A COMPARISON OF MORPHO-SEMANTIC COMPLEXITY IN COMPOUNDING: AKAN VS. JAMAICAN CREOLE

Annife Lance Campbell, Mr
ABSTRACT

A Comparison of Morpho-semantic Complexity in Compounding
Akan vs. Jamaican Creole

Annife Lance Campbell

The main purpose of the research reported here is to present findings to determine whether the noun + noun compounding system of Jamaican Creole is linguistically simpler or more complex than Akan.

To bring clarity to the matter, first, compound words were taken from two historical dictionaries representing both languages to create morpho-semantic compound databases. Secondly, a rubric of complexity was designed based on the semiotic notions of indexicality, transparency and diagrammaticity associated with Natural Morphology. Thirdly, using these semiotic principles as evaluative devices, the compounds were designated as simple, moderately complex or complex. Following this classification system, comparisons between the languages based on the following properties were carried out: assessing the number of compounds in each complexity category displaying a preference for a particular compound morphological pattern and identifying the number of semantic categories used to express compounds within each of these morphosemantic categories.

Based on the comparisons undertaken, the results indicated that Jamaican Creole displayed more morphosemantically complex compounds, preferred more complex morphological patterns to create compound words and expressed more semantically opaque categories than Akan. These results suggest that noun + noun compounding in Jamaican Creole is more complex than Akan.

Therefore, these findings support the observation that language complexity is not an entirely time dependent notion given the extent to which semantic opacification emerging from metaphoricity, exocentricity and folk etymology was abundant in Jamaican Creole, a much younger language, in comparison to Akan.

Keywords: Annife Lance Campbell; Jamaican Creole, Compound words; Language Complexity, Natural Morphology; Jamaican Creole morphology; Akan.
ACKNOWLEDGMENTS

First and foremost, I want to express gratitude to my supervisor, Dr. Michele Stewart, for her critical assessments of my ideas, scholarly exchanges on the topics explored and insightful recommendation that I should always make myself ‘savvy with the ideas’ before making any pronouncements. These factors have all contributed to making this thesis a reality.

Thanks are also due to Duku Osei for reviewing the database of Akan compound words to ensure accuracy. I am also indebted to the expertise of Wilton Baxter in helping to format the document and the extensive revisions in compiling the databases.

Lastly, I am grateful for the support from everyone who contributed in any way whatsoever to the completion of this project.
DEDICATION

I lovingly dedicate this thesis to my mother, Velora Lawson, for her support and patience throughout the writing of this thesis.
## TABLE OF CONTENTS

Abstract

i

Acknowledgements

ii

Dedication

iii

List of tables

vii

List of figures

ix

### Chapter 1: Introduction / Assessing complexity in grammars

1.1. Measuring overall complexity in Creole grammars

1

1.2. Weaknesses in the measurement of overall Grammatical complexity

2

1.3. If not overall complexity, then restricted complexity?

10

1.4. Research Questions

12

1.5. Value of an investigation into linguistic complexity

13

1.6. Summary of chapter one

15

1.7. Outline of the study

16

### Chapter 2: Towards a definition of complexity

2.1. What is language complexity?

17

2.2. Identifying an appropriate morphological theory to define absolute complexity

26

2.3. Natural Morphology and its relationship to Natural Phonology

40

2.4. Natural Morphology and its relationship to Markedness Theory

42

2.5. A basic framework of Natural Morphology

46
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6. A Natural Morphology account of Compound word formation</td>
<td>52</td>
</tr>
<tr>
<td>2.6.1. The principle of diagrammatic iconicity</td>
<td>53</td>
</tr>
<tr>
<td>2.6.2. The principle of indexicality</td>
<td>61</td>
</tr>
<tr>
<td>2.6.3. The principle of morphosemantic and morphotactic transparency</td>
<td>66</td>
</tr>
<tr>
<td>2.7. Creating a definition of complexity from Natural Morphology</td>
<td>69</td>
</tr>
<tr>
<td>2.8. Creating a rubric of complexity in compounding from naturalness scales</td>
<td>75</td>
</tr>
<tr>
<td>2.9. Criticisms against Natural Morphology</td>
<td>80</td>
</tr>
<tr>
<td>2.9.1. Universal preferences criticism</td>
<td>80</td>
</tr>
<tr>
<td>2.9.2. Implication / prediction criticisms</td>
<td>81</td>
</tr>
<tr>
<td>2.9.3. Deriving morphological or linguistic properties from semiotics criticism</td>
<td>83</td>
</tr>
</tbody>
</table>

**Chapter 3: Methodology and Data**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Description of the data sources</td>
<td>87</td>
</tr>
<tr>
<td>3.2. Deciding what is a compound word based on the historical sources</td>
<td>91</td>
</tr>
<tr>
<td>3.3. Methodology – Collecting and coding compound words</td>
<td>99</td>
</tr>
<tr>
<td>3.3.1. Step one – Selecting words from the dictionaries for the databases</td>
<td>99</td>
</tr>
<tr>
<td>3.3.2. Step two – Organising the words into semantic categories in the database</td>
<td>100</td>
</tr>
<tr>
<td>3.3.3. Step three – Classifying words as least complex, moderately complex and complex using rubric and scales of preference</td>
<td>115</td>
</tr>
</tbody>
</table>
Chapter 4: Data Analysis

4.1. Determining the basis for the comparison of complexity between the two systems 147

4.2. Comparison of least complex compounds 161

4.3. Comparison of moderately complex compounds 162

4.4. Comparison of complex compounds 167

4.5. Research questions and complexity 169

4.6. Conclusion 171

4.7. Limitations and the way forward 174

References 178

Appendices

Appendix 1: Database of Jamaican Creole compound words 187

Appendix 2: Database of Akan compound words
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Comparison of complexity between Tsez and Saramaccan</td>
<td>2</td>
</tr>
<tr>
<td>Table 2</td>
<td>Hypothetical comparison using McWhorter’s (2005) rubric of complexity</td>
<td>8</td>
</tr>
<tr>
<td>Table 3</td>
<td>moderately complex compounds</td>
<td>78</td>
</tr>
<tr>
<td>Table 4</td>
<td>Complex compounds</td>
<td>79</td>
</tr>
<tr>
<td>Table 5</td>
<td>Universal scale of morphotactic transparency</td>
<td>84</td>
</tr>
<tr>
<td>Table 6</td>
<td>Application of rubric and preference scale to compound word classification  (least complex compound words)</td>
<td>117</td>
</tr>
<tr>
<td>Table 7</td>
<td>Moderately complex compound words type 1 (metaphor acting on modifier alone)</td>
<td>129</td>
</tr>
<tr>
<td>Table 8</td>
<td>Moderately complex compound words type 2 (double headed)</td>
<td>132</td>
</tr>
<tr>
<td>Table 9</td>
<td>Moderately complex compound words type 3 (clipped head)</td>
<td>135</td>
</tr>
<tr>
<td>Table 10</td>
<td>Moderately complex compound words type 4 (exocentric head)</td>
<td>138</td>
</tr>
<tr>
<td>Table 11</td>
<td>moderately complex compound words type 5 (exocentric head)</td>
<td>141</td>
</tr>
<tr>
<td>Table 12</td>
<td>Completely complex compound words</td>
<td>144</td>
</tr>
<tr>
<td>Table 13</td>
<td>Comparison of least complex compounds between Akan and JC</td>
<td>156</td>
</tr>
<tr>
<td>Table 14</td>
<td>Comparison of least complex compounds between Akan and JC</td>
<td>157</td>
</tr>
<tr>
<td>Table 15</td>
<td>Comparison of least complex compounds between Akan and JC</td>
<td>158</td>
</tr>
</tbody>
</table>
Table 16: Comparison of least complex compounds between Akan and JC 159
Table 17: Comparison of moderately complex compounds between Akan and JC 160
Table 18: Comparison of moderately complex compounds between Akan and JC 163
Table 19: Comparison of moderately complex compounds between Akan and JC 164
Table 20: Comparison of moderately complex compounds between Akan and JC 165
Table 21: Comparison of complex compounds between Akan and JC 168
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Flow chart depicting how Markedness Theory labels a linguistic structure</td>
<td>43</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Sub – theories of Naturalness in Natural Morphology</td>
<td>47</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Universal preferences in compound word formation</td>
<td>53</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Diagrammatic representation of linguistic signs using bars</td>
<td>55</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Naturalness scale of diagrammaticity in compounding</td>
<td>60</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Naturalness scale of indexicality in compounding</td>
<td>65</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Naturalness scale of morphotactic and morpho-semantic transparency in compounding</td>
<td>68</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Complexity scale of diagrammatic iconicity in compounding</td>
<td>71</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Complexity scale of indexicality in compounding</td>
<td>72</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Complexity scale of morpho-semantic and morpho-tactic transparency in compounding</td>
<td>73</td>
</tr>
</tbody>
</table>
1.1 Measuring overall complexity in Creole grammars

McWhorter (2001) designed a metric of complexity to quantify and compare the overall grammatical complexity of the Creole language, Saramaccan with Tsez, Lahu and Maori. To determine the overall grammatical complexity of each language, the following steps were undertaken. (1) Count the number of parts making up each grammatical subsystem in each language. (2) Determine which language has more parts in each grammatical subsystem. The language that has more parts in each subsystem is seen as more complex. (3) The language that has the highest numerical value or number of parts overall is seen as the more complex of both languages. Below I present a summarized version in table format of the McWhorter (2001) comparison between Saramaccan and Tsez.
<table>
<thead>
<tr>
<th>Subcomponent of Grammar</th>
<th>Tsez</th>
<th>Saramaccan</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonemic inventory</td>
<td>42 segments</td>
<td>30 segments</td>
<td>Tsez (complex) Saramaccan (simple)</td>
</tr>
<tr>
<td>Inflectional Morphology</td>
<td>4 noun classes 4 verb classes Agreement prefixes Inflectional suffixes</td>
<td>0 inflectional affixes</td>
<td>Tsez (complex) Saramaccan (simple)</td>
</tr>
<tr>
<td>Derivational Morphology</td>
<td>14 derivational affixes</td>
<td>3 derivational affixes</td>
<td>Tsez (complex) Saramaccan (simple)</td>
</tr>
<tr>
<td>Semantic / Pragmatics</td>
<td>1 evidential marker</td>
<td>0 evidential marker</td>
<td>Tsez (complex) Saramaccan (simple)</td>
</tr>
<tr>
<td>Suppletive morphology</td>
<td>? suppletive pairs</td>
<td>2 suppletive pairs</td>
<td>Tsez (complex) Saramaccan (simple)</td>
</tr>
<tr>
<td>OVERALL</td>
<td>65 parts</td>
<td>35 parts</td>
<td>Tsez is overall more complex than Saramaccan</td>
</tr>
</tbody>
</table>

Having designated Saramaccan as representative of Creole languages and using Tsez as an older language, based on the table above, McWhorter came to the bold conclusion that ‘the world’s simplest grammars are Creole Grammars’.

1.2 Weaknesses in the measurement of overall grammatical complexity

I wanted to use this metric of complexity to replicate this comparison with my native language, Jamaican Creole (hereafter JC) against an older language Akan, purported to be the major substrate language in the formation of Jamaican Creole (Alleyne, 1986: 313; Patrick, 2007: 127 cited in Bartens (2011: 202).
The legitimacy of this comparison is enhanced by data presented by Farquharson (2007:26) which demonstrate that JC similarly to an older language like English also displays compounds whose overall meanings are not predictable from the meaning of their constituents combined. These JC compounds are based on the English morphological pattern of ‘pick – pocket’. Examples include:

(a) [brok ‘break’ + vaibz ‘vibes’] = ‘a killjoy’
(b) [tai ‘tie’ + ed ‘head’] = ‘a scarf which is used to tie the head’
(c) [chruo ‘throw’ + wod ‘word’] = ‘a quarrel’

In the examples above (a) indicates a type of person, (b) identifies a tangible object while (c) defines an action that has been nominalized or in other words the action has been transformed into a noun. Farquharson also indicates that this pattern of verb + noun compounding remains peripheral in the compounding system of English. This is in stark contrast to Jamaican in which the pattern covers a larger semantic range; thus indicating a greater usage.

These compounds highlight the need for a sustained investigation and comparison of other compounding patterns in its purported lexifier language to see if compounds whose meanings are unpredictable and possessing other characteristics which may be labeled as complex might be thrown up.
If the comparison alluded to in the beginning of this section had been undertaken and the results had demonstrated that like Saramaccan, JC was less complex when compared to an older language; such results would have been compelling evidence for McWhorter’s claim. Unlike Saramaccan, JC had remained in contact with English, its superstrate language. This would imply that JC has had the opportunity to obtain many complexifying features from English. Therefore, Jamaican Creole being a member of the subset of Creoles that have had continuous co-existence with their older superstrate, still being overall less complex than an older language would have provided strong support for McWhorter’s position. However, the application of McWhorter’s rubric to a comparison of the overall grammatical complexity of Akan versus JC was difficult to replicate because of two weaknesses associated with his rubric. These weaknesses derived mainly from the rubric’s inability to quantify the notion ‘overall / global grammatical complexity’, which I address.

Overall grammatical complexity is associated with an attempt to take a language’s grammar and break it into discrete subcomponents such as morphology, syntax, phonology, semantics, and pragmatics. Each subcomponent is then subdivided into discrete units that are quantifiable. Based on contrasting views offered by Deutscher (2009) and Miestamo (2008), I realised that the computation of ‘overall grammaticality’ had these
two inherent weaknesses, (a) the problem of representativity and (b) the problem of comparability.

The problem associated with representativity according to Miestamo (2008) is that ‘no metric can pay attention to all aspects of grammar that are relevant for measuring global complexity … because it is difficult to exhaustively count all grammatical details of the languages being studied’ (p. 30). Examples of grammatical details that do not yield easily to segmentation include, the use of a prosodic phenomenon such as intonation in morphosyntax to indicate a distinction in meaning between sentences based on melody and intonational patterns (Hagege, 2001: 73).

Miestamo (2008: 31) recommends overcoming the difficulty associated with representativity by preselecting for comparison, those subcomponents of the grammar that can serve as being representative of the entire grammar and that will enable one to have clear complexity differences between the languages. In an updated version of his rubric, McWhorter (2005) offered some guidelines for selecting these subcomponents based on the following criteria: overspecification, structural complexity and structural elaboration.
However, in stark contrast to Miestamo and McWhorter, Deutscher (2009:250) demands that to prove overall complexity requires ‘showing that for every single subdomain of grammar (not just for an eclectic range of subdomains) one language scores lower or equal consistently to the other language’ (italics mine). Deutscher also points out that a list of subdomains cannot be selected to be representative of the entire grammar of a language based on these notions for the primary reason that the preselection on these grounds constitutes a certain level of discrimination. Prior to the comparison, one has to knowingly select areas of the grammar that are segmentable to provide vast quantitative differences. Deutscher recommends instead that the measurement of overall complexity should simply count the number of parts in each segmented subdomain of both languages exhaustively and unambiguously. Despite what appears to be a way out of the representativity problem by McWhorter’s pre-selection criteria, I concur with Deutscher’s insight that these criteria would compromise the objectivity of the comparison. Additionally, any attempts to apply McWhorter’s selection guidelines to the grammars of both Jamaican Creole and Akan would require access to fine-grained, detailed, complete and unambiguous grammatical descriptions of both languages. To overcome the prejudice of pre-selection by adopting Deutscher’s recommendation would equally require undertaking a daunting if not an impractical task to count ‘exhaustively and unambiguously’ all aspects of both Jamaican Creole and Akan. Whichever solution I embarked on would not have solved the problem of representativity in any satisfactory
way due to the paucity of detailed grammatical descriptions and the inability
to quantify prosodic features which are prominent in both Jamaican Creole
and Akan.

The second limitation associated with a quantitative measurement of overall
language complexity is the issue of comparability. Miestamo (2008) asserts
that comparability is concerned with ‘the impossibility of quantifying each
individual aspect of a grammar’s contribution to the overall complexity of the
language’ (p. 31). This can become problematic in circumstances where the
languages being compared display a significant quantity of one criterion but
display an equal amount in the other criteria.

Using notions later advanced by McWhorter (2005) as an update to his 2001
(metric), I capture hypothetically in table 2 below what such a scenario might
entail.
Table 2: Hypothetical comparison using McWhorter’s (2005) rubric of complexity

<table>
<thead>
<tr>
<th>criterion</th>
<th>Language A</th>
<th>Language B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overspecification</strong></td>
<td>‘the marking of semantic categories left to context in many or most other languages such as evidential marking’</td>
<td><strong>A lot</strong></td>
</tr>
<tr>
<td><strong>Structural elaboration</strong></td>
<td>‘number of rules mediating underlying forms and surface forms such as morphophonemics’</td>
<td>minimal</td>
</tr>
<tr>
<td><strong>Irregularity</strong></td>
<td>‘the more irregularities a grammar contains, the more complex it is overall’</td>
<td><strong>Large amount</strong></td>
</tr>
</tbody>
</table>

If the scenario depicted in table B was obtained based on a comparison between Akan and JC, the question then is which criterion should weigh more so as to decide on overall complexity? Having not carried out a comparison between Akan and JC makes it speculative whether or not such a scenario would have arisen. However, the point being made is that if such conflicting results were obtained, coming to a definitive stance would require making a subjective decision on the matter. This decision would require that perhaps I might have to privilege irregularity over overspecification in order to resolve the deadlock. But on what basis would I make such an important decision?
In addition to the issue discussed above, Deutscher (2009: 249) identifies another issue relevant to comparability. Comparability is also required between the different criteria being used to quantify complexity within each specific subdomain of the grammar. He indicates that in McWhorter’s (2001) rubric ‘the parts to be counted are of a different nature’. In one instance (semantics / pragmatics complexity), one is required to count the number of obligatory or optional semantic distinctions made by the grammar in particular functional domains. In this circumstance, one is really counting explicit distinctions. Secondly, for (inflectional morphology), one is required to count the number of distinct forms (these are not semantic distinctions any more) that one is counting, but rather irregular versus regular forms. Thirdly, one is required to count the number of items making up an inventory (phonemic inventory). This is counting the micro units that build up the component. Are all these measures comparable or in other words, is counting a semantic distinction the same as counting the number of phonemes in an inventory? McWhorter’s rubric does not address this issue nor does it view as a problem, the question of what constitutes ‘parts’ to be counted in the grammar. However, I strongly endorse Deutscher’s observation here and disagree completely with McWhorter’s stance. The privileging of some criteria for comparison based on subjective and arbitrary decisions and the inability to ensure that the parts of the grammar being compared between both JC and
Akan are of the same quality are challenges that would have been difficult to overcome.

1.3 If not overall complexity, then restricted complexity?

Being unable to replicate McWhorter’s work using JC and Akan means that this thesis does not seek to challenge McWhorter’s hypothesis that ‘the world’s simplest grammars are Creole grammars’. Instead of focusing on overall grammatical complexity which undergirds McWhorter’s claim and which I have concluded is far too broad to meaningfully investigate, difficult to quantify as a variable and requires access to detailed analyses of both languages, I chose instead to investigate the concept of language complexity from a more constrained perspective known as restricted complexity.

Restricted complexity focuses an investigation into language complexity on just one particular subdomain of the grammar or to just one aspect of a particular subdomain. I have narrowed the scope of my investigation to the morphological component of both Akan and JC. Morphological complexity is an area that has been the subject of more investigations than has syntactic complexity. Examples of investigations into morphological complexity include Kusters (2008), comparing the complexity of the verbal inflectional systems of the varieties of Argentinian, Ecuadorian and Bolivian Quechua;
McWhorter (2001) whose rubric in the morphological section was largely biased towards a comparison of inflectional morphology and Miestamo (2008) comparing and contrasting the use of negative affixes in Kiowa and Kennant and Kayrabora and Karok. Most investigations into morphological complexity involving Creole languages have been predominantly focused on segmentable morphosyntactic exponents such as derivational and inflectional affixes along with their allomorphs. McWhorter (2011) has now broadened the scope of morphological complexity to include other segmentable morphemes such as classifiers, particles, clitics and reduplicated morphemes.

In most of these investigations complexity is analyzed in terms of segmentable that is countable morphemes. Strikingly, in comparisons involving a Creole language, the Creole at all times displays fewer of these countable grammatical morphemes than do older languages. From this observation, the conclusion is drawn that Creoles tend to be morphologically less complex languages overall than older languages. I have noted that these comparisons operate from an a priori basis, that is, one is knowingly comparing a poorly inflected system against a highly inflected system or a system with a paucity of classifiers against a system rich with derivational morphemes.
Based on what I contemplate to be unfair comparisons, I have decided to make the scope of my investigation even narrower by focusing on just one aspect of the morphological component of both JC and Akan. This is the compounding system. Unlike other morphological processes such as affixation, compounding is a highly productive morphological process in both Jamaican Creole and Akan. Therefore, the selection of compounding for investigation eliminates the issue of comparability which arises in the quest for overall grammatical complexity. Specifically, in this thesis, I will compare the noun+ noun compounding systems of a Creole language and an older language. Reasons for this restriction is based on pragmatic grounds and reflects precedence observed in the literature wherein noun + noun compounds are the most widely studied type of compounds. Furthermore, this comparison eliminates the bias of comparing unequal morphological systems and ensures that it is the same morphological phenomenon under investigation in both languages. Additionally, investigations into morphological complexity have rarely taken compounding through a sustained and thorough investigation (see Chapter 2 for a discussion).

1.4 Research Questions

The focus of this investigation into language complexity will be geared towards restricted complexity. The narrow scope of the investigation will be morphological complexity. Within the morphological domain, word compounding has been selected as an even narrower focus for an investigation
into morphological and morphosemantic complexity. Therefore, the questions that this thesis seeks to answer are:

(a) Are there any differences in the noun + noun compounding systems of JC and Akan?

(b) Do these differences contribute to differences in complexity between the compounding systems of both languages?

(c) Is the noun + noun compounding system of JC simpler / less complex than Akan?

1.5 Value of an investigation into linguistic complexity

(a) Theoretical value

Being able to provide a satisfactory response to these questions will contribute to a deeper understanding of the notion of linguistic complexity and its value to linguists as a variable in carrying out investigations into areas such as language description, typology and comparison. The primary aim of McWhorter’s study of linguistic complexity was to affirm his position that Creole languages form a typological class that makes them distinct from other languages synchronically on the basis of their level of language complexity. A study of this nature can also contribute to this typological debate. The essential question that could be asked is, from a typological perspective, is the compounding system of JC among the simplest of compounding systems in the world?
(b) Defining and describing complexity

The definition of complexity is still unsettled. This emerges primarily from the fact there are two main perspectives from which complexity is viewed (see chapter 2 for discussion). Defining linguistic complexity as an absolute property of a language system through the morphological framework of Natural Morphology (NM) will provide support for the absolutist’s position on linguistic complexity. Quantifying overall complexity might not be the most useful way of investigating linguistic complexity; therefore this investigation aims to bring clarity on whether or not a more constrained definition of complexity should come to prevail in linguistics as opposed to one based on global / overall grammatical complexity?

Variables for measuring morphological complexity primarily consist of segmentable morphological devices. However, a more multidimensional approach (quantitative and qualitative) might be more insightful in providing a more robust description of complexity. The rubric of complexity (see chapter 2) constructed for this thesis is primarily built on qualitative notions that can be subjected to quantification. The rubric will be of value to future researchers in compounding as they will have descriptive categories for describing varying levels of complexity in compound words.
(c) **Description of the Jamaican Creole noun + noun compounding system**

The current work will also undertake a detailed semantic examination of the noun + noun compounding patterns and semantic categories of JC which has not been undertaken elsewhere in the literature. This will contribute greatly to improving the paucity of detailed literature on the compounding system of JC.

**1.6 Summary of chapter one**

The previous discussion highlighted that overall / global complexity can become difficult to quantify objectively because of the problems of representativity and comparability. These two difficulties contributed greatly to the challenges that would be necessary to overcome to replicate McWhorter’s (2001) comparison of the overall grammatical complexity between a Creole language and an older language. Despite solutions being advanced to overcome the problems associated with computing the overall grammatical complexity of a language, I have concluded that these recommendations are insufficient. Therefore, a shift in focus from overall / global complexity to restricted complexity might be a more fruitful way for examining the issue of linguistic complexity. Morphological complexity will become the focus of this investigation and more narrowly noun + noun compounding, which will involve a description and comparison of morphological and semantic complexities associated with compounding.
1.7 Outline of the study

Chapter 2 will offer a description of the relative and absolute approaches in defining language complexity. From this discussion my motivation for adopting the absolute approach for defining complexity will be revealed. Since the absolute approach requires a theoretical framework, I will offer criteria for adopting Natural Morphology (NM) as the theoretical framework for defining complexity in compounding. I will offer a description of this framework, outline how it approaches the process of compound formation and demonstrate how I will define complexity from the NM framework itself. A rubric of complexity which originates from the framework will also be put forward. This rubric will be used to diagnose a compound word as least complex, moderately complex and complex based on this morphological framework. Chapter 3 discusses Methodology. I will present the two historical dictionaries from which the compound words selected were drawn, the basis on which the data were selected from the sources to create two comparable databases of compound words from both languages and apply the rubric of complexity designed in chapter 2 to the data. This will be followed by an analysis of the data in Chapter 4.
CHAPTER 2: TOWARDS A DEFINITION OF COMPLEXITY IN COMPOUNDING

In this chapter, I will first review the literature on defining complexity, making a distinction between absolutist and relativist approaches to complexity. I will then apply the absolutist approach to complexity to the theory of NM while outlining the theoretical framework of NM itself, in order to show how compounding operates in this framework. This will lead me to create a definition for complexity and a rubric of complexity.

2.1 What is language complexity?
Defining language complexity is difficult and challenging. Miestamo (2008:24) has proposed two major perspectives from which language complexity can be defined. These are the absolute approach and the relative approach. In his view, in the relative approach, complexity is reflected in the cost / difficulties that linguistic features cause for different types of language users such as L1 acquirers, L2 learners, speakers and hearers. In contrast, in the absolute approach, complexity is an objective property or an autonomous aspect of a linguistic object.
Similarly, Kusters (2008: 8) describes the relative approach as one that evaluates complexity from the perspective of various language users as to whether or not a linguistic phenomenon is experienced as complex or simple to encode or decode. On the other hand, Miestamo (2008: 24) advocates that the absolutist evaluate whether or not a linguistic phenomenon can be described as complex or simple by using the framework of a theoretical model or linguistic theory.

I will now consider the main problems faced by each approach. The relativist approach has three main problems. I now present these as important questions with which a relativist has to contend:

(a) Which type of language user’s perspective is taken?

(b) Is it possible to have a general account of complexity that takes all users’ perspectives into account simultaneously?

(c) How do we determine the difficulty of a linguistic structure to a particular language user?

Question (a) is based on the observation that the particular language user’s perspective that the relativist chooses becomes problematic because of trade off effects among the various types of language users (Miestamo 2008: 25).
One will have to make a distinction between L1 (native speakers) versus L2 (second language learners) or speakers as producers of languages versus hearers as language processors. Based on the type of language user selected, what might be simple for native speakers, for instance, parsing tri-constituent compounds such as \([\textit{uol} \; \textit{‘old’} + \textit{man fren} \; \textit{‘man friend’}] = \textit{‘a male friend known for a long time’}\) or \([\textit{uol man} \; \textit{‘old man’} + \textit{fren} \; \textit{‘friend’}] = \textit{‘a friend who happens to be old’}\) into binary structures might be complex for L2 learners because of bracketing paradoxes. Similarly, questions what might be difficult for hearers might facilitate speakers. Additionally, one’s early linguistic habits might facilitate or impede foreign language learning as acknowledged by (Riddle, 2008: 145).

On the other hand, question (b) highlights another important concern. The relativist is unable to present a neutral, all encompassing perspective of complexity that reflects all types of language users simultaneously. And finally, question (c) reminds us that the relativist has to strongly rely on performance data from psycholinguistic testing, acquisition data and language processing to be able to provide a description of the difficulty / complexity experienced by various language users.
Similarly, Miestamo (2008) and Kusters (2008) discuss challenges which the absolutist approach presents. These are phrased below as questions:

(a) Are the theoretical, conceptual and heuristic tools used by a theory, a part of a language user’s innate knowledge, and by extension do they describe the performance of language users?

(b) Does absolute complexity correlate with actual user difficulty?

(c) On what basis do I select a particular linguistic theory?

I open this section of the discussion with the issues raised by question (a), followed by (b) and close by examining the concerns highlighted in (c). Question (a) is concerned with the relationship between data and theory and in my view, is more of a philosophical debate. The relativist subscribes to the position that a linguistic theory aims to account for how native speakers use their language. Therefore, the theoretical concepts (rules, levels of representations, cognitive structures such as a language acquisition device etc.) used by the theory are psychologically real in the sense that they form a part of the speakers’ innate linguistic knowledge (Kusters 2008:5). The empirical evidence for the existence of these conceptual notions must be reflected in the performance data of actual speakers. Being guided by this perspective, the relativist concludes that linguistic theories provide an empirical description of language use and by extension the language user’s experiences of complexity / difficulty.
On the other hand, the absolutist holds the view that a linguistic theory uses notions such as rules, length of descriptions, complexity, and typological patterns to describe some ‘emergent properties’ or ‘linguistic reality’ (Miestamo 2008: ) of the linguistic data. Grammatical rules, for instance, are not a part of the language user’s innate linguistic knowledge. Instead, the language user has the genetic ability to learn and create them. In this view, then, rules, length of descriptions, complexity etc. are properties that emerge from the linguistic objects and have become heuristic tools or theory internal concepts (Miestamo 2008) used by the linguist to study the data. Therefore, the absolutist is not convinced that a theoretical linguist is actually describing the neural activities involved in language production.

I share the absolutist view that a separation between the linguistic object (phonemes, words, clauses etc) produced by the language user and his conceptualization and experiences of them is necessary. By holding to this view, the absolutist concludes that a description of complexity is a reflection of properties of a linguistic object and not actual experiences of complexity.

Question (b) addresses the relationship between what the absolutist claims to be complex and actual experiences of complexity. In this regard, the absolutist might have to concede that without external investigation, this relationship between complexity and difficulty remains uncertain and would require
external validations to see if structures labelled as unpredictable, complex, irregular, rare and non-transparent actually present any processing challenges to actual language users. As discussed earlier, examining complexity from language users’ processing and analysis of complexity has its own sets of challenges. Therefore, this work is concerned only with descriptive complexity and will not assess experimental complexity.

Question (c) surrounds the observation that different linguistic theories will provide dissimilar accounts of complexity on the basis of the same language data. Kusters (2008: 7) demonstrates how contrasting views of ‘complexity’ were put forward by Distributed Morphology (Kerke 1996) and a ‘pre syntactic lexicalist theory’ (Lakamper and Wunderlich 1998). Both theories were being used to account for complexities in relation to the inflectional systems of Bolivian and Cuzco Quechua. Distributed Morphology had opacity of some affixes resulting from reinterpretation of their functions as the property that made one inflectional system complex and non-transparent and the other simple and transparent. Despite looking at the same data, the pre syntactic lexicalist theory designated one system as complex based on the number of allomorphs, homonyms and fused morphemes each system possessed. The Distributed theory had a qualitative approach to complexity while the pre syntactic lexicalist theory had a quantitative approach. This difference in what counts as ‘complex’ makes the absolutist appears to be a
relativist in disguise. The absolutists’ notion of complexity is relative to a particular theoretical framework parallel to the relativists’ description of complexity as relative to a user’s perspective.

To overcome this impression, the absolutists resort to the selection of general concepts to define complexity. Miestamo (2008: 28) taking McWhorter’s (2001) position as a precedence reiterated the view that one should select ‘widely accepted theoretical concepts …. that are likely to be accepted by linguists of different theoretical persuasions’. Some of the more generalised concepts that absolutists have relied on are:

1. Economy – (How many categories are expressed?) (Kusters 2008)
2. Transparency of meaning – (One meaning – One form associations) (Kusters 2008, Miestamo 2008)
3. Predictability - (consistency, regularity etc.) (McWhorter 2001)
4. Redundancy - (Is there extra information being provided? - lexical elaboration / overspecification) (Riddle 2008, McWhorter 2005)
5. Description length - (Based on the theory being used – how many rules, levels, constraints) are needed to define the object or explain a phenomenon? (Juola 2008, Progovac 2010)
(6) Universal preference - (marked vs. unmarked / least preferred vs. universally preferred, frequent vs. non frequent, rarity vs. prevalence).

(Dahl 2004)

The absolutist has to take these general notions and use them to define complexity within a particular framework. The aim is to treat a notion such as transparency as being sufficiently general so that regardless of the theoretical approach selected, it can be consistently applied. For example, an endocentric compound word must be seen as transparent in any theory of compounding. However, the definition of transparency associated with a particular theory will be dependent on the relevant framework.

The discussion above highlighted the positions on complexity along with their major challenges. A relativist’s account of complexity has the most difficult hurdles to overcome.

If I were to select the relativist approach, one relevant empirical question would be: are compounds with an external morphosemantic head (exocentric) vs. compounds with an internal morphosemantic head (endocentric) more difficult for children acquiring JC as an L1 or as an L2 vs. Akan? Another equally important question could be: are phonological processes such as prefix
deletion and tone assignment in Akan compounding a challenge to L1 acquirers of Akan vs. JC L2 learners of Akan? These and other questions posed by the relativist require psycho-linguistic testing, acquisition and processing data to determine and validate if these processes are complex for the various language users and why. Experimental literature on compound processing in both JC and Akan do not, however, currently exist to satisfactorily answer such questions. Therefore, adopting the absolute approach makes empirical questions on complexity become a separate issue to be validated through future research on compound processing.

I will adopt the absolutist’s tradition and attempt to operationalize some of these generalised notions through the morphological framework of NM in order to create a definition of complexity. The criteria for the selection of a morphological theory are discussed in Section 2.3 below along with the generalised notions in relation to compounding.

Since the absolutist approach will be adopted, the next section will examine the notions used by absolutists. I will discuss each notion in relation to complexity in compounding (Section 2.3). The discussion indicates that complexity in compounding can be assessed from a multi-dimensional perspective. However, given that this thesis is concerned with restricted
complexity, it becomes necessary to restrict the number of complexity criteria. Issues of oscillating between complexity and simplicity, indeterminacy of complexity status and weighting the contribution of each notion to complexity can be avoided by working with fewer notions. The generalised notions selected as criteria for describing complexity in this thesis will be operationalized through a linguistic theory. I will therefore offer criteria for selecting such a theory (Section 2.4). The discussion concludes that NM best satisfies these requirements.

2.2 Identifying an appropriate morphological theory to define absolute complexity

The generalised concepts used by the absolutist are broad enough to be applied to various types of morphological processes. I will take each concept and examine its relevance in examining complexity in compounding. This is a useful activity as it allows the researcher to see the myriad ways in which complexity in compounding can be perceived. At the end of the discussion, universal preferences and transparency were the two main concepts selected for describing complexity. Finally, I conclude that a theory of preference is compatible with descriptions of absolute complexity.
(1) Economy – How many categories are expressed? The fewer categories expressed, the simpler a linguistic system. The more categories expressed, the more complex the system.

Economy, broadly speaking within the sub-field of language complexity is primarily based on a principle of quantity. This notion is most profitably applied to an investigation of affixal inflectional morphology. The investigator would be interested in finding out if grammatical categories such as tense, aspect, pluralisation, gender, mood etc. are expressed. An inflectional system that expresses all of these categories would be less economical. In contrast, one that expresses fewer of these distinctions would be seen as more economical.

This notion cannot be applied analogously to compound words since compounds are primarily created to serve naming functions. However, one could make this notion relevant to compounding by examining the number of semantic categories that compound words express in a particular language.

A language which expresses more semantic categories (location, possession, material etc.) can be seen as less economical when compared to one that expresses fewer semantic categories. Another useful metric from this notion
would be to examine the number of word categories (adjectives, adverbs, prepositions etc.) used to form compound words. Compound systems which restrict the types of word class would be complex. On the other hand, a simpler system is less restrictive.

(2) Transparency of meaning – (One meaning – One form associations).

Violations of transparency will result in opacity. Opacity makes a system complex.

This notion, when applied to compounding, focuses the investigation on whether or not the elements in the compound display more or less a consistency of meaning with the independent forms that they originated from. The constituents in the compound can obscure their original, transparent meanings through the processes of idiomaticity, metaphoricity, pragmatic context (cultural knowledge), metonymy and specialisation. Constituents that retain their original semantic meanings are transparent while those that do not are opaque.

Transparency is equivalent with semantic compositionality. Compositionality accounts for the transparency of meaning by examining the overall meaning of a compound word to determine if it is a transparent combination of the
meanings of its individual constituents and their arrangement within the compound.

Another concept closely related to transparency in compounding is analysability. Analysability is concerned with whether one can take a compound word and break it into recognizable parts. For example, the English compound word, humbug can be viewed as unanalysable since breaking it into ‘hum’ and ‘bug’ is not helpful in understanding its meaning. In this thesis, transparency will be operationalized as a combination of consistency of meaning, analysability and semantic compositionality since these conceptualizations all overlap.

(3) Predictability - (consistency, regularity, prototypicality etc.) Linguistic systems that operate in an invariant manner are seen as simple. Deviations from this create irregularities. More irregularities make a system more complex.

The application of this notion to compounding assumes that there are some processes in compounding that are predictable, consistent and regular. If compounds are found that do not conform to these expected patterns, they have to be treated as irregularities.
I present a scenario based on English compounding. In English, compound stress assignment can be predicted for many noun + noun compound words. The principle for assigning stress can be formulated as ‘place stress on the first member of the compound as in DRUG store and WEATHER man’. (Capitals represent stress). Noun + noun compound words in English that do not exhibit this pattern of stress assignment are considered to have unpredictable stress. For example, in the minimal pair ‘MADISON street’ and ‘Madison AVENUE’ provided by Kunter (2011:1), the former exhibits stress on the first element in the compound; hence, one would find it quite odd that the latter does not conform to this expected pattern. Instead stress is placed on the second constituent of the compound. If English were to be compared with a language that has stress assignment in compounding, and it was revealed that English has more noun + noun compounds that do not conform to the principle of stress assignment, then English would be seen as a compound system with more unpredictable and irregular forms in this regard, hence more complex.

Predictability appears to be sufficiently entailed in other areas such as universal preferences and transparency and will not be used as an independent criterion for examining complexity in this thesis.

(4) Redundancy - (lexical elaboration / overspecification) - Is there repetitive information being provided? Linguistic systems or objects
that provide unnecessary information can be seen as complex. The lack of redundant information can be seen as simple.

The measurement of complexity via this notion is more applicable to inflectional morphology in areas such as double aspect marking using morphemes which both provide essentially the same information. An example is seen in the sentence below from a serial verb construction in the Portuguese based Creole language spoken on the island of Sao Tome in the Gulf of Guinea.

Zon ka dese ba poson
Zon ASP go.down ASP go.city
Zon used to go down to the city of S. Tome

Hagemeijer (2001:417)

The data above illustrate that aspect marking occurs twice on the same verb form ‘go.down’ and produces an iterative meaning which in essence is more of the same.

In relation to compound words, redundancy is illustrated when compounds are formed through the compounding of synonymous forms that provide repetitive information.
This is highlighted with the JC example below.

\[ \text{ruusta} \text{ ‘rooster’} + \text{kak} \text{ ‘cock’} \] = a rooster/cock.

In the compound, both words provide essentially the same information whether one moves in either a leftward or rightward direction. Redundancy makes these compounds semantically complex. Additionally, redundancy is a specific semantic notion, and would not be worth investigating on its own, and so is not of interest in this study.

(5) Description length - (Based on the linguistic theory being used – how many rules, levels, constraints etc.) are needed to define the object. Linguistic objects that require additional stipulations to offer a complete description are seen as complex in comparison to those that require fewer.

Some compound words would require a longer length of description to account for their internal structure than others. For instance, a description of the internal structure of compounds might rely on examining hierarchical layers. If the compound requires several hierarchies to account for its internal structure, particularly if it was formed through recursion from either right branching or left branching, then it is a complex compound.
The examples below will serve as illustrations.

‘tea party’

[tea party] + committee] (right branching)

(left branching) [Students + [tea party + committee]

The internal structure of the compounds having undergone either right or left branching will require accounting for more recursive layers, embedded compounds and a longer length of description are needed to capture their formation from an existing structure in comparison to compounds formed from simple juxtaposition.

Description length is best used to account for syntactic complexities resulting from recursivity, bracketing paradoxes and compound forms created from derivational processes. For example, in English several rules have been posited for deriving synthetic compounds. Synthetic compounds conform to the pattern V + N + ‘er’ as in ‘truck + drive + er’. A derivational rule, an affix rule and a compounding rule have all been advanced to account for its formation. In contrast, a root compound such as ‘car key’ with no derived words will not require such elaborate mechanisms and could be regarded as a more economical compound.
Phonological complexity is also best suited to a measurement of description length and economy. Phonological complexity will require a longer description to account for morpho-phonological processes such as vowel harmonization, tone and stress assignment. Compound systems in which similar processes are marginal or non-existent will require a shorter description length.

Also, many phonological theories describe morpho-phonological processes in compound formation by positing the existence of phonological rules. In some languages compound words are not treated as word islands in which their phonological structures are insulated from each other. Where word boundaries do not prevent phonological interaction between the constituents, morphophonological processes such as vowel harmonization, deletion, neutralization etc. may apply. Compound systems with these phonological processes will require a longer description length to outline their formation and by extension could be described as less economical than ones in which these morphophonological processes do not obtain.
As syntactic and phonological complexities of compounds remain largely unexplored in this thesis, the notions description length and economy which are more appropriate for investigating phonological and syntactic complexities within these areas are ignored.

(6) Universal preference - Unmarked, most preferred and frequent linguistic objects are seen as simple in comparison to marked, highly dispreferred and rare objects, which are seen as complex. A cross linguistic study done on universals in compounding by Guevara and Scalise (2009: 20) show that word based or two member / binary compounds, endocentric compounds, transparent compounds and compounds made from major word class categories are among the world’s frequent type of compounds.

In applying this notion to compound words several questions can be asked:

- Which is most preferred in the world’s languages? Word based, stem based or root based compounding?

- Are binary compounds most preferred to three, four, five, six morpheme compounds cross linguistically?
- Are endocentric compounds more preferred among the world’s languages than exocentric compounds?

- Do transparent compounds occur more frequently than non-transparent compounds across languages?

- Are compounds created from major class words among the most frequent type of compound words in the world’s languages?

Compounds that have properties that make them among the most frequent types are unmarked while those that occur less frequently are marked.

Of the criteria for assessing the complexity of compound words discussed above, I select universal preferences and transparency. I base my selection on the restricted scope of my investigation. Recall (Section 1.6) that morphological and semantic complexities of compound words are my primary interests.

Since, I am making universal preferences largely criterial for assessing complexity, a theory of preference becomes most appropriate for the investigation to be undertaken. Preference theories generally attempt to
account for the language user’s preferences in creating linguistic structures. In relation to theories of grammar, a preference theory is viewed by Dressler (1985:38) as ‘evaluative devices’ by this, he means that the descriptive language used to characterise linguistic objects in a theory inherently provides concepts which can themselves be used to assess the properties of the linguistic structures. Preference theories generally make use of concepts such as marked / unmarked to evaluate language user’s preferences. To be useful for this investigation, such a theory should have the following characteristics:

- Provide an adequate description and explanation of the linguistic properties that contribute to preferences highly valued cross linguistically in compounding. These properties can be treated as commensurate with simplicity.
- Provide an adequate description and explanation of the linguistic properties that contribute to compound structures being highly dispreferred in the world’s languages. Dispreferences can be linked with complexity.
- Since the theory has to account for preferred and dispreferred properties, a universal typology of compound structures must be accounted for. Therefore, the theory must be typologically based.
- Show sensitivity to the fact that preferences really reflect a gradient nature. This makes the property of complexity a matter of degree rather than absence or presence.
Predict relationships between preferences and processing. Do preferences aid the acquisition of compounds, compound interpretation etc.? This aspect of the theory will not come under scrutiny in this thesis since descriptive complexity is the focus of this thesis.

Apart from these general properties of preference theories, there are properties associated with linguistics structures within preference theories. These are listed below.

- “flexibility, i.e., it does not establish inflexible decisions about structures
- relativity, i.e., it provides a relative degree of preferences among a set of logically possible ones
- scalarity, i.e., the more preferences a structure satisfies, the more preferred it will be
- cancellation i.e., any preference may be overridden under the influence of competing factors” Tzanidaki (1998: 240)

The theory of NM fulfils all the requirements of a preference theory. NM provides an adequate description and explanation of the semiotic properties that make compound words preferred and dispreferred cross
linguistically. The theory is deeply embedded in typology, it accounts for a variety of compound types cross linguistically. Additionally, it treats the notion of naturalness as a gradient phenomenon, hence providing a less rigid definition of complexity. NM makes some testable claims on the relationship between naturalness and complexity which are useful for future research.

In summary, universal preferences and semantic transparency were the criterial notions selected for examining complexity. These notions will be operationalized in a preference theory. Natural morphology satisfies the requirements of such a theory.

In the next section, I present a description of the NM framework. I begin by outlining its historical relationship to Natural Phonology and Markedness theory. Following that, I provide a schematic outline of the three sub-theories that encompasses the framework. I then indicate how this theory accounts for compound word formation and finally, I put forward my definition for complexity.
2.3 Natural Morphology and its relationship to Natural Phonology

I open the discussion here by describing the historical relationship that Natural Morphology shares with Natural Phonology. Dressler, Mayerthaler, Panagl and Wurzel (1987) decided to exploit the notion of ‘naturalness’ developed by Stampe (1979) in a phonological theory called Natural Phonology (NP) to create principles for a morphological theory to be called NM. Stampe’s (1979) notion of ‘naturalness’ was defined within the context of a notion called ‘phonological processes’ that was central to NP. It is therefore important to understand what is meant by ‘phonological processes’ first, before one can grasp Stampe’s intended use of the notion ‘naturalness’. De Cuypere (2008: 125) defines Stampe’s (1972) notion of a phonological process as:

> A potential operation changing a class of sounds / sequences of sounds involving a specific common strain on the human language capacity to a class of sounds / sound sequences differing from the initial class (only) in that it lacks the straining sound property.

In essence, a phonological process is a motivated operation which is necessary for simplifying sequences of sounds that may prove difficult to produce or might strain the human speech organs (i.e. the articulatory and perceptual organs). A phonological process that is perceived to put less strain on the language capacity is therefore assigned the label of ‘naturalness’ by Stampe. The notion ‘naturalness’ as used by Stampe in effect comes to mean a phonological process that puts very little cognitive and articulatory demand on the language capacity.
Since there are phonological processes that strain the language capacity, by analogy Natural Morphologists such as Dressler, Mayerthaler and Wurzel began their search for ‘morphological processes’ that would also strain the language capacity. The Natural Morphologists therefore, adopted Stampe’s (1979) notion of ‘phonological naturalness’ and applied this to Morphology, dubbing it ‘morphological naturalness’.

Similar to Stampe (1979), who concluded that phonological processes can be assigned the label ‘naturalness’, Natural Morphologists concluded that morphological structures and operations can be assigned ‘values of naturalness’. Naturalness values are defined as ‘predicates of evaluation to which certain unconscious evaluation procedures of speakers correspond on the object level; they are not a part of the language system itself’ (Wurzel 1998: 9). The idea here is that language users can use these ‘naturalness values’ to unconsciously assess the extent to which a morphological structure or process ‘strains the language capacity’ (Wurzel 1998: 63).

To summarize, the concept of ‘naturalness’ in NM owes its genesis to the concept of ‘naturalness’ in NP. Naturalness in NM parallels naturalness in NP and simply means that the morphological structure or process ‘does
not strain the language capacity’. The term ‘naturalness’ is therefore being used synonymously by these authors with terms such as optimal, unmarked, highly preferred, cognitively simple, less difficult and least complex.

2.4 Natural Morphology and its relationship to Markedness Theory

Natural Morphologists adopted several ideas from another phonological theory called The Phonological Theory of Markedness proposed by Bailey and Mayerthaler (1976). Theories of markedness generally describe binary oppositions between linguistic objects. The element in the pair described as unmarked is usually associated with properties such as frequency, simplicity, presence of X, default, preferences, etc. The entity labeled as marked carries the opposite of these properties such as infrequency, complexity, absence of X and dispreferences. Below I present Faingold’s (1996: 23) outline of the Phonological Theory of Markedness based on presentations made in Bailey (1977, 1984) and Mayerthaler (1981). What follows is just a brief simplified account using a flow chart (figure1) of a rather complex theory on markedness and I present it here to show the origin of ideas within NM.
(1) Examination of linguistics structures in areas such as: First language acquisition, baby talk, error analysis, language pathology, language history, pidgins and creoles, language typology, frequency, analogy and neutralization.

(2) Identifying data as either marked or unmarked structures

(3) Mechanisms of phonological development

(a) Biological mechanisms: marking values in different positions in the syllable, higher level phonetic processes, (e.g. assimilation, dissimilation and prosodic reconstructing) perception and production of constraints

(b) Sociocommunicational mechanisms: pragmatic principle of markedness – reversal, sociohistorical constraints (e.g. borrowing of prestige rules, hypercorrection, spelling pronunciations)

Flow chart depicting how Markedness Theory labels a linguistic structure
In section (1), this aspect of the theory deals with the kinds of data that one has to examine in order to answer the question - How to assign the labels ‘less marked’ or ‘more marked’ to phonological structures found in various areas of language production?

This question is extremely important, since the Phonological Theory of Markedness had dispensed with an earlier interpretation of markedness. This earlier definition (Jespersen 1927) saw markedness as rigid and binary, where a linguistic structure was either marked or unmarked.

In section (2), the Phonological Theory of Markedness, markedness was reinterpreted in a more gradient sense with the labels ‘less marked’ and ‘more marked’. In order to assign these gradient labels, Bailey and Mayerthaler (1977/84) suggest that one needs to examine the linguistic data in the areas mentioned above to find common patterns that can be labeled with these descriptors. Again, to simplify this aspect of the theory, it aims to explain how phonological structures that are labeled as ‘less marked’ or ‘more marked’ are changed, acquired and perceived.

In section (3), the assumption is that there are two mechanisms responsible for shaping linguistic structures: a biological and a sociocommunicational
mechanism. Biological constraints are believed to shape linguistic structures (i.e. ‘less marked’ and ‘more marked’ structures) into human physiological constraints (i.e. the auditory, articulatory and mental systems) on which they are based. The Sociocommunicational mechanism operates primarily on social relationships.

The creators of NM accepted the basic outline of this framework. However, they dispensed with the notion of ‘markedness’ and replaced it with the term ‘naturalness’ previously adopted from Stampe (1969). Natural morphologists now use the term ‘natural’ in a relativised manner. That is, a morphological structure or process is now viewed as ‘more’ or ‘less natural’ similar to how phonological structures were seen as ‘more’ or ‘less marked’. Having adopted the basic outline of the Phonological Theory of Markedness, Natural Morphologists in essence also adopted the idea that morphological naturalness must be derived from various areas in language production such as aphasia and child language data.

In review, NM adopts two main ideas from the Phonological Theory of Markedness, firstly, the notion that ‘naturalness’ is best interpreted as a gradient notion and secondly, the idea that ‘morphological naturalness
principles’ must be motivated from extra-linguistic bases similar to ‘markedness’ properties.

2.5 A basic framework of Natural Morphology

Based on the historical background provided earlier, it is no surprise that Dressler (2000), a leading Natural Morphologist, defines NM as a theory ‘based on an integration of concepts taken from markedness and Natural Phonology’ (pg. 267). To explain how morphological structures are created, Natural Morphologists advanced the following schematic outline based on the work of (Kilani-Schoch (1988), Dressler et al. (1987), Dressler (2000)).

The theory of NM is made up of three sub-theories: (A) Universal naturalness otherwise called System Independent Naturalness, (B) Typological Naturalness also known as System Dependent Naturalness and (C) Language Specific Naturalness referred to otherwise as System Adequacy Naturalness. I will provide a brief description of each sub – theory based on figure 2.
Figure 2

<table>
<thead>
<tr>
<th>Sub – theories of Naturalness</th>
<th>Sources of linguistic data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Universal Morphological Naturalness</strong>&lt;br&gt;(e.g., the principle of the least effort)</td>
<td><strong>Extra - linguistic</strong>&lt;br&gt;(cognitive, phonetic, semiotic psychological, sociological etc.)</td>
</tr>
<tr>
<td>A. [Universal Naturalness ]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Typological or System Dependent Naturalness&lt;br&gt;(e.g., a preference for simple CV structures)</th>
<th><strong>Linguistic typology</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>B. [Typological Naturalness ]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language Specific System Naturalness&lt;br&gt;(absence of clusters in a language)</th>
<th><strong>Linguistic preferences</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>C [Language Specific Naturalness]</td>
<td></td>
</tr>
</tbody>
</table>

Sub – theories of Naturalness in Natural Morphology

Based on (Dziubalska-Kolaczyk 2001: 76) Modifications were done by me to achieve clarity.
Sub – theory A – Universal morphological naturalness

The process of word formation begins with sub – theory (A), which is primarily associated with Mayerthaler’s (1981) work on markedness. However, it has been recast as a preference theory by Dressler (1999). The aim in this sub – theory is to establish deductively, universal naturalness morphological principles and their naturalness scales of preferences / parameters from extra – linguistic bases and non – morphological components of the grammar. Much like the Phonological Theory of Markedness, Natural Morphologists attempt to derive these principles from performance and theoretical data in areas such as Psycholinguistics, Neurolinguistics, Language Acquisition, Language disorders (aphasic studies), Semiotics, Communication Studies, Language Change, Pidgin and Creole language formation, Cognition, Sociology and Typology (implicational universals and statistical generalizations). In this respect, Natural Morphologists subscribe to the view that similar cognitive principles underlie all these areas. Potentially, Universal Morphological Naturalness principles and their scales of preferences / parameters can come from any of the bases mentioned earlier; however, Natural Morphologists have predominantly derived them from the discipline of Semiotics (Peirce: 1965). Generally, the principles and their preference / parameter scales in NM are claimed to be independent of the morphological system of any particular language. In the diagram above,
the principle of least effort is derived from the extra-linguistic bases mentioned there.

Sub-theory B – Typological or System Dependent Naturalness

After identifying universal morphological principles, one has to contemplate how these principles are going to be reflected in individual languages along with their naturalness scales of preferences/parameters; hence, the need for sub-theory (B). Sub-theory (B) was inspired by Skalicka’s (1979) theory of ideal language types. These ideal constructs agglutinating, fusional, analytical and polysynthetic have been reinterpreted by Dressler (1985: 337–348) as responses to naturalness conflicts. Naturalness conflicts arise from the observation that different areas in the language display preferences for different types of naturalness. For example in phonology, naturalness aims at optimal articulation and perception while naturalness in inflectional morphology aims at optimal symbolization of grammatical categories. These components are part of the same language system and must produce a single set of language forms; however, these forms cannot reflect maximum phonological naturalness and maximum morphological naturalness all the time. There are times when one type of naturalness has to be sacrificed for greater naturalness over the other. Therefore, the aim of this sub-theory is to offer an explanation as to why languages do not always elect the more
natural preferences / parameters from the preference scales. The non-
selection of the most natural preferences / parameters is claimed to be as a
result of naturalness conflicts. The typology of the language can help to
understand this conflict. Additionally, the sub-theory focuses on how the
morphological universal principles formulated in the theory of Universal
Naturalness, can be expressed cross-linguistically by the specific
typological properties of different languages. The diagram indicates that
the linguistic outcomes from sub theory B results in a preference for
simple phonotactics in the form of a CV structure.

Sub – theory C – Language Specific System Naturalness

Another useful way of explaining why a language does not select the
‘most natural’ unit on the preference / parameter scale of a morphological
principle is offered by sub – theory (C) (Wurzel 1984 Dressler and
Ladanyi 2000), if sub – theory (B) cannot offer a satisfactory account.
This sub-theory is best used to explain rare morphological structures and
idiomatic constructions. Language specific naturalness may contradict
some universal morphological preferences / parameters as well as some
typological preferences / parameters. The implications for these
contradictions impact the distribution of rare morphological structures in a
language and even in specific cases, explains the absence of a certain
morphological feature in a language.
Summary of NM

The sub-theories of NM though presented autonomously in the diagram, interact. Natural morphologists conceptualize NM more as an evaluative account of word formation, rather than one that relies on formal mechanisms to explain how morphological structures are derived. Additionally, they observe that languages do not always display the most natural preferences / parameters on the scalar universal morphological principles. At times, that which is preferred universally by other languages may be rendered dispreferred or even disallowed by typological adequacy in another language. It should also be noted that the outline presented above represents a hierarchical description in which universal preferences are seen at a higher level and typological adequacy and system-dependent adequacy are at a lower level. System adequacy is also considered to be at a lower level than typological adequacy. Each lower level can specify, overturn or filter out the preferences / parameters that are expressed at the higher level. The diagram indicates, for example, that a linguistic consequence of the preference for CV structures is the absence of clusters in a given language.
Having outlined the basic assumptions and machinery of this morphological framework, the next section of this thesis will demonstrate in a procedural fashion how the theory accounts for compound word formation. On completion of this account, I will demonstrate how a definition of complexity can arise from the theory.

2.6 A Natural Morphology account of Compound word formation

In Natural Morphology, the evaluation of compound formation begins at Universal Naturalness with the search for morphological universal principles and their parameters / preferences scales in a non-linguistic component. The non-linguistic component advanced by Dressler (2005) for compound word formation is the discipline of Semiotics. Semiotics is concerned with the theoretical and practical study of verbal and non-verbal signs. From Semiotics, Dressler (2005) provided several universal word formation principles which guide the formation of compound words. Among these, the following form the basis of discussions in this work: the principle of diagrammatic iconicity, the principle of indexicality and the principle of morpho-semantic and morpho-tactic transparency. Below, is an illustration of the principles and the extra-linguistic bases of sub-theory A, and in the sub-sections which follow, an explanation of these principles, indicating how they apply to compound word formation in NM.
Universal preferences in compound word formation

2.6.1 The principle of diagrammatic iconicity

In this section, I first discuss the semiotic notion of iconicity in its most generalized form. Secondly, I point out how morphologists have drawn parallels between the creation of non-linguistic signs and linguistic structures. Finally, I demonstrate how various semiotic properties can be reflected in compound words to varying degrees.

An iconic relationship is based on similarity between an object and its representation. Iconicity captures three relationships, namely: images, diagrams and metaphors. Images are based on referents appearing like the real objects they are based on due to simple qualities found in both, for example, an architect’s model of a building versus the actual building. On the other hand, diagrams capture relational similarities between various elements in the object and its representation. A clear example of semiotic diagrammatic iconicity would be a chart illustrating population growth. The illustration will indicate a relationship between the number of people
and the population size. Finally, metaphors indicate a less transparent iconic relationship between an object and its representation as in comparing a foot to the ‘foot of the mountain’.

The diagram is a non-linguistic sign. In the creation of diagrams such as a chart as indicated above, semioticians propose that the human agent through analogy and other cognitive processes created a visual representation of an abstract situation by capturing similarities in size between population growth and the corresponding increase in the number of people.

In constructing linguistic signs, Natural morphologists, equally assume that the human agent by analogy to non-linguistic signs create a diagrammatic iconic representation of the referent using phonemes. I attempt to illustrate this relationship below using bars and their linguistic counterpart below.
Diagrammatic representation of linguistic signs using bars

These bars represent the difference in unemployment figures between two years. The relative difference is conventionally reflected by the size of the bars. In constructing verbal linguistic signs with phonemes, an analogy is drawn between the length of a morphological structure and the difference in size between the bars. For example, the semantic concept of pluralisation is encoded in morphological structures diagrammatically wherein; boy (singular) would reflect the lower bar while the bar that extends further would reflect boy + s (plural). Pluralization takes place in a constructional diagrammatic iconic way by increasing the morphological sign using more phonemes than their singular counterparts (Dressler 2000). Hence, for him, in a linguistic sense, diagrammatic iconicity represents an analogy (comparison) between the relation of the morphosemantic and morphotactic properties of a morphological form and the parts making up the form itself.
In respect to compound word formation, an analogy is made between the relations of head versus non-head as represented by the constituents in the compound structure and the relative difference in the size of the bars. Another conventionalized way of interpreting the difference in the size of the bars is to suggest that the lower bar remains subordinate to the higher bar that dominates it. Linguistically, this subordinating / dominance relationship is captured diagrammatically in compound words by examining the head / non – head relationship between the constituents in the compound and their morphosemantic and morphosyntactic properties. Compounds overall typically inherit their morphosemantic and morphosyntactic properties from the members making up the compound.

The element responsible for contributing its semantic properties to the compound overall is known as its morphosemantic head. As the morphosemantic head, it semantically dominates the non-head by allowing only its semantic features such as animacy, abstractness, reference etc. to percolate to the overall compound. On the other hand, the element from which it will inherit its syntactic properties is known as its morphosyntactic head. As the morpho-syntactic head, this element syntactically dominates the non-head by allowing only its own syntactic features such as pluralization, gender assignment, distributional properties etc. to percolate to the overall compound.
The notion head as being used in this thesis represents a coalescing of a morphosemantic and a morphosyntactic head as one element in a compound word. When such a head is located in the morphological structure, it is called an endocentric head. However, there are compounds in which the morphosemantic head does not coalesce with the morphosyntactic head, when this happens, the morphosemantic head resides outside the morphological structure and is called an exocentric head. Compounds with exocentric morphosemantic heads are not regarded as headless structures since the ‘place holder’ of the exocentric morphosemantic head in the morphological structure will function as a morphosyntactic head.

In many compound words both types of headhood are encapsulated in only one element. Whenever, this situation is obtained, the element functioning as both the morphosemantic and the morphosyntactic head is seen as dominating its non-head which remains subordinate. Given that the semantic and syntactic properties of the non-head are restricted from being passed on to the overall referent of the compound, the non-head will always remain completely dominated semantically and syntactically.
Natural Morphologists have deduced from the observation of data that in compound word formation, there is a universal preference for compound words such as these in which one element completely dominates the other element. This type of compound is known as a subordinated compound. These are the most natural or universally preferred compounds and maximally reflect the principle of diagrammatic iconicity. However, NM sees its principles as scalar properties wherein some compounds will not maximally reflect diagrammaticity iconicity. This is captured in the naturalness scale of diagrammaticity (figure 5) which I have created below using JC compound words as examples.
Figure 5 shows that most natural diagrammaticity requires that the morpho-semantic head and the morpho-syntactic head coincide. This allows the overall semantic and syntactic properties of the compound to be inherited from one dominant element in an endocentric compound. Less diagrammaticity results when more than one constituent contributes their semantic and syntactic properties overall to an endocentric compound. Dominance no longer exists. In exocentric compounds, dominance gets lost since the morpho-semantic head existing outside the morphological structure still continues to contribute its semantic properties while the syntactic properties are contributed to the element acting as a place holder (morpho – syntactic head) in the morphological structure. Again complete dominance by one compound constituent is lost.
Most natural diagrammaticity
The morpho–semantic and the morpho–syntactic head coincide as one element in the endocentric compound.
Hence, this element semantically and syntactically dominates the non-head.
Jamaican: aam + tik ‘crutches’ (lit. ‘arm’ + ‘stick’)

Less natural diagrammaticity
There are two morpho–semantic heads.
The dominance relationship is lost because the two semantic heads are equipollent.
Jamaican: ed + skul ‘the skull’ (lit. ‘head’ + ‘skull’)

Least natural diagrammaticity
The morpho–semantic head is outside, while the morpho–syntactic head is internal.
The dominance relationship is lost. The external morpho–semantic head dominates semantically, while the morpho–syntactic head dominates syntactically.
Jamaican: kichin + bich ‘lamp’ (lit. ‘kitchen’ + ‘bitch’)

Naturalness scale of diagrammaticity in compounding
2.6.2 The principle of indexicality

Another universal principle that guides the evaluation of compound words in NM is the principle of indexicality. While an iconic relationship is based on similarity between an object and its representation, an indexical relationship is based on contiguity, proximity or adjacency between an object and its representation. In other words, indexicality captures a relationship in which the object and its representation are conceptually dependent on each other for full interpretation. A non linguistic sign which is formed on the basis of indexicality is a signpost pointing in the direction of the place it designates. The direction of the signpost is a symbolic representation of the actual place the signpost identifies.

To draw an analogy, between an indexical non linguistic sign such as the pointing signpost and a linguistic sign of the form [base] + [affix], Carstairs-McCarthy (2002: 225) said that, Dressler thinks that it’s a simplification to analyse the ‘-s’ in ‘cats’ as only demonstrating plural. For Dressler, the affix cannot occur except bound to some noun stem, therefore, in some metaphorical (italics mine) sense the plural affix points towards the noun stem (base) similarly to how a signpost points towards a neighbouring town.
Since indexicality evokes a principle of adjacency, indexicality when expressed in bimorphemic morphological structures requires that ‘the two concepts are considered as conceptually close because they share semantic properties, one affects the other, and both are perceived as inseparable’ (Haiman 1983: 783).

Similarly, in compound words, indexicality is not only concerned with the linear arrangements of elements in the morphological structure, it also aims to capture a relationship of dependency between its non-head / head constituents. In noun + noun compound words, indexicality holds best between the nominal constituents when there is a relationship of modification. In this relationship one member of the compound can be said to modify the other member, the modified. Since the element functioning as both the morphosemantic and morphosyntactic head modifies the non head, a relationship of dependency holds between them. This requires that the modifier is placed as closely as possible to the constituent that it modifies.

Therefore, indexicality means that the element designated as the morphosemantic and morpho-syntactic head needs to be adjacent to the non-head which it dominates in the morphological structure of the compound word
if they share a relationship of dependency. In endocentric compounds, where the modification relationship occurs, the morphosemantic and morphosyntactic head must be adjacent to the non head. Indexicality is lost between compound members sharing a copulative or coordinating relationship. These types of compounds are regarded as double headed compounds in which the constituents in the compound are simply juxtaposed beside each other without any dependency holding between them, because both components can be regarded as heads. Indexicality is also lost when a relationship of dependency exists between the constituents in a compound, but there are intervening elements such as infixes which prevent the modifier from being adjacent to the element that it modifies. A morphosemantic head which is exocentric diminishes indexicality. In exocentric compounds, the internal compound constituents could still be seen as modifying this exocentric morphosemantic head; however, instead of being found in the morphological structure, this head remains outside of the compound structure thereby violating the principle of indexicality.
I have captured this principle and its parameters / preferences in the naturalness scale of indexicality below using JC examples.

Figure 6 points out that the principle of indexicality operates on a principle of adjacency based on a dependency relationship holding between the members of the compound. In endocentric compounds, the constituent acting as the dominating morphosemantic and morphosyntactic head has to be adjacent to the non head that it modifies and which remains subordinate to it. This results in most natural indexicality. Indexicality becomes obscured by linguistic material such as infixes that intervene between the dominating head and the non head. Further erosion of indexicality occurs when the morpho-semantic head remains outside the compound structure. The external location prevents the head from being adjacent to the non-head in an exocentric compound. In an exocentric compound a modification relationship still holds between the external morphosemantic head and the internal constituents of the compound. While the principle of subordination / dominance seeks to capture the percolation of morphosemantic and morphosyntactic properties of the head to the overall compound despite its location, indexicality reflects location of the head in relation to its non head based on a dependency relationship.
Figure 6

**Most natural indexicality**
Morpho-semantic head is internal and adjacent to the non head.
Jamaican: iej + piepa (lit ‘age + paper’) = birth certificate

**Less natural indexicality**
Morpho – semantic head is internal but not adjacent to its non head.
It is separated by extra material such as infixes.
English: ‘gas o meter’ (No Jamaican examples found)

**Least natural indexicality**
Morpho-semantic head is external.
Jamaican: baal + pliet (lit ‘a bade’ + plate’)
= ‘a bird’

Naturalness scale of indexicality in compounding
2.6.3 The principle of morphosemantic and morphotactic transparency

Apart from diagrammatic iconicity and indexicality, the principles of morphosemantic and morphotactic transparency are also universal principles that guide the evaluation of compound words in NM. Natural morphologists advocate that a transparent sign is an iconic representation of the referent in the sense that the sign maintains many of the features of the referent and allows one to see continuities. The notion of transparency is captured in compounding on two levels: Morphotactic transparency which comes down to morpho-phonological transparency and morphosemantic transparency. In morphotactic transparency, the free/ bound forms from which the compound originates must not have undergone morphophonological alterations that obscure the phonological relationship between the forms in the compound and their free/bound sources. In semantic transparency, the words are assigned their original meaning in the compound structure. Morpho-semantic transparency and morpho-tactic transparency are complementary principles and when they coincide, the compound is expected to be transparent or most natural. Using ideas first advanced by Libben (1998), I have created a scale of morpho-semantic and morpho-tactic transparency illustrated below.
In figure 7, complete transparency of both morphemes creates endocentric compounds. Opacification may affect either the head or the non head of the compound resulting in transparency being diminished. A less natural compound results from partial opacification. Total opacification creates exocentric compounds.
Figure 7

Naturalness scale of morphotactic and morpho-semantic transparency in compounding

Most natural transparency
Transparency of both members of the compound
Jamaican: ‘waata + buut’ (lit. ‘water’ + ‘boot’) = water boots

Less natural transparency
Transparency of the head member, opacity of the non – head member
OR
Transparency of the non – head member, opacity of the head member
Jamaican: jakfruut + fut’ (lit. ‘jackfruit’ + ‘foot’) = a swollen diseased leg (elephantiasis)

Least natural transparent
Opacity of both members of the compound
Jamaican: ‘griev + diga’ (lit. ‘grave’ + ‘digger’) = an insect
In conclusion, compound words are typically formed within Sub – theory (A), the universal theory of morphological naturalness. The universal principles of word formation and their parameter / preference scales can account for the typological observation that many compound forms will ‘naturally’ deviate from the optimal end of the preference scales.

Sub – theory (B) of the NM framework only becomes necessary to explore if one wants to provide an account for the non-selection of the universally preferred compound types in a particular language. This can be done by examining the typology of the language. Similarly, Sub-theory (C) only becomes applicable if language specific naturalness is needed to explain the presence or absence of a particular compound type in a specific language. Having described how the theory of NM accounts for compound formation, I will now offer my definition of complexity in compounding based on these semiotically driven concepts.

2.7 Creating a definition of complexity from Natural Morphology

Following Natural morphologists, I will adopt the concept that complexity is a gradient aspect of language, and use the scales of naturalness outlined earlier (Section 2.7) as the basis of this definition. Therefore, I will take each universal word formation principle along with its parameter / preference scales. I will then
equate each naturalness parameter / preference point along each universal scale as an inverse indicator of complexity. This reinterpretation of the points along the universal scales of preferences will mean that the optimal end of each universal scale which is now labelled as ‘most natural’ will be equated with ‘simple’ or ‘least complex’. The points labelled ‘less natural’ are now equated with ‘moderately complex’ and the points labelled ‘least natural’ are now equated with ‘complex’. This is indicated below with all naturalness scales.
Figure 8

(MOST NATURAL)
Least complex diagrammaticity
The morpho – semantic and the morpho – syntactic head coincide as one element in the compound.
This element semantically and syntactically dominates the non - head.
Jamaican: aam + tik ‘crutches’

(LESS NATURAL)
Moderately complex diagrammaticity
There are two morphosemantic heads.
The dominance relationship is lost because the two semantic heads are equipollent.
Jamaican: kitni + pus ‘a pus who is a kitn or a kitn who is a pus’

(LEAST NATURAL)
Complex diagrammaticity
The morpho – semantic head is external, while the morpho – syntactic head is internal.
The dominance relationship is lost. The external morpho - semantic head dominates semantically, while the morpho - syntactic head dominates syntactically.
Jamaican: bad + yai ‘a look of contempt’

Complexity scale of diagrammatic iconicity in compounding
Least complex indexicality
Morpho-semantic head is internal and adjacent to the non head
Jamaican: iej + piepa ‘birth certificate’

(MOST NATURAL)

Moderately complex indexicality
Morpho-semantic head is internal but not adjacent to its non head.
It is separated by extra material such as infixes.
English: gas + ‘o’ + meter.
No Jamaican examples found.

(LESS NATURAL)

Complex indexicality
Morpho-semantic head is external
Jamaican: baal + pliet ‘a bird’

(LEAST NATURAL)

Complexity scale of indexicality in compounding
Complexity scale of morpho-semantic and morpho-tactic transparency in compounding
Based on the typological scales of preferences presented above, the non-optimal end of each word formation principle indicates complexity. On the other extreme end of each scale, the most optimal pole indicates simplicity/least complex, while the mid points on each scale correlate to moderate complexity. Based on these observations, complexity is defined as the extent to which a compound word shows conformity or non-conformity with universal preferences of word formation. A least complex compound word conforms maximally with the desired characteristics of the word formation principles, while a moderately complex compound word conforms to some of these characteristics maximally while departing from some. On the other hand, a complex compound typically departs from most or all of the desired characteristics in comparison to a moderately complex compound word. This may be summarised as follows:

(a) A simple compound is optimally diagrammatic, shows indexicality maximally and is completely morpho-phonologically and morpho-syntactically transparent.
(b) A moderately complex compound departs from some of the optimal preferences.
(c) A complex compound word departs from most or all the optimal preferences.

In the next section, I will create a rubric for measuring complexity based on the definitions provided earlier.
2.8 Creating a rubric of complexity in compounding from naturalness scales

By relying on the three semiotically inspired principles of word formation and the scales of preferences that emerge from them, I can design a rubric with a set of diagnostics for identifying a compound word in a language as either complex, moderately complex or simple / least complex. In order to accomplish this, I take the description of each simple / least complex end of each morphological scale and use them as guiding questions.

**Rubric for assessing complexity**

**Preference for diagrammaticity**

(1) Does morpho - semantic and morpho - tactic dominance reside internally in one element (head) of the compound? **YES**

**Preference for indexicality**

(2) Does the compound have a modification relationship between its constituents so that its morphosemantic and morphosyntactic head need to be adjacent to its non head in the compound? **YES**

**Preference for morphosemantic and morphotactic transparency**

(3) Does the compound have both its non head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? **YES**
If the compound returns a yes in all three categories, based on the definition of complexity being advanced in this thesis, then the researcher would classify such a compound as simple or least complex.

Since a moderately complex compound is expected to display some departure from the preferred end of the scales, a compound that returns a combination of least complex, moderately complex and complex properties must be labelled as an overall moderately complex compound. This is demonstrated below.

**Moderately complex compounds**

**Diminished diagrammaticity**
(1) Does morphosemantic and morphosyntactic dominance reside internally in one element (head) of the compound? **NO. Check if the compound has both constituents acting as the morpho-semantic and morpho-syntactic head simultaneously.**

**Diminished indexicality**
(2) Does the compound have a modification relationship between its constituents so that its morphosemantic and morphosyntactic head is adjacent to its non head in the compound? **NO. Check to see if the internal relationship between the constituents in the compound is one**
of simple juxtaposition without any dependency. Or check if a modification relationship exists, but extra material such as infixes prevents the morphosemantic and morphotactic head from being adjacent to the nonhead.

**Diminished transparency**

(3) Does the compound have both its non head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? **NO. Check to see if the compound displays transparency of the head member and opacity of the non head member or vice versa.**

In table 3, I present a summarised version of all the possibilities for a moderately complex compound found in the data used for this thesis.
Table 3: Moderately complex compounds

<table>
<thead>
<tr>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>Morphosemantic transparency</th>
<th>OVERALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TYPE 1 – least complex</td>
<td>Least complex</td>
<td>Moderately complex</td>
<td>Moderately complex</td>
</tr>
<tr>
<td>TYPE 2- moderately complex</td>
<td>Least complex</td>
<td>Least complex</td>
<td>Moderately complex</td>
</tr>
<tr>
<td>TYPE 3- complex</td>
<td>complex</td>
<td>Moderately complex</td>
<td>Moderately complex</td>
</tr>
</tbody>
</table>

Based on the table above, semantic transparency is treated as the most ideal property that a compound should possess, therefore for compounds to be complex, the minimal requirement is that both constituents must be totally non – transparent.

At the other extreme, compounds that show non conformity to the optimal preferences will be labelled as complex. The table below indicates what obtains for a compound to be labelled as complex in this thesis.
Table 4: Complex compounds

<table>
<thead>
<tr>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>Morphosemantic transparency</th>
<th>OVERALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>complex</td>
<td>complex</td>
<td>Complex (Both constituents opaque)</td>
<td>complex</td>
</tr>
</tbody>
</table>

In summary, my definition for complexity is based on the theoretical descriptions of NM. The essence of complexity in my thesis is that linguistic complexity is a gradient phenomenon which expresses itself in the extent to which compound words optimally display conformity with the universally preferred characteristics of various semiotic universal principles of word formation. Keeping in mind the gradient nature of the definition of complexity, a rubric becomes useful to assist the researcher in identifying and classifying compounds as least complex, moderately complex and complex. The rubric offered here is a series of questions designed to achieve this purpose.

Before closing this chapter, I address some frequent criticisms that are levelled against NM. I will discuss some of the major critiques of the Naturalness theory with the intention of showing that they do not weaken my position on complexity or make the project superficial.
2.9 Criticisms against Natural Morphology

2.9.1 Universal preferences criticism

One criticism that has been made of NM comes in the form of the question ‘if languages prefer more natural forms over less natural ones, why is it then that the languages of the world have not ended up employing only the most natural techniques a long time ago’? Dalton-Puffer (1998: 54). Put another way to be of some relevance to this thesis, the question would be, if the compounding systems of languages universally prefer more natural / least complex compound forms, why are subordinated compounds not the only compound forms, i.e. those in which the morphosemantic and morphosyntactic head coincide?

This question is often posed to NM theorists to recognise that complex morphological forms which are largely dispreferred should actually be non-existent given natural languages’ overwhelming preference for simple forms. Natural Morphologists have responded by pointing out that they conceptualize language as a system of conflicts. The different goal of each subsystem conflicts with the naturalness tendencies of other subsystems. This conflict prevents the optimal end of the preference scale from always prevailing.
By accepting the Natural Morphologists response, I recognise that this critique undermines a fundamental assumption of the theory which is the feature of conflicts. Conflicts can block the optimal end of the preference scales from dominating; therefore, this thesis benefits from a theory that anticipates a typology of compound forms to include optimal and non optimal forms.

2.9.2 Implication / prediction criticisms

Another criticism often levied against the NM framework has been against the validity of some of its predictions. Natural Morphologists often claim that there are implications for labelling morphological structures as more and most natural. By applying these predictions to the notion of complexity, similar implications arise. Some of these predictions are:

(a) **Frequency prediction**
A more natural and by extension a less complex compound form will be statistically more frequent and widespread in a language and across the world’s languages.

(b) **Language change prediction**
A more natural and by extension a less complex compound form will be fairly resistant to language change.

(c) **Language acquisition prediction**
A more natural and by extension a less complex compound form will be acquired early in L1 acquisition.
(d) **Language disorder prediction**

A more natural and by extension a less complex compound form will be relatively unaffected by language disorders (aphasia).

In empirical studies undertaken in some of these areas, NM claims have not always been substantiated. Natural morphologists point out that these predictions are not to be seen as absolutes or ‘laws’ but rather as degrees of probability, tendencies, possibilities or statistical generalizations.

This relativised response might make predictive claims made by NM appear to be potentially unfalsifiable. However, my purpose for adopting the theory of NM for this project is primarily descriptive and not empirical. The theory provides me with some descriptive categories along a continuum that are useful in classifying and describing compounds as simple, moderately complex and complex structures. Therefore, my definition of complexity is not undermined by the empirical data that seemingly in some instances contradict the expectations of some naturalness predictions. Empirical evidence in future research might actually prove that the compound words that I have labelled as complex are indeed less frequent in the world’s languages, more likely to undergo language change, difficult for children to acquire and likely to be affected
by aphasia. Complexity as I have defined it in this thesis is purely a descriptive label with no evaluation of difficulty attached to it. Recall the discussion in section 2.1 concerning the absolutist and relativist approach to complexity.

2.9.3 Deriving morphological or linguistic properties from semiotics criticism

Another serious objection levied against NM is the contention that Natural Morphology has made no systematic connection between a universal parameter and its extralinguistic base in a straightforward way. Antagonists have asked, can precise linguistic concepts be derived from vague semiotic principles such as diagrammaticity and indexicality? (Dalton-Puffer: 19). Natural morphologists have responded by providing empirical justifications for their claims by subjecting the semiotic principle of transparency for example to psycholinguistic testing (Katamba 2004:181).

Table 5 shows the results which were obtained from a Psycholinguistic test conducted on the principle of morphotactic transparency in word formation (Mackay 1978 cited in Katamba (2004). Phonological rules (PR’s) such as allophony, resyllabification, neutralization, as well as morpho-phonological rules (MPR’s) such as suppletion were examined for
their effects on phonological transparency in word derivation. Mackay’s findings are as follows:

Table 5: Universal scale of morphotactic transparency

<table>
<thead>
<tr>
<th>Rank</th>
<th>Transparency Level</th>
<th>Process Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Most transparent</td>
<td>PR’s rule interfere</td>
</tr>
<tr>
<td>2.</td>
<td>PR’s intervene</td>
<td>Resyllabification interferes</td>
</tr>
<tr>
<td>3.</td>
<td>Neutralizing PR’s</td>
<td>Rider → rid + er (flapping American)</td>
</tr>
<tr>
<td>4.</td>
<td>MPR’s (no fusion)</td>
<td>Electric → electric +ity</td>
</tr>
<tr>
<td>5.</td>
<td>MPR’s interfere</td>
<td>Conclude → conclusion</td>
</tr>
<tr>
<td>6.</td>
<td>Allomorphic MPR’s</td>
<td>Dec[ai]de → dec[i]sion</td>
</tr>
<tr>
<td>7.</td>
<td>Weak suppletion</td>
<td>Child → Childr + en</td>
</tr>
<tr>
<td>8.</td>
<td>Least transparent</td>
<td>Strong suppletion (no rules)</td>
</tr>
</tbody>
</table>

The experimental data come from psycholinguistic experiments done by Mackay (1978) cited in Katamba (2004). The experiment indicated that word formation can be best processed in 1 and worst in 8. For instance, the stem [excite] can best be perceptually isolated from [excite + ment], while the most difficult perceptual isolation is found in [am, are, is, was] from [be]. These results indicate that resyllabification, neutralization and suppletion contribute to diminishing transparency.
In support of the claims made above, Katamba (2004:181) goes on to point out that a universal principle such as morphotactic transparency should fulfil the following three main conditions:

(a) It must have a linguistic function (motivation).

(b) It must have a function in performance.

(d) It must have an experimentally tested extralinguistic foundation in performance.

The universal principle of transparency fulfils all these criteria. From the experiment above, transparency facilitates word formation and word processing; hence it has a linguistic motivation. In actual psycholinguistic testing, performance data reveal that transparency as understood as analysability, compositionality and semantic continuities is a crucial factor in morphological analysis. One therefore can conclude that from the semiotic notion of transparency, can emerge from the universal linguistic principle that derived morphological forms will aim to preserve phonetic and morphological continuities as much as possible.

Using transparency as an example, Katamba shows clearly that it is indeed possible to justify the claim that universal linguistic principles can be deductively generated from semiotic principles by subjecting the
principles to performance evaluation. I accept this position, but provide no further psycholinguistic evaluations of the semiotic principles.
CHAPTER 3: METHODOLOGY AND DATA

In this chapter, I will provide brief descriptions of the data sources, followed by a description of how compound words are represented in these two sources. I then provide a step by step account of the methodology used to choose particular compound words from the dictionaries to build representative compound databases in both languages. I then proceed to reclassify this database into semantic categories, before they can be used to demonstrate how the rubric of complexity previously developed in chapter 2 can be used to classify different compounds along the various complexity preference scales that I developed from the NM framework.

3.1 Description of the data sources

My aim is to evaluate the relative complexity of the noun + noun compounding system in JC and Akan. Thus, the appropriate data for such a comparative study of complexity consist of noun + noun compounds from Akan and JC.

The compound words for this study were selected from two historical dictionaries. The JC data were obtained from Dictionary of Jamaican English (hereafter DJE) by F.G. Cassidy and R.B. Le Page ([1967], 1980)
while the Akan data were taken from *A Dictionary of the Asante and Fante Languages called Tshi (Chwee, Tw?i)* (hereafter DAFL) by Christaller (1881). To the best of my knowledge, these are the only existing historical dictionaries for these two languages. I will begin with a brief description of the DJE followed by a similar treatment to the DAFL.

According to Cassidy and Le Page, ‘The Dictionary of Jamaican English is a historical, descriptive dictionary of the English Language in all the forms it has taken in Jamaica since 1655’. However, by Cassidy and LePage’s own admission, this description is somewhat misguided as a great deal of the material recorded actually represents Creole rather than English forms. The dictionary labels these as ‘dialect forms’ with the abbreviation (dial); therefore, ‘dialect’ is best interpreted as JC. These JC words were the main focus of this thesis.

The data in the dictionary were primarily collected by Cassidy (1952) through the following techniques:

(1) The identification of Jamaicanisms in the historical literature on Jamaica up to the year 1850.
(2) The collection of words through questionnaires used with various types of informants such as cane cutters, school teachers, social workers, skilled and unskilled workers etc.

(3) Tape recordings of informants’ free speech, songs, biographies and stories.

Additional data came largely from a ‘Dialect word and phrases competition’ sponsored by the country’s leading newspaper, the Daily Gleaner in 1943. From this, valuable word lists were drawn.

The words collected in the DJE are classified to reflect the geographical location of the informants, who represent an islandwide distribution. The words are orthographically represented by a phonemic spelling system now known as the Cassidy-Le Page writing system for JC. The dictionary stands out as a quality repository of lexical information from which a substantive database of noun + noun compounds in JC can be built.

The DAFL was produced primarily for missionary purposes. The primary author, a Wesleyan Methodist, wanted to create a document to facilitate literacy development in Akan for both the native Africans and the missionary Europeans. The data were primarily collected by Christaller from personal conversations with informants across a wide geographical
area. The data reflect folklore, proverbs, stories etc. and cover a number of Akan / Twi dialects which include Ashante, Fanti, Akyem and Akuapem.

Based on conversations with native speakers of Akan, it was pointed out that the dialectal differences among the Akan / Twi dialects are minor. Therefore, there was no need to overly narrow my search in the DAFL to ensure that only Akan / Twi words were selected seeing that Asante and Fante forms were also included. This dictionary dates over 130 years. It is for these reasons that the DAFL can be considered a useful source from which to obtain noun + noun compounds in Akan to create a suitable database.

In rounding off this section of the discussion, let me reemphasize that both historical dictionaries provided an excellent, readily available and representative source for noun + noun compound words in their respective languages. Despite both dictionaries being created historically to serve different purposes, they have become standard sources of information for researchers working on both languages. Selecting the data from historical dictionaries rather than contemporary dictionaries or compilations allowed me to identify both diachronic (compounds labeled as obsolete) and synchronic patterns of compounding.
I will now turn to a description of how noun + noun compound words are defined and represented in both historical dictionaries.

3.2 Deciding what is a compound word based on the historical sources

The DJE, through orthographic devices distinguishes a compound word from phrases in two ways. A compound word is written with a hyphen separating the constituents for example, **ije** ‘age’ + **piepa** ‘paper’ = ‘birth certificate’ (pg. 6) or sometimes with the absence of one, for example, **aligeta** ‘alligator’ + **apl** ‘apple’ = ‘apples that alligators feed on’ (pg. 8).

Both words are explicitly labeled with the designation ‘**sb**’ which means substantive and identifies the overall syntactic category of the compound word as a noun. Constructions deemed to be phrasal compounds such as **redi** ‘ready’ + **moni lan** ‘money land’ = ‘land bought with cash’ (pg. 376) are labeled as ‘**sb phr**’ which means the construction is of a phrasal nature.

There were also noun + noun combinations which overall produced a word of a different syntactic category as in **bod** ‘bird’ + **mout** ‘mouth’ (pg.43) which overall functions like a pre-modifier similarly to a descriptive adjective. For example, ‘the bird-mouth mango’, which means
‘a mango, based on its appearance, that looks as if it had been pecked by a bird’. However, an alternate interpretation is available, in the sense that [bird-mouth] + mango] could be a complex compound in which the noun mango is pre-modified by the noun bird – mouth. These types of combinations were labeled as ‘adj’ based on syntactic behaviors which overlapped with adjectives.

In examples such as **baitn** ‘biting’ + **ans** ‘ants’ = ‘a type of black ants’ (pg.45), confusion exists around whether or not the syntactic category of the first constituent ‘**baitn**’ is a present participle verb or a gerund functioning as a nominal form. Cassidy offers no resolution on the matter. I therefore excluded these from the JC database and similar forms from the Akan database. Future work needs to be undertaken to determine the true syntactic category of these constituents.

I also left out high frequency forms such as **bakra** ‘backra’+ **cabij** ‘cabbage’ = ‘a cabbage of superior fine quality’. In these compounds, the modifier supplies the meaning ‘superior’ as in ‘**bakra** ‘backra’ + **calalu** ‘calaloo’ (pg. 18), ‘**bakra** ‘backra’ + **pain** ‘pine’ (pg. 19) and **bakra** ‘backra’ + **yam** ‘yam’ built by analogy based on the morphological pattern provided by **bakra cabij**. Forms like these, I consider better
treated as cases of affixation. Much like the high frequency forms in Akan that were also not included (see below). A closer examination of the JC form, ‘bakra’ indicates that the literal lexical meaning ‘white man’ is overshadowed by the extended meaning ‘superior’. This extended meaning remains transparent in combinations consisting of the form ‘bakra’. A distinction can be drawn between a transparent compound word like dakta ‘doctor’ + shap ‘shop’ = a dispensary or drugstore and words containing the high frequency form ‘bakra’. In dakta ‘doctor’ + shap ‘shop’, both forms maintain their transparent meanings consistently. On the other hand, high frequency forms provide a transparent meaning from a semantic extension of the word itself. This extended meaning comes to predominate and is supplied lexically like a prefix to any combination that the high frequency form becomes a part of. Another example of high frequency forms in JC would be ‘man’ and ‘uman’ to indicate lexical gender distinctions. In combinations involving these words, they do not supply the transparent literal meaning of man or woman, but provide instead the adjectival extended meanings ‘male and female’. An example is ‘man ‘man’ + ginep ‘guineap’ = ‘a guinep tree that does not bear the actual fruit’. Here the word ‘man’ indicates masculine gender characteristics embodied by the plant based on its inability to reproduce and not that the tree is literally a man or owned by a man etc.
In review, my designation of items as lexical prefixes is based on the following properties. (1) The word supplies an extended yet transparent meaning in combinations that it occurs rather than the literal lexical meaning. (2) A lexical prefix differs from traditional conceptualizations of prefixes in one major way. Unlike a language such as English, where derivational suffixes like ‘-ion’, ‘-ness’ etc., prototypically do not function independently of the bases they are attached to, lexical prefixes can. (3) These high frequency forms show similar semantic behaviours to English combining forms such as ‘poly’ as in polyglot, polyandry etc., that are considered to be ambiguous between a derivational form and a compound word.

Where Cassidy and Le Page clearly indicated that a word has an Old English origin, I left it out as in ‘bort’ ‘birth’ + nait paati ‘night party’ = ‘a party given on a baby’s birth’. They identified this word as having its earliest citation in 1849 in the (Oxford English Dictionary (OED). I treat these as not true JC compounds.

Christaller identified compound words in Akan using the abbreviated form (cpds). He suggests that some high frequency forms such as ‘ba’ in words like bákònː [basa, kònː] = ‘wrist’ (pg 4) really represents
compounding. I disagreed with this suggestion since these forms appear to behave more like derivational prefixes or the category called combining forms in English such as ‘poly’ as in polyandry, polybasic, polyglot, polymorphous etc. It is still unresolved whether these forms represent compounding or not in English. I have not included them in my database for Akan.

The high frequency forms in both JC and Akan can be considered as derivational morphemes given their similarity in semantic behavior to English combining forms.

The most explicit way in which Christaller presents a compound word is the following – abakyere [basa, kyere], ‘gold and costly beads tied round the wrist in honour of one’s birth or deed’ (pg. 4). In this example, Christaller has provided the two forms that the compound word originated from in brackets. Since Akan compound formation is subjected to phonological modifications such as affix deletion, syllable deletion, vowel harmonization etc., it is important for him to highlight the original forms prior to the application of these processes. The non-bracketed form represents the compound word after any of these processes apply.
In other instances, Christaller presents a compound word as two constituents separated by a hyphen as in ‘ो-बेन-तिया ‘horn’ + ह्ये-ने-फो ‘player’ = ‘the king’s horn player’ (pg. 13). This compound word is presented with the noun prefix (ो) on the first constituent. When compounds are presented like this, they are still placed in my database of Akan compound words. They have not been excluded despite the phonological processes involved in transforming them into compound constituents were not made explicit.

Despite orthographic representations being used as the main criterion for compound status, the researchers did not explicitly indicate whether they had made use of traditional syntactic, morphological, phonological and semantic tests to determine if these words were compounds or not. Nonetheless, both authors’ description of a compound word seems to coalesce with some prototypical properties identified with compound words in the literature. By examining the structures put forward as compound words, they typically exhibit the following properties: binarity, respect for lexical integrity and headedness. I will elaborate these in turn.

1. The Jamaican and Akan noun + noun compound words in the databases are prototypically bimorphemic structures. The bases or constituents
within the compounds are composed of at least two nominal words or two nominal stems respectively. This does not mean that JC and Akan cannot have tri-constituents and longer compounds. Given the property of recursivity, left and right branching expansion can take place to create novel noun + noun compound words in both languages. Hence, compounds which take the superficial form N1 + N2 + N3, can in fact be analyzed as containing substructures: N1 + [N2 + N3] or [N1 + N2] + N3. An example in JC would be **dakta** ‘doctor’ + **shap** ‘shop’ = ‘a dispensary, chemist’s shop or drugstore’. Through right branching a novel compound such as *[dakta ‘doctor’ + shap ‘shop’] + kaana ‘corner’ = ‘the corner of the doctor’s shop’ can be created. Even in this instance, binarity is preserved.

2. These bimorphemic morphological structures demonstrate strong syntactic atomicity or lexical integrity. Lexical integrity is defined by Anderson (1992:8) cited in Mithun (2010:48) as ‘the syntax neither manipulates nor has access to the internal structures of words’ The assumption guiding this principle is that the morpho – syntactic properties of individual constituents in the compound cannot be manipulated by an external linguistic unit or syntactic rules such as agreement principles. Both Jamaican and Akan noun + noun compound words do not allow individual constituents to be modified by an adjective or allow any anaphoric relation
between a compound constituent and an external element. Additionally, syntactic manipulation of the internal word structure of noun + noun compounds of both languages is not tolerated by processes such as extraction and fronting, focusing, coordination and ellipsis.

3. The noun + noun compounds are prototypically morphosemantic and morphosyntactic headed structures. The presence of an endocentric morphosemantic head tends to coincide with the internal morphosyntactic head into one dominant element in the compounds. Sometimes the morphosemantic head is an external referent and in this instance the morphosemantic head does not coincide with the morphotactic head. There were instances in which two heads existed. The majority of the compounds were right headed structures. Compounds in Akan and Jamaican were not treated as headless structures in this thesis.

In summary, we can conclude that Cassidy and Le Page along with Christaller consider a compound to be a prototypical bimorphemic morphological unit that displays strong syntactic atomicity and are generally head-modifier structures. I note however that the researchers did not indicate whether use was made of the traditional morphological, phonological and syntactic tests for compound status.
Based on Cassidy and Le Pages’ and Christaller’s descriptions of what a compound word is, I engaged in the following steps to build two separate yet comparable databases of noun + noun compound words in both Jamaican Creole and Akan.

3.3 Methodology – collecting and coding compound word

3.3.1 Step one – Selecting words from the dictionaries for the databases

I went through DJE page by page and selected all the noun + noun compound words using Cassidy’s criterion for what he labels as a compound word. I adopted a similar approach to the DAFL, with some modifications. Not being a native speaker of Akan presented the challenge of assigning each constituent in the compound its relevant meaning. For example, o\h\émm\a [\oh\ene ‘chief’ + \oba ‘wife’] = ‘wife of a chief or queen’ (pg. 173), On encountering this compound word, I had to search individually for each compound member as a monomorphemic item, taking into careful consideration the presence of tones and the overall meaning of the compound word. I could then assign each component its lexical meanings as in ‘\oh\ene’ = ‘chief’ and ‘\oba’ = ‘wife’. To overcome
the possibility of having misassigned a meaning, I had each meaning checked by a native speaker of Akan / Twi. The native speaker also provided me with information on the cultural and social context which was not provided by the dictionary to help me understand the meanings of the words more clearly.

3.3.2 Step two—Organising the words into semantic categories in the database

In order to work out the semantic relation between the constituents of each compound in a systematic way, I adopted Gange and Shobens’ Thematic Relation model (1997) for nominal compound interpretation. The model is known as CARIN (Competition Among Relations in Nominals). The creators of this model make three critical assumptions. (1) The language user interprets a compound word by making use of their knowledge about both the meaning and the use of the concepts combined. (2) Information about the semantic relation is stored directly with the modifier concept and thereby more reliably than the head guides the search for the correct interpretation of a combined concept. (3) Competition among the various thematic or semantic relations between the constituents will make sure that the one that gets selected is the one most frequently used, with the greatest strength and most readily available.
The first observation to note is that both dictionaries simply supplied a
word based description of the overall meaning of the compound entries.
What was not always provided was information pertaining to social and
cultural context of usage. Hence, one pitfall was that the semantic relation
between the constituents could not always be derived from the lexicalized
meaning provided.

I use an Akan example, to demonstrate what I mean here. In the
compound word béemma ‘arrow’ + hama ‘string’ = ‘bow string’ (pg. 12).
Without appropriate recourse to discourse, encyclopedic and possible
cultural knowledge, deciding which semantic relation below applies to the
lexicalized definition might be problematic.

- Location / container – ‘a string located on a bow’
- Use – ‘a string used to operate a bow’
- Possession – ‘the string of the bow’

In applying the model to the compound word above, I noted the
lexicalized meaning provided by the dictionary. The meaning indicates
that the overall object being named is primarily a ‘string’. However, I
place the focus on the modifier element as being mainly responsible for
providing cues according to CARIN, as to how language users might have
created this word in the first instance. I therefore conclude that the modifier ‘bow’ indicates a possessor of the head ‘string’ and that the semantic relation in this instance is one of possession.

I maintained this approach to the other words in the database and based on my application, I derived the semantic relations/categories below from the data.

One crucial assumption guiding my classification is the existence of an underlying modifier + head relationship template guiding the creation of compound words in JC and Akan. Compounds in their overt forms may not maintain this underlying relationship due to equi-pollence, metaphoricity, exocentricity etc.
(a) **Possessive compounds**

Modifier is a possessor of the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

<table>
<thead>
<tr>
<th>N1</th>
<th>‘possessor of’</th>
<th>N2</th>
</tr>
</thead>
</table>

JC – [**dakta** ‘doctor’ + **shap** ‘shop’] = ‘a dispensary, chemist’s shop or drugstore’

Akan - **On** wèm – mó [n wene ‘potter’ + obo ‘clay’] = ‘potter’s clay’

Modifier is an entity possessed by the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

<table>
<thead>
<tr>
<th>N1</th>
<th>‘possessed by’</th>
<th>N2</th>
</tr>
</thead>
</table>

JC – [**wait** ‘white’ – **beli** ‘belly’+ **rat** ‘rat’] = ‘the cane rat’

Akan – Pattern not found in dictionary

(b) **Location compounds**

Modifier indicates a location or habitat of the head

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

<table>
<thead>
<tr>
<th>N1</th>
<th>‘is location / habitat of’</th>
<th>N2</th>
</tr>
</thead>
</table>

JC – [**kotn** ‘cotton’ – chrii ‘tree’] + **wuurm** ‘worm’] = ‘the timber worm’

Akan - n kwa – nōmā [kwae ‘forest’ + anoma ‘a bird or fowl’] = ‘a bird from the primeval forest’
Modifier identifies the entity that occupies a location indicated by the head.

\[[\text{modifier}] + \text{semantic relation} + \text{[head]}\]

\begin{align*}
\text{N1} & \quad \text{‘is located in’} & \quad \text{N2}
\end{align*}

JC – [jinga ‘ginger’ + piis ‘piece’] = ‘field where gingers are grown’

Akan - owupá [owu ‘death’ + mpa ‘bed’] = ‘the bed on which someone dies’

(c) **Material compounds** – modifier identifies material that constitutes the head

\[[\text{modifier}] + \text{semantic relation} + \text{[head]}\]

\begin{align*}
\text{N1} & \quad \text{‘constitutes’} & \quad \text{N2}
\end{align*}

JC – [krokus ‘crocus’ + bag ‘bag’] = ‘a bag made from crocus’

Akan - mmekwan [abe ‘palm – nut’ + nkwan ‘soup’] = ‘palm – nut soup’

(d) **Container compounds** – Modifier identifies a container in which the head is contained.

\[[\text{modifier}] + \text{semantic relation} + \text{[head]}\]

\begin{align*}
\text{N1} & \quad \text{‘is a container containing’} & \quad \text{N2}
\end{align*}


Akan – asà + hiná [nsà ‘palm wine’+ ahina ‘pot’] = ‘pot for palm – wine’
(e) **Instrument compounds**

Modifier identifies an object associated with the instrument created by the head

\[ \text{[modifier]} + \text{semantic relation} + \text{[head]} \]

N1 ‘creates an instrument with’ N2

JC – [faiya ‘fire’ + stuon ‘stone’] = ‘the stone on which the pot is rested while cooking’

Akan – Pattern not found in dictionary

(f) **Agent nominal compounds**

Modifier identifies the object of the nominal agent indicated by the head

\[ \text{[modifier]} + \text{semantic relation} + \text{[head]} \]

N1 ‘undergoes the activities of’ N2

JC – [dopi ‘duppy’ + kecha ‘catcher’] = duppy (ghost) catcher

Akan – [o\,ke\,te\, ‘mat’ – soafo ‘carrier’] = ‘carrier of the king’s mat’

(g) **Body part compounds** – Modifier identifies a body part which combines with the head to produce a name for a section of the body.

\[ \text{[modifier]} + \text{semantic relation} + \text{[head]} \]

N1 ‘is a body section of’ N2

JC – [fut ‘foot’+ batam ‘bottom’] = ‘the bottom of the feet (sole)’

Akan – Pattern not found in dictionary
(h) **Food compounds** - Modifier identifies animal that eats the food item indicated by the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

\[\text{N1} \quad \text{‘is consumed by’} \quad \text{N2}\]

JC – [**bod** ‘bird’ + **pepa** ‘pepper’] = ‘a variety of peppers favoured by birds’

Akan – Pattern not found in dictionary

(i) **Time compounds** - Modifier indicates a time associated with the head

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

\[\text{N1} \quad \text{‘is a time for’} \quad \text{N2}\]

JC – [**mie** ‘may’ + **bog** ‘bug’] = ‘beetles which are evident around the month of May’

Akan – [**adekyee** ‘the whole day i.e. morning, noon afternoon and evening’ + **soroma** ‘star’] = ‘morning star or day star’

(j) **Collective compounds** – Modifier identifies a cover term for a set of items identified by the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

\[\text{NI} \quad \text{‘is a general category for’} \quad \text{N2}\]

JC – [**bred** ‘bread’/ **fuud** ‘food’ + **kain** ‘kind’] = ‘various kinds of starchy foods such as dasheen, yams, pumpkins etc’

Akan – Pattern not found in dictionary
(k) **Hybrid compounds** - Modifier identifies the object that is produced through cross breeding.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

N1 ‘is a half breed as indicated by’ N2

JC – [muul ‘mule’ + raial ‘royal’] = ‘a cross bred horse of donkey and horse ancestry’

Akan – Pattern not found

(l) **Quantity compounds** - Modifier indicates entity which is in the quantity indicated by the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

N1 ‘occurs in the quantity of’ N2

JC – Pattern not found in dictionary

Akan - berefi [bere\(\tilde{\text{w}}\) ‘palm branches / leaves’ + efi ‘bundle’] = ‘bundle or heap of palm branches’

(m)**Source / origin compounds** - Modifier indicates the source from which the head originates.

\[
\text{[modifier]} + \text{semantic relations} + \text{[head]}
\]

N1 ‘is the source for’ N2

JC – Pattern not found in the dictionary

Akan – [abus\(\text{ùa}\) ‘family’ + y\(\text{àrê}\) ‘sickness, illness, disease’] = ‘hereditary disorder’
(n) **Characteristic compounds**

Modifier identifies a type of person who displays the ‘general characteristics’ identified by the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

\[
\begin{align*}
\text{N1} & \quad \text{‘displays’} \quad \text{N2}
\end{align*}
\]

JC – Pattern not found in the dictionary

Akan - **akwā – kyére** [akoa ‘servant, subject, slave’ + akyere ‘wretch worthy or destined to be killed’] = ‘a rascal’

Modifier identifies a non – human entity associated with the ‘general characteristics’ indicated by the head.

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

\[
\begin{align*}
\text{N1} & \quad \text{‘displays’} \quad \text{N2}
\end{align*}
\]

JC – Pattern not found in the dictionary

Akan - **oguammuruwá** [oguan ‘sheep’ + aburuwa ‘that has not produced any young’] = ‘a grown sheep that had not had any young’ (infertility)

Modifier identifies a ‘general characteristic’ associated with the head

\[
\text{[modifier]} + \text{semantic relation} + \text{[head]}
\]

\[
\begin{align*}
\text{N1} & \quad \text{‘displays’} \quad \text{N2}
\end{align*}
\]

JC – Pattern not found

Akan - **osiâm – má** [siane ‘disgrace’ + oba ‘child’] = ‘a child or person who brings mischief’
(o) **Property compounds** - Modifier identifies an object or situation that receives the benefits / properties provided by the head.

[modifier] + semantic relation + [head]

N1  ‘benefits from / is cured by’  N2

JC – Pattern not found in dictionary

Akan - **akeka + duru [okekaw ‘tooth ache’ + aduru ‘medicine’]** =

‘medicine for toothache’

(p) **X is about Y compounds** – Modifier identifies the object that is the major focus of the head.

[modifier] + semantic relation + [head]

N1  ‘provides focus for’  N2

JC – [anansi ‘a spider’ + tuori ‘story’] = ‘a fairytale with anansi (the spider) as its prominent figure

Akan – Pattern not found in dictionary

(q) **Metaphorical compounds** – the modifier provides the metaphorical entity that the head is compared with based on some metaphorical resemblances such as colour, size, appearance etc.)

[modifier] + semantic relation + [head]

N1  ‘metaphorically resembles’  N2

JC – [kraab ‘crab’ + spaida ‘spider’] = ‘a spider that looks like a crab in its general form’
Akan – [n\_kura ‘mice’ + n\_hwi ‘hair’] = ‘soft, down hair on babies or face
when they first appear’

The metaphorical link is reflected in the semantic relation.

1. **Possessive metaphorical compounds** – The modifier is a possessor of the exocentric metaphorical head.

   ![diagram](modifiers-and-semantic-relations)

   \[\text{[modifier]} + \text{semantic relation} + \text{[exocentric head]}\]

   \[\text{N1 ‘is a possessor of’ N2}\]

   JC – [rieza ‘razor’ + mout ‘mouth’] = ‘the cutting edge of a razor’

   Akan - o\_dam + p\_rè [o\_dan‘house’ + aparew ‘rib’] =

   ‘framework of a roof’

   The metaphorical link lies between the exocentric head and its
   ‘place holder’, ‘mouth’

2. **Locational metaphorical compounds** – The modifier indicates a location or habitat of the exocentric metaphorical head.

   ![diagram](modifiers-and-semantic-relations)

   \[\text{[modifier]} + \text{semantic relation} + \text{[exocentric head]}\]

   \[\text{N1 ‘is the location / habitat of’ N2}\]

   JC – [stuon ‘stone’ + ier ‘hair’] = ‘lichen on rock’

   Akan – Pattern not found in dictionary
The metaphorical link lies between the exocentric head and its place holder ‘ier’

3. **Complete metaphorical compounds** – Both constituents metaphorically describe some property of the exocentric head.

\[ \text{[modifier]} + \text{semantic relation} + \text{[exocentric head]} \]

\[ N1 \text{ ‘is the possessor of’ N2 } \]

JC – \[ \text{Gad } \text{‘God’ + aas ‘horse’} \] = ‘the praying mantis / stick insect’

Akan – \[ \text{nne\-n\-kye\-n – sè [o\-den\-kye\-m ‘alligator / crocodile’} + esè ‘lower teeth’] = ‘a prickly plant’

The metaphorical relationship of resemblance exists between both constituents and the exocentric head. This non-transparent reading overrides the underlying transparent ‘possessive’ interpretation.

(r) **Equipollent compounds (double headed)** – modifier identifies an entity which is a hyponym of the head. (Double semantic head)

\[ \text{[modifier]} + \text{semantic relation} + \text{[head]} \]

\[ N1 \text{ ‘is a type of’ N2 and vice versa} \]

JC – \[ \text{ruusta ‘rooster’ + kak ‘cock’} \] = rooster

Akan – Pattern not found in dictionary
(s) **Idiomatic compounds** – Both constituents are opaquely related to the exocentric head though nicknaming, euphemism, humour, perjoration etc.

[modifier] + semantic relation + [head]

N1 ‘various semantic relations underlingly’ + [exocentric head]

JC – [saal ‘salt’ + daag ‘dog’] = ‘nickname for fritters (johnnycakes)’

Akan – [guan ‘sheep’ + tiri ‘head’] = ‘a fool’s errand’

The idiomaticity rises from the relationship between the exocentric morphosemantic head and the internal compound constituents. The transparent readings are masked by distance away from the social settings, discourse context and pragmatic markers that gave rise to the humour, basis for the nickname and the euphemism used in creating a name.
(t) **Clipped head compounds** – Modifiers are the possession of an external clipped head.

JC – [**lang-tiel** ‘long -tail’ + (**homin-bod** ‘humming-bird’) = ‘the streamer tail humming bird’

Akan – Pattern not found in dictionary

These compounds become a complex metonymical type when the head is clipped. For example, if the head ‘homin bod’ is clipped, you would be left with ‘lang tiel’. ‘Lang tiel’ would still be understood as referring to a humming bird. I contrast these with compounds like ‘strienja ‘stranger’ + flai ‘fly’ = ‘a moth / a firefly’. When these compounds are clipped you end up with a monomorphemic item which can no longer be analysed as a compound.
(u) **Part – whole compounds** - Both constituents (not just a single modifier) make up the external whole

\[ \text{[modifier]} + \text{semantic relation} + \text{[exocentric head]} \]

\[ \text{N1 'constitutes' N2} \]

JC – \([\text{buod 'board'} + \text{angl 'handle'}]\) = ‘a cutlass with a handle made from board’.

Akan – Pattern not found in dictionary

The relationship between the exocentric head and the internal constituents is important here. Unlike previous compounds where an exocentric head was associated with partial or completely non-transparent constituents, the constituents remain completely transparent.

(v) **Partial Folk etymology** – compounds in which only one constituent has undergone phonological reinterpretation and consequently its meaning remains unmotivated in the compound.

JC – \([\text{gungo 'congo'} + \text{piiz 'peas'}]\) = ‘peas’

The element ‘gungo’ is a phonological reinterpretation of ‘congo’ according to Cassidy and Le Page (2002: 118)
Complete Folk etymology – Compound was created from the phonological and morphological reinterpretation of borrowed, technical or scientific vocabulary associated with both constituents.

JC – [baal ‘ball’ + plet ‘plate’] = a large blue pigeon with a white head.

Akan – Pattern was not found in dictionary

The words [baad + pate] were phonologically reinterpreted to create the compound above. Cassidy and Le Page (2002)

3.3.3 Step three – Classifying words as least complex, moderately complex and complex using rubric and scales of preference

I applied the rubric of complexity and the preference scales developed in chapter 2 to assign each compound word in each semantic category a complexity status. Below I present explicit demonstrations of this application using Jamaican Creole examples starting with the least complex category.

Table 6, below indicates in a left to right reading how I proceeded in assigning the label least complex to compound words. The process begins with taking a compound word from a semantic category, in this example; the category is instrument compounds. Recall that in section (3.3.2) I made the assumption that compound formation entails an underlying
morphological form in Jamaican and Akan which is a ‘Modifier + Head’ structure.
Table 6

Application of rubric and preference scale to compound word classification (least complex compound words)

<table>
<thead>
<tr>
<th>Compound</th>
<th>Underlying form</th>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>aam</strong> ‘arm’ + <strong>tik</strong> ‘stick’ = ‘crutches’</td>
<td>Modifier + head</td>
<td>‘tik’ is the morphosemantic head of the compound.</td>
<td>modifier + head Aam tik</td>
<td>How is meaning assigned?</td>
</tr>
<tr>
<td>Semantic category: instrument compound</td>
<td>What is it? The Compound IS A = ‘tik’ The compound IS NOT An = ‘aam’</td>
<td>‘tik’ is the morphosyntactic head of the compound.</td>
<td>Dependency relationship exists between modifier and modified, therefore head is required to be adjacent to the non head.</td>
<td>Check referent in real world. aam tik</td>
</tr>
<tr>
<td></td>
<td>The compound on word formation meaning is ‘a tik associated with an aam’</td>
<td>Both morphosemantic and morphosyntactic head coincide.</td>
<td>Both forms are transparent.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The morphosemantic head = ‘tik’ The morphosyntactic head = ‘tik’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This structure assumes that from a linear perspective, heads are always located in noun + noun compounds in both languages to the right of their modifiers. Recall that my argumentation on the notion of head in a compound is guided by my adoption of Scalise and Guevara’s (2009) idea that there are no headless compounds. This claim suggests that in the compound pattern ‘Modifier + Head’, a head is a co-indexation of a morphosemantic and a morphosyntactic head as one element in a compound word. Such a head is located in the morphological structure and is called an endocentric head. However, there are compounds in which the morphosemantic head is separated from the morphosyntactic head, when this happens, the morphosemantic head resides outside the morphological structure and is called an exocentric head. The compound is not regarded as a headless structure since the ‘place holder’ of the exocentric head in the morphological structure will function as a morphosyntactic head. Recall this discussion in section 2.6.1 on diagrammaticity and headedness.

I applied Allen’s (1978) test ‘IS A?’ to identify the morphosemantic head. The entity in the compound that identifies the referent in the real world that the compound overall refers to indicates the morphosemantic head.

For example, in the JC compound word ‘aam ‘arm’ + tik ‘stick’ as indicated in table 6, after applying the ‘IS A’ test, the answer would be a type of ‘stik’ and not a type of ‘aam’. The morphosemantic head would be an internal constituent and the head would be seen as endocentric.
On the other hand, when the test is applied to a JC compound word like, kyaat ‘cart’ + wiil ‘wheel’, the answer would be ‘a dumpling’, this means that no constituent in the compound identifies the real referent of the compound. When this kind of result is obtained, the morphosemantic head is treated as an external entity and the morphosemantic head would be labeled as exocentric. But the compound is not considered headless since the element ‘wiil’ continues to function as a morpho-syntactic head.

A morpho-syntactic head of a compound contributes relevant language specific morphosyntactic properties to the overall compound. In languages with inflectional affixes to indicate pluralization, gender assignment, etc. morpho-syntactic headhood can be readily identified by locating the constituent in the compound to which the inflectional morpheme attaches itself. In JC, due to the non-existence of these types of morphemes, distributional properties become the most useful way to identify the morpho-syntactic head. A noun in JC can typically occupy three syntactic positions in a sentence: subject, direct object and indirect object. A nominal morphosyntactic head is expected to contribute these syntactic properties to the overall compound. In an endocentric compound, the morphosemantic head typically coincides with its morphosyntactic head. Therefore, having determined that ‘tik’ is the morphosemantic head of ‘aam tik’, I selected it as the potential morphosyntactic head. In the first
group of sentences (1), tik and ‘aam tik’ can occupy the same syntactic distributions. In sentences (2), ‘aam’ presents the same distributional properties. The test at this point appears unrevealing, however, in sentence (2) there is a degradation of the original meaning. The inability of ‘aam’ to preserve the original meaning is taken as an indicator that ‘tik’ is the morpho-syntactic head.

1. Di tik de iina im bag
   The stick COP in his bag
   ‘The stick is in his bag’

Di aam tik de iina im bag
   The arm stick COP in his bag
   ‘The crutches are in his bag’

2. Di aam de iina im bag
   The arm COP in his bag
   ‘The arm is in the bag’

Di aam tik de iina im bag
   The arm COP in his bag
   ‘The crutches are in his bag’
These two tests were applied to all words in the database to determine morphosemantic head and morphosyntactic headhood.

I started with the preference for diagrammaticity; having identified the morphosemantic and the morphosyntactic head, the next issue is to determine if they coincide. After determining that they do, the next step was to indicate where along the scale of diagrammaticity this compound would fall. This is indicated below with (X).

**Preference for diagrammaticity**

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? **YES**
Least complex diagrammaticity
The morpho-semantic and the morpho-syntactic head coincide

Moderately complex diagrammaticity
There are two morphosemantic heads

Complex diagrammaticity
The morpho-semantic head is external, while the morpho-syntactic head is internal

Naturalness scale of diagrammaticity (repeated)
I then proceeded to examine indexicality and transparency in a similar fashion as illustrated below.

**Preference for indexicality**

(2) Does the compound have a modification relationship between its constituents so that its morphosemantic and morphosyntactic head is adjacent to its non head in the compound? **YES**
Least complex indexicality
Morpho-semantic head is internal and adjacent to the non head

Moderately complex indexicality
Morpho-semantic head is internal but not adjacent to its non head.
It is separated by extra material such as infixes.

Complex indexicality
Morpho-semantic head is external

Naturalness scale of indexicality in compounding (repeated)
Preference for morpho – semantic and morpho – tactic transparency

(3) Does the compound have both its non–head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? YES
Least complex transparent
Transparency of both members of the compound

Moderately complex transparent
Transparency of the head member, opacity of the non – head member
OR
Transparency of the non – head member, opacity of the head member

Most complex transparent
Opacity of both members of the compound

Naturalness scale of morpho – semantic and morphotactic transparency in compounding (repeated)
In sum, from the application of the rubric above, one can conclude that simplicity is reflected in the compound since it displays maximal conformity to all the semiotic principles. These compounds are classified as least complex, therefore, with regard to all principles. These compounds are subordinate that is they exhibiting a subordinate-dominant relationship between the head and the non-head. They are also indexical since the endocentric head shares a dependency relation with the non head and is required to be adjacent to it in the morphological structure. Additionally, these compounds are totally transparent hence no semantic opacity. All compounds labelled as least complex are expected to conform to this pattern. A complete listing of all least complex compound words in JC and Akan is available in the database included as Appendix 1 and 2 respectively.

I now apply the rubric to ‘jangkrou lamp’ as an example of compounds which are moderately complex. Emerging from this application were 5 types of moderately complex compound words where complexity arises from opacification processes such as equipollence, metaphoricity and idiomaticization and semantic clipping that have not totally obscured the transparency of all constituents in the compound.
In type-1 moderately complex compounds, as shown in Table 7 below using ‘Jangkrou lamp’ as an example, complexity was achieved by semantic opacity resulting from metaphoricity acting on the modifier alone. The head remained totally transparent in this compound. As indicated in the explanation of transparency, the metaphorical link emerged from mapping properties from the domain of appliances (the instrument ‘lamp’) onto properties from the domain of birds (‘Jangkrou’). In this particular instance based on perceived similarities in appearance, the black smoke of the lamp was mapped onto the black plumage of the bird because of the similarity in colour. The bright, red flame of the lamp was mapped onto the red head of the bird. This link is still based on similarities of colour. By identifying resemblances between these two distinct entities, the language users have created a metaphorical link which might be largely lost to present day speakers, despite, the head remaining immediately transparent.
### Table 7: Moderately complex compound words type - 1 (metaphor acting on modifier alone)

<table>
<thead>
<tr>
<th>compound</th>
<th>Underlying form</th>
<th>Diagrammaticity</th>
<th>indexicality</th>
<th>transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jangkruo</strong>&lt;br&gt;‘carrión’ + lamp&lt;br&gt;‘lamp’ = ‘a type of lamp’&lt;br&gt;Semantic category: Metaphorical compounds</td>
<td>Modifier + head</td>
<td>‘lamp’ is the morphosemantic head of the compound.</td>
<td>Modifier + head</td>
<td>How is meaning assigned?&lt;br&gt;Check referent in real world.</td>
</tr>
<tr>
<td></td>
<td>What is it?&lt;br&gt;The Compound IS A = ‘lamp’&lt;br&gt;The compound IS NOT A = ‘jangkruo’&lt;br&gt;The compound on word formation meaning is ‘a lamp associated with a jangkruo’&lt;br&gt;The morphosemantic head = ‘lamp’&lt;br&gt;The morphosyntactic head = ‘lamp’</td>
<td>'lamp’ is the morphosyntactic head of the compound.</td>
<td>Jangkruo + lamp</td>
<td>jangkruo lamp&lt;br&gt;non literal literal (opaque) (trans)&lt;br&gt;Metaphorical&lt;br&gt;black plumage&lt;br&gt;red head red flame&lt;br&gt;Opaque (modifier) + transparent (head)</td>
</tr>
</tbody>
</table>
Preference for diagrammaticity

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? **YES.** This means the compound would fall on least complex end of the naturalness scale of diagrammatic iconicity in diagram 4.

Preference for indexicality

(2) Does the compound have its morpho–semantic and morpho–syntactic head adjacent to its non head in the compound? **YES.** On the naturalness scale of indexicality, this compound would fall on the least complex pole in diagram 5.

Preference for morpho – semantic and morpho – tactic transparency

(3) Does the compound have both its non – head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? **NO,** the modifier is opaque hence compromising maximal transparency. On the naturalness scale of morphotactic and morphosemantic transparency, this compound would fall on the moderately complex point along the scale on diagram 6.
Moderately complex Type 2 compounds are exemplified by ‘ruusta kak’ in Table 8. Unlike moderately complex type 1 compounds, in which complexity arose mainly from diminishing transparency through metaphoricity on the modifier, these compounds have their complexity largely associated with violations of diagrammaticity resulting in equipollence and loss of indexicality. This violation eradicated the dominance / subordinated relationship between the head and the non head. A relationship of equity exists between both constituents in the sense that each constituent can equally identify the referent of the compound in the real world. This has resulted in the provision of redundant information, ensuring transparency, but violating indexicality. The presence of the two morphosemantic heads violates the expectation that only one head should exist in a modifier + head structure in its overt representation.
### Table 8: Moderately complex compound words type 2 (double headed)

<table>
<thead>
<tr>
<th>Compound</th>
<th>Underlying form</th>
<th>Diagrammaticity</th>
<th>Indexicality</th>
<th>Transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ruusta</strong> ‘rooster’ + <strong>kak</strong> ‘cock’ = rooster / cock</td>
<td>Modifier + head</td>
<td>‘ruusta’ and ‘kak’ are the morphosemantic heads of the compound.</td>
<td>Modifier + head ruusta + kak</td>
<td>How is meaning assigned? Check referent in real world. ruusta kak</td>
</tr>
<tr>
<td></td>
<td>What is it? The Compound IS A = ‘kak’ The compound IS ALSO A = ‘rooster’</td>
<td>‘kak’ is the morphosyntactic head of the compound.</td>
<td>Modification relationship is lost, therefore no dependency relation hold between the constituents despite adjacency satisfied.</td>
<td>literal literal (trans) (trans)</td>
</tr>
<tr>
<td></td>
<td>The compound on word formation meaning is ‘cock that is a rooster’ with the modifier supplying redundant information</td>
<td>With two morphosemantic heads, dominance is lost.</td>
<td>Redundant information provided from underlying modifier.</td>
<td>trans (modifier) / trans (head)</td>
</tr>
<tr>
<td></td>
<td>The morphosemantic head = ‘kak’ and ‘ruusta’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Preference for diagrammaticity

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? NO. Given the existence of two morphosemantic heads, this compound would fall on the midpoint labeled moderately complex in diagram 4.

Preference for indexicality

(2) Does the compound have a modification relationship between its constituents so that its morpho-semantic and morpho-tactic head is adjacent to its non-head in the compound? (NOT APPLICABLE). The modification relationship has been lost due to the presence of the two morphosemantic heads. Therefore, the constituents in the compound share an adjacency relationship with no dependency between them. This compound would be placed at the moderately complex point in diagram 5.

Preference for morpho-semantic and morpho-tactic transparency

(3) Does the compound have both its non-head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? YES. On the scale of morphosemantic and morphotactic transparency, this compound would fall on the least complex end of the scale displayed in diagram 6.
With regard to Type 3 moderately complex compounds, I use *fed* (‘feather’) *fu* (‘foot’) + *fo* (‘fowl’) = ‘a fowl’, as a canonical example. Table 9 sets out the pattern for these compounds. Here, moderate complexity arises from the fact that the morphosemantic head need not be pronounced. When it is not pronounced, an exocentric metonymical compound is produced. For example, if the head ‘fowl’ is clipped from the compound, ‘fed fu’ = ‘feather foot’ is still understood to be a fowl. However, now it becomes an exocentric compound with transparent morphemes. Indexicality is violated, since the metonymical head ‘fowl’ is no longer adjacent to its modifier. On the other hand, when the head is produced a transparent possessive endocentric compound exists.
Table 9: Moderately complex compound words type 3 (clipped head)

<table>
<thead>
<tr>
<th>compound</th>
<th>Underlying form</th>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>transparency</th>
</tr>
</thead>
</table>
| **Feda fut**
‘feather foot’ +
**foul** ‘fowl’
Semantic category:
Clipped headed compounds | Modifier + head
What is it?
The Compound IS A = ‘fowl’
The compound IS NOT A = ‘feda fut’
The compound on word formation meaning is ‘a fowl’ with ‘feda fut’
The morphosemantic head = ‘fowl’
The morphosyntactic head = ‘fowl’ | The clipped external referent ‘fowl’ is the morphosemantic head of the compound.
‘fut’ is the morphosyntactic head of the compound.
The morphosemantic head and the morphosyntactic head do not coincide, dominance is lost. | Modifier + head
Feda + fut
The morphosyntactic head is not adjacent to the non head in the morphological structure. | How is meaning assigned?
check referent in real world. |
Preference for diagrammaticity

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? NO. In relation to diagram 4, this compound would fall on the least complex end of the continuum in regard to diagrammaticity.

Preference for indexicality

(2) Does the compound have its morpho-semantic and morpho-syntactic head adjacent to its non head in the compound? NO. The morphosemantic head remains outside the morphological structure. This results in violation of indexicality based on exocentricity. This compound will fall on the complex end of the naturalness scale of indexicality displayed in diagram 5.

Preference for morpho-semantic and morpho-tactic transparency

(3) Does the compound have both its non-head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? YES. On the naturalness scale of morphosemantic and morphotactic transparency, this compound would be located on the least complex section in diagram 6.
Type 4 moderately complex compounds are exemplified using beli ‘belly’ + ‘cork’ kaak’ = ‘navel’. As seen in Table 10, this type of compound can be contrasted with type 1 compounds where metaphoricity acts on the modifier alone. In contrast, in type 4 compounds, metaphoricity is a property of the exocentric morphosemantic head only. In this example, the metaphorical association arises from the observation that the navel, the real referent of the compound, looks like a cork in the sense that they both close or cover some entity. The literal interpretation of this compound would be a ‘cork for the belly’.
Table 10: Moderately complex compound words type 4 (exocentric head)

<table>
<thead>
<tr>
<th>compound</th>
<th>Underlying form</th>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beli ‘belly’ +</td>
<td>Modifier + head</td>
<td>The external referent ‘A part of the human body – the navel - that metaphorically resembles a cork’ is the morphosemantic head of the compound.</td>
<td>Modifier + head beli + kaak</td>
<td>How is meaning assigned?</td>
</tr>
<tr>
<td>kaak ‘cork’</td>
<td></td>
<td>‘kaak’ is the morphosyntactic head of the compound.</td>
<td></td>
<td>Check referent in real world.</td>
</tr>
<tr>
<td>Semantic category: Clipped head compounds</td>
<td></td>
<td>The morphosemantic head and the morphosyntactic head do not coincide, dominance is lost.</td>
<td></td>
<td>Beli kaak</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>literal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>non trans</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transparent modifier / opaque head</td>
<td></td>
<td>(opaque)</td>
</tr>
</tbody>
</table>
Preference for diagrammaticity

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? **NO.** Given that the morphosemantic head is external, while the morphosyntactic head is internal; this compound will fall on the complex end of the naturalness scale of diagrammaticity outlined in diagram 4.

Preference for indexicality

(2) Does the compound have its morpho-semantic and morpho-syntactic head adjacent to its non-head in the compound? **NO.** This compound will fall on the complex point along the naturalness scale of indexicality in diagram 5.

Preference for morpho-semantic and morpho-tactic transparency

(3) Does the compound have both its non-head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? **NO.** Since only one member of the compound is non-transparent, this compound falls on the moderately complex point along the naturalness scale of morphosemantic and morphotactic transparency. See diagram 6.
The final type of moderately complex compound words is illustrated using ‘kuo’ cow + cod ‘kad’ = ‘a whip’ as an example in table 11. The complexity associated with these compounds is primarily associated with exocentricity. However, these compounds reveal that exocentricity does not necessarily entail transparency as both constituents in the compound are transparent despite the compound having an exocentric morphosemantic head.
Table 11: Moderately complex compound words type 5 (exocentric head)

<table>
<thead>
<tr>
<th>compound</th>
<th>Underlying form</th>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kou 'cow' + kad ‘cod’= ‘whip’</strong></td>
<td>Modifier + head</td>
<td>The external referent ‘A whip made from the cow’s penis’ supplies the morphosemantic properties.</td>
<td>Modifier + head kuo + kad</td>
<td>How is meaning assigned?</td>
</tr>
<tr>
<td>Semantic category: both constituents compose the external modifier</td>
<td>What is it? The Compound IS NOT A = ‘kuo’ The compound IS NOT A = ‘kad’</td>
<td>‘kad’ is the morphosyntactic head of the compound.</td>
<td>The morphosemantic head is not adjacent to the non head in the morphological structure.</td>
<td>Check referent in real world.</td>
</tr>
<tr>
<td></td>
<td>However, the lexicalized meaning is ‘a whip made from the cow’s penis’</td>
<td>The morphosemantic head and the morphosyntactic head do not coincide, dominance is lost.</td>
<td></td>
<td>kuo</td>
</tr>
<tr>
<td></td>
<td>The morphosemantic head = a whip</td>
<td></td>
<td></td>
<td>literal</td>
</tr>
<tr>
<td></td>
<td>The morphosyntactic head = ‘kad’</td>
<td></td>
<td></td>
<td>non literal (trans)</td>
</tr>
</tbody>
</table>

Diagram:

- **Modifier + head**
- **kuo** (literal)  
- **kad** (non literal (trans))  
- Transparent (modifier) / + Transparent (head)
Preference for diagrammaticity

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? **NO.** This compound has an external morphosemantic head despite a modification relationship holding between it and the internal constituents of the compound. See diagram 4.

Preference for indexicality

(2) Does the compound have a modification relationship between its constituents, so that, the constituent functioning as both the morphosemantic and morphosyntactic head needs to be adjacent to its non head in the compound? **YES.** However, the morphosemantic head remains outside the morphological structure and hence this compound will fall at the complex section along the scale of indexicality outlined in diagram 5.
Preference for morpho – semantic and morpho – tactic transparency

(3) Does the compound have both its non – head constituent and its head constituent as morphosemantically and morphotactically transparent morphemes? **YES.** Despite having an exocentric head, both morphemes in the compound remain totally transparent. This compound will be found at the least complex transparent point of the naturalness scale presented in diagram 6.

I now classify the category labelled as complex. **Niigl** ‘needle’ + ‘case’ **kies** = ‘dragonfly’, is used as an example to show how the rubric was applied to such compounds. Complex compounds are expected to violate maximally all three semiotic word formation principles. In ‘niigl kies’, diagrammaticity, indexicality and transparency are all violated maximally. Complexity results from both constituents providing metaphorical references for naming the external referent an insect. ‘Niigl kies’ provides a domain of instruments from which the pointed ends of the needles were mapped onto the slender shape of the insect’s body.
### Table 12: Completely complex compound words

<table>
<thead>
<tr>
<th>Compound</th>
<th>Underlying form</th>
<th>diagrammaticity</th>
<th>indexicality</th>
<th>transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Niigl</strong> 'needle' + kies 'case'</td>
<td>Modifier + head</td>
<td>'insect that looks like a case of needles' is the morphosemantic head of the compound.</td>
<td>'kies' is the morphosyntactic head of the compound.</td>
<td>The head is not adjacent to the non head in the morphological structure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The compound on word formation meaning is ‘a kies’ of ‘niigl’</td>
<td>The morphosemantic head and the morphosyntactic head do not coincide, dominance is lost.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The morphosemantic head = ‘kies’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The morphosyntactic head = ‘kies’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Diagrammaticity** refers to the way the meaning is derived from the parts of the compound.
- **Indexicality** refers to the relationship between the compound and its external referent.
- **Transparency** explores the relationship between the morphosemantic and morphosyntactic heads.
Preference for diagrammaticity

(1) Does the compound have its morpho-semantic and morpho-syntactic head coincide internally in the compound? **NO.**

Therefore, on the naturalness scale of diagrammatic iconicity presented in diagram 4, this compound falls at the complex end.

Preference for indexicality

(2) Does the compound have a modification relationship between its constituents, so that, the constituent functioning as both the morphosemantic and morphosyntactic head needs to be adjacent to its non head in the compound? **NO.** Despite, a modification relationship holding between the exocentric morphosemantic head ‘insect’ and the internal constituents of the compound, the principle of adjacency between dependent elements is violated. This compound will fall at the complex end of the scale shown in diagram 5.

Preference for morphosemantic and morphotactic transparency

(3) Does the compound have its morpho-semantic and morpho-syntactic head adjacent to its non head in the compound? **NO.**

Given that both members of the compound are opaque, this
compound has to be represented at the complex end of the transparency scale in diagram 6.

In summary, using exemplars, I have outlined how the rubric of complexity and the preference scales were applied in classifying compounds as simple/least complex, moderately complex and complex in the databases.
CHAPTER 4: DATA ANALYSIS

The databases of noun + noun compounds in JC and Akan are presented in appendices 1 and 2. I have arranged the words in both databases into similar semantic categories based on the various semantic relations between the constituents in the compound. Additionally, Gange and Shobens’ CARIN model of compound processing and Thematic Relations model were adopted to ensure systematicity in this process. (Recall section 3.3.2.) I have assigned the words in each semantic category different complexity categories using the rubric of complexity and the various naturalness scales as presented in Chapter 3. At this point, I am now prepared to compare the two noun + noun compounding systems to evaluate their complexity relative to each other.

4.1 Determining the basis for the comparison of complexity between the two systems

This thesis gives prominence to preference as an indicator of complexity on the word level. Recall from section 2.3 that preference theories aim to account for properties of linguistic structures based on characteristics such as scalarity, relativity and flexibility when compared to a canonical prototype. Preference will also be used to indicate complexity on a system level. The system level is composed of three categories. First, compounds that display maximum conformity with the universally preferred semiotic characteristics of diagrammatic iconicity, indexicality and
morphosemantic and morphotactic transparency form a category within the compound system referred to as the ‘least complex’ or universally preferred compounds. Secondly, compounds that reflect a mixture of moderate and maximum conformity with the universally preferred semiotic characteristics highlighted above create a category known as moderately complex or universally moderately dispreferred compounds. Finally, the category ‘complex’ or universally dispreferred compounds is composed of compounds that display complete dispreferences for the universally preferred semiotic characteristics.

Intuitively, it might seem plausible to designate one compounding system as more complex if it possesses FEWER ‘least complex’, MORE ‘moderately complex’ and MORE ‘complex’ compound words in comparison to another that shows a reversal.

However, this would be a quantitative measurement of complexity on the word level. This would simply be an actual counting of the number of compound forms making up each category. A quantitative measurement of complexity based on the number of compound words is unhelpful in this thesis for two reasons (a) Representativity and (b) Productivity. The issue of representativity is concerned with the limitations of working with
corpuses and dictionaries. My databases are built from two historical dictionaries that contain compound words recorded at particular times in the past. I am unable to validate if the compounds recorded in the dictionaries is an accurate representation of the number of compounds in actual existence when the data was being collected and compiled.

In relation to productivity, universally highly preferred compounds tend to be very transparent and hence, provide an easy, predictable and economical way of generating new compound words. The challenging question is, were those potential but unrecorded compounds excluded from the dictionaries? If they were, then my count would not have been an exhaustive one.

To overcome the problems associated with a quantitative analysis based on the word level, the comparison of complexity is shifted to the system level. The system level in my conceptualization is composed of semantic categories, morphological patterns and universal semiotic properties like diagrammaticity, indexicality and morphosemantic / morphotactic transparency. The solution embarked on requires that the morphological patterns used to create compound words of the various categories (least complex, moderately complex and complex) be noted. Some
morphological patterns are classified as universally preferred, universally moderately dispreferred and dispreferred. My observation of the database indicates that the universally preferred morphological patterns are always used to create ‘least complex’ compounds. Paralleling this observation, moderately complex and complex compounds are always created by moderately dispreferred and dispreferred morphological patterns also.

I list below the patterns of compound formation that emerge from the databases. These are:

Universally preferred morphological pattern
(a) Transparent modifier + transparent morphosemantic / transparent morphosyntactic head – endocentric

Universally moderately dispreferred morphological patterns
(b) Transparent modifier + (opaque) metaphorical morphosemantic head - exocentric
(c) (opaque) Metaphorical modifier + transparent morphosemantic/ morphosyntactic head - endocentric
(d) Transparent morphosemantic / morphosyntactic head + transparent morphosemantic / morphosyntactic head- endocentric
(e) transparent (Metonymical) modifier + transparent (metonymical) morphosemantic head - exocentric
Universally dispreferred morphological patterns

(f) Opaque (metaphorical) modifier + opaque (metaphorical) morphosemantic head - exocentric

(g) opaque modifier + opaque morphosemantic head - exocentric

By focusing on the morphological patterns, I do not need to count the actual number of compounds in each category. The patterns indicate the type of compound words that potentially can be created within the noun + noun compounding systems of both languages.

The next point to contemplate was how these morphological patterns were used to create compounds to express various semantic categories within each compounding system. (See chapter 3 for a complete listing of these categories). The semantic categories are not treated as preferred or dispreferred.

Before I do an actual comparison, I present three hypothetical situations below. These situations are intended to capture how one compounding system can be labelled overall complex or simple. Complexity is indicated
by preferences for particular types of compounding morphological patterns and compounds to express a variety of semantic categories.

Hypothetical scenario 1

A compound system which expresses **ALL** its semantic categories using a transparent modifier + transparent morphosemantic / morphotactic head head exhibits the universally preferred morphological pattern wherein only ‘least complex’ compounds occur. These are optimally diagrammatic, indexical and transparent. Hence, this constitutes a simple compound system. This compound system shows a preference for very transparent, highly productive and predictable compound words. There would be no opaque forms and by extension language users are not expected to encounter any difficulties in interpreting and creating compound words.

Hypothetical scenario 2

A compound system which expresses **ALL** its semantic categories using an opaque modifier + transparent morphosemantic / morphosyntactic head head or vice versa displays moderately dispreferred morphological patterns so that only moderately complex compounds emerge. These are compounds that are moderately diagrammatic, indexical and partially transparent and will reflect a moderately complex compounding system. A
compound system conforming to the description above has a preference for compound words that violate some semiotic preferences but still provide the language user with access to transparent morphemes to facilitate interpretation and processing. Novel compound word formation is possible though productivity would be restricted in this system. Redundancy is likely to exist and predictability will hold in some instance while not guaranteed in others.

Hypothetical scenario 3
A compound system which expresses **ALL** its semantic categories using an opaque modifier + opaque morphosemantic head accommodates universally dispreferred morphological patterns, hence only complex compounds which are non-diagrammatic, non-indexical and totally non-transparent are created from this compound system. This system shows a preference for unproductive and opaque compound words. Novel compound words based on existing morphological patterns would be difficult to create. All compound forms would have to be learnt and stored in memory. All compound words would require working through metaphorical, specialized or idiomatic inferences to interpret their meanings. Without access to discourse context and cultural knowledge, most L2 learners of such a compounding system would find it difficult to predict compound meanings.
These three extreme hypothetical situations are non-existent compounding systems. In any productive compounding system such as Akan and Jamaican all three patterns described above exist simultaneously but not necessarily equally. Therefore in order to compare the two systems, a way must be found to split each category. This is indicated below.

1. A compound system which expresses **MORE** semantic categories using a transparent modifier + transparent morphosemantic / morphosyntactic head which reflects the universally preferred morphological pattern that constitutes only ‘least complex’ compounds since they display optimally diagrammaticity, indexical and transparency is a **SIMPLER** compounding system than one that uses **FEWER**.

2. A compound system which expresses **MORE** semantic categories using a transparent modifier + opaque morphosemantic head and vice versa accounts for moderately universally dispreferred morphological patterns and by extension creates only ‘moderately complex’ compounds that are moderately diagrammatic, indexical and partially transparent is a more **COMPLEX** compounding system than one that uses **FEWER**.
3. A compound system which expresses **MORE** semantic categories using an opaque modifier + opaque head which reflects a universally dispreferred morphological pattern and exhibits only ‘complex’ compounds that are non-diagrammatic, non-indexical, and totally non-transparent is a more **COMPLEX** compounding system than one that uses **FEWER**.

Having now established a basis for designating one noun + noun compounding system as overall more complex than another, below I present an explicit application of this rubric. That application will make reference to tables 13 – 17 which are provided here.
Table 13: Comparison of least complex compounds between Akan and JC

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple/ least complex</td>
<td>Maximally diagrammatic</td>
<td>1. Possessive&lt;br&gt;- Modifier is a possessor of the head.</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Modifier is an entity possessed by the head.</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Optimally indexical</td>
<td>2. Location&lt;br&gt;- Modifier indicates location / habitat of the head.</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Completely transparent</td>
<td>- Modifier identifies the entity that occupies a location indicated by the head.</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Material&lt;br&gt;Modifier identifies material that constitutes the head</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>
Table 14: Comparison of least complex compounds between Akan and JC (cont’d)

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple/ least complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximally diagrammatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimally indexical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completely transparent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Container</strong></td>
<td>Transparent modifier + transparent head</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td><strong>Instrument</strong></td>
<td>Modifier identifies an object associated with the instrument created by the head.</td>
<td>NO</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td><strong>Body part</strong></td>
<td>One or both elements identify a body part.</td>
<td>NO</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td><strong>Food</strong></td>
<td>Modifier identifies animal that eats the food item indicated by the head.</td>
<td>NO</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td>Modifier indicates a time associated with the head.</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>
Table 15: Comparison of least complex compounds between Akan and Jamaican (cont’d)

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple/ least complex</strong></td>
<td>Maximally diagrammatic</td>
<td><strong>9. Collective compounds</strong>&lt;br&gt;The head identifies various types of the modifier.</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Optimally indexical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Completely transparent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10. Hybrid compounds</strong></td>
<td></td>
<td><strong>The head identifies the modifier that is produced through cross breeding.</strong></td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>11. Quantity</strong></td>
<td></td>
<td><strong>Modifier indicates entity which forms the quantity indicated by the head.</strong></td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>12. Source / Origin</strong></td>
<td></td>
<td><strong>Modifier indicates the source from which the head originates.</strong></td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>
Table 16: Comparison of least complex compounds between Akan and Jamaican (cont’d)

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
</table>
| **Simple/ least complex** | Maximally diagrammatic       | 13. **Characteristics**
Modifier identifies a type of person who displays the ‘general characteristics’ identified by the head. | YES  | NO       |
|                      | Optimally indexical           |                                                        |      |          |
|                      | Completely transparent        |                                                        | YES  | NO       |
| 14. **Property**     | Modifier identifies an object or situation that receives the benefits / properties provided by the head. | YES  | NO       |
Table 17: Comparison of least complex compounds between Akan and JC (cont’d)

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple/ least complex</td>
<td>Transparent modifier + transparent head</td>
<td><strong>15. X is Y</strong>&lt;br&gt;X is about Y compounds – Modifier identifies the object that is the major focus of the head.</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Maximally diagrammatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimally indexical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completely transparent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td><strong>YES</strong></td>
<td><strong>21</strong></td>
<td><strong>21</strong></td>
</tr>
<tr>
<td>YES</td>
<td></td>
<td></td>
<td><strong>14</strong></td>
<td><strong>13</strong></td>
</tr>
<tr>
<td>NO</td>
<td></td>
<td></td>
<td><strong>7</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>
4.2 Comparison of least complex compounds

Table 13 indicates that the category ‘least complex’ from the databases of both Akan and JC are being compared. The universally preferred morphological pattern ‘transparent modifier + transparent head’ is noted. I then noted the various semantic categories expressed by the compound type and the morphological pattern.

The tables indicate that there are twenty semantic categories that were expressed using a transparent modifier + transparent head (universally preferred morphological pattern) to create ‘least complex’ (universally preferred compounds). Akan expressed thirteen of these semantic categories. On the other hand, JC expressed eleven.

The upshot from the comparison is that Akan expresses MORE semantic categories than JC, and is therefore the SIMPLER compounding system of the two in this category. Akan demonstrates a slightly higher preference for least complex morphological patterns and compound words than JC. At this point, Akan appears to be the simpler of the two compounding systems, but this conclusion would be premature considering that other types of compounds like those labeled as moderately complex and complex have not been similarly compared.
4.3 Comparison of moderately complex compounds

To broaden the scope of the comparison, moderately complex compounds across both languages are now being compared shown in Table 14.
Table 18: Comparison of moderately complex compounds between Akan and JC

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Moderately complex</strong>&lt;br&gt;Partial erasure of diagrammatic indexicality semantic and morphotactic transparency</td>
<td>(opaque)Metaphorical modifier + transparent head (endocentric head)</td>
<td>1. <em>Metaphorical compound</em>&lt;br&gt;Modifier provides the metaphorical entity that the head is compared with based on some metaphorical resemblances such as colour, size, appearance etc.)</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Transparent modifier + (opaque) metaphorical head (exocentric head)</td>
<td>2. <em>Locational metaphorical</em>&lt;br&gt;Modifier indicates location / habitat of the metaphorical head.</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. <em>Possessive metaphorical</em>&lt;br&gt;Modifier is a possessor of the metaphorical head.</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>
Table 19: Comparison of moderately complex compounds between Akan and JC (cont’d)

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderately complex</td>
<td>Partial erasure of diagrammaticity, indexicality, semantic and morphotactic transparency</td>
<td>4. Clipped head</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Transparent modifier + transparent modifier. (exocentric head)</td>
<td>Modifiers are the possession of an external clipped head. Part – whole compounds - Both constituents (not just a single modifier) make up the external whole</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Equipollence</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Transparent head + transparent head. (Double morpho - semantically headed compounds)</td>
<td>Both constituents equally identify the same referent of the compound</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(endocentric)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 20: Comparison of moderately complex compounds between Akan and Jamaican (cont’d)

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic relations</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderately complex</td>
<td>Partial erasure of diagrammatic indexicality semantic and morphotactic transparency</td>
<td>6. Partial folk etymology One constituent has undergone phonological reinterpretation</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TOTAL</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>YES</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

TOTAL: 7 (Akan) 7 (Jamaican)
In the moderately complex category of compound words, Akan used only two of the morphological patterns and only two of the semantic categories. Jamaican Creole used all morphological patterns and all semantic categories. The results indicate that JC expresses MORE semantic categories and morphological patterns than Akan. JC is therefore classified as a more COMPLEX compounding system than Akan in this category. JC displays a higher preference for moderately dispreferred compounds and morphological forms than Akan.

The difficulty at this point is that Akan is simpler in one category while Jamaican is complex in one category; this makes it difficult to draw any overall comparison of complexity between the two systems. As a result, it is important to compare the two systems in terms of the category complex compounds.
4.4 Comparison of complex compounds

Table 15 examines how the noun + noun compounding system of both JC and Akan make use of those compounds labeled as complex. These compounds were expressed by two morphological patterns.
Table 21: Comparison of complex compounds between Akan and Jamaican

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Morphological pattern</th>
<th>Semantic categories</th>
<th>Akan</th>
<th>Jamaican</th>
</tr>
</thead>
<tbody>
<tr>
<td>complex non-diagrammatic</td>
<td>Opaque (metaphorical) modifier + opaque (metaphorical) head. exocentric head</td>
<td>1. complete Metaphorical Both constituents metaphorically describe some property of the external head.</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Non-indexicality semantic and morphotactic non-transparency</td>
<td>opaque modifier + opaque head exocentric head</td>
<td>2. Idiomatic Both constituents are opaquely related to the external head through (perjoration, humor, euphemism etc.)</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Folk etymology</td>
<td>Compound was created from the phonological and morphological reinterpretation of borrowed, technical or scientific vocabulary.</td>
<td>3. Folk etymology</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>YES</strong></td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>NO</strong></td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The table shows that Akan expressed only two of these semantic categories and only one of the morphological patterns. In contrast, JC expressed all semantic categories and all morphological patterns. JC expresses MORE semantic categories; therefore, JC is a more COMPLEX compounding system than Akan in this category. JC displays a slightly higher preference for expressing semantic categories with more universally dispreferred morphological patterns and universally dispreferred compounds.

{indicate that this was not a counting of occurrence but of patterns and categories}

Before I come to an overall conclusion, I return to the research questions posed in chapter 1. The thesis aims to provide a definitive response to the following questions.

4.5. Research questions and complexity

(a) Did the comparison of complexity reveal any differences between the noun + noun compounding systems of Jamaican Creole and Akan?
The restricted comparison undertaken in this thesis revealed three differences in preferences between the two systems. They are:

(1) The noun + noun compounding system of Akan displays a slightly higher preference for using the (simpler / least complex) universally preferred morphological pattern and more semantic patterns to create universally preferred compounds than JC.

(2) The noun + noun compounding system of JC shows a greater preference for moderately dispreferred morphological patterns and more semantic categories to form moderately dispreferred compounds than Akan.

(3) The noun + noun compounding system of JC indicates a slightly higher preference for using highly dispreferred morphological patterns and more semantic categories to form universally dispreferred compounds than Akan.

(b) Do these differences contribute to differences in complexity between the two noun + noun compounding systems?

These differences in preferences indicate that the JC noun + noun compounding system is more tolerant of dispreferred morphological patterns and ‘moderately complex and complex compound words’ in
expressing more semantic categories than Akan. Based on the stipulation for complexity being used in this thesis, the numerical differences though not massive can be interpreted as indicative of differences in complexity preferences between the two systems. There are five moderately complex patterns not exhibited by Akan.

(c) Which compounding noun + noun system overall is least complex of the two?

Overall, JC consistently indicated a higher preference for the dispreferred section in each category (least complex, moderately complex, complex) compared, I would conclude that the JC system overall is a more complex system than Akan.

4.6 Conclusion
It is important to be able to advance an explanation for the results obtained from the comparisons. Based on Natural Morphology predictions, it is expected that any language with word compounding as a productive morphological device for expanding its vocabulary stock, would have most of its compound words appearing in the least complex category. This prediction is borne out in the data presented, since the majority of the
compound words listed for both Jamaican Creole and Akan (Twi) fall within this category (whether the count is representative or not).

The heavy reliance on a completely transparent, universally preferred compound pattern to express the highest number of semantic categories with the ‘least complex’ compounds in both languages might be related to the demand of productivity and word formation. High productivity requires that forms are kept as transparent as possible to assist in the relative ease of forming new compound words based on existing morphological patterns.

I am assuming that while the morphological component of Akan has evolved to allow some high frequency compound constituents to become suffixes, JC has continued to rely more extensively on noun + noun compound forms. The presence of derivational suffixes in Akan provides an alternative morphological process to extend its vocabulary than compounding.

Given JC’s greater reliance on noun + noun compounding to Akan’s, a greater preference for the exploitation of metaphorical, metonymical and idiomatic references to express a wider semantic range is not surprising in
JC. These processes result in violations of diagrammaticity, indexicality and morphosemantic and morphotactic transparency in a gradient manner, giving an increase in greater complexity in its compounding system.

**Implications for a historical view of complexity**

McWhorter (2001) expressed the view that complexity requires millennia of years to develop and become evident in languages as ornamental adornments. This view of linguistic complexity has been explicitly formulated in relation to morphological complexity. A closer examination of complexity from this perspective reveals that complexity is conceptualized as a time dependent notion. Based on the rubric for diagnosing complexity in this thesis, I do not support the view that complexity is necessarily a time dependent notion. This was most clearly indicated with clipped morphosemantic headed compounds. By speakers omitting the morphosemantic heads of transparent compounds, one can instantaneously create an exocentric form. However, with access to crucial discourse and cultural contexts, this head can be understood. Similarly, the notion of exocentricity as an entailment of non-transparency was not supported with part-whole compounds. Complexity arising from redundancy in compounding was equally impressive in a Creole language contra McWhorter’s 2001 claim. Semantic drifts into metaphoricity and idiomaticization do not require ‘millennia of change’. Some of these
metaphorical links are hardly penetrable to current day speakers and have complexified the noun + noun systems of both languages.

Since McWhorter claims that older languages by and large should express greater complexity than younger / Creole languages, one should expect Akan to demonstrate more abundantly than Jamaican, more moderately complex and complex morphological patterns (the universally dispreferred properties). However, this pattern did not emerge from the investigation.

4.7 Limitations and the way forward

This thesis has attempted to make a contribution to our understanding of complexity in word formation specifically as it regards noun + noun compounding. In doing so, it has also contributed to our knowledge of semantic patterns in compounding, as I found it necessary to develop semantic categorization presented in chapter 3. Nevertheless, many questions remain unanswered, some arising from limitations of the work itself, some thrown up by my own findings. I summarized the main ones here in the hope that this may motivate further research.
(a) Expanding the database

While the current investigation tells us about descriptive complexity differences between the noun + noun compounding system of Akan and JC, it reveals nothing about the compounding systems overall of both languages. Hence the scope of this work can be broadened to include other major word class compounds such as verb + noun compounds or adjective + noun compounds. Incidentally, Akan uses more adjective + noun compounds, the extent to which that reality might throw a different light on the conclusions I have drawn here is worthy of future exploration.

(b) Refining the rubric of complexity

One preliminary observation, I have noted is that verb + noun compounds in JC can productively create exocentric compounds in JC as in buor ‘bore’ + nuoz ‘nose’ ‘a person who bores the nose’ (pg. 60), bos ‘burst’ + belt ‘belt’ = ‘a paunchy man’ (See chapter 1). This has implications for the rubric used to classify compounds as complex and non complex. The rubric will come under greater scrutiny once the database has been expanded since it will now have to address a new range of semantic and syntactic properties associated with verbs and adjectives. The extent to which the rubric will need further refinement for diagnosing complexity in compounding cross-linguistically needs further investigation.
(c) **Defining compounding**

The vexing issue of what is a compound word in both Jamaican and Akan needs to be seriously addressed. Despite accepting Cassidy and Le Page’s as well as Christaller’s labels as indicative of compound status, orthographic representations of compound words can hardly be seen as the most reliable diagnostics for compound hood. An exploration of syntactic, morphotactic and semantics tests and properties need to be undertaken to offer a more definitive definition.

(d) **Does descriptive complexity coincide with actual complexity?**

Performance was not examined in this thesis. Therefore, one crucial limitation of this investigation into complexity is that it offers no insight into whether or not what was described as complex will actually be experienced as complex in performance by language users. This is an important question in the evolving sub-field of language complexity. Having done descriptive work on complexity in compounding, for future work, I suggest that validation can be achieved through experimental tests on compound processing in JC. This can include experiments to measure reaction times in using an existing compounding morphological pattern to create novel compound forms, interpreting exocentric compounds and analyzing compounds into their constituent morphemes. Such results can be analyzed in Libbens’ (1998) theory of compound processing.
The results will facilitate an objective definition of complexity being put forward.
REFERENCES

University of Connecticut, Storrs, Mansfield, CT, USA.


In N. Corver & H. V. Reimsdjik (Eds.), Semi – lexical categories: the function of content words and the content of function words (pp. 415 - 446) Berlin: Walter de Gruyter


