). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

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

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

Taft, M. (1979). Lexical access via an orthographic code: The basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning & Verbal Behavior, 18*, 21-39.

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

van Jaarsveld, H. J., & Rattink, G. E. (1988). Frequency effects in the processing of lexicalized and novel compounds. *Journal of Psycholinguistic Research, 17*, 447-473.

Wisniewski, E. J. (1996). Construal and similarity in conceptual combination. *Journal of Memory and Language, 35*, 434-453.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes, 9*, 341-368.

Appendix I

Word frequencies, response times and accuracy for word stimuli in Experiments 1 and 2

Note. RT = response time; Acc = accuracy; Letter = number of letters; Syl = number of syllables; Surf = surface frequency; Lemma = lemma frequency; Tran = transparency rating; Mono = monomorphemic word.

		Experiment 1		**Experiment 2**						
No.	**Compound**	**RT**	**Acc**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**	**Tran**
1	eyewitness	734	96	829	96	10	3	0.95	1.12	4
2	newsletter	698	91	781	96	10	3	2.79	3.41	5.5
3	stepladder	879	87	900	84	10	3	0.73	0.78	6.1
4	goalkeeper	853	100	932	96	10	3	1.40	1.51	4.7
5	bandwagon	802	96	918	100	9	3	0.89	0.89	2.9
6	bartender	682	100	721	100	9	3	2.29	2.40	6
7	fairyland	744	100	838	88	9	3	0.73	0.78	5
8	handiwork	861	61	1084	60	9	3	1.90	1.90	4.4
9	northwest	730	100	825	100	9	2	0.73	0.73	6.5
10	oversight	694	91	841	96	9	3	1.06	1.12	4.2
11	paperback	723	96	786	100	9	3	3.63	5.53	3.8
12	racehorse	745	100	813	96	9	2	1.06	1.68	6
13	riverside	804	96	770	96	9	3	0.78	0.78	5.9
14	shipowner	965	87	1303	36	9	3	0.73	1.01	5.7
15	sunflower	613	100	745	100	9	3	1.12	1.73	3.7
16	tightrope	849	100	1073	92	9	2	1.23	1.34	5.9
17	waterhole	667	100	787	100	9	3	0.73	1.01	5.8
18	anteater	851	52	1109	92	8	3	0.61	1.01	4.5
19	baseball	693	96	721	100	8	2	6.15	6.15	3.8
20	billfold	805	70	967	40	8	2	0.84	0.84	3.6
21	bookshop	696	100	740	92	8	2	3.97	5.81	6
22	campfire	621	100	701	92	8	3	1.23	1.40	5.4
23	charcoal	729	91	784	100	8	2	5.36	5.36	3.9
24	darkroom	700	100	739	88	8	2	1.01	1.01	5.8
25	eggshell	689	96	722	100	8	2	1.45	1.73	6.5
26	flagpole	688	96	846	100	8	2	0.89	1.06	5.8
27	folklore	770	96	828	92	8	2	2.46	2.46	2
28	footwear	711	100	728	100	8	2	1.51	1.51	5.8
29	goatskin	762	87	931	76	8	2	0.78	1.01	5.9
30	haystack	644	96	958	92	8	2	0.95	1.17	6
31	hometown	700	96	653	100	8	2	0.95	1.06	6.7
32	madhouse	726	91	807	92	8	2	1.12	1.28	3.4
33	password	635	100	697	100	8	2	1.28	1.40	5.6
34	pussycat	721	91	758	92	8	3	0.61	0.61	3.5

35	showdown	717	96	862	100	8	2	2.29	2.29	1.7
36	starfish	636	96	791	100	8	2	0.89	0.89	4.9
37	daydream	693	96	832	96	8	2	1.28	2.74	5.3
38	archway	811	91	832	96	7	2	2.23	2.68	3.9
39	bathtub	693	100	678	100	7	2	1.73	1.90	6.3
40	bulldog	674	96	728	100	7	2	0.95	0.95	3.1
41	coconut	688	96	832	96	7	3	2.23	2.85	2.5
42	doormat	729	100	740	92	7	2	0.84	0.84	5.8
43	earring	758	91	928	96	7	2	0.89	3.30	6.5
44	gearbox	757	96	1113	68	7	2	0.61	0.84	4.8
45	hearsay	875	87	1018	88	7	2	1.17	1.17	3.3
46	oatmeal	709	96	669	100	7	2	2.07	2.07	5.5
47	pancake	672	100	679	100	7	2	1.34	2.57	3.5
48	pickaxe	1048	52	1131	60	7	2	0.73	0.84	4.6
49	plywood	734	96	804	92	7	2	1.68	1.68	3.8
50	popcorn	649	96	663	100	7	2	0.78	0.78	4.6
51	rosebud	717	96	812	96	7	2	1.23	1.73	5.3
52	skyline	660	91	808	100	7	2	3.18	3.35	3.9
53	teargas	935	78	958	84	7	2	1.12	1.12	4.6
54	topsoil	820	100	849	92	7	2	1.79	1.79	5.8
55	cobweb	781	91	1035	88	6	2	1.06	2.18	3.3
56	jigsaw	733	100	735	100	6	2	2.46	2.74	2.1
57	layman	741	83	1084	72	6	2	5.87	6.87	2.7
58	seabed	805	100	1005	80	6	2	1.34	1.34	3
59	teacup	705	100	791	96	6	2	1.17	1.79	6.3
60	pinup	832	87	1025	84	5	2	0.73	1.06	3.6
No.	**Mono**	**RT**	**Acc**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**	
61	guillotine	1032	74	1060	92	10	3	1.17	1.23	
62	silhouette	878	96	780	100	10	3	2.57	3.63	
63	tourniquet	1165	57	1093	60	10	3	0.61	0.61	
64	thermostat	800	96	932	96	10	3	1.40	1.62	
65	sophomore	1044	91	759	92	9	3	0.67	0.78	
66	badminton	761	96	696	100	9	3	2.51	2.51	
67	persimmon	1061	35	1246	36	9	3	0.61	0.73	
68	tarpaulin	931	39	1168	24	9	3	1.62	2.18	
69	turquoise	733	74	994	100	9	2	0.73	0.78	
70	narcissus	1131	48	1173	56	9	3	0.89	1.12	
71	crocodile	776	96	743	96	9	3	3.91	5.59	
72	scoundrel	946	83	882	100	9	2	1.23	1.68	
73	albatross	916	83	1058	84	9	3	0.78	0.78	
74	porcupine	776	96	842	96	9	3	0.73	1.01	
75	stratagem	1237	26	1301	24	9	3	1.12	1.68	
76	sphincter	1055	57	981	68	9	2	0.95	1.12	
77	mannequin	900	61	993	80	9	3	0.50	0.78	
78	sentinel	959	83	1094	72	8	3	0.61	1.17	
79	mackerel	910	78	1118	72	8	2	6.26	6.26	

80	mongoose	828	100	817	92	8	2	0.78	0.89
81	tortoise	782	74	863	92	8	2	3.97	5.64
82	carousel	843	96	806	100	8	3	1.23	1.34
83	porridge	936	74	1013	92	8	2	5.31	5.31
84	cashmere	877	87	941	96	8	2	1.17	1.17
85	throttle	711	96	730	96	8	2	1.68	1.84
86	plankton	877	91	947	92	8	2	1.06	1.06
87	billiard	805	100	786	100	8	2	2.57	2.57
88	ordnance	931	22	983	44	8	2	1.40	1.40
89	trombone	744	87	915	88	8	2	0.61	0.78
90	parlance	1445	26	1110	28	8	2	1.12	1.12
91	scaffold	845	83	862	96	8	2	1.17	1.28
92	jaundice	955	74	898	84	8	2	1.34	1.34
93	couscous	1090	48	1144	44	8	2	1.40	1.40
94	hibiscus	1048	65	1119	76	8	3	0.61	0.61
95	shrapnel	1010	57	952	76	8	2	2.18	2.18
96	rickshaw	991	70	1043	76	8	2	0.61	0.78
97	skirmish	924	83	948	88	8	2	1.51	2.46
98	treacle	897	22	1577	12	7	2	2.51	2.51
99	syringe	831	100	824	100	7	2	1.62	2.12
100	sputnik	1121	43	1240	72	7	2	0.84	1.01
101	panacea	1067	22	1314	8	7	3	2.07	2.57
102	gristle	1140	70	1010	72	7	2	0.89	0.89
103	bailiff	1302	17	1361	56	7	2	0.78	3.13
104	gentian	1787	4	1522	16	7	2	0.67	0.78
105	satchel	1056	74	1063	72	7	2	1.06	1.17
106	console	810	96	969	92	7	2	1.90	2.23
107	liqueur	984	74	981	56	7	2	1.40	2.51
108	parquet	1176	52	1169	44	7	2	0.78	0.78
109	gazelle	803	83	876	100	7	2	1.45	1.90
110	sucrose	969	96	880	96	7	2	0.78	0.78
111	giraffe	787	100	869	96	7	2	1.23	1.56
112	compost	787	91	888	100	7	2	3.24	3.24
113	chinook	841	91	849	100	7	2	0.89	1.06
114	hatchet	796	96	876	100	7	2	1.56	1.79
115	fresco	995	48	1165	52	6	2	0.95	2.35
116	anthem	920	87	817	92	6	2	2.46	2.74
117	guitar	681	96	698	100	6	2	5.70	6.65
118	herpes	859	91	824	100	6	2	1.34	1.34
119	falcon	792	96	780	100	6	2	1.23	1.73
120	bugle	957	65	1278	72	5	2	0.67	1.12

Appendix II

Word frequencies, response times and accuracy for word stimuli in Experiment 3

Note. Cond = condition; com_hh = compound words with high-frequency first constituent and high-frequency second constituent; com_hl = compound words with high-frequency first constituent and low-frequency second constituent; com_lh = compound words with low-frequency first constituent and high-frequency second constituent; com_ll = compound words with low-frequency first constituent and low-frequency second constituent; RT = response time; Acc = accuracy; Letter = number of letters; Syl = number of syllables; Surf = surface frequency; Lemma = lemma frequency; C1 = lemma frequency of first constituent; C2 = lemma frequency of second constituent; Tran = transparency rating; Mono = monomorphemic word.

Cond	No	Compound	RT	Acc	Letter	Syl	Surf	Lemma	C1	C2	Tran
com_hh	1	afterlife	748	85	9	3	1.68	1.68	1157	834	5.58
com_hh	2	airbed	814	74	6	2	1.06	1.06	253	273	4.81
com_hh	3	banknote	834	96	8	2	0.11	0.78	178	176	4.81
com_hh	4	bearskin	850	96	8	2	0.22	0.34	123	106	6.19
com_hh	5	bluebird	690	93	8	2	0.34	0.45	132	103	6.54
com_hh	6	bottleneck	747	100	10	3	0.45	1.06	122	82	4.46
com_hh	7	brainwave	764	96	9	2	0.28	0.34	75	117	3.08
com_hh	8	campsite	658	100	8	2	0.84	0.95	96	83	6.12
com_hh	9	casebook	814	78	8	2	0.22	0.50	496	451	4.19
com_hh	10	checklist	710	96	9	2	0.22	0.28	114	114	5.85
com_hh	11	clothesline	799	85	11	2	0.67	0.84	124	316	5.85
com_hh	12	cornerstone	842	93	11	3	1.34	1.62	126	123	4.35
com_hh	13	eggcup	741	78	6	2	0.34	0.50	86	81	5.19
com_hh	14	footfall	745	63	8	2	0.22	0.56	331	343	3.50
com_hh	15	freestyle	742	93	9	2	0.50	0.50	223	110	5.12
com_hh	16	godson	832	93	6	2	1.28	1.28	239	206	2.50
com_hh	17	heartbreak	638	93	10	2	0.73	0.73	164	259	4.04
com_hh	18	mainstay	773	74	8	2	1.34	1.51	193	254	3.00
com_hh	19	matchbox	645	93	8	2	0.73	0.84	88	104	6.38
com_hh	20	quarterfinal	807	89	12	4	0.17	0.67	65	116	5.12
com_hh	21	racehorse	708	89	9	2	1.06	1.68	109	133	6.58
com_hh	22	showplace	745	89	9	2	0.22	0.28	538	743	5.23
com_hh	23	starfish	664	100	8	2	0.89	0.89	106	195	4.73
com_hh	24	postmark	781	96	8	2	0.22	0.34	94	104	4.96

com_hh	25	turntable	834	93	9	3	0.78	0.89	683	236	5.23
com_hh	26	dogleg	965	59	6	2	0.11	0.11	119	176	3.65
com_hh	27	milkshake	724	96	9	2	0.39	0.45	110	136	4.15
com_hh	28	playsuit	697	81	8	2	0.06	0.17	542	99	3.96
com_hh	29	meatball	685	100	8	2	0.17	0.28	75	112	6.81
com_hh	30	countrywoman	758	93	12	4	0.28	0.34	589	851	5.77
com_hl	31	bloodlust	816	78	9	2	0.28	0.28	143	10	4.08
com_hl	32	grassroots	729	96	10	2	0.89	0.89	91	7	3.27
com_hl	33	hairpin	703	93	7	2	1.12	1.34	200	32	5.27
com_hl	34	handcuff	679	100	8	2	0.06	0.28	796	8	5.19
com_hl	35	earlobe	737	93	7	2	0.56	0.61	88	4	5.85
com_hl	36	heatstroke	865	93	10	2	0.22	0.22	138	44	3.54
com_hl	37	landslide	754	100	9	2	1.34	1.45	288	47	5.88
com_hl	38	lifebuoy	911	56	8	2	0.17	0.17	834	1	3.92
com_hl	39	nameplate	849	81	9	2	0.22	0.34	394	56	4.65
com_hl	40	nightshirt	777	100	10	2	1.12	1.28	464	61	5.19
com_hl	41	nosebleed	695	89	9	2	0.34	0.61	83	24	6.54
com_hl	42	paintbrush	723	96	10	2	0.56	0.78	98	40	6.31
com_hl	43	pickaxe	848	52	7	2	0.73	0.84	191	9	4.46
com_hl	44	pushcart	724	93	8	2	0.17	0.45	137	14	5.92
com_hl	45	seagull	681	100	7	2	1.01	1.51	173	5	4.15
com_hl	46	shipmate	726	93	8	2	0.06	0.22	77	30	4.96
com_hl	47	shoelace	701	96	8	2	0.34	1.23	80	17	5.69
com_hl	48	snowflake	624	100	9	2	0.28	1.62	66	11	6.46
com_hl	49	stepladder	846	85	10	3	0.73	0.78	160	16	6.46
com_hl	50	sunbeam	719	96	7	2	0.56	0.78	153	23	5.96
com_hl	51	teakettle	848	74	9	3	0.17	0.17	91	12	6.35
com_hl	52	toothache	689	89	9	2	1.28	1.40	88	17	6.46
com_hl	53	doorknob	675	96	8	2	1.01	1.45	387	7	6.42
com_hl	54	topsoil	768	93	7	2	1.79	1.79	257	53	6.00
com_hl	55	wastebasket	784	89	11	3	0.95	1.06	103	24	6.42
com_hl	56	watchtower	893	89	10	3	0.22	0.67	290	61	5.35
com_hl	57	watermelon	668	100	10	4	0.28	0.34	470	3	3.38
com_hl	58	windpipe	740	100	8	2	0.45	0.45	138	35	4.00
com_hl	59	goldfinch	908	67	9	2	0.06	0.17	88	1	3.42
com_hl	60	wildfowl	874	78	8	2	1.06	1.06	94	3	5.38
com_lh	61	busybody	841	89	8	4	1.06	1.51	61	364	3.77
com_lh	62	madhouse	784	93	8	2	1.12	1.28	49	621	4.00
com_lh	63	lamplight	676	89	9	2	1.84	1.84	35	406	6.54
com_lh	64	pinewood	723	96	8	2	0.67	0.84	18	97	5.88
com_lh	65	wormhole	744	89	8	2	0.45	1.34	18	94	4.88
com_lh	66	barnyard	761	96	8	2	0.45	0.61	13	87	5.42
com_lh	67	ragbag	707	59	6	2	0.50	0.50	13	82	5.15

com_lh	68	alleyway	804	78	8	3	0.50	1.06	12	1345	4.96
com_lh	69	anthill	905	74	7	2	0.28	0.34	12	119	5.65
com_lh	70	salesgirl	769	85	9	2	0.39	0.45	11	439	5.88
com_lh	71	knitwear	842	70	8	2	0.39	0.39	9	251	5.58
com_lh	72	cloakroom	799	89	9	2	1.12	1.34	9	517	5.00
com_lh	73	broomstick	754	93	10	2	0.34	0.89	8	133	6.35
com_lh	74	choirboy	786	74	8	3	0.28	0.73	7	359	6.15
com_lh	75	skateboard	667	100	10	2	0.45	0.84	7	112	4.46
com_lh	76	cloverleaf	790	89	10	3	0.11	0.11	6	83	5.35
com_lh	77	dewdrop	865	85	7	2	0.06	0.11	5	175	6.12
com_lh	78	masthead	1007	70	8	2	0.39	0.39	3	560	3.38
com_lh	79	sandpaper	773	96	9	3	0.34	0.34	59	227	5.31
com_lh	80	flashpoint	842	81	10	2	0.22	0.45	48	567	4.00
com_lh	81	witchdoctor	765	89	11	3	0.67	1.06	32	186	3.73
com_lh	82	cornfield	798	96	9	2	0.56	1.01	26	197	6.54
com_lh	83	forklift	721	85	8	2	0.28	0.39	18	110	3.69
com_lh	84	dyestuff	994	26	8	2	0.22	0.22	13	86	3.04
com_lh	85	harelip	1022	30	7	2	0.06	0.11	9	78	1.81
com_lh	86	ferryboat	674	81	9	3	0.56	0.61	11	78	5.35
com_lh	87	latchkey	891	78	8	2	0.28	0.28	4	89	4.62
com_lh	88	cheerleader	673	100	11	3	0.39	0.89	21	144	6.04
com_lh	89	panhandle	682	85	9	3	0.50	0.50	27	86	3.04
com_lh	90	inkwell	904	70	7	2	0.50	1.01	10	1638	5.19
com_ll	91	beefsteak	807	85	9	2	0.28	0.34	17	12	6.42
com_ll	92	candlewick	757	78	10	3	0.39	0.39	16	4	5.54
com_ll	93	thunderbolt	759	96	11	3	1.23	1.62	16	18	4.31
com_ll	94	foxhound	692	85	8	2	0.11	0.56	16	9	4.19
com_ll	95	haystack	818	96	8	2	0.95	1.17	15	23	6.46
com_ll	96	frostbite	685	89	9	2	0.45	0.45	12	42	3.12
com_ll	97	hovercraft	840	96	10	3	0.78	0.78	12	19	4.62
com_ll	98	chinaware	875	74	9	3	0.22	0.22	12	2	3.23
com_ll	99	innkeeper	774	96	9	3	0.39	0.50	11	9	5.85
com_ll	100	scarecrow	760	100	9	2	0.78	1.01	11	8	4.92
com_ll	101	dairymaid	780	85	9	3	0.28	0.39	11	17	4.35
com_ll	102	peppermint	669	93	10	3	1.06	1.28	11	7	3.31
com_ll	103	grapevine	731	100	9	2	1.12	1.28	10	6	4.77
com_ll	104	videotape	796	100	9	4	0.39	0.39	7	37	5.35
com_ll	105	candyfloss	826	81	10	3	0.11	0.11	7	1	2.73
com_ll	106	oxcart	816	63	6	2	0.06	0.11	6	14	4.69
com_ll	107	corkscrew	785	93	9	2	1.23	1.34	6	28	5.27
com_ll	108	sawmill	719	100	7	2	0.34	0.50	5	21	4.08
com_ll	109	pawnbroker	889	78	10	3	0.61	1.12	5	6	3.04
com_ll	110	toenail	723	89	7	2	0.45	1.56	30	31	6.38

com_ll	111	beanstalk	741	96	9	2	0.34	0.34	22	17	5.77
com_ll	112	honeycomb	709	100	9	3	0.84	1.01	21	14	3.31
com_ll	113	gumboot	886	52	7	2	0.06	0.28	10	40	2.23
com_ll	114	hearthrug	1028	26	9	2	0.34	0.39	5	15	4.12
com_ll	115	pallbearer	878	85	10	3	0.06	0.39	3	5	2.96
com_ll	116	wingspan	816	89	8	2	0.34	0.39	59	10	6.08
com_ll	117	moonshine	703	93	9	2	0.95	0.95	59	38	2.65
com_ll	118	fleshpot	879	63	8	2	0.06	0.39	52	37	2.58
com_ll	119	tightrope	783	93	9	2	1.23	1.34	49	44	5.08
com_ll	120	cowshed	903	67	7	2	1.01	1.12	42	28	5.77

Cond	**No**	**Mono**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**
Mono	121	scoundrel	836	89	9	2	1.23	1.68
Mono	122	guillotine	1077	89	10	3	1.17	1.23
Mono	123	sacrosanct	1030	26	10	3	0.89	0.89
Mono	124	sphincter	1004	78	9	2	0.95	1.12
Mono	125	plankton	835	89	8	2	1.06	1.06
Mono	126	gristle	933	70	7	2	0.89	0.89
Mono	127	sophomore	872	93	9	3	0.67	0.78
Mono	128	turquoise	677	100	9	2	0.73	0.78
Mono	129	trombone	824	93	8	2	0.61	0.78
Mono	130	rickshaw	887	85	8	2	0.61	0.78
Mono	131	plantain	962	48	8	2	0.45	0.73
Mono	132	marguerite	1239	44	10	3	0.50	0.73
Mono	133	persimmon	1088	48	9	3	0.61	0.73
Mono	134	Sanskrit	966	37	8	2	0.61	0.61
Mono	135	tourniquet	1010	52	10	3	0.61	0.61
Mono	136	hibiscus	934	59	8	3	0.61	0.61
Mono	137	vermouth	1038	48	8	2	0.56	0.56
Mono	138	riffraff	1139	48	8	2	0.56	0.56
Mono	139	mandolin	888	74	8	3	0.56	0.61
Mono	140	nautilus	905	59	8	3	0.56	0.61
Mono	141	pancreas	780	93	8	2	0.50	0.50
Mono	142	ginseng	917	81	7	2	0.50	0.50
Mono	143	trapeze	916	85	7	2	0.45	0.50
Mono	144	protegee	1042	63	8	3	0.45	0.45
Mono	145	limerick	855	93	8	3	0.39	0.39
Mono	146	camphor	832	22	7	2	0.39	0.39
Mono	147	tungsten	861	78	8	2	0.39	0.39
Mono	148	bludgeon	947	59	8	2	0.39	0.39
Mono	149	dandruff	800	100	8	2	0.34	0.34
Mono	150	pastrami	899	59	8	3	0.34	0.34

Appendix III

Word frequencies, response times and accuracy for word stimuli in Experiment 4

Note. Cond = condition; Transp = transparent compounds; Mono = monomorphemic words; RT = response time; Acc = accuracy; Letter = number of letters; Syl = number of syllables; Surf = surface frequency; Lemma = lemma frequency; Tran = transparency rating.

No	**Cond**	**Word**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**	**Tran**
1	Opaque	magpie	901	81	6	2	0.73	1.06	1.56
2	Opaque	mayfly	972	78	6	2	0.67	0.78	3.00
3	Opaque	linseed	1001	33	7	2	1.06	1.06	3.11
4	Opaque	backlog	941	48	7	2	1.01	1.01	3.22
5	Opaque	cockpit	756	93	7	2	3.30	3.30	2.00
6	Opaque	headway	800	93	7	2	1.90	1.90	3.33
7	Opaque	lawsuit	735	100	7	2	1.06	1.56	3.56
8	Opaque	pigtail	749	100	7	2	0.78	1.90	3.44
9	Opaque	wedlock	892	81	7	2	1.40	1.40	3.89
10	Opaque	alderman	1006	81	8	3	1.62	2.01	3.56
11	Opaque	wormwood	922	81	8	2	0.67	0.67	3.00
12	Opaque	ironwork	756	93	8	3	1.23	1.23	6.22
13	Opaque	namesake	779	70	8	2	0.61	0.61	3.11
14	Opaque	pullover	707	93	8	3	2.23	2.57	5.67
15	Opaque	smallpox	768	100	8	2	1.51	1.51	4.00
16	Opaque	trapdoor	764	96	8	2	0.89	1.01	5.00
17	Opaque	undertow	869	81	8	3	0.84	0.84	4.67
18	Opaque	honeycomb	747	100	9	3	0.84	1.01	4.33
19	Opaque	quicksand	756	100	9	2	0.61	0.78	5.22
20	Opaque	scarecrow	797	96	9	2	0.78	1.01	5.22
21	Opaque	peppermint	717	96	10	3	1.06	1.28	4.22
22	Opaque	turtleneck	775	93	10	3	1.34	1.51	3.78
23	Opaque	honeysuckle	787	96	11	4	1.62	1.62	3.33
24	Opaque	hollyhock	992	41	9	3	0.39	0.73	3.11
25	Opaque	dumbbell	830	93	8	2	0.34	0.50	2.33
26	Opaque	hogwash	889	85	7	2	0.11	0.11	3.00
27	Opaque	ragtime	892	78	7	2	0.50	0.50	2.56
28	Opaque	jailbird	847	100	8	2	0.17	0.17	3.22
29	Opaque	heatwave	732	85	8	2	0.34	0.45	6.22
30	Opaque	hearsay	952	78	7	2	1.17	1.17	2.89
31	Transp	payday	682	96	6	2	0.89	0.89	6.44
32	Transp	airbed	846	81	6	2	1.06	1.06	5.67
33	Transp	teargas	1006	81	7	2	1.12	1.12	6.33
34	Transp	keyhole	695	96	7	2	0.78	1.12	6.67
35	Transp	sunrise	763	96	7	2	2.85	2.91	6.78

36	Transp	oatmeal	726	100	7	2	2.07	2.07	6.67
37	Transp	rosebud	742	100	7	2	1.23	1.73	6.33
38	Transp	soybean	716	93	7	2	1.12	1.96	5.56
39	Transp	hairpin	715	100	7	2	1.12	1.34	5.44
40	Transp	eggshell	679	93	8	2	1.45	1.73	6.44
41	Transp	starfish	722	100	8	2	0.89	0.89	5.33
42	Transp	newsroom	696	96	8	2	1.28	1.34	6.67
43	Transp	inkstand	1046	70	8	2	0.67	0.67	4.67
44	Transp	bathrobe	701	100	8	2	2.23	2.40	6.00
45	Transp	footwear	726	100	8	2	1.51	1.51	6.56
46	Transp	wildfowl	888	63	8	2	1.06	1.06	4.11
47	Transp	flagpole	705	93	8	2	0.89	1.06	7.00
48	Transp	notepaper	702	89	9	3	1.12	1.12	6.78
49	Transp	riverside	822	96	9	3	0.78	0.78	6.11
50	Transp	shipwreck	734	96	9	2	0.61	0.73	6.22
51	Transp	pocketbook	774	100	10	3	1.01	1.01	5.67
52	Transp	goalkeeper	824	93	10	3	1.40	1.51	6.11
53	Transp	stomachache	952	52	11	3	1.40	1.79	6.56
54	Transp	innkeeper	905	93	9	3	0.39	0.50	5.78
55	Transp	dunghill	1175	48	8	2	0.34	0.34	3.78
56	Transp	handgun	721	96	7	2	0.11	0.22	6.00
57	Transp	dustpan	724	81	7	2	0.45	0.50	6.44
58	Transp	meatball	707	100	8	2	0.17	0.28	6.78
59	Transp	gatepost	931	89	8	2	0.34	0.61	6.22
60	Transp	seagull	737	100	7	2	1.01	1.51	4.11
61	Mono	poodle	677	100	6	2	0.73	1.12	
62	Mono	larynx	1023	85	6	2	0.78	0.78	
63	Mono	sputnik	1309	44	7	2	0.84	1.01	
64	Mono	gristle	1044	56	7	2	0.89	0.89	
65	Mono	tempest	869	89	7	2	3.07	3.30	
66	Mono	phoenix	770	100	7	2	1.90	1.96	
67	Mono	giraffe	780	96	7	2	1.23	1.56	
68	Mono	scallop	770	96	7	2	0.95	1.84	
69	Mono	jasmine	858	89	7	2	1.40	1.40	
70	Mono	throttle	657	89	8	2	1.68	1.84	
71	Mono	mongoose	951	89	8	2	0.78	0.89	
72	Mono	charisma	940	74	8	3	1.34	1.34	
73	Mono	trombone	1027	70	8	2	0.61	0.78	
74	Mono	graffiti	924	96	8	3	2.46	2.46	
75	Mono	pantheon	913	26	8	2	1.62	1.62	
76	Mono	scaffold	907	93	8	2	1.17	1.28	
77	Mono	broccoli	700	96	8	3	0.89	0.89	
78	Mono	porcupine	798	96	9	3	0.73	1.01	
79	Mono	sophomore	980	93	9	3	0.67	0.78	
80	Mono	turquoise	832	100	9	2	0.73	0.78	
81	Mono	guillotine	1060	81	10	3	1.17	1.23	
82	Mono	thermostat	897	96	10	3	1.40	1.62	
83	Mono	cholesterol	829	96	11	4	1.40	1.40	

84	Mono	mannequin	1092	85	9	3	0.50	0.78
85	Mono	dandruff	738	100	8	2	0.34	0.34
86	Mono	raccoon	779	100	7	2	0.28	0.34
87	Mono	cheetah	824	96	7	2	0.56	0.67
88	Mono	nocturne	1014	44	8	2	0.28	0.39
89	Mono	vermouth	1019	59	8	2	0.56	0.56
90	Mono	lozenge	1024	37	7	2	0.78	1.28

Appendix IV

Word frequencies, response times and accuracy for word stimuli in Experiments 5 and 6

Note. Word stimuli were given the way they were presented in Experiment 6 when spaces were inserted; Transp = transparent compounds; Mono = monomorphemic words; RT = response time; Acc = accuracy; Letter = number of letters; Syl = number of syllables; Surf = surface frequency; Lemma = lemma frequency; Tran = transparency rating.

		Experiment 5		**Experiment 6**						
No	**Opaque**	**RT**	**Acc**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**	**Tran**
1	bill fold	883	48	933	63	8	2	0.84	0.84	3.68
2	pussy cat	783	97	768	100	8	3	0.61	0.61	3.76
3	ear shot	695	84	847	95	7	2	2.63	2.63	2.73
4	base ball	671	97	707	95	8	2	6.15	6.15	3.86
5	coco nut	724	97	871	100	7	3	2.23	2.85	2.70
6	folk lore	823	97	764	89	8	2	2.46	2.46	3.49
7	cob web	846	90	787	79	6	2	1.06	2.18	2.59
8	band wagon	862	100	856	89	9	3	0.89	0.89	2.62
9	show down	848	97	974	100	8	2	2.29	2.29	2.59
10	pawn broker	1050	48	851	53	10	3	0.61	1.12	3.81
11	moon shine	776	100	828	100	9	2	0.95	0.95	3.27
12	life buoy	1184	39	1265	32	8	2	0.17	0.17	3.89
13	fork lift	752	90	890	95	8	2	0.28	0.39	3.86
14	pan handle	844	90	871	95	9	3	0.50	0.50	3.32
15	flesh pot	903	19	790	26	8	2	0.06	0.39	2.54
16	foot fall	736	61	868	74	8	2	0.22	0.56	3.19
17	god son	890	58	1005	79	6	2	1.28	1.28	2.65
18	may fly	958	58	1008	63	6	2	0.67	0.78	3.16
19	back log	1001	65	1049	84	7	2	1.01	1.01	2.65
20	cock pit	786	97	767	100	7	2	3.30	3.30	2.00
21	head way	842	94	952	100	7	2	1.90	1.90	3.03
22	law suit	707	100	787	95	7	2	1.06	1.56	2.92
23	pig tail	786	94	828	95	7	2	0.78	1.90	2.76
24	wed lock	908	77	923	84	7	2	1.40	1.40	3.70
25	worm wood	845	71	920	89	8	2	0.67	0.67	2.84
26	name sake	925	84	958	79	8	2	0.61	0.61	2.78
27	small pox	763	94	833	100	8	2	1.51	1.51	3.16
28	under tow	847	74	1020	79	8	3	0.84	0.84	3.89
29	honey comb	688	100	750	95	9	3	0.84	1.01	3.54
30	pepper mint	711	100	728	100	10	3	1.06	1.28	3.84
31	turtle neck	729	97	818	95	10	3	1.34	1.51	2.97
32	honey suckle	826	90	793	89	11	4	1.62	1.62	2.70
33	dumb bell	845	94	827	89	8	2	0.34	0.50	1.70
34	hog wash	874	71	975	79	7	2	0.11	0.11	2.08

35	rag time	802	65	863	63	7	2	0.50	0.50	1.86
36	jail bird	766	90	787	95	8	2	0.17	0.17	2.97

No	**Transp**	**RT**	**Acc**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**	**Tran**
37	home town	702	100	772	95	8	2	0.95	1.06	6.35
38	ant eater	964	77	879	100	8	3	0.61	1.01	5.76
39	arch way	841	90	882	100	7	2	2.23	2.68	4.84
40	man power	791	94	807	95	8	3	7.43	7.43	5.43
41	hair cut	700	97	675	95	7	2	2.35	3.07	6.24
42	eye sight	721	97	777	89	8	2	2.74	2.74	6.35
43	tea cup	683	100	642	95	6	2	1.17	1.79	6.22
44	fairy land	737	90	905	84	9	3	0.73	0.78	5.16
45	farm yard	738	87	748	100	8	2	2.29	2.51	6.14
46	hover craft	892	87	875	95	10	3	0.78	0.78	4.78
47	north west	715	97	786	95	9	2	0.73	0.73	6.51
48	bear skin	823	87	871	84	8	2	0.22	0.34	5.68
49	barn yard	769	100	746	100	8	2	0.45	0.61	5.35
50	ferry boat	816	81	961	95	9	3	0.56	0.61	5.03
51	fox hound	775	77	792	74	8	2	0.11	0.56	4.11
52	wind pipe	862	100	830	95	8	2	0.45	0.45	4.51
53	pay day	740	97	808	100	6	2	0.89	0.89	6.24
54	air bed	930	81	837	68	6	2	1.06	1.06	5.14
55	key hole	770	100	821	95	7	2	0.78	1.12	6.03
56	sun rise	724	97	758	89	7	2	2.85	2.91	6.46
57	oat meal	678	97	852	100	7	2	2.07	2.07	5.59
58	rose bud	792	100	776	100	7	2	1.23	1.73	5.22
59	soy bean	786	100	790	100	7	2	1.12	1.96	5.68
60	hair pin	699	97	720	95	7	2	1.12	1.34	5.97
61	star fish	693	100	741	89	8	2	0.89	0.89	4.97
62	ink stand	984	71	1156	47	8	2	0.67	0.67	4.11
63	foot wear	735	100	725	95	8	2	1.51	1.51	6.11
64	flag pole	766	97	785	100	8	2	0.89	1.06	6.24
65	note paper	812	90	822	95	9	3	1.12	1.12	5.81
66	pocket book	753	97	892	100	10	3	1.01	1.01	5.62
67	goal keeper	838	90	778	89	10	3	1.40	1.51	5.43
68	stomach ache	917	71	908	100	11	3	1.40	1.79	6.41
69	gate post	854	65	914	74	8	2	0.34	0.61	5.68
70	hand gun	711	100	757	95	7	2	0.11	0.22	5.59
71	dust pan	820	100	835	95	7	2	0.45	0.50	5.51
72	meat ball	673	94	728	100	8	2	0.17	0.28	6.46

No	**Mono**	**RT**	**Acc**	**RT**	**Acc**	**Letter**	**Syl**	**Surf**	**Lemma**
73	plank ton	895	84	1045	89	8	2	1.06	1.06
74	hibis cus	968	68	1138	68	8	3	0.61	0.61
75	ecli pse	858	97	770	100	7	2	2.46	2.74
76	mac kerel	953	68	1145	74	8	2	6.26	6.26
77	cri pple	834	97	736	100	7	2	2.07	2.79
78	bil liard	762	94	859	95	8	2	2.57	2.57
79	buc kle	737	94	714	95	6	2	1.01	1.90
80	sopho more	815	97	826	95	9	3	0.67	0.78

81	shrap nel	1069	58	977	63	8	2	2.18	2.18
82	tourni quet	1030	48	958	63	10	3	0.61	0.61
83	tur quoise	728	90	802	100	9	2	0.73	0.78
84	tung sten	998	58	1130	68	8	2	0.39	0.39
85	man dolin	898	74	1105	79	8	3	0.56	0.61
86	manne quin	889	81	898	79	9	3	0.50	0.78
87	lime rick	867	90	1263	47	8	3	0.39	0.39
88	pan creas	768	100	859	84	8	2	0.50	0.50
89	poo dle	676	100	767	100	6	2	0.73	1.12
90	la rynx	839	81	925	68	6	2	0.78	0.78
91	scar let	778	90	819	95	7	2	0.56	0.84
92	tem pest	817	81	1050	63	7	2	3.07	3.30
93	phoe nix	799	97	869	95	7	2	1.90	1.96
94	gira ffe	748	100	813	100	7	2	1.23	1.56
95	sca llop	762	94	826	100	7	2	0.95	1.84
96	jas mine	864	97	766	95	7	2	1.40	1.40
97	mon goose	854	74	882	89	8	2	0.78	0.89
98	trom bone	971	81	919	84	8	2	0.61	0.78
99	cash mere	780	87	945	89	8	2	1.17	1.17
100	bro ccoli	711	97	787	95	8	3	0.89	0.89
101	porcu pine	800	97	751	100	9	3	0.73	1.01
102	guillo tine	963	77	990	79	10	3	1.17	1.23
103	thermo stat	812	97	860	95	10	3	1.40	1.62
104	chole sterol	741	100	832	84	11	4	1.40	1.40
105	dan druff	788	97	812	95	8	2	0.34	0.34
106	rac coon	723	100	759	100	7	2	0.28	0.34
107	chee tah	768	90	770	95	7	2	0.56	0.67
108	chip munk	769	97	884	95	8	2	0.00	0.11

Appendix V

Word frequencies, response times and accuracy for word stimuli in Experiment 7

Note. Transp = transparent compounds; Mono = monomorphemic words; RT_B = response time of words whose first constituents were presented in black color and second constituents were presented in red color; Acc_B = accuracy of words whose first constituents were presented in black color and second constituents were presented in red color; RT_R = response time of words whose first constituents were presented in red color and second constituents were presented in black color; Acc_R = accuracy of words whose first constituents were presented in red color and second constituents were presented in black color; Letter = number of letters; Syl = number of syllables; Surf = surface frequency; Lemma = lemma frequency; Tran = transparency rating.

No	Opaque	RT_B	Acc_B	RT_R	Acc_R	Letter	Syl	Surf	Lemma	Tran
1	billfold	792	47	996	35	8	2	0.84	0.84	3.68
2	pussycat	682	100	757	100	8	3	0.61	0.61	3.76
3	earshot	768	82	827	88	7	2	2.63	2.63	2.73
4	baseball	610	100	692	88	8	2	6.15	6.15	3.86
5	coconut	697	100	650	100	7	3	2.23	2.85	2.70
6	folklore	817	71	778	100	8	2	2.46	2.46	3.49
7	cobweb	866	88	868	94	6	2	1.06	2.18	2.59
8	bandwagon	785	94	806	100	9	3	0.89	0.89	2.62
9	showdown	881	100	872	94	8	2	2.29	2.29	2.59
10	pawnbroker	977	53	871	24	10	3	0.61	1.12	3.81
11	moonshine	778	94	751	100	9	2	0.95	0.95	3.27
12	lifebuoy	1057	35	1038	47	8	2	0.17	0.17	3.89
13	forklift	837	94	838	82	8	2	0.28	0.39	3.86
14	panhandle	882	88	798	71	9	3	0.50	0.50	3.32
15	fleshpot	956	12	1293	24	8	2	0.06	0.39	2.54
16	footfall	623	35	773	41	8	2	0.22	0.56	3.19
17	godson	841	76	989	82	6	2	1.28	1.28	2.65
18	mayfly	821	35	1102	71	6	2	0.67	0.78	3.16
19	backlog	1071	53	856	47	7	2	1.01	1.01	2.65
20	cockpit	723	100	833	100	7	2	3.30	3.30	2.00
21	headway	914	88	814	88	7	2	1.90	1.90	3.03
22	lawsuit	802	100	747	100	7	2	1.06	1.56	2.92
23	pigtail	811	100	818	100	7	2	0.78	1.90	2.76
24	wedlock	894	76	959	71	7	2	1.40	1.40	3.70

25	wormwood	912	47	1056	41	8	2	0.67	0.67	2.84
26	namesake	980	82	970	71	8	2	0.61	0.61	2.78
27	smallpox	798	100	720	100	8	2	1.51	1.51	3.16
28	undertow	991	76	1092	71	8	3	0.84	0.84	3.89
29	honeycomb	773	100	713	100	9	3	0.84	1.01	3.54
30	peppermint	654	100	670	100	10	3	1.06	1.28	3.84
31	turtleneck	699	100	728	100	10	3	1.34	1.51	2.97
32	honeysuckle	754	88	708	82	11	4	1.62	1.62	2.70
33	dumbbell	847	94	885	88	8	2	0.34	0.50	1.70
34	hogwash	1010	82	950	76	7	2	0.11	0.11	2.08
35	ragtime	871	71	924	53	7	2	0.50	0.50	1.86
36	jailbird	820	100	802	100	8	2	0.17	0.17	2.97

No	**Transp**	**RT_B**	**Acc_B**	**RT_R**	**Acc_R**	**Letter**	**Syl**	**Surf**	**Lemma**	**Tran**
37	hometown	676	100	677	94	8	2	0.95	1.06	6.35
38	anteater	825	100	762	88	8	3	0.61	1.01	5.76
39	archway	940	76	927	82	7	2	2.23	2.68	4.84
40	manpower	757	88	743	94	8	3	7.43	7.43	5.43
41	haircut	644	100	634	94	7	2	2.35	3.07	6.24
42	eyesight	757	100	704	100	8	2	2.74	2.74	6.35
43	teacup	613	100	725	100	6	2	1.17	1.79	6.22
44	fairyland	779	76	809	82	9	3	0.73	0.78	5.16
45	farmyard	735	94	806	100	8	2	2.29	2.51	6.14
46	hovercraft	821	88	823	100	10	3	0.78	0.78	4.78
47	northwest	814	94	775	82	9	2	0.73	0.73	6.51
48	bearskin	745	100	820	82	8	2	0.22	0.34	5.68
49	barnyard	669	94	797	88	8	2	0.45	0.61	5.35
50	ferryboat	1066	65	865	76	9	3	0.56	0.61	5.03
51	foxhound	798	71	796	76	8	2	0.11	0.56	4.11
52	windpipe	859	94	762	100	8	2	0.45	0.45	4.51
53	payday	666	100	719	94	6	2	0.89	0.89	6.24
54	airbed	871	65	857	71	6	2	1.06	1.06	5.14
55	keyhole	773	100	712	100	7	2	0.78	1.12	6.03
56	sunrise	687	100	692	94	7	2	2.85	2.91	6.46
57	oatmeal	663	100	708	100	7	2	2.07	2.07	5.59
58	rosebud	755	100	763	100	7	2	1.23	1.73	5.22
59	soybean	748	94	697	94	7	2	1.12	1.96	5.68
60	hairpin	720	100	775	100	7	2	1.12	1.34	5.97
61	starfish	713	94	700	100	8	2	0.89	0.89	4.97
62	inkstand	1017	41	967	59	8	2	0.67	0.67	4.11
63	footwear	699	100	666	100	8	2	1.51	1.51	6.11
64	flagpole	774	100	709	100	8	2	0.89	1.06	6.24
65	notepaper	767	88	776	82	9	3	1.12	1.12	5.81
66	pocketbook	866	100	801	100	10	3	1.01	1.01	5.62
67	goalkeeper	883	100	822	94	10	3	1.40	1.51	5.43
68	stomachache	793	71	985	94	11	3	1.40	1.79	6.41
69	gatepost	911	47	833	41	8	2	0.34	0.61	5.68
70	handgun	745	100	738	100	7	2	0.11	0.22	5.59
71	dustpan	817	100	738	94	7	2	0.45	0.50	5.51

72	meatball	721	100	698	100	8	2	0.17	0.28	6.46

No	Mono	RT_B	Acc_B	RT_R	Acc_R	Letter	Syl	Surf	Lemma
73	plankton	776	88	847	88	8	2	1.06	1.06
74	hibiscus	1096	71	976	53	8	3	0.61	0.61
75	eclipse	745	88	767	100	7	2	2.46	2.74
76	mackerel	936	76	913	59	8	2	6.26	6.26
77	cripple	735	94	736	100	7	2	2.07	2.79
78	billiard	737	94	779	100	8	2	2.57	2.57
79	buckle	784	88	730	94	6	2	1.01	1.90
80	sophomore	912	88	788	88	9	3	0.67	0.78
81	shrapnel	856	47	1033	47	8	2	2.18	2.18
82	tourniquet	984	53	1102	29	10	3	0.61	0.61
83	turquoise	782	100	755	100	9	2	0.73	0.78
84	tungsten	920	82	968	53	8	2	0.39	0.39
85	mandolin	1033	65	854	82	8	3	0.56	0.61
86	mannequin	781	47	813	76	9	3	0.50	0.78
87	limerick	943	82	869	71	8	3	0.39	0.39
88	pancreas	790	94	810	100	8	2	0.50	0.50
89	poodle	715	100	669	94	6	2	0.73	1.12
90	larynx	813	88	812	100	6	2	0.78	0.78
91	scarlet	744	94	787	100	7	2	0.56	0.84
92	tempest	805	88	833	88	7	2	3.07	3.30
93	phoenix	785	94	780	100	7	2	1.90	1.96
94	giraffe	730	100	660	100	7	2	1.23	1.56
95	scallop	724	94	740	100	7	2	0.95	1.84
96	jasmine	803	100	723	94	7	2	1.40	1.40
97	mongoose	833	65	870	82	8	2	0.78	0.89
98	trombone	963	76	854	82	8	2	0.61	0.78
99	cashmere	917	94	772	88	8	2	1.17	1.17
100	broccoli	669	94	682	100	8	3	0.89	0.89
101	porcupine	714	100	693	100	9	3	0.73	1.01
102	guillotine	936	47	1021	53	10	3	1.17	1.23
103	thermostat	823	94	843	100	10	3	1.40	1.62
104	cholesterol	770	94	760	94	11	4	1.40	1.40
105	dandruff	807	94	730	94	8	2	0.34	0.34
106	raccoon	686	94	679	94	7	2	0.28	0.34
107	cheetah	732	94	768	100	7	2	0.56	0.67
108	chipmunk	720	100	710	100	8	2	0.00	0.11