LINGUISTIC EVIDENCE: GRAMMATICALITY JUDGMENTS AND SERIAL VERB CONSTRUCTIONS IN THAI

TERRY ADAMS

Presented in Partial Fulfillment of the Requirements for the Degree of MASTER OF ARTS IN LINGUISTICS

Payap University
June 2012
Title: Linguistic Evidence: Grammaticality Judgments and Serial Verb Constructions In Thai

Researcher: Terry Adams

Degree: Master of Arts in Linguistics

Advisor: Larin Adams, Ph.D.

Approval Date: 20 June 2012

Institution: Payap University, Chiang Mai, Thailand

The members of the thesis examination committee:

1. ________________________________ Committee Chair
   (Associate Professor, Saranya Savetamalya, Ph.D.)

2. ________________________________ Committee Member
   (Larin Adams, Ph.D.)

3. ________________________________ Committee Member
   (Peter Freeouf, Ph.D.)
ACKNOWLEDGEMENTS

I owe the greatest thanks to Dr. Larin Adams who besides sharing my surname also shares a true and deep interest in pursuing unknowns and encouraged me to follow my interests into these deep issues. Without his advice and focused mind I doubt this thesis would have been completed. I also owe many thanks to Dr. Peter Freeouf who has been a most valuable intellectual resource. His wife, Associate Professor Dr. Nitida Adipattaranan, and his daughter Saranta Freeouf helped me explore subtle semantic nuances that only Thai native speakers with a genuine interest in linguistics could do.

The choice of topic for this thesis was not an easy choice to make since so many interesting ideas have presented themselves over the previous four years of coursework in this graduate program. I was advised early on to choose a thesis topic that interested me personally, that was narrow, researchable and at the same time feasible. My initial simple idea of investigating “grammaticality judgments as data” in Thai serial verb analysis led me however, toward a deep examination of unresolved foundational issues in linguistics.

My college roommate from 1968-1970 was Vanich Khanthatham from Prae, Thailand. Our friendship, based on helping each other learn each other’s language, started my interest in Thai. I became interested, not just learning to speak and understand the language as a practical matter, but fascinated also as to why learning the language presents so many particular difficulties for English speakers. It was not the tones, or the alphabet, but the great semantic and syntactical differences between the two languages that made learning each other’s language so challenging. My experiences with learning Russian, German and Spanish posed far less difficulty.

My interest in Thai led to an interest in linguistics which persisted over a lifetime spent teaching physics. The past five post-retirement years at Payap University have been an intellectual adventure rewarding beyond measure. Having professors that
respect the intelligence of their students and talk “to” and not “at” their students has made this an ideal setting to study linguistics. Being the oldest member of the class as well as being older than all of my professors seemed to bother no one. I found that age and life experience can make up somewhat for having less energy and needing more sleep than my classmates. I thank all my teachers for their patience and willingness to always respectfully listen and respond to my questions and encourage my curiosity.

I feel a special debt to my wife, Nancy, who supported my desire to pursue my lifelong interest in linguistics and changed her own life greatly to help make this dream happen.

Terry Adams
ABSTRACT

This thesis is broadly concerned with the status of data used as evidence in linguistics and specifically concerned with testing grammaticality judgments which have appeared in the recent literature in Thai linguistics. A survey using questionnaire and interview techniques checked linguistic claims made by researchers regarding Thai serial verb constructions. The results show that concerns about the problematic nature of simple binary grammaticality judgments used as evidence for theoretical claims are warranted. Various approaches to linguistics are contrasted in their varying interpretations of data as evidence, grammar and the goals of linguistics. Reasons for believing that Thai offers important corrective knowledge for linguistic theorists attempting to make universal generalizations are discussed in detail.
ข้อเรื่อง: ข้อเบิกทางภาษาศาสตร์ ความถูกต้องทางไวยากรณ์และหน่วยสร้าง กริยาเรียงในภาษาไทย
ผู้วิจัย: เทอร์รี่ อดัมส์
ปริญญา: ศิลปศาสตรมหาบัณฑิต สาขาวิชาภาษาศาสตร์
อาจารย์ที่ปรึกษา: ดร. ลาวิท อดัมส์
วันที่อนุมัติผลงาน: 20 มิถุนายน 2555
สถาบันการศึกษา: มหาวิทยาลัยพายัพ จังหวัดเชียงใหม่ ประเทศไทย
จำนวนหน้า: 94
ค่าสำคัญ: ข้อเบิกทางภาษาศาสตร์ ความถูกต้องทางไวยากรณ์ กริยาเรียง

บทคัดย่อ
วิทยานิพนธ์ฉบับนี้ศึกษาสถานะของข้อมูลในฐานะข้อเบิกทางภาษาศาสตร์ และทดสอบความถูกต้องของกริยาเรียงความถูกต้องทางไวยากรณ์ของหน่วยสร้างกริยาเรียงที่ปรากฏในงานวิจัยโดยเครื่องมือที่ใช้ในงานวิจัยชิ้นนี้ได้แก่แบบสอบถามและการสัมภาษณ์จากผลการศึกษาพบว่ากริยาเรียงความถูกต้องทางไวยากรณ์มีความถูกต้องทางไวยากรณ์เชิงทวิภาคแบบไม่ซับซ้อนโดยอาศัยการเปรียบเทียบของแนวทางการศึกษาภาษาศาสตร์ที่หลากหลายในการตีความข้อมูลที่เป็นทั้งข้อเบิกทางภาษาศาสตร์ไวยากรณ์และจุดมุ่งหมายของการศึกษาภาษาศาสตร์ได้อย่างถูกต้อง โดยผู้วิจัยเชื่อว่างานวิจัยภาษาไทยได้นำเสนอองค์ความรู้ที่ถูกต้องสำหรับนักทฤษฎีภาษาศาสตร์ในความพยายามทางข้อสรุปในเชิงแบบลักษณ์ภาษาต่าง

v
# TABLE OF CONTENTS

Acknowledgements .................................................................................................................. ii
List of Tables .......................................................................................................................... viii
List of Figures ......................................................................................................................... ix

Chapter 1 Introduction .......................................................................................................... 1
   1.1 Introduction ..................................................................................................................... 1
   1.2 Thesis overview ............................................................................................................. 2
   1.3 Introduction to the problem .......................................................................................... 3
   1.4 Basic introduction to Thai ............................................................................................. 4
   1.5 Summary of Methodology ............................................................................................ 6
   1.6 Contribution of the thesis ............................................................................................. 6
   1.7 Limitations and scope ................................................................................................... 7
   1.8 Similar research ............................................................................................................ 7

Chapter 2 Theoretical background behind this research ......................................................... 9
   2.1 Introduction ..................................................................................................................... 9
      2.1.1 Landscape of Contemporary Linguistics ................................................................. 9
   2.2 Data in the Externalist approach .................................................................................. 11
   2.3 Data in the Essentialists approach .............................................................................. 14
   2.4 Summary ....................................................................................................................... 17

Chapter 3 Data Collection, Methodology and Results ............................................................ 19
   3.1 Overview ....................................................................................................................... 19
   3.2 Methodology ............................................................................................................... 19
   3.3 Sentences Surveyed ..................................................................................................... 20
   3.4 Summary Chart of Results ........................................................................................... 23
   3.5 Discussion of individual sentences .............................................................................. 24
       3.5.1 Sentence #1: Thepkanjana ................................................................................... 25
       3.5.2 Muansuwan and serial verbs ............................................................................... 26
       3.5.3 Sentences #2-#5: Muansuwan ............................................................................ 28
       3.5.4 Sentences #6-#7: Muansuwan ............................................................................ 33
       3.5.5 Sentences #8-#9: Muansuwan ............................................................................ 35
LIST OF TABLES

Table 1: Three Approaches to the Study of Language ........................................10
Table 2: Informant and Survey Summary .............................................................20
Table 3: Summary of Survey Results .................................................................24
Table 4: Expanded Summary of Results .............................................................47
Table 5: English Grammaticality Judgments .......................................................52
Table 6: Thai verbs scrambled word order ..........................................................70
Table 7: Common collocations of jɔɔŋ .................................................................76
Table 8: Common collocations of klap: (of 28 listed) .........................................76
Table 9: Common collocations of troy: (of 29 listed) .........................................77
Table 10: Common collocations of kхаам (of 16 listed) ....................................77
LIST OF FIGURES

Figure 1: Tai Language Family.................................................................5
Figure 2: Schematic of the Essentialist view of language .........................15
Figure 3: Muansuwan part of speech diagram.........................................26
Figure 4: Thepkanjana SVC Flat SVC Structure.....................................27
Figure 5: Muansuwan Suggested SVC Structure.....................................28
Chapter 1
Introduction

1.1 Introduction
This thesis is broadly concerned with the nature of data in linguistics and more specifically with the role of grammaticality judgments offered as evidence in recent Thai linguistic research. Linguistic research of any type involves data but the definition and meaning of what is considered “data” varies depending on the approach taken to the study of language within the broader discipline of linguistics. Data used as evidence for hypothesized syntactic or semantic structures often consists of simply the self-generated sentences offered by the researcher doing the research. Preliminary steps to verify the grammaticality or acceptability of these sentences are, as a matter of course, not considered necessary. However in recent years research has been done that shows the inadequacy and unreliability of single informant grammaticality judgment data (Featherston, 2007; Schutze, 1996; Cowart, 1997).

Research on Thai serial verbs has yielded a range of acceptability judgments that draw into question the problematic nature of clear grammaticality judgments. As this phenomena became apparent in the researcher’s own investigation, it became desirable to check the data underlying several major works on Thai serial verbs and see how consistently the grammaticality judgments of a larger pool of subjects correlated with the grammaticality judgments that those works were based on. The Thai language along with many other languages in this region like Chinese, Khmer, Vietnamese and Lao exemplifies “an extreme of pragmatically-oriented grammar” (Enfield, 2007: 272). One important consequence of this fact in regard to testing sentences for grammaticality is that a decontextualized sentence can be open to a wide range of interpretations or to no immediate accessible interpretation without appropriate supplemental contextual information. A second fact about languages in this region that complicates syntactic analysis in general is that sentences are frequently one clause made up of multiple verbs. As will be discussed in detail in Chapter 5, this structural feature has resisted efforts of linguists to come to analytical agreement or even descriptive agreement in characterizing the
phenomena of serial verb constructions. The resistance to description or analysis may stem from the fact that sentences in Thai simply do not have to have a main verb as a clausal core feature. The syntactical role of the main verb is to mark clause boundaries but serial verbs in Thai can cohere as one event, describe a temporal series of events, an event and an evaluation of that event or an event and result of that event.

This investigation was undertaken using a basic written survey with bilingual native Thai speakers and personal interviews with monolingual Thai speakers. The results show that the purported acceptability or non-acceptability of some of the presented tokens failed to be recovered when tested with native speakers. In some cases sentences when tested did not fall into the predicted categories of acceptable or unacceptable for a majority of my informants. Additionally, claimed semantic contrasts between certain paired sentences were not found.

1.2 Thesis overview

In Chapter 2, I discuss three broad perspectives to the study of language that have been termed Externalist, Emergentist, and Essentialist approaches to linguistics. Linguistics has emerged historically as a field of study with broadly different conceptions as to what the subject of study should be, what the goal of study should be and what counts as explanation. The concepts of grammar in each of the three approaches are contrasted and the special status of grammaticality judgments in the Essentialist approach emerges as the main problematic issue regarding the status of empirical data used as evidence in theory construction.

In Chapter 3, I present the results of a survey that I performed using questionnaire methods and interview techniques to check the acceptability claims of 20 sample Thai sentences. Each of these sentences has appeared in the literature as tokens that supported specific theoretical claims in Thai linguistic research. The claims are reviewed briefly before the results of the survey are given for each claim. The sentences were surveyed for acceptability by native speakers, meaning recovery of the English free translations and in some cases claimed meaning contrasts in paired sentences. The results of the survey revealed that concerns about the empirical basis of much linguistic research are warranted.

The problem of reliability of grammaticality judgment data and the purported difference between acceptability and grammaticality as distinct concepts is discussed
in some depth in Chapter 4. I tested English sentences with 24 native English speakers that had been offered as evidence for syntactical claims to show the reader the difficulty of obtaining reliable consistent data which can be considered empirical. The normative nature of grammaticality judgments as well as the normative or prescriptive notion of grammar itself is an issue at the core of linguistic research and I review the Essentialist perspective on this issue. Recent empirical research on English syntax by computer linguists developing Natural language processing Technology reveals that the results of Essentialist syntactical research is of little use in this effort.

In Chapter 5 specific issues concerning linguistic theory and the languages of Southeast Asia are discussed. The languages of SE Asia seem to be defined in terms of what their grammars lack rather than how their grammar actually structures their languages. The lack of a core verb is perhaps the issue that seems most perplexing to the Eurocentric view of sentence structure. In this chapter the issues of serial verbs and the confusion surrounding efforts to fit these structures into an analytical framework is reviewed. I return to the data in Chapter 3 and show in detail how certain sentences were problematic due to semantic confusion. I offer the notion of a semantic versus a syntactic grammar and show how sentences in Thai lack the syntactic grid of an English sentence where grammatical morphemes leave slots for open class word substitutions.

1.3 Introduction to the problem
The etymological meaning of the word “data” is “(things that are) given”. When one researcher takes the analysis of another researcher as an example or evidence for argumentation, analysis has now become “data”. Sentences can be used as data and an analysis of a sentence can be data as well. Something is not, in and of itself data, since data is connected to a function that it serves (Lehmann, 2004: 4). Data can be more or less abstract as the two cases above illustrate. In this thesis evidence presented as data in the research of others was checked for the most basic empirical claims and not the abstract claims made for the sentences as evidence for syntactic and semantic argumentation. To make this point more clear, any sentence presented by a linguist as evidence has in Lehmann’s terms become “semiotic object of a higher order”. It is an expression in language coupled with a statement of the meta-language which predicates a certain property over that object. The question posed and answered was, can these sentences stand up to a check with native speakers for
acceptability, meaning translation and contrast claims between certain pairs of sentences?

1.4 Basic introduction to Thai

Thai is the most studied and described language in the family of Tai-Kadai languages. In fact, over half of the approximately 90 million speakers of Tai-Kadai languages speak Thai as a first or second language (Diller, 2008: 31).

In Thailand the linguistic situation is extremely complex and by far the best explication of the linguistic situation in Thailand is Smalley’s *Linguistic diversity and national unity: Language ecology in Thailand* (1994). Thais themselves are often unaware of the fact that a majority of Thais learn Standard Thai in school. Thais with a high school education will say that they are Thai and that they speak Thai. Further questioning is usually required to elicit the fact that they also speak a regional language and that language is actually their first language. The prestige of Standard Thai is unquestioned by all Thais. In fact as Smalley points out, Northern Thais will insist that their language, Khammuang, is not a language in the sense that Thai is a language. It is rather, “village speech”, or just a version of Thai that country people speak. An education, a career, and economic success are dependent upon mastery of Standard Thai. The language variety of Thai that is studied in this thesis is Standard Thai but differentiating Standard Thai from Central Thai is difficult. An interesting definition of what Standard Thai is perceived to be is quoted from Beebe (1974) in Smalley (1994).

Standard Thai has been called a national language, a regional dialect, a language for use in schools, an equivalent of Central Thai, an equivalent of Bangkok Thai, an equivalent of Central or Bangkok Thai, a language not equivalent to, but closely resembling Central Thai, a prestige dialect of educated speakers regardless of origin, and a model of what Central Thai is supposed to be (Beebe, 1974: 74).

The Central dialect, to which Standard Thai belongs, is one of the four major dialects and has 20 to 25 million speakers. The other major dialects are the Northeastern dialect (Isaan or Lao) with about 23 million speakers, the Northern dialect (Kham Muang, Lan Na or Yuan) with 6 million speakers, and the Southern dialect with 5 million speakers (Summer Institute of Linguistics website).
Besides these major dialects, a number of related languages are spoken in the country. Almost all of these languages belong to the Southwestern branch of Tai (Li, 1977). Speakers of these Tai languages comprise about 90% of the whole population. Non-Tai languages include Khmer (Austro-Asiatic), Malay (Austronesian), Karen (Tibeto-Burman), and Hmong (Hmong-Mien). The 2000 census reports that 2.3% of the population speaks Khmer and 2.3% Malay. In addition, in cities and towns southern varieties of Chinese such as Teochiu (or Swatow) and Hakka are commonly found. Languages of the Southwestern branch of Tai are found not only in Thailand but also in Laos, northern Vietnam, Myanmar, India, and also in southern China. The other two branches are the Central and Northern branches. The Central branch includes languages spoken in northern Vietnam and southern China (e.g. Nung, Tay). The Northern branch includes other languages of southern China (e.g. N. Zhuang, Bouyei, Seak). The Tai language family with these three sub-groups is related to other sister and parent branches to make up a larger language stock called Kadai or Tai-Kadai.

The figure below shows the position of Thai in a language family tree. This tree is from The Tai – Kadai Languages (Diller, Edmundson and Luo, 2008: 7) where it is described as “a tentative diagram for reference”.

![Diagram of Tai Language Family](image)

**Figure 1: Tai Language Family**
1.5 Summary of Methodology

The core data for this thesis comes from taking examples of serial verb sentences from the linguistic literature on Thai and asking native Central Thai speakers if they find these sentences acceptable. Both questionnaire methods and interviews were used to gather the data for this thesis. Respondents were asked to rate the sentences for acceptability as OK, confusing but OK, unacceptable, or uninterpretable. I asked these bilingual students to also translate the sentences in English and make comments on the sentences regarding the structure or vocabulary of each sentence. In many cases where the sentences were judged to be unacceptable, reasons for their unacceptability were discussed in subsequent interviews. Approximately half of the informants were fluent to some degree in English and many of their comments appear in Chapter 3. Others were Thai speakers and tapes were made and notes taken on their responses during the interviews.

1.6 Contribution of the thesis

The use of linguist self-generated sentences used as data has been controversial for decades as will be discussed in detail in Chapters 2 and 4. The criticism has focused on Essentialist (generative) methodology in particular but the practice is still being defended by linguists. Jackendoff believes that “it would cripple linguistic investigation to require that all judgments of ambiguity and grammaticality be subject to statistically rigorous experiments on naive subjects” (Jackendoff, 2010). But in this modern age where information is collated, processed, sifted, and stored at a lower cost than ever before, can this lack of interest in the details and specifics of data and data collection be upheld in a field which purports to be an empirical science?

Critics of linguists’ use of intuitive data are not recommending that detailed statistical analysis be done on informant grammaticality judgments. Featherston argues that at an utter minimum for the basis of advancing a linguistic argument multiple informants along with multiple lexical variants of the structures investigated should be required and that this information should be presented as part of the research (Featherston, 2007: 10).

In Chapter 3 I present in detail the actual reality of attempting to access the postulated linguistic competence of Thai native speakers asked to judge acceptability of 20 sentences. The result reveals in detail that often informant judgments are not
reflexive, binary or unproblematic. The more experience one has with this process, the less one can believe that syntactic and semantic levels of interpretation are cognitively separable.

1.7 Limitations and scope
The concept of data in the various approaches to linguistics is a wide domain to explore and much more could be brought out especially with regard to the revolution in data processing capability that is occurring today in corpus and computer linguistics. Trends indicate that our recent ability using computers to examine, sort, collate and analyze language will revolutionize the subject of natural language research. My research was aimed only at answering a limited number of basic questions regarding the reproducibility of basic data. What this thesis does not do is:

1. Attempt a re-analysis of the original work.
2. Argue the correctness of the original syntactical or semantic claims.

1.8 Similar research
The Thai wordว่า originally meaning “say” can serve as a complementizer for verbs of expression, perception and judgment as shown in the example below.

(1)

\[
\text{Nuan} \text{ คิด} \text{ ว่า} \text{ (zero) ที่} \text{ ภาษา} \text{ รอบ} \text{ วัน} \text{ นี้} \\
\text{Nuan} \text{ think} \text{ that} \text{- should go school day this} \\
\text{“Nuan thinks that (he) should go to school today”}
\]

In answering the question,”Can the zero element in this sentence refer to a subject outside of the sentence or is this ambiguity not allowed?” different answers are given by two different Western trained Thai linguists in their PhD dissertations. Pingkarawat (1989) claims that this is a case of syntactic binding and sentences like #0 are not ambiguous while Hoonchamlong (1991) says that referents outside the sentence are allowed.
This intriguing situation is discussed in detail by Diller and Khanittanan (2002) who attempt to test this disagreement by surveying other Thai native speakers. The point is made that this is not a trivial disagreement since questions of zero reference in languages such as Thai, Lao, and Chinese are of particular interest since generative approaches have sought to formulate a universal taxonomy of zero elements (e.g. pro, trace, PRO, and the like) and to establish cross-linguistic parameters to account for the distribution of zero elements. Diller and Khanittanan attempted to test sentences with zero elements with 45 native speakers in an effort to derive a possible taxonomy of zero elements for Thai. Their conclusion was that “success in imagination” was responsible for many of the positive acceptability judgments and they raise the critical question of whether creating a taxonomy of zero elements using grammaticality judgments is, in fact, a possible goal.
Chapter 2
Theoretical background behind this research

2.1 Introduction
The meanings of key terms in linguistic theory are tied to theoretical assumptions and the word “grammar” is perhaps the most important yet at the same time most theory laden term in linguistics. Linguists themselves have completely different conceptions of what this term means within the greater community of linguists. When a linguist uses the term “grammaticality judgment” the meaning will certainly differ from the normative judgment meant by a teacher, but depending on the theoretical approach taken by a linguist the term can have very different meanings within the field of linguistics.

There has been a struggle of ideas as to what should be the proper subject of a true scientific study of language since Saussure and confusion about terminology then and now seems still to be an issue. Saussure was concerned about, ‘the absolute ineptness of current terminology, the necessity to reform it, and, in order to do that, to show what sort of subject language in general is’ (Ferdinand de Saussure, 1916: 44).

2.1.1 Landscape of Contemporary Linguistics
The confusion and unexamined assumptions held by scholars and researchers working with different approaches have led to a situation where linguists are talking past one another, meaning different things, while using the same terminology. In order to help bring some order to the different theoretical assumptions implicit in the research of linguists that I will test, I have included Table 1 below. It is copied from Stanford University Department of Philosophy website (2012) and summarizes the three broad approaches to the study of language that exist today. Each approach has a history and a philosophy of what the goals, appropriate methodology, and proper subject matter of linguistics should be. Terms such as “syntax”, “grammar”, and “data” have different meanings in these approaches and much confusion results from this fact of using the same word with different meanings. The names for these
three different approaches will be used to describe methodologies and perspectives that Thai researchers have followed in presenting their data and analyses.

To most linguists the mere mention of the names of Bloomfield, Sapir, and Chomsky bring to mind what could be termed philosophies of language and these three figures are respectively the “fathers” of these three approaches-Externalists, Emergentists, and Essentialists.

The point of this chapter is to briefly review the concept of linguistic data in the different approaches to linguistics and in particular, how the concept of a grammaticality judgment works with each approach. The expressed goal of linguists working in all these approaches seems to be to make linguistics ‘an empirical science’. I will briefly review the differing concepts of what counts as data and how grammaticality judgments figure into in each of the three approaches outlined in Table 1.

<table>
<thead>
<tr>
<th>Table 1: Three Approaches to the Study of Language</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary phenomena</strong></td>
</tr>
<tr>
<td>Actual utterances as produced by language users</td>
</tr>
<tr>
<td><strong>Primary subject matter</strong></td>
</tr>
<tr>
<td><strong>Aim</strong></td>
</tr>
<tr>
<td><strong>Linguistic structure</strong></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Values</th>
<th>Abstract productive types</th>
<th>Language users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurate modeling of linguistic form that accords with empirical data and permits prediction concerning unconsidered cases</td>
<td>Cognitive, cultural, historical, and evolutionary explanations of phenomena found in linguistic communication systems</td>
<td>Highly abstract, covering-law explanations for properties of language as inferred from linguistic intuitions</td>
</tr>
<tr>
<td>Children's language</td>
<td>A nascent form of language, very different from adult linguistic competence</td>
<td>A series of stages in an ontogenetic process of developing adult communicative competence</td>
</tr>
<tr>
<td>What is acquired</td>
<td>A grasp of the distributional properties of the constituents of expressions of a language</td>
<td>A mainly conventional and culturally transmitted system for linguistic communication</td>
</tr>
</tbody>
</table>

2.2 Data in the Externalist approach

If you are an Externalist (structuralist) the main goal of a linguistic theory is to develop accurate models of the ‘structural properties of language’. The great success in early linguistic research came from the discovery of the phoneme as a basic linguistic unit (Goldsmith and Laks, 2010: 120). This same methodology was thought to be extendable to research on morphemes and distributional studies of words in sentences, analogous to the study of the distribution of phonemes within words became the research paradigm. One of the basic assumptions of structuralist approaches to linguistics is that “nothing may be called ‘linguistic’ that is not manifest or manifested one way or another between the mouth of the speaker and the ears of the listener (Martinet, 1960: 1). Bloomfield was attacked for being anti-mentalism in his approach to linguistics and Chomsky gained many converts to the Essentialist approach who agreed with him that linguistics should also be concerned with what was happening inside the mind.
Today many linguists who do not accept the Essentialist approach and who do not believe that linguistics is a psychological science, still consider the study of language structure their primary interest. Many are now involved in corpus research since data and data collection has been revolutionized by new technologies using computers. Language can be explored in ways never before possible with technologies that allow data organization, syntactic pattern recognition and data search capabilities that promise a revolution in linguistics. Computers could bring about changes within linguistics similar in scope to that which allowed physical sciences to extend their field of study into microscopic and macroscopic domains and which created completely new fields of study.

The issue of grammaticality is of interest in this approach since large amounts of actual language usage data are now available for syntactic study. A three-million-word collection of recorded and transcribed spontaneous telephone conversations known as the Switchboard corpus was made available in 1992 and two now decades later Google N-Grams Corpus with 155 billion English words is freely available to study the distributions of word patterns in English. For the first time in history what is actually used as language as data can be “data mined” and the Essentialist contention that a native speaker’s intuitive access to “competence” is a better source of data can be tested and answered empirically.

One kind of work that is representative of the Externalist tendency is nicely illustrated by Bresnan et al. (2007) and Bresnan and Ford (2010). They contrasted double-object and recipient-PP constructions in English and found that a number of types of expressions that linguists have often taken to be ungrammatical do in fact turn up in actual use. They conclude that they are these constructions are not ungrammatical or even unacceptable in usage, but merely dis-preferred. This issue of the relationship between acceptability and grammaticality has been held to be a “competence” versus “performance” issue in the Essentialist approach but it continues to be at the heart of an unresolved debate outside of this approach and will be discussed in detail in Sec 4.3-4.5.

In summary, Externalists regard linguistic data is actual language usage, available in a corpus of text, recordings of actual speech or data recorded in linguists’ notebooks as attested usage. What is not even considered as primary empirical data is “introspective data”. It is not the linguist’s position to have an intuition as to the acceptability of a sentence. Acceptability is an observation about a speaker’s usage and is not an intuition in the mind of the linguist.
A common way of acquiring data by asking a bilingual informant to translate sentences from English language into the target language might seem like a very obvious and easy way to gather pertinent data. This is called Reverse Translation Elicitation (Samarin, 1967: 114–115). The problem is that many informants, if not most, cannot do this reliably, and will translate a word for word equivalent in their language. The results can be unnatural and unidiomatic. The worst problem is that the linguist will have no way of really knowing which translations are natural and idiomatic, and which ones are not. The result is a sentence grammar straight from English, and ultimately the grammar written on the basis of these data will be, in Mary Haas’ words (p.c. to de Reuse), “a grammar of translation into English” (Chelliah and de Reuse, 2011: 371).

Linguistics in the Emergentist approach is to be understood in the total context of communication. This approach has also been called functionalist or cognitive, and linguists working with this approach do consider their work part of general psychological theory.

Syntax in this approach is not autonomous since pragmatic and semantic functions can be linked with grammatical form. Thus the same grammatical structures that are of interest to Essentialists can be studied in this approach but their origin and explanation lie not in postulated abstract Universal grammar but in more specific communicative functions of language. In Construction grammar which can be considered as part of this same approach to syntactical analysis, constructions of words can form a sort of semantic template and restrictions and constraints on sentence structure derive from semantic meaning. This approach sees the data of linguistics as “packets of syntactic, semantic, and syntactic information tied together by constraints” (Goldberg, 1995: 16). Data in this approach could be idealized sentences abstracted from a context of use but what cannot be abstracted away in the analysis of these sentences are factors of communicative intent, short term memory and processing limitations and any other factors that are involved in language as communication.

Data in this view is limited to language used in communication and not data derived from a linguist’s self invented example sentences. Created sentences are held to be “bad data points” and not empirical data. The criticism from this perspective is that formal linguistic analysis is armchair theorizing and that it is not empirical data gathering at all. The use of intuition to simply create data is strongly criticized by many linguists working in other approaches. “Take a look at any recent article on
formal syntax and see whether anything other than the theorist's judgments constitutes the data on which the arguments are based”. (Ferreira, 2005: 372).

In summary both the Externalists and Emergentists agree that the subject matter of linguistics should come from language usage. Arguments as to whether sentences are grammatical or not under these two approaches are not major issues because data as usage has an empirical basis. On the other hand the issue of grammaticality judgments is a major issue in the next approach to linguistics that I will discuss.

2.3 Data in the Essentialists approach

The term “Essentialists” in the above table refers to what is often termed formal linguistics most associated with Chomsky and the program of generative linguistics. If a cline can be drawn between a conception of language as vernacular in all its idiomatic richness and complexity at one extreme and logical/propositional and rule governed at the other extreme, the view from the Essentialist perspectives is strongly formal. Grammar takes as its model of natural language from formal languages such as mathematics and propositional logic. From this original metaphor, many other ideas flow naturally. Among the most important assumptions is that the distinction between syntax and semantics is necessary and must be maintained because the purpose of formal languages is to allow users to manipulate abstract symbols in algorithmic ways without regard for their semantic meaning or communicative use.

In Figure 2, a schematic picture of this structural view of language from Jackendoff along with his explanations as to the structure of language will make this point clearly. I will also refer to this schematic view of language as presented by Jackendoff in greater detail in the following chapters and question some of the assumptions made in his approach. The postulated structures that underlie the English clause “the little star is beside the big star” appear below.
Figure 2: Schematic of the Essentialist view of language

Consider next the syntactic structure. This is a tree diagram of the familiar sort. The largest constituent, the sentence (S), divides into a noun phrase (NP) (which serves as the subject) and a verb phrase (VP) (which serves as the predicate); the NP divides into a Determiner, a modifying adjective phrase (AP), and a head noun (N), which carries the features 3rd person count singular. (If this were a French or German sentence, the Det and A
would also have number, and all constituents of NP would have grammatical
gender as well.) The VP divides into a head verb (V) and a prepositional
phrase (PP), the PP divides into a preposition and its NP object, and the NP
divides like the subject NP. Attached to the V is an Inflection which includes
present tense plus the features 3rd person singular, which agree with the
subject. The way I have notated this tree differs from standard convention in
two respects. First, it is customary to put the words of the sentence at the
bottom of the tree. I have omitted the words for reasons to be discussed more
fully in Chapter 5. The basic reason is that things like “the” and “star” are
actually pieces of phonology, not syntax. The only aspects of words that play
a role in syntax are the part of speech (Det, N, etc.) and syntactic features
such as 3rd person singular and present tense. (Jackendoff, 2002: 6).

Jackendoff makes the point that at every level in language description there are
structures that can be studied separately. The fact that words are not actually part
of syntax in his view of language, but actually “pieces of phonology”, is important to
understand in the following chapters regarding grammaticality judgments of
sentences and reasons for acceptability or unacceptability. The assumed strict
compositionality of words as tokens of categories is what is challenged in
grammatical analysis of languages like Thai.

The goal of an Essentialist is not to find and document structures in particular
languages but to discover universal principles providing explanations for deep and
cross-linguistically constant linguistic properties. Linguists are involved in a
psychological project, the latest term is biolinguistics to distinguish this higher goal.
The belief that one could be contributing to the discovery of an innate language
faculty and of the child’s initial cognitive state through detailed formal analysis of
certain structures and languages, is very appealing to many linguists and
psycholinguists (Hawkins, 2004: 4).

The original ambition of generative grammar as it was explained to me when I
enrolled in a graduate class in Transformational Linguistics in 1969 was to explain
in terms of the uniform syntactic theory how the infinite set of grammatical
sentences could be generated by a finite list of algorithms that “must exist in the
brain”. This high ambition was unfortunately limited by the only available
methodology to find these algorithms. Syntacticians had to derive a classification
scheme for all linguistic phenomena and posit an enormous set of very complicated
syntactic rules which, and this is the most difficult problem, are all mutually
interdependent. The only source of evidence to find and in a sense “reverse engineer” our inner syntactic structures was the intuitive acceptability of some sentences and unacceptability of others. Chomsky claimed that linguistic competence can tell us what is grammatical and what is not; and that we have a sound access to this competence: we can make reliable grammaticality judgments (intuitively), even if our performance does not always match them. The difference – says Chomsky – is due to ‘such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, errors in applying knowledge of the language in actual performance.’ (Chomsky, 1965: 3–4)

A native speaker has knowledge of language that underlies both speaking and understanding and this linguistic competence is what grammars are supposed to be—characterizations of this competence. What Mieszko Talasiewicz (2010) calls “the haughtiness of linguistics” comes from the grandiosity of this claim. The status of acceptability judgments is the core of much of the current debate and the claim that grammaticality judgments can be the methodological route of discovery to this inner competence is what is at issue in this thesis.

2.4 Summary

In summary, there is yet no uniform methodology that could be termed “the practice of linguistics”. There are today many competing “grammars” or frameworks of linguistic description that may provide solutions to some problems while being inadequate for others. Common ground on which one could compare and evaluate these conceptions, theories and paradigms does not exist (Talasiewicz, 2010: 174).

But if syntax and grammar lack a common consensus view, there seems to be an emerging consensus in all approaches to linguistics that primary linguistic data in the form of language descriptions should be preserved. Lehmann writes:

This moment in the history of the discipline happily coincides with new and urgent demands being made on it from outside, viz. from the speech communities of languages threatened by extinction. As if awakening from sleep in a scientific greenhouse, the discipline has suddenly become aware of the fact that its capacity is urgently needed for the documentation and description of most of the languages of the world, both for the sake of their speech communities and their interest in their cultural tradition and for the sake of the very database of the discipline itself. Language documentation has become a slogan in today’s linguistics. As is usual in such cases, some
members of the scientific community who are more flexible in publicizing the work they had always been doing than in adapting to urgent demands from outside have adopted the new term as a more effective label under which to sell traditional linguistic description. Most of us, however, have understood that the new situation demands a rethinking of our methodological bases. In endangered languages, data constitute a value for their own sake because they are irreplaceable. Consequently, we need to develop methodological standards for their scientific and practical treatment so that future generations can make the best possible use of them (Lehman, 2004: 207).

Lehman makes the valid point that today linguists are urgently needed to help save the database of the discipline. The field notes and word lists of the few linguists working a century ago with languages that have since disappeared are now of great value. The rise in influence of the Essentialist approach in linguistics research has changed the focus of research in mainstream linguistics. But it is possible that linguists working in these three different approaches can unite at least in the safeguarding of data from endangered languages—the primary evidential basis of linguistics.
Chapter 3
Data Collection, Methodology and Results

3.1 Overview
In this chapter the method of collecting data is reviewed and then the particular sentences for which grammaticality judgments were sought are discussed. An overview of the types of sentences chosen is provided along with a rationale as to why they were chosen. Then the specific claims made by other researchers for each sentence will be discussed. Some of the claims are more complex and interwoven with other claims that involve some explication of some theoretical background to be understood. Others are relatively straightforward and involve a meaning or usage difference between two contrasting constructions. Some of the sentences are claimed to be unacceptable while most are claimed to be acceptable. I will give a percentage test response and provide specific details of each test since they were done at different times.

3.2 Methodology
Both questionnaire methods and interviews were used to gather the data for this thesis. A graduate class of English majors at Chiang Mai University volunteered to answer a questionnaire which consisted of 18 sentences in Thai. They were asked to rate the sentences for acceptability as OK, confusing but OK, unacceptable, or uninterpretable. These students were also asked interpret the sentences in English and make comments on the sentences regarding the structure or vocabulary of each sentence.

In separate elicitation sessions I used these 18 sentences along with two more and used interview techniques with 12 monolingual Thais which I tape recorded and transcribed. A group of 5 office workers volunteered to rate the sentences for acceptability and this occurred in a group setting. There was much group discussion in each case as some of the sentences were difficult to process due to various possible semantic construals. The other 7 monolingual Thai speakers were interviewed individually and we went through the interviews question by question in some cases taking up to 30 minutes to finish the list.
After this data was collated and analyzed, I interviewed 5 more native speakers and had them explain how the sentences in #2 - #5 could be meaningfully construed. These informants took pencil and paper and discussed various scenarios of the positions of the speakers, whether the agent “crossed” the bridge, and what information was assumed and what seemed clearly stated.

Table 2: Informant and Survey Summary

<table>
<thead>
<tr>
<th>Informant Descriptions</th>
<th>No.</th>
<th>Written Survey</th>
<th>Interview</th>
<th>Group Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiang Mai University Students</td>
<td>18</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Office Workers</td>
<td>5</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Office Workers</td>
<td>7</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

3.3 Sentences Surveyed

In (2)-(21) are each of the sentences surveyed. Each sentence is discussed in detail in the sections below.

(2) Survey Sentence #1

Maalii wîj troj jîon klàp khâw paj
Malee run go.straight reverse return enter go
“Malee ran straight back in, away from the speaker”

(3) Survey Sentence #2

Maalii wîj jîon klàp troj khâam saphaan paj
Malee run reverse return go.straight cross bridge go
“Mali ran back straight, crossing the bridge, away from the speaker”

(4) Sentence #3

Maalii jîon khâam saphaan troj ?ôk paj
Malee reverse cross bridge go.straight exit go
“Mali ran back straight, crossing the bridge, away from the speaker”
(5) Sentence #4

Maalii trọŋ ṭɔɔk jɔɔn khām saphaan paj
Malee go.straight exit reverse cross bridge go

“Mali ran back straight, crossing the bridge, away from the speaker”

(6) Sentence #5

Maalii ṭɔɔk trọŋ khām saphaan jɔɔn paj
Malee exit go.straight cross bridge Reverse go

“Mali ran back straight, crossing the bridge, away from the speaker”

(7) Sentence #6

naj sāam naathii Nikorn kin khâaw le ṭīm
in three minute Nikorn eat rice and full

“After three minutes, Nikorn ate rice and he was full.”

(8) Sentence #7

naj sāam naathii Nikorn kin khâaw ṭīm
In three minute Nikorn eat rice full

“Nikorn ate rice until he was full in three minutes.”

(9) Sentence #8

Piti duu nāŋ cōb paj
Pitit watch Movie End PFTV(go)

“Piti finished watching a movie.”

(10) Sentence #9

*Wijada kamlaŋ chīa rīaj phīl
Wijada PROG believe issue ghost

“Wijada is believing in ghosts.”

(11) Sentence #10

dōn paj thīŋ rāan
walk go arrive shop

“(He) walked away and reached the shop.”

(12) Sentence #11

dōn paj yaŋ rāan
walk go (to) shop
“(He) walked to the shop.”

(13) Sentence #12
khin  paŋ  sùu  chán bon
Ascend  go  arrive.stay  upstairs
“He went up and got to the upstairs and stayed.”

(14) Sentence #13
khin  paŋ  yaj  chán bon
ascend  go  (to)  upstairs
“He went up to the upstairs.”

(15) Sentence #14
* dəən
Walk  reach  home
(no free translation deemed unacceptable)

(16) Sentence #15
? thenit  kʰi  kila  thiি  chan  chɔːp  thiি sùt
tennis  COP  sport  REL  PRO  like  most
“Tennis is the sport that I like best.”

(17) Sentence #16
* kʰaan  rian  kʰi  sɨj  thiি  sãamkhaan  thiি sùt  sãmrɔp  nák  rian
NOMLZ  study  COP  thing  REL  important  most  for  student
“To study is the most important thing for students.”

(18) Sentence #17
māj  yâak  paŋ  tham ḳaañ
NEG  want  go  work
“do not want to go to work.”

(19) Sentence #18
yâak  māj  paŋ  tham ḳaañ
want  NEG  go  work
“do not want to go to work.”

(20) Sentence #19
ø  tʃŋ  rũp  klãp  paŋ  rĩak  hãj  ø  maa  bɔk  ø
(you) must hurry return go call cause (him) come tell (me)
“(you) must hurry back (and) summon (him) to come and tell (me).”

(21) Sentence #20
?
aw  ลำย  pai  ทาม  kin  duu  sì
take  try  go  make  eat  see  PART
“go ahead and take them and try cooking them to eat”

3.4 Summary Chart of Results
The following table summarizes the results of the Thai serial verb grammaticality survey. This brief summary is intended to guide first the detailed discussion of each sentence and second the more elaborate table of results in Section 3.6.
Table 3: Summary of Survey Results

<table>
<thead>
<tr>
<th>Example Number</th>
<th>Number of Grammar Responses</th>
<th>Responses that agree with researcher in finding example 'un')acceptable'</th>
<th>Number of Semantic Responses (English translations)</th>
<th>Responses that match the meaning proposed in the research paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28</td>
<td>28</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>9</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>15</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>9</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>10</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>24</td>
<td>10</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>12</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>11</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>12</td>
<td>28</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>28</td>
<td>7</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>28</td>
<td>25</td>
<td>25</td>
<td>22</td>
</tr>
<tr>
<td>16</td>
<td>28</td>
<td>0</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>17</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>18</td>
<td>28</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>20</td>
<td>12</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

3.5 Discussion of individual sentences
The phenomena of verb serialization or multi-verb constructions are common in many languages. Diller calls Thai “a verb-loving language” where serial verbs can be built up into larger structures of serial verbs which he terms “amalgam constructions” (Diller, 2006). Thai has many of these with different “recipes” for grammatical assembly. Diller notes that constraints to build up amalgam constructions are complex and involve interplay between the lexical semantics of the individual words as well as word pairs along with pragmatic factors. These are important points to remember in the discussion that will follow.
3.5.1 Sentence #1: Thepkanjana

The earliest and probably the most well known amalgam recipe for Thai motion event descriptions is from Thepkanjana’s thesis (1986). It would be difficult to understand the forthcoming claims made by Muansuwan or the sentences that she presents as evidence without more background explication to put them into the context of an ongoing debate and effort to understand serial motion event verbs in Thai.

The sentence below is to my knowledge the first appearance of Malee in motion-event sentences in the Thai serial verb literature (Thepkanjana, 1986: 136).

(22) Survey Sentence #1

Maalii wîŋ troŋ jôn klâp khâw paj
Malee run go.straight reverse return enter go
“Malee ran straight back in, away from the speaker”

Thepkanjana claims that when more than one Thai motion verb is used to represent a single motion event, the linear order should be as in (23):

(23)
(a) initial verb: mode of locomotion or travel
(b) characteristic shape of path
(c) change of direction
(d) motion relative to reference object
(e) orientation of path based on moving subject
(f) orientation based on perspective of speech-act participants

A single amalgam construction would rarely include all six components. But for the verbs present, argues Thepkanjana, the relative ordering in (23) applies. She calls a verb of the slot A ‘initial verb’ and verbs of the slots B to E ‘serial verbs’ which modify the initial verb. However, it is not obligatory that a verb of every slot be present, and therefore a verb of any slot has a chance to become the initial verb.

Results of acceptability test:
28 out of 28 respondents found the sentences acceptable.

Claim of acceptability by the researcher is strongly supported by the survey results.
Semantic Meaning Results:
Meanings recovered by those questioned was the same as Thepkanjana in 10 out of 15 instances.

Some examples of alternate meanings provided by the survey are in (24).

(24)
   a) Malee ran straight turns around and goes in
   b) Malee runs turns around and goes inside

3.5.2 Muansuwan and serial verbs
Muansuwan's thesis on DSVC's is, in a sense, a hypothesized better re-categorization of semantic and syntactic structure for DSVC constructions.

She models Thai Directional SVCs by “proposing a few lexical specifications, two lexical rules and three syntactic rules or schemata. I start by presenting the relevant portion of the Thai lexical type hierarchy I assume in Figure 3 (Muansuwan, 2002: 60, Figure 2.1).

\[
\begin{array}{c}
\text{head} \\
\ldots \\
\text{Part of speech (POS)} \\
\ldots \\
\text{preposition} \quad \text{noun} \quad \text{verb} \\
\text{main verb} \quad \text{serial verb}
\end{array}
\]

Figure 3: Muansuwan part of speech diagram

According to Muansuwan there are two types of verbs in Thai: main verbs and serial verbs. ‘Main verb’ can refer to the union of verbs which do not occur in SVCs. Simply understood, if a verb is part of a SVC it is not the main verb. For example, the main verb in a SVC is the verb or verbs that must appear in the initial position of a Directional SVC sequence. In other words, “main verbs are verbs which occur where heads are expected to occur in a head-initial language”.

26
Muansuwan further argues that Thai Directional serial verb constructs include two kinds of syntactic structures: a recursive co-headed structure and a complementation structure. She believes that she can show that Directional SVCs in Thai are distinct from those found in other languages in terms of:

(i) the number of verbs that can occur in a SVC,
(ii) their constituent structure.

Since, in Thai, adverbs occur VP final she assumes that an adverb marks the boundary of a VP. In other words, an adverb indicates that what comes before it in the sentence is a VP. The placement of an adverb can, thus, be used to test where the VP break. Using this as a tool to check constituency she claims that within DSVC the ordering of all verbs is free except for initial and final verb.

She claims that previous analyses of Directional SVCs (Thepkanjana, 1986) cannot account for the syntactic behavior of Thai Directional SVCs, particularly the presence of ordering constraints among non-sister constituents, and that previous analyses of Thai Directional SVCs are also incorrect about their constituent structure and the nature of the ordering constraints.

In flat tree structure, like Thepkanjana assumes (shown in (25)) (Muansuwan, 2002: 43), a conventionally created verb phrase (VP1) binds all motion verb phrases of the six classes (VP2 through VP7) together. This verb phrase structure is of iteration, but not of recursion.

(25)

```
VP1
  VP2  VP3  VP4  VP5  VP6  VP7  VP*
    wiŋ   troŋ  jɔɔn  klap  khaw  paj
```

**Figure 4: Thepkanjana SVC Flat SVC Structure**

She concludes that the structure of the Thai Directional SVC is as shown below in (26)(Muansuwan, 2002: 53).
Muansuwan’s very first example of a Directional SVC in (27) is closely modeled on Thepkanjana’s #1 above in (22).

Sentence #2 is claimed to consist of six verbs, which share a common subject: Malii. Semantically, Malii is the figure of the complex motion event encoded by the sequence of verbs. The first verb in denotes a manner of motion and the non-initial verbs encode a directed motion, which includes information about the path, direction, and/or deictic center of the motion event. Each verb in a Directional SVC describes the same, single event from a different perspective.

(27) Survey Sentence #2

Maalii wíŋ j̪ɔŋ kláp troŋ khâam saphaan paj
Malee run reverse return go.straight cross bridge go

“Mali ran back straight, crossing the bridge, away from the speaker”

There are 3 variations on sentence #2 that were surveyed in addition to sentence #2. Each is discussed below.

3.5.3 Sentences #2-#5: Muansuwan

The 3 variations on sentence #2 are listed below in examples (4-6).
As shown in examples (27)-(30), the free translations are claimed by Muansuwan to be the same for test sentences #2 through #5.

The data from sentences #2-5 is used to claim that reordering is possible and preserves meaning contrary to Thepkanjana’s claim of ordering restrictions. Crucially for interpreting the survey data, all sentences are claimed by Muansuwan to be acceptable and all have the same meaning. At issue is the claimed paradigmatic amalgam recipe for Thai motion event descriptions offered by Thepkanjana where sequence of verbs in slots has a semantic sequential ordering. Muansuwan’s specific claim is that Thepkanjana is wrong and that other than the first and last slots the other slots are free ordered.

3.5.3.1 Sentence #2: Muansuwan

(31) Survey Sentence #2

Maa-lii  wíŋ jɔn  klàp t론  ʔɔk  paj
Malee  run reverse return go.straight cross bridge go

“Mali ran back straight, crossing the bridge, away from the speaker”

Results of acceptability test:

For sentence #2, 9 out of 24 native speakers agreed with Muansuwan and found the sentence acceptable.

Semantic Meaning Results:
The same meaning as the researcher was found by 2 out of 9 respondents.

Some examples of alternate meanings provided by the survey are in (32)

(32)  
   a) Malee ran back and crossed the bridge  
   b) Malee turned her back and ran straight across the bridge  
   c) Malee turned back at the bridge  
   d) Ran back at the opposite bridge  

Notably this sentence was considered bad enough that most informants would not attempt a translation. This means one could interpret these semantic results as “2 out of 24” with most respondents not even trying for an interpretation.  

**Comments collected:**  
   a) This is a bad sentence.  
   b) This is sort of two things that could not happen together.  
   c) I don’t understand it’s like Malee is at one end running to the other.  
   d) This is like two sentences here.  

### 3.5.3.2 Sentence #3: Muansuwan  

(33) Sentence #3  

```
Maalii j5n khâam saphaen troŋ ʔsoŋ paj
Malee reverse cross bridge go.straight exit go
```

“Mali ran back straight, crossing the bridge, away from the speaker”  

**Results of acceptability test:**  
For Sentence #3 15 out of 24 native speakers agreed with the linguistic literature and found the sentence acceptable.

**Semantic Meaning Results:**  
The same meaning as the researcher was found by 3 out of 11 respondents.

The meanings in (34) were provided by the survey respondents when they tried to interpret the sentence.

(34)  
   a) Malee walked down the bridge and returned to cross the bridge.  
   b) Malee makes a U-turn crossing the bridge.  
   c) Malee turned back and crossed the bridge.
d) Malee turned then crossed the bridge and then goes straight forward.
e) Malee turns across the bridge and goes straight away.
f) Malee went back across the bridge.
g) Malee U-turn to cross a bridge and straight forward.

Comments collected:
  a) Very confusing sentence
  b) Could mean several meanings

3.5.3.3 Sentence #4: Muansuwan

(35) Sentence #4

\[ Maalii \quad tro\gamma \quad ?\ddot{\omega}k \quad j\ddot{\alpha}n \quad kh\ddot{a}am \quad sap\ddot{a}an \quad paj \]

Malee go.straight exit reverse cross bridge go

“Mali ran back straight, crossing the bridge, away from the speaker”

Results of acceptability test:
For sentence #4, 18 out of 24 native speakers disagreed with Muansuwan and found the sentence unacceptable.

Semantic Meaning Results:
Out of the 6 who found the sentence acceptable, only 2 provided the same meaning as in Muansuwan. The meanings respondents provided are in (36).

(36)

  a) Meanings recovered by testing: 2 out of 6
  b) Malee goes straight then U-turn
  c) Malee walked straight to cross the bridge.
  d) Malee went back across the bridge.
  e) Malee turn back to cross the bridge.
  f) Malee makes a U-turn to cross the bridge and then straight forward.

Comments collected:
  a) Thinking about this gives me a headache.
  b) Bad sentence.

3.5.3.4 Sentence #5: Muansuwan

(37) Sentence #5

\[ Maalii \quad ?\ddot{\omega}k \quad tro\gamma \quad kh\ddot{a}am \quad sap\ddot{a}an \quad j\ddot{\alpha}n \quad paj \]
Malee exit go.straight cross bridge reverse go
“Mali ran back straight, crossing the bridge, away from the speaker”

Results of acceptability test:
For sentence #5, 19 out of 24 native speakers disagreed with Muansuwan and found the sentence unacceptable.

Semantic Meaning Results:
Of 9 respondents who hazarded a meaning, only 1 respondent provided the same meaning as in Muansuwan. Other meanings are given in (38).

(38)
  a) Malee gets out then turns back at the opposite bridge.
  b) Malee walked to the other side of the bridge and then turned back.
  c) Malee crossed the bridge and then turned after that.
  d) Malee goes out opposite the bridge.

Comments collected:
  a) Confusing
  b) Could mean several things

3.5.3.5 Summary of claims
The full implication of these results is in the conclusion but already some things are clear. While Thepkanjana’s example was closely aligned with the grammaticality and semantic judgments of the survey respondents, the same is not true of the DSVC examples from Muansuwan. In fact there are three contravening claims. One is that the majority of respondents found all examples unacceptable. The second is that only a tiny portion of respondents agreed with the meaning claims associated with those sentences. The third, less obvious point, is that the free ordering claims about serial verbs as exemplified in sentences #3-#5 is not supported.

These first four sentences came from Muansuan’s discussion of DSVC. The next two sentence come from her discussion of and Adjoining Constructions (AJC), which are a subset of Sequential Constructions (Muansuwan,2002: 209-214).
3.5.4 Sentences #6-#7: Muansuwan

Muansuwan begins with some terminology issues and makes a series of distinctions between construction types.

Her terms are: Sequential Constructions, Resultative Constructions (as occur in English), Resultative Serial Verb Constructions (Thepkanjana, 1986), and Adjoining Constructions (which are a subset of Sequential Constructions).

She claims that Thai has two subsets of sequential constructions, ordinary and AJC and in both constructions the first verb phrase encodes an event that precedes the event coded by the resulting verb but that there is a key difference between the two. The constructions differ in that “whereas there can be no interval of time between the temporal traces of the events described by the verbs in Adjoining Constructions there can be an interval of time between the events described by the verbs in Ordinary Sequential Constructions”. She shows this with two examples shown below:

(39) Sentence #6

\[
\text{naj sāam naathii Nikorn kin khâaw lé ʔím}
\]

in three minute Nikorn eat rice and full

“After three minutes, Nikorn ate rice and he was full.”

(40) Sentence #7

\[
\text{naj sāam naathii Nikorn kin khâaw ʔím}
\]

in three minute Nikorn eat rice full

“Nikorn ate rice until he was full in three minutes.”

She claims that there is a possible contrast of meaning where #7 is considered as one event and not as two events, whereas #6 could be interpreted as a reduced conjoined clause with ambiguity in interpretation. She claims that adverbial phrases can combine with an Adjoining Construction with no ambiguity since the adverbial phrase means “the eventuality X holds after three minutes”. “X” is the event meaning that this adverbial phrase has scope over. Thus with the ADJ construction there is no ambiguity in distinguishing between starting to eat in three minutes versus eating and getting full in three minutes. Both sentences are claimed as acceptable and further that a meaning difference exists in there being no ambiguity in interpretation between the two sentences.
3.5.4.1 Sentence #6: Muansuwan

(41) Sentence #6  
\[\text{naj sāam naathii Nikorn kin khāaw lē ?ǐm}\]  
in three minute Nikorn eat rice and full  
“After three minutes, Nikorn ate rice and he was full.”

Results of acceptability test:  
For sentence #6, 12 out of 28 native speakers agreed with Muansuwan and found the sentence acceptable.

Semantic Meaning Results:  
Of 12 respondents who agreed with the acceptability of the sentence, 10 provided the same meaning as in Muansuwan.

Comments collected:  
None

Based on the face that the majority of respondents disagreed with the acceptability of sentence #6, Muansuwan’s claim of acceptability is not supported.

3.5.4.2 Sentence #7: Muansuwan

(42) Sentence #7  
\[\text{naj sāam naathii Nikorn kin khāaw ?ǐm}\]  
In three minute Nikorn eat rice full  
“Nikorn ate rice until he was full in three minutes.”

Results of acceptability test:  
For sentence #7, 28 out of 28 native speakers agreed with Muansuwan and found the sentence acceptable.

Semantic Meaning Results:  
Of 28 respondents who agreed with the acceptability of the sentence, 16 provided the same meaning as in Muansuwan. No respondent found any contrast in meaning between sentence #6 and sentence #7

Comments collected:  
1) Strange to say it this way.
b) Putting three minutes last is better.
c) I never hear “and” between these two words.
d) I understand but saying it this way is strange.
e) This means he eats and is full in three minutes so maybe in two minutes to make him feel like that.

### 3.5.4.3 Summary of claims

While all speakers found sentence #7 acceptable, most did not find #6 acceptable. In those cases where a speaker could find a contrast in between the meaning of #6 and #7 they did not. This means that while the grammatical acceptability claim holds for sentence #7 (though not for #6), the claim about semantic contrast does is not supported.

### 3.5.5 Sentences #8-#9: Muansuwan

The next two sentences come from Muansuwan’s study of what she terms “aspect morphemes”. She calls these constructions aspectual constructions (AC) and proceeds to list and attempt to describe their meanings, possible positions, and their combinatorial possibilities. In the slot immediately following the “base verb phrase” she identifies three verbs among the 12 verbs in her table (p. 99) which she calls superlexical morphemes. This term comes from Smith (1997). These verbs have predicative meaning whether in base verb slot or the morpheme slot. Super-lexical morphemes can co-occur with and precede several aspect morphemes which do not have predicative meaning. The reverse order is not allowed. Sentence #8 is predicted to be acceptable. In sentence #9 the claim is that the stative verb chīa “believe” cannot co-occur with the progressive marker kamlaj.

(43) Sentence #8

\[
Piti \text{ duu nāŋ cōb paj}
Pitit \text{ watch movie end PFTV(go)}
\]

“Piti finished watching a movie.”

(44) Sentence #9

\[
*Wijada \text{ kamlaj chīa rāŋ phǐ}
Wijada \text{ PROG believe issue ghost}
\]
“Wijada is believing in ghosts.”

3.5.5.1 Sentence #8: Muansuwan

(45) Sentence #8

Piti duu nãŋ còb paj
Piti watch movie end PFTV(go)
“Piti finished watching a movie.”

Results of acceptability test:
For sentence #8, 10 out of 24 respondents found this sentence acceptable. This means that the claim of acceptability is not supported since a majority disagreed.

Semantic Meaning Results:
Of 15 respondents who attempted to give a meaning, only 9 provided the same meaning as in Muansuwan. Other meanings are given in (46).

(46)

a) Understandable as finished watching movie.
b) If you add lezw at the end it means finished watching movie.
c) He just finished watching the movie.

Comments collected:
a) Remove paj and Ok
b) Change paj to lezw and OK
c) No one would say it this way
d) Còp paj means to end up with negative feeling
e) People say còp paj when they break up
f) You have to add another word
g) Nobody would use paj

3.5.5.2 Sentence #9: Muansuwan

(47) Sentence #9

*Wijada kamlan chila riŋ phi
Wijada PROG believe issue ghost
“Wijada is believing in ghosts.”

Results of acceptability test:
For sentence #9, 12 out of 28 respondents found this sentence acceptable. In this case Muansuwan claims this sentence is unacceptable. The claim of unacceptability is only partially supported since a majority disagreed.

**Semantic Meaning Results:**
In this case, 12 respondents who attempted to give a meaning, and 8 provided the same meaning as in Muansuwan. Other meanings are given in (48).

(48)

a) Wijada is starting to believe in ghosts.
b) Now Wijada is believing in ghost.
c) Wijada is believing ghost story.
d) Wijada believes that there are ghosts.

**Comments collected:**

a) Leave out kamlaŋ it is OK
b) chūa is rarely used with kamlaŋ since it is a fact not an action

Takahashi has authored several dozen papers taking a cognitive linguistic approach to linguistic analysis which combines aspects of Externalist and Emergentist approaches. The next nine sentences come from four different papers she has written concerning Thai prepositions, copula choices, and verb negation.

### 3.5.6 Sentences #10-#13 Takahashi – allative and arrivals (thīŋ, yaŋ and sūu)

Takahashi claims that the Thai word yang derived from a verb meaning “maintain”, “sustain”, “exist” or “bring about” and is the only allative preposition in Thai. Although there are other Thai verbs that can serve as prepositions, yang has lost all verbal or in her term “non-contentful” meaning and is now the only true allative grammatical marker in Thai. Takahashi then contrasts thīŋ with sūu, “a representative of arrival verbs” that she admits has a complex meaning and includes the “sense of sharing and staying”. Takahashi has searched corpus data and claims that sūu which can serve as an allative preposition as well as thīŋ still maintains its semantic verb-like meaning of not just “arrive” but “arrive and stay”. Takahashi claims that all sentences are acceptable and the contrast between yaŋ and thīŋ and
sùu is a contrast between a two verb construction and a verb and allative preposition construction. She claims the yag constructions are simplex events while the two verb constructions (thĩŋ and sùu) show an arrival event following a cause event.

3.5.6.1 Sentence #10: Takahashi
(49) Sentence #10
    
    "(He) walked away and reached the shop."

Results of acceptability test:
For sentence #10, 28 out of 28 respondents found this sentence acceptable. This supports Takahashi claims this sentence is acceptable.

Semantic Meaning Results:
In this case, 28 respondents attempted to give a meaning, and they all agreed with the meaning provided by Takahashi.

Comments collected:
    a) OK Walking on the way to the
    b) Meaning arrived at shop already.
    c) Already there.
    d) You are walking to the shop and you are now at the shop.
    e) The speaker or the person is at the shop.

3.5.6.2 Sentence #11: Takahashi
(50) Sentence #11
    
    "(He) walked to the shop."

Results of acceptability test:
For sentence #11, 28 out of 28 respondents found this sentence acceptable. This supports Takahashi claims this sentence is acceptable.

Semantic Meaning Results:
In this case, 28 respondents attempted to give a meaning, and they all agreed with the meaning provided by Takahashi. Thus the claim about a contrast in meaning is supported. Some of the meanings provided are in (51):

(51)

a) Walking to get to the shop.
b) Somebody is walking towards a shop but not arrived yet.
c) You are still walking to the shop.
d) You are not there yet maybe in between.

Comments collected:

a) This is more formal.

3.5.6.3 Sentence #12: Takahashi

(52) Sentence #12

\[
\text{khin \ baj \ sùu \ chán bon}
\]

Ascend go arrive stay upstairs

“He went up and got to the upstairs and stayed.”

Results of acceptability test:

For sentence #12, 2 out of 28 respondents found this sentence acceptable. This is opposed to Takahashi’s claim this sentence is acceptable.

Semantic Meaning Results:

In this case, only 2 respondents attempted to give a meaning, and only 1 agreed with the meaning provided by Takahashi.

Comments collected:

a) inappropriate preposition
b) formal, strange
c) not used
d) should replace with thì
e) correct with heaven but not this way
f) only used with heaven or abstract noun

3.5.6.4 Sentence #13: Takahashi

(53) Sentence #13

\[
\text{khin \ paj \ yaj \ chán bon}
\]
ascend go (to) upstairs
“He went up to the upstairs.”

Results of acceptability test:
For sentence #13, 7 out of 28 respondents found this sentence acceptable. This is opposed to Takahashi’s claim this sentence is acceptable.

Semantic Meaning Results:
In this case, only 7 respondents attempted to give a meaning, and only 3 agreed with the meaning provided by Takahashi. Thus, the claim of contrasting meaning is not supported.

Comments collected:
a) not used this way
b) omit yaŋ
c) sounds formal and strange
d) unacceptable preposition
e) can understand but strange

3.5.7 Sentence #14: Takahashi
Takahashi claims that dəən expresses the simplest and most general sense of locomotion of animate beings. It has simply a generalized “walk” meaning and that it needs another verb as a path marker or end point marker with it to convey a complete event description. This explains why dəən cǎak bāan is grammatical since it marks a point for the action to start from. In the case of dəən a path marker would be paj or maa. Her claim is that sentence below is ungrammatical because thíŋ needs a starting point marker or a path marker like paj or maa.

(54) Sentence #14
* dəən thíŋ bāan
walk reach home
(no free translation deemed unacceptable)

Results of acceptability test:
For sentence #14, 28 out of 28 respondents found this sentence acceptable. The result directly contravenes and opposes Takahashi’s claim that this sentence is unacceptable.
Comments collected:

a) Emphasis is on how far he walked
b) Usually has a subject
c) Should have lew at the end

3.5.8 Sentences #15-#16 Takahashi – copulas khi and pen

Takahashi attempts in this analysis to incorporate “cognitive science insight into linguistic analyses”. She references studies which attempt to draw distinctions between types of cognition, either immediate and reflexive or deliberative and thoughtful.

She believes that the difference between khi and pen can be explained as follows: If ‘fast/sensation-like’ processing is taken, then khi is used, while pen is chosen if ‘slow/thought-like’ processing is employed”. She believes that evidence to support khi’s association with ‘fast/sensation-like’ processing is the fact that in the NP1(referential)-copula-NP2 structure, the more ‘referential’ NP1 is, the more acceptable khi is as a copula. Takahashi claims that khi is an identifier and requires a concrete NP1 and that when NP1 refers to a concrete specific object with a ‘face’ khi is more acceptable than pen. In other words, khi is associated with a ‘fast/sensation-like’ processing so when the object referred to is abstract as “studying” in sentence #16, khi would be unacceptable usage. The sport of tennis which is less abstract would be less unacceptable but still “iffy” as in sentence #15.

(55) Sentence #15
?

thenit khi kila thii chan chɔɔp thii sùt

tennis COP sport REL PRO like most

“Tennis is the sport that I like best.”

(56) Sentence #16
* kaan rian khi siŋ thii s˚amkhaan thii sùt s˚amràp n˚ak rian

NOMLZ study COP thing REL important most for student

“To study is the most important thing for students.”

3.5.8.1 Sentence #15: Takahashi

(57) Sentence #15
?

thenit khi kila thii chan chɔɔp thii sùt

tennis COP sport REL PRO like most

“Tennis is the sport that I like best.”
Results of acceptability test:
For sentence #15, 25 out of 28 respondents found this sentence acceptable. This is opposed to Takahashi’s claim this sentence is unacceptable.

Semantic Meaning Results:
25 respondents attempted to give a meaning, and 22 agreed with the meaning provided by Takahashi.

Comments collected:
   a) Sounds better if start with “kila thii chan choop”

3.5.8.2 Sentence #16: Takahashi
(58) Sentence #16
   * kaan rian khi siŋ thi thii sāamkhaan thi süt səmrap nāk rian
   NOMLZ study COP thing REL important most for student
   “To study is the most important thing for students.”

Results of acceptability test:
For sentence #15, 28 out of 28 respondents found this sentence acceptable. This is opposed to Takahashi’s claim this sentence is unacceptable.

Semantic Meaning Results:
28 respondents attempted to give a meaning, and 28 agreed with the meaning provided by Takahashi.

It is fairly clear that these examples do not support the claims of Takahashi.

3.5.9 Sentences 17-18 Takahashi – negation before and after verbs
Takahashi claims that both patterns are sometimes acceptable. She claims that a class of complement taking verbs follows this pattern that the negation of first verb negates the feeling towards the action of the second verb (NEG VP1 VP2) (18) while (VP1 NEG VP2) (19) which negates the second verb means the feeling of the first verb is not negated but the second action verb carries the feeling of negation. In other words the meaning of #18 is a sarcastic comment that the speaker “has a desire to not go to work” and #17 that the speaker has “no desire to go to work”.

(59) Sentence #17
mâj yàak paj tham ɲaan
NEG want go work
“do not want to go to work.”

(60) Sentence #18
yàak mâj paj tham ɲaan
want NEG go work
“do not want to go to work.”

3.5.9.1 Sentence #17: Takahashi

(61) Sentence #17
mâj yàak paj tham ɲaan
NEG want go work
“do not want to go to work.”

Results of acceptability test:
For sentence #17, 28 out of 28 respondents found this sentence acceptable. This supports Takahashi’s claim that this sentence is acceptable.

Semantic Meaning Results:
28 respondents attempted to give a meaning, and all 28 agreed with the meaning provided by Takahashi.

3.5.9.2 Sentence #18: Takahashi

(62) Sentence #18
yàak mâj paj tham ɲaan
want NEG go work
“do not want to go to work (sarcastic).”

Results of acceptability test:
For sentence #18, only 6 out of 28 respondents found this sentence acceptable. This is opposed to Takahashi’s claim that this sentence is acceptable.

Semantic Meaning Results:
Only 6 respondents attempted to give a meaning, and of those only 1 agreed with the meaning provided by Takahashi. Since only 1 out of 28 speakers found the meaning used for the contrast claimed by Takahashi, the claim cannot be supported.
Comments collected:

a) In case someone really loves his job but he wishes he does not want to go.
b) Nobody will ever say this
c) Sounds bad and strange

3.5.10 Sentence #19: Smythe Reference Grammar

These last two sentences come from reference grammar books which offer them as examples of colloquial speech where interpretation would require context clues. There were no context clues provided in the survey.

This sentence appears in Smythe’s Thai reference grammar where it is used as a typical example of serial verb usage. What appear to be eight verbs are really only four separate actions since the verbs are grouped semantically into four “chunks”. The problem in interpretation is retrieving pronouns from context. I wanted to test this common register of Thai speech for acceptability.

(63) Sentence #19

ø tɔŋ rûp klàp pəj riak hâj Ø maa bɔk ø
(you) must hurry return go call cause (him) come tell (me)
“(you) must hurry back (and) summon (him) to come and tell (me).”

Results of acceptability test:

For sentence #19, 8 out of 13 respondents found this sentence acceptable.

Semantic Meaning Results:

Of the 8 respondents attempting to give a meaning, they all gave a similar meaning to that provided by Smythe. Some of the meanings provided are in (64).

(64)

a) I have to rush back and call the person to come back and tell us.
b) You have to call that person to come back to tell us more detail.
c) We must call her to tell us.
d) Have to get someone to tell something real quick.
e) We need to go back to call him to tell us.
f) Have to hurriedly go back to tell her to tell.

Comments collected:
3.5.11 Sentence #20: Enfield (adaption)
This sentence (65) is a Thai translation of a sentence in Lao offered by Enfield (2007) as an example of a serial verb construction in L(ow) register speech. Enfield’s comment was that this ability to use serial verbs is “part of the genius of the language” and I wanted to test this example of phasaa talâat ‘market speech’ to check if meaning and normative attitudes toward usage would surface in the survey.

(65) Sentence #20

ʔaw ֩ mɔŋ 押 tham ֨ ʁat ʁum ֩ ʁi “go ahead and take them and try cooking them to eat”

Results of acceptability test:
For sentence #20, 9 out of 12 respondents found this sentence acceptable.

Semantic Meaning Results:
Of the 9 respondents attempting to give a meaning, they all gave a similar meaning to that provided by Enfield. Some of the meanings provided are in (66).

(66)

a) Try and make to eat.
   b) Take and cook and eat, go ahead.

Comments collected:
   a) A little confusing
   b) Remove the word ʔaw.

3.6 Discussion of Summary Table of Results

Data Table 4 is an attempt to display in summary form the overall results of this survey. The purpose of this thesis is to test data offered as evidence in various approaches to linguistic theory. Thus the issue of a contrast between the claims made by Muansuwan and Takahashi and those claims being supported or not
supported by this survey will be highlighted in the table. Claims were made as to acceptability, meaning in English, and contrast in meaning between the sentence pairs in the cases of #6 and #7, #10 and #11, #12 and #13, #15 and #16, #17 and #18. In the three sentences #3 - #5 the claim was made that the meaning was the same.

Quantitative measurement of meaning between translations resists any true empirical measurement so for the purposes of this survey the meaning of the sentences were subdivided into events. For example in the Malee sentences the three actions deemed integral to the event descriptions were a reversal of direction of Malee, the crossing of the bridge, and the motion of Malee with respect to the speaker. If the English translation had only one or two of these actions described then the meaning recovery was considered to be one third or two thirds respectively.

Since the sentences were tested at various times and with different numbers of native speakers the table will list the absolute number of respondents in the denominator of a fraction. The numerator will provide the number of respondents that agreed with the claim of the linguist. In many cases a minority of informants offered translations so the denominator reflects the number of attempted translations.
<table>
<thead>
<tr>
<th>No.</th>
<th>Tested Sentence</th>
<th>Source + Result</th>
<th>Survey Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Maalii wîŋ troŋ jɔɔn klæp khâw paj</td>
<td>Thepkanjana 1986, 28/28 acceptable</td>
<td>28/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>Malee run go straight reverse return enter go</td>
<td>Meaning match, 10/15 match</td>
<td>10/15 match</td>
</tr>
<tr>
<td></td>
<td>“Malee ran straight back in, away from the speaker”</td>
<td>Total, 38/43 total match</td>
<td>38/43 total match</td>
</tr>
<tr>
<td>2</td>
<td>Maalii wîŋ jɔɔn klæp troŋ khâam saphaan paj</td>
<td>Muansuwan 2002, 9/24 acceptable</td>
<td>9/24 acceptable</td>
</tr>
<tr>
<td></td>
<td>“Mali ran back straight, crossing the bridge, away from the speaker”</td>
<td>Total matches, 11/33 total match</td>
<td>11/33 total match</td>
</tr>
<tr>
<td>3</td>
<td>Maalii jɔɔn khâam saphaan troŋ ?ɔɔk paj</td>
<td>Muansuwan 2002, 15/24 acceptable</td>
<td>15/24 acceptable</td>
</tr>
<tr>
<td></td>
<td>Mali reverse cross bridge go straight exit go</td>
<td>Meaning match, 3/11 match</td>
<td>3/11 match</td>
</tr>
<tr>
<td></td>
<td>“Piti went back straight, crossing the bridge, out away from the speaker”</td>
<td>Total, 18/35 total match</td>
<td>18/35 total match</td>
</tr>
<tr>
<td>4</td>
<td>Maalii troŋ ?ɔɔk jɔɔn khaam saphaan paj</td>
<td>Muansuwan 2002, 6/24 acceptable</td>
<td>6/24 acceptable</td>
</tr>
<tr>
<td></td>
<td>Mali go straight exit reverse cross bridge go</td>
<td>Meaning match, 2/6 match</td>
<td>2/6 match</td>
</tr>
<tr>
<td></td>
<td>“Piti went back straight, crossing the bridge, out away from the speaker”</td>
<td>Total, 8/30 total match</td>
<td>8/30 total match</td>
</tr>
<tr>
<td>5</td>
<td>Maalii ?ɔɔk troŋ khâam saphaan jɔɔn paj</td>
<td>Muansuwan 2002, 9/28 acceptable</td>
<td>9/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>Mali exit go straight cross bridge reverse go</td>
<td>Meaning match, 1/9 match</td>
<td>1/9 match</td>
</tr>
<tr>
<td></td>
<td>“Piti went back straight, crossing the bridge, out away from the speaker”</td>
<td>Total, 10/37 total match</td>
<td>10/37 total match</td>
</tr>
<tr>
<td>6</td>
<td>naj sâam naathii Nikorn kin khâaw lê ʔím</td>
<td>Muansuwan 2002, 12/28 acceptable</td>
<td>12/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>In three minute Nikorn eat rice and (Nikorn/he) be full</td>
<td>Meaning match, 10/12 match</td>
<td>10/12 match</td>
</tr>
<tr>
<td></td>
<td>“After three minutes, Nikorn ate rice and he was full.”</td>
<td>Total, 22/40 total match</td>
<td>22/40 total match</td>
</tr>
<tr>
<td></td>
<td>English Text</td>
<td>Muansuwan 2002</td>
<td>Match</td>
</tr>
<tr>
<td>---</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td>7</td>
<td><strong>naj sǎam naathii Nikorn kin khâaw ṭǐm</strong> in three minutes Nikorn eat rice be full</td>
<td>Muansuwan 2002</td>
<td>28/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>“Nikorn ate rice until he was full in three minutes.”</td>
<td>Total</td>
<td>44/56 total match</td>
</tr>
<tr>
<td>8</td>
<td><strong>Piti duu nάŋ còb paj</strong> Piti watch movie end PFTV (go)</td>
<td>Muansuwan 2002</td>
<td>10/24 acceptable</td>
</tr>
<tr>
<td></td>
<td>Piti finished watching a movie.</td>
<td>Total</td>
<td>25/39 total match</td>
</tr>
<tr>
<td>9</td>
<td><strong>“Wijada kamlāŋ chú ràng phi</strong>** Wijada PROG believe issue ghost**</td>
<td>Muansuwan 2002</td>
<td>12/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>Wijada is believing in ghosts</td>
<td>Total</td>
<td>20/40 total match</td>
</tr>
<tr>
<td>10</td>
<td><strong>daɔŋ paj thǐng raάn</strong> walk go arrive shop</td>
<td>Takahashi 2005</td>
<td>28/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>“(He) walked away and reached the shop.”</td>
<td>Total</td>
<td>56/56 total match</td>
</tr>
<tr>
<td>11</td>
<td><strong>daɔŋ paj yaŋ ráán</strong> walk go (to) shop</td>
<td>Takahashi 2005</td>
<td>28/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>“(He) walked to the shop.”</td>
<td>Total</td>
<td>56/56 total match</td>
</tr>
<tr>
<td>12</td>
<td><strong>khin paj sùu chúán bon</strong> ascendent go arrive/stay upstairs</td>
<td>Takahashi 2005</td>
<td>2/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>“He went up and got to the upstairs and stayed.”</td>
<td>Total</td>
<td>3/30 total match</td>
</tr>
<tr>
<td>13</td>
<td><strong>khin paj yaŋ chúán bon</strong> ascend go (to) upstairs</td>
<td>Takahashi 1997</td>
<td>7/28 acceptable</td>
</tr>
<tr>
<td></td>
<td>He went up to the upstairs.”</td>
<td>Total</td>
<td>10/35 total match</td>
</tr>
<tr>
<td>14</td>
<td>* dən th 사람이 반</td>
<td>walk reach home</td>
<td>no translation offered</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>?? thenit khi kila thii chan chôp thii sût ??</td>
<td>tennis COP sport REL PRO like most</td>
<td>“Tennis is the sport that I like best.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>kaan rian khi siŋ thii sâmkaanan thii sût sâmràp nák rian</td>
<td>NOM study thing R.PRON important for student(s)</td>
<td>“To study is the most important thing for students.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>màŋ yàak paj tham ɲaan</td>
<td>NEG want go work</td>
<td>“do not want to go to work”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>yàak màŋ paj tham ɲaan</td>
<td>want NEG go work</td>
<td>“want to go not to work” (sarcastic)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>tàn riip klâp paj riâk háj maa bôk</td>
<td>(you) must hurry return go summon cause (him) come tell</td>
<td>“You must hurry back and summon him to come and tell me.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>?aw lɔɔŋ paj tham kin duu sì</td>
<td>Take try go make eat see PART</td>
<td>“go ahead and take them and try cooking them to eat”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total claims</td>
<td>Meanings recovered total 175/264</td>
<td>434/752</td>
<td>Acceptability or unacceptability claims supported 253/488</td>
</tr>
<tr>
<td>Data from claims of contrast between pairs</td>
<td>Muansuwan 2002 0/26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claimed identical meanings for #3, #4, #5 out of 26 meanings offered none matched another</td>
<td>Muansuwan 2002 0/40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data from #6 and #7 Claimed meaning contrast 0/40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claimed meaning contrast between</td>
<td>Takahashi 2005 28/28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#10 and #11 contrast supported</td>
<td>0/9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#12 and #13 contrast not supported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#15 and #16 contrast not supported (iffy and unacceptability claims not supported)</td>
<td>n/a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#17 and #18 contrasted not supported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total contrast claims supported 29/134</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4
Relationship of this survey of grammaticality judgments to the field of linguistics

4.1 Overview of Chapter
This chapter will begin in Section 4.2 by showing some examples of grammaticality judgments in English as a way of introducing the problem of addressing data as evidence in current linguistic theory.

In Section 4.3 the current idea of grammatical competence as being a rule governed formal system accessible through intuitive grammaticality judgments will be discussed.

In Section 4.4 the explication of the formal interpretation of grammaticality contrasted with the notion of acceptability will be investigated.

In Section 4.3 problems associated with linguist-generated assumed acceptable sentences used as data are revealed to be a major problem in some research approaches including the research tested in Chapter 3.

In Section 4.6, the inseparability of normative judgments from grammaticality judgments will be reviewed.

In Section 4.7 a re-interpretation of linguistic data as social normative in its very nature is offered by Itknonen (1981) as a way to resolve the major dilemma of empirical data in language.

In Section 4.8 the Emergentist approach to data and syntax is discussed with an emphasis on the communicative aspects of language.

In Section 4.9 data from the perspective of Computer Natural language processing is reviewed and new questions arise as to how to conceptualize “grammar” in a formal, useful and productive way.
4.2 English and Grammaticality

Since the Thai examples are interpretable to English speakers only to the extent that they can trust the free translation and glossing, this section looks at similar variable grammaticality judgments in English. These sentences all appeared in the journal Natural Language and Linguistic Theory and were collected by Riemer (2009). The claims made for each of these sentences involved issues of A-movement, coordination and extraction from NPs. They are included to show examples of what is regarded as evidence by linguists working within the Essentialist approach.

(67) Each other's supporters attacked the candidates.

(68) What subject was the student hoping to pass the examination on?

(69) These books are hard for Bill to decide when to read.

(70) The more you think about it, the problem compounds.

(71) Mary said that Joe liked these pictures of herself.

All five of these sentences were judged as ungrammatical. I sent e-mails to friends and relatives, all native speakers of American English some from the southern United States and some from the northeastern United States and asked for grammaticality judgments. I received 24 replies and the results appear in Table 5 below:

<table>
<thead>
<tr>
<th>Example Sentences</th>
<th>OK</th>
<th>*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(67) Each other's supporters attacked the candidates.</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>(68) What subject was the student hoping to pass the examination on?</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>(69) These books are hard for Bill to decide when to read.</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>(70) The more you think about it, the problem compounds</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>(71) Mary said that Joe liked these pictures of herself.</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

The conclusions that one can draw from this simple poll are not unimpeachable but do illustrate a simple observation that “ill-formedness” or intuitive ungrammaticality is not uniformly shared among all speakers of a common language. The supposed
ungrammatical of these sentences was evidence for various claims made about English syntax. If judgments vary as to acceptability to this extent then is data truly “empirical”? Linguists working in this Essentialist approach make a distinction between performance data versus competence data that allows a way out of this dilemma. This distinction is controversial to say the least. This distinction is examined in the next section.

4.3 Ungrammaticality and intuition

A claim of unacceptability (the performance meaning according to Essentialists) should not be confused with claims of ungrammaticality. In an effort to clarify this distinction we will look at the difference between these concepts as explained by a linguist working within this approach.

In a 2010 college text “Grammar as Science”, Larson explains what it means to consider a sentence ungrammatical versus unacceptable. “Determining whether a sentence is ungrammatical/ill-formed is much trickier than determining whether it’s unacceptable. In the case of unacceptability, we simply ask the speaker whether a sentence is good or not. In the case of ungrammaticality, we must find out whether the unacceptability arises from a particular source. Sentences can be unacceptable for many reasons. Ungrammaticality is a narrower concept” (Larson, 2010: 58).

Larson considers the distinction between ungrammatical and unacceptable separable in principle but gives no guidance how this separation can be performed in practice. In case the reader is still unclear as to the distinction Larson goes on to spell out in greater detail. “In saying that a speaker has internalized a set of syntactic rules, we’re claiming that those are the rules the speaker draws on in judging well-formedness and ill-formedness. We’re saying that those rules account for the judgments. Accordingly, when we attribute a set of rules to a person, we expect that the person will judge sentences generated by the rules to be well-formed and sentences not generated by the rules to be ill-formed. Sentences generated by the rules will be judged by the person to be well-formed. Sentences not generated by the rules will be judged by the person to be ill-formed (Larson, 2010: 58).”
Let us assume that a sentence that a native speaker judges as unacceptable such as (72) which is taken from Larson’s textbook as an example of a grammatical but unacceptable sentence.

(72) That that that that Bart saw Homer surprised us is doubtful puzzled Lisa pleased Marge annoyed.

Larson explains,

Remember unacceptable does not mean ungrammatical!

One way to deal with this conflict might be to reconsider the basic data and ask whether they are really ungrammatical at all. Surely we could not say them and expect to be understood, nor could we understand them readily if they were spoken to us. But recall that this is a different thing from saying that they are ungrammatical: that their sentence structure doesn’t conform to the proper patterns of English. In fact, it has been proposed that examples ((like this)) are not really ungrammatical but instead grammatical and simply very hard to process. (Larson, 2010: 152)

So if we make the effort to keep an open mind and consider this interpretation then the word “grammar” must now be interpreted differently. Grammar must now refer to the abstract postulated rules regarding syntactical patterns that have been identified by Essentialists to exist in the language under study. The hypothesized patterns could generate an infinite number of sentences that would be interpreted as nonsense by native speakers but these sentences would be grammatical just not acceptable. They are grammatical because the linguist hypothesizing the grammatical pattern decided that they match a postulated syntactic rule and not because a native speaker claims the sentence to be unacceptable.

This seems rather hard to process for someone who believes that the data should control the theory rather than the theory dictating what should be the data. A sentence can be completely nonsensical but still be grammatical if it follows certain allowable syntactical “forms” such as the tree diagram in Figure 2. Recall that words in Jackendoff’s terms were “pieces of phonology” and now we are to think of sentences as possible “tree structures”. In other words, English sentences are just the set of possible
tree structures allowed but not realized. Thus the finite list of forms comprises the
grammar.

The famous sentence ‘colorless green ideas sleep furiously” that is held by many to
unequivocally show the separation of semantic and syntactic meanings is widely quoted
and widely misunderstood. This sentence was claimed by Chomsky to show three ways
that do not determine whether a sentence is grammatical or not grammatical.

a. its inclusion in a corpus
b. it being meaningful
c. it being statistically probable.

Chomsky points out that even though the sentence is grammatical, it is not included in
any known corpus at the time and is neither meaningful nor statistically probable. He
concludes that "grammar is autonomous and independent of meaning, and probabilistic
models give no particular insight into some of the basic problems of syntactic structure
(Chomsky, 1957: 15).” So this sentence as an exemplar of the separability of syntax and
semantics was actually a point of argumentation regarding the source of data for
linguistic study. Chomsky claimed that intuitive judgments were better evidence for
theory construction than data from corpus or usage. The question as to whose
grammaticality judgments should be used is the subject of discussion in the next
section.

The idea that this internalized grammar exists and is represented by rules in the mind
can seem more a matter of an ideological commitment than an outline of a falsifiable
theory. The presence of these rules and whether they are conscious or unconscious has
never been clarified by Chomsky and clarification of this point is not in the literature
(Devitt, 1999: 5). Over the decades since his original claims, Chomsky has never made
this point clear but the following quote is certainly interpretable as assuming that this
knowledge is conscious knowledge.

Clearly, the actual data of linguistic performance will provide much
evidence...along with introspective reports (by the native speaker, or the linguist
who has learned the language). This is a position that is universally adopted in
practice (Chomsky,1965: 18).
Various scholars working in this approach to linguistics have also avoided further explication of exactly how this knowledge could be or might be represented. Notice Larson’s wording “the speaker draws on this knowledge” says nothing about how this happens in fact. Other linguists assume that this knowledge is unconscious and that they have to work diligently to uncover this unconscious knowledge.

4.4 Ungrammaticality and unacceptability

For now we accept the fact that there exists no consensus and substantial debate on how linguistic rules could be represented or even whether these rules are unconsciously or consciously available to a native speaker. But the problem as to how we can clarify the distinction between the notions of acceptability and grammaticality is still an unresolved issue. Larson states that sentences can be grammatical if generated/formed/predicted by syntactical rules but not be acceptable. We must then be able to distinguish what “acceptable” means for this statement to have meaning.

This quote comes from “Linguistic Intuitions” by Gareth Fitzgerald (2010):

Acceptability and interpretability as data sources are to be distinguished from the theoretical notion of grammaticality, and what is generated by a grammar. Speakers have no intuitions about what a grammar mandates, in the theoretical sense of a grammar that concerns linguists. (Fitzgerald, 2010: 130).

In a book entirely devoted to the issue of empirical data in linguistics, Carson Schütze claims that “people are incapable of judging grammaticality – it is not accessible to their intuitions” (1996: 26). His stance on the possibility of distinguishing the two concepts is given in the following quote:

The intuition we have reflects itself in acceptability judgments, not grammaticality judgments. Acceptability is a far cry from grammaticality and linguists might construct arguments about the grammaticality of a sentence, but all that a linguistically naïve subject can do is judge its acceptability... I [Schütze] will follow the existing literature in treating grammaticality judgment and acceptability judgment as synonyms [...] with the understanding that the former is unquestionably a misnomer, and only the latter is a sensible notion (Schütze, 1996: 26–27).
I assume that by “sensible” he means “apprehendable by the senses” and not “reasonable”, but perhaps not, and the notion of a firm distinction between the two notions is not forthcoming. All we can do is measure sentence acceptability, in any case, and when grammaticality has been so resistant to being defined by those who use the word the most (it was coined by Chomsky in 1965) the confusion between the meaning difference, if it can be made, continues today.

4.5 The linguists judgments as data
The data in existing literature conflates acceptability and grammaticality and Schutze’s book lists many suggestions as to how acceptability judgments should be studied and brought into conformity with research standards in the behavioral sciences. I note here that Schutze is a linguist committed to the Essentialist approach to linguistics and the goal of his book is to make this approach more empirical. The following factors are just some of the real world issues that have been shown in studies to interfere or affect results of grammaticality testing. The results of testing have been shown to depend on:

a. linguistic training (p.17)
b. the number of repetitions (p.18)
c. parsability (length or the number of embeddings) (p. 18)
d. the register of the given usage (p. 18)
e. whether the subjects were asked to judge the well-formedness of some expressions before or after the expressions were given (p. 90)
f. whether the anomaly is closer to the beginning of the sentence (pp. 75–76)
g. overall literacy and education (p. 127)

Far from being a reflexive binary judgment, the act of judging sentence grammaticality is revealed to be a “complex behavioral performance’ (Bever and Carroll, 1981). “In many ways, intuition is less regular and more difficult to interpret than speech” (Labov, 1972). Schütze believes that the methodology of linguistics can improve, provided that linguists simply follow the data protocols of other social sciences.

In regard to linguists using their own self-generated sentences as data Featherston (2007) criticizes a practice of what appears to have occurred with some of the sentences that were tested in Chap. 3.
The most fundamental part of the problem is that a significant number of linguists are still, in spite of all the warnings to the contrary, using as the basis of their work what we might call linguist’s judgments. In the worse cases these are introspective judgments given by linguists themselves as the data base of their own theoretical work, on the basis of a single example sentence, not checked against the intuitions of other independent informants, and often idealized to a dichotomy of good or bad. (Featherston, 2007: 272)

Featherston questions whether this activity can properly be called “research”. The focus “is simply hypothesis building using these tokens of intuitively derived grammatical sentences as empirical data.” The selected and chosen examples from this “research” are offered up as evidence for the construction of UG. “[...] much work in the generative tradition of grammar is fatally undermined by its oversimplified assumptions about the patterns that the data which it is attempting to model actually exhibit” (Featherston, 2007: 270). The subjectivity of a linguist’s intuitive grammatical judgments compromises the objectivity of the data which is then offered as evidence. Linguists simply should not offer their linguistic intuition as evidence in theoretical debates. “For similar reasons it is inadequate, too, to take judgments uncritically from the literature. Data referred to in the literature are often dubious because they are based solely on the given author’s linguistic intuition”(Featherston, 2007: 278).

The debate against using linguist self-generated sentence grammaticality judgments has heated up in recent years as Externalists in particular have challenged some of the data that has appeared in journal articles. This quote from Sampson shows how serious other linguists consider the problem to be. “Looking back at the syntax published a couple of decades ago makes it rather clear that much of it is going to have to be redone from the ground up just to reach minimal levels of empirical accuracy. Faced with data flaws of these proportions, biology journals issue retractions, and researchers are disciplined or dismissed” (Pullum, 2007: 3). Today with easy corpus search available to all researchers, claims of the existence of various syntactic and semantic patterns in languages like English can be checked in a real way. Theory needs data to evolve and develop and data poor theory can be poor theory.

I believe that sample sentences #2 through #9, in the last chapter, were linguist self-generated examples offered up as models to support a theoretical construction. These
sentences were difficult if not impossible for my informants to process and when questions were asked of the informants the various interpretations proved that these sentences did not match the free translations offered by the linguist. The specific reasons for why these sentences were difficult to understand will be discussed in detail in Chapter 5.

4.6 The inseparability of normative judgments from grammaticality judgments

In the next section I will briefly discuss the notion of grammaticality from a normative viewpoint. Linguists feel that the study of language as science should have nothing to do with normative grammar issues. All three approaches to language agree that language is a natural phenomenon and that what we should or should not say has no part in the study of language. But, since asking for acceptability judgments entails a speaker's subjective viewpoint on what is acceptable in speech this connection between proper and improper speech is impossible to avoid in language research.

In fact, as a linguist, my view is that if you're a native speaker of English, no matter what your dialect, then you already know English grammar perfectly. And if you're a native speaker of a different language, then you know the grammar of that language perfectly. By this, I don't mean that you know (consciously) a few prescriptive rules, such as those mentioned in the last paragraph, but that you know (unconsciously) the much more impressive mental grammar of your own language – as do all its native speakers. Although we've all learnt this grammar, we can think of it as knowledge that we've never been taught, and it's also knowledge that we can't take out and examine (Tallerman, 2011: 2).

The next quote is from the preface of an English Grammar text that is marketed worldwide.

“No variety of English – including standard English – is inherently better or worse than any other. However, the standard variety is the one that has the greatest value in social terms as a means of communication, especially for public and professional communication. The notion of standard English is especially important to learners of the language. Because of its high social value, learners
are justifiably anxious to ensure that the English they learn is standard English” (Nelson, 2001: 3).

Note the change to “standard English” from “proper English” which has become stigmatized usage in today’s academic world. The notion of normativity, notions of good or bad, are not discussed or acknowledged in Essentialist approaches to grammar. The formal approach to modeling natural language supposedly stems from an objective scientific perspective that describes language as a natural object. In philosophy normative statements are not considered positive or logical statements and questions of what should or should not be are considered normative and thus beyond a rational approach. They are connected with feelings not rational descriptions. This quote gives the flavor of the scientific linguistic attitude towards prescriptive grammar:

Linguistics is concerned with studying languages and language in general, much as biology studies living species and life in general. From this scientific perspective, the norms of prescriptive grammar are to linguistics as the American Kennel Club's breed standards are to biology: arbitrary evaluative standards of no relevance to objective description (Wasow, 2003: 296).

The influence of the norms of “proper” speech, however form the core of the layman’s concept of grammar. For example sentence (73) is considered to have five grammatical errors for an English speaker in Standard British English (Wasow, 2003: 295).

(73) Hopefully, we will be able to easily figure out who to talk to.

Most American English speakers would find it acceptable speech while (74) might sound strange and not acceptable at all in colloquial speech.

(74) I hope that we shall be able easily to figure out to whom to talk.

Tracing the switch from “will” to “shall” in proper American English was the subject of a PhD study for the linguist Charles Carpenter Fries nearly a century ago. His insistence on the value of a neutral language description made enemies and his position was attacked by many in the English teaching community. Jacques Barzun of Columbia University described him as "the theorist who engineered the demise of grammar in the American schools." What continually thwarted Fries was the popular idea that matters concerned with language are simple binaries of good and bad, correct or incorrect. The facts of linguistic usage were not of interest to many and his simple message was
seldom understood. In his textbook on English grammar he tried to make the point that “the differences of usage among native speakers of English are much greater and much more intricate than is usually believed”. Critics of his approach wrote in a review of his text, The Structure of English, “even in this Age of Science, there are areas of our culture about which the disclosure of the cold facts is unwelcome, and our language is one of those areas.” (Fries, 2007).

So in the light of this attitude which has not diminished or changed, how is a researcher to believe that grammaticality judgments have no normative nature? After asking for the grammaticality judgments for the English sentences referred to in 4.2, the most common e-mailed remark was, “am I right?” The assumed inner voice of competence seems an abstract concept indeed. Normativity simply cannot be separated from language research or language description.

4.7 Itkonen’s resolution to the major dilemma of empirical data in language

This section deals with a possible resolution to this dilemma of how to regard linguistic data as empirical data. Itkonen (1980) is a linguist and theorist who argues that linguistics is a normative subject and once the basis for an intuitive science is delineated linguistics can logically proceed to study language and not pretend it is engaging in a program modeled on the paradigm of the physical sciences.

Esa Itkonen has written a great deal about intuitive versus introspective versus observational data. His work is a serious attempt to distinguish what linguistics does versus what linguists think they do by building up a sensible ontology of what an intuitive science truly is. Rather than attempt to summarize 700 pages of argumentation I will paraphrase some points that bear upon the social nature of language that seems missing in all three approaches that have been discussed earlier.

The term “autonomous linguistics” refers to the abstract formal logical approach to language which is “metalinguistic” in nature. Itkonen believes that a formal approach is necessary and possible but that the refusal to make clear the differences between intuitive grammaticality judgments as empirical physical spatio-temporal data versus intuitive judgments as intuitive data has created a crisis in linguistics.
Autonomous linguistics rests upon an intuitive base but has to maintain a fiction that it is in some sense comparable to a physical science which can measure physically real objects. The Essentialist work following the example of Chomsky’s research

….has resulted in the curious situation that people who do nothing but analyze their own intuitive knowledge of self-invented sentences or of imaginary speech acts claim to be investigating informant behavior according to the strictest cannons of experimental science. In this context the linguists' capacity for self-contradiction seems almost unlimited (Itkonen, 1981: 130-131).

Itkonen’s views are important for understanding how the issue of empirical data can be reconciled with the facts of language research.

The basic tenet of Itkonen (1978) and (1983) is that language is primarily a normative entity and that language exists as an object of common knowledge. The prescriptive grammarian does not describe what is said or how it is understood, but rather what ought to be said or how it ought to be understood. And because the norms (or rules) of language that determine these 'ought'-aspects cannot be individual they must be social. Thus, language is a social entity (in addition to being a normative entity). Social norms do not exist in a vacuum, but are rather 'supported by' individual persons and, thus, by individual minds. Language as a social and normative entity is investigated by 'autonomous linguistics'. Language as a social and non-normative entity is investigated by 'sociolinguistics'. Language as a non-social that is individual psychological and non-normative entity is investigated by 'psycholinguistics'.

In summary form, the separation between normative and scientific does not need to be made a dichotomy. Grammar can be considered normative at its core and still be studied by autonomous linguistics in a scientific manner.

### 4.8 Emergentist approach to data and syntax

In this section I discuss the Emergentist approach towards syntactical structure and the possibility that universal grammar does not need to be postulated a priori in order to study abstract structures in natural languages. Many linguists in this approach have turned away totally from clause structure analysis and toward pragmatic approaches to language used communicatively. In this approach grammar is interesting insofar as it aids in the construction of an interpretive theory. Noam Chomsky (1992) sees the wish
for a theory of the interpreter as tantamount to a ‘demand for a theory of everything’, the pursuit of which will lead to a theory of nothing.” Chomsky goes on to say, “The interpreter, presented with an utterance and a situation, assigns some interpretation to what is being said by a person in this situation.” But this, the topic of successful communication, “is far too complex and obscure to merit attention in empirical inquiry” (Chomsky, 1992: 120). Some linguists do not believe that this is the case and that some grammatical categories and structures have pragmatic explanations.

One example of a functional approach to grammar analysis is a reanalysis of imperfect versus preterit tensed Spanish verbs. In this view the imperfect tense might be better analyzed as verb marking of “off topic” back grounded verbs versus the preterit marking “on topic” in focus event verb (DeJong, 2009). The complex issues of coordination and subordination can be understood as examples of narration and elaboration. Coordinate constructions show narration while clauses have subordinating structures to show elaboration (Verstraete, 2009). Some linguists working in the Emergentist approach accept the fact that natural languages have a syntactical structure that needs an explanation. The formal approach has assumed that these categories exist a priori and innately. Other linguists have assumed that these categories must be learned in some sense and the most common approach is to assume they are semantically based (Langacker, 1987; Croft, 1991; O’Grady, 1997). Other approaches are that the surface grammatical forms in sentences are simply due to efficiency and processing design features.

In this theory sentence structure has to be understood as part of a theory of sentence processing (O’Grady, 2005; Hawkins, 2004). The processor has no “grammar” rather it has two features that acting together give the appearance of an innate grammar. The processor has just two computational features-a propensity to operate on pairs of elements and a propensity to combine functors with their arguments at the first opportunity. These two processing principles working together create a hierarchal branching structure that appears as if grammatical rules are at work. The appearance of an architectural design to grammar is just that- an “appearance”. Formal syntactic representations are however, “nothing more than the residual record of how the computational system goes about combining words to form sentences” (O’Grady, 2005: 3).
Abstract and fundamental properties of syntax are argued to be derivable from simple principles of processing efficiency. The emerging correlation between performance and grammars exists because grammars have “conventionalized the preferences of performance, in proportion to their strength and in proportion to their number, as they apply to the relevant structures in the relevant language types” (Hawkins, 2004: 2). Hawkins believes grammars are just ‘frozen’ or ‘fixed’ performance preferences. What are “acceptable usages” changes over time and this dynamic functional view of language allows for language change.

Intuitive grammaticality judgments in this approach have little role as primary evidence for theory building. Language as communication in whatever form— as sentence fragments, pauses between words, intonation patterns, all are considered as data.

4.9 Data from computer natural language processing

Linguistics has entered a new more empirical era and probabilities of specific usages are now open to empirical verification through the use of large databases. New methods of research are possible and demands are increasing for semantic data and methods to organize and systematize that data. One of the great discoveries of linguistics has been how complex natural languages really are, as compared to assumptions made decades ago when formal linguistics had only assumptions about the complexity of the task (Steedman, 2011).

If we think of a grammar of a natural language as a type of software, analogous to today’s metadata markup processing software, then the grammars of the size that are needed to read the newspaper are very large indeed. For example, “the context-free phrase-structure (CF-PS) base grammar that Collins induced from the human-annotator-labeled Penn Wall Street Journal Treebank, using around 50 Part-of-Speech (POS) labels as lexical categories, has around 12,000 PS rule types” (Steedman, 2011).

Steedman notes that facts like this from natural language processing technology make one ask the question are “grammars proposed by theoretical linguists yet ready to help computational NLP ”. The answer is obviously “no”. According to Steedman, from the 1960s until the mid 1970s, there was almost complete theoretical unanimity among linguists, psycholinguists, and computational linguists. The consensus soon fell apart, however, largely because of early disagreements about the role of semantics in the competence theory (Chomsky, 1972), the recognition of the unconstrained power (and
consequent weak explanatory force) of structure-dependent transformational rules (Peters and Ritchie, 1973), and the realization of the huge amount of syntactic ambiguity inherent in human-scale grammars.

In spite of wide spread public recognition of the Essentialist program and in particular the fame of Chomsky as the creator of this newer scientific approach to “doing linguistics” the generative program has not succeeded in successfully modeling any natural language. Furthermore, very little of the work done in the style of formal algebraic modeling of language, as done in the earlier Chomsky tradition, is usable in computational linguistics. Another recent textbook in Computational Semantics puts the situation in an even clearer perspective.

Still, it is fair to say that the practical achievements of computational semantics have so far been quite limited. The reasons for that, I think, are two-fold. Automated symbolic processing of natural language is notoriously brittle: even where it is clear what the system should compute, it often lacks the necessary resources, in particular wide coverage lexicons with substantive semantic information and world knowledge in accessible form. But in many cases the problem goes deeper. We still haven’t even properly understood yet what it is that should be computed (Eijck, Jan Van & Unger, Christine, 2010).

Steedman makes the case that the structure of grammar is still an open question. Wider coverage of rules means making lists of exceptions and lexical items seem to resist simple semantic or syntactic categorization. His conclusion is that the true nature of grammar might be that it is structured more like traditional grammars—“large lexically and morphologically anchored and licensed for idiosyncrasy and exception” (Steedman, 2011: 5). To bring theoretical linguistics back into the real world so that it can be of practical use Steedman thinks that the theory of syntax needs radical overhaul. The past practice of simply choosing some off the shelf logical scheme like case frames or first order logic and then trying to force fit sentences into these molds means that we have to tolerate syntactical complexity in order to get a formula to build up the real sentence structure. This is the wrong way and the reverse of the way it should be done. The real world need is for grammars that directly support low-complexity derivation of a considerable variety of surface constructions. The work of data bank parsing has taught computational linguists that there can be thousands of syntactically legal parses of moderately long sentences. This is a fact generally not known or appreciated in the
Essentialist community of theorists who have increasingly isolated themselves from real world applications of linguistic theory.

4.10 Summary discussion

In summary, linguistics today is in flux and challenges to foundational issues make choosing a methodology and approach to data collection and data analysis difficult. We do not know now for certain what data will serve as evidence for a future "science of language". But to move any debate or research program beyond rhetoric and argumentation to clear hypothesis testing requires at a minimum interpersonally objective data. Linguistic researchers in any of the three approaches so far discussed should not be producing the data that they analyze.

New technologies and new ideas mean change which some researchers welcome and others may not. Language changes slowly and theory changes rapidly so the collected data is the real empirical data. Data collection can become data destruction in the sense that in some approaches to analysis data that is “theoretically inconvenient” disappears from the database (Talasiewicz, 2010: 10). There appears to be a wide gulf between the perceptions of the state of “progress” in linguistic theory between different communities of working linguists. Contrary to the picture painted in introductory linguistic textbooks such as the Larson text in Sec 4.3, there is little evidence in the practice of field linguistics, computer linguistics or corpus linguistics to show that we currently have methods of formal analysis of grammar able to successfully model any natural human language. The existing successes of machine translation, search algorithms, and speech recognition software have all been made possible not by the application top down formal theory but rather by bottom up massive data collection using statistical techniques. The simple fact is that the serious study and description of language requires massive amounts of data collection. On the positive side, the ability to collect, store, sift, and analyze large data bases has never been easier or less costly than it is today. Technology has always resulted in change and these changes could allow the goal of building an empirical base for linguistics possible far beyond the limited dreams of linguists just decades ago.
Chapter 5
Implications for the linguistic analysis of Thai

5.1 Overview of the chapter
In Section 5.2, the features that appear to make Eurocentric analysis problematic are discussed.

In Section 5.2, examples are used to show how sentence acceptability seems less syntax driven than it is by semantically recoverable meaning.

In Section 5.4, I discuss the concept of a main verb and the theory driven need to find such a verb in Thai.

In Section 5.5, an example of a typical Thai sentence is contrasted with the English translation in terms of the number of grammatical morphemes.

In Section 5.6, the example from Ahom, an extinct Tai language, is described by a linguist working before the Essentialist universalist approach to language description became the norm.

In Section 5.7, a semantic analysis of verbs used in the Malee bridge sentences reveals the confusion and primacy of semantic information as necessary preliminary empirical data.

In Section 5.8 normative grammaticality judgments are a feature of all languages and a brief look at Thai with its complicated speech registers and their impact on grammaticality judgments is discussed.

5.2 Eurocentric linguistics
The following is the mission statement of an organization that has done a great deal for the community of linguists in this region. I have underlined what I believe is the most important point that will be discussed in detail in this chapter.
The Southeast Asian Linguistics Society (SEALS) was conceived by Martha Ratliff and Eric Schiller in 1990 as a needed forum for the linguists who have the languages of mainland and Pacific Southeast Asia as their primary research focus. It is our hope that the activities of the Society will lead to:

1. greater communication within this group of scholars, especially across the gap which has heretofore divided researchers of mainland Southeast Asian languages and the Austronesian languages of the Pacific;

2. needed publication of descriptive, theoretical and historical accounts of these languages, in the first instance in the form of these proceedings volumes; and

3. greater awareness of these languages by non-specialist linguists, many of whom attempt to make universal and typological generalizations about the human language faculty without the important corrective which knowledge of Southeast Asian languages provides.

I am not certain what specific “important corrective” these authors had in mind but I would surmise that issues such as those below are involved in the “important corrective knowledge”. The points involve a seeming unawareness that languages can do without structures and forms that are considered core features of grammar in the Eurocentric approach to linguist analysis. Features such as:

a. Verb concatenations where identifying a main verb is problematic
b. Topic comment rather than SVO order as organizational principle of sentence order
c. Participant roles recoverable largely by context and pragmatics
d. No morphology to help with word categorization into syntactic classes
e. No marking for grammatical relations (no case no agreement).

The fact that this list contain the words “no” (4 times), as well as “rather than” and “problematic”, means that the syntax and grammars of languages in this region are being defined in terms of what they do not possess rather than what features they do exhibit (Minnegishi, 2011).
The focus of this thesis is to examine the concept of empirical data in linguistics and the point has been made that there is no data without a theory. To accept the fact that data is theory driven one must also acknowledge that a commitment to one theory can blind us and make us ignore data. Thai has many wonders and lessons to show us if we can be open to new ways to look at what a grammar needs to be. The existence of a syntactic grid whereby grammatical morphemes force the positioning of words as category tokens into slots is not a generalized characteristic of languages in this region.

These characteristics in the list above are areal features and this same description would pertain for Vietnamese, Khmer, Hmong, and many other languages in the region which belong to various language families. Often the same sentence can be glossed word for word in sequence in these languages. The World Atlas of Language Structures presents data on hundreds of languages and makes quantitative comparisons between languages possible by measuring how similar they are on a range of structural both phonological and grammatical features. The convergence among MSEA languages is so thoroughgoing that the typologist Östen Dahl calls the area ‘the ultimate Sprachbund’ (Dahl, 2008: 218). The list below shows the measure of typological distance between specific languages.

Dutch versus German = 10  
German vs. English = 21.1  
Polish vs. Russian = 12.8  
English versus Persian = 42  
Thai vs. Khmer = 12  
Thai vs. Vietnamese = 11.4

The fact that Thai, Khmer and Vietnamese belong to three separate language families yet share greater typological affinity than English and German is remarkable and suggests that linguistic descriptions better suited to Thai will be easily extensible to other languages of Southeast Asia.

5.3 Implicit syntactic assumptions challenged
First some examples are offered that show characteristics of Thai that contrast strongly with the single verb monoclusal constructions of English. Word boundaries are problematic in Thai to an extent difficult to communicate to speakers of English. I will
show an example of the semantically productive technique of re-ordering verbs by taking three common verbs and scrambling them to see how many are meaningful and considered as acceptable sentences by Thai native speakers. The point of this exercise is to illustrate the lack of semantic or grammatical stability outside the phrase for the individual “words”. Traditionally Thai words have been termed “free morphemes” but this term merely contrasts morphemes as bound or free. Thai “words” can change meaning and function in different contexts. By this I mean that very specific lexical pairings or groupings of words have different meanings. Strict constituency and words as stable units outside of syntax cannot be assumed.

Table 6: Thai verbs scrambled word order

<table>
<thead>
<tr>
<th>Thai</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>ʔaw paj dâj take go can</td>
<td>(you) can take it with you,</td>
</tr>
<tr>
<td>ʔaw dâj paj take can go</td>
<td>Unacceptable</td>
</tr>
<tr>
<td>paj dâj ʔaw go can take</td>
<td>Unacceptable</td>
</tr>
<tr>
<td>paj ʔaw dâj go take can</td>
<td>go take it</td>
</tr>
<tr>
<td>dâj paj ʔaw can go take</td>
<td>got a chance to take something</td>
</tr>
<tr>
<td>dâj ʔaw paj can take go</td>
<td>took and left</td>
</tr>
<tr>
<td>ʔaw paj take go</td>
<td>take, (too many meanings to list)</td>
</tr>
<tr>
<td>ʔaw dâj take can</td>
<td>take it if you want it</td>
</tr>
<tr>
<td>dâj paj can go</td>
<td>went, got to go</td>
</tr>
<tr>
<td>paj dâj go can</td>
<td>(we) can go,</td>
</tr>
<tr>
<td>paj ʔaw go take</td>
<td>go and take it</td>
</tr>
<tr>
<td>dâj ʔaw</td>
<td>got to take, took something</td>
</tr>
</tbody>
</table>
The list above of three words scrambled and arranged in every possible combination form acceptable meaningful clauses in 13 out of 15 arrangements. Many linguists when reading glossed words strings as in the examples above draw conclusions regarding sentence structure as if this were in fact some simplified English sentence or just a list of verbs. The fact that some meaning recovery is possible with this technique leads them to believe that they are seeing a pared down English sentence. The fact that a verb can constitute a full clause in these languages is somehow not credited as a fact. Linguists writing descriptive grammars of languages in this region try to make that clear in case after case as in a recent grammar of Garo a Tibeto-Burman language.

**Word Order.** Garo sentences require nothing except a verb. Even though Garo has no hint of verb agreement, neither agent, patient, nor any other actor needs to be explicitly mentioned so long as the larger verbal and nonverbal context makes the intended meaning clear. Nouns, noun phrases, pronouns, and adverbs are, of course, frequently used to flesh out the meaning of a sentence, and noun phrases can have great internal complexity. Nevertheless, they are not essential. (Burling, 2009).

All a sentence needs to be acceptable is a verb or series of verbs. This is a true description of many languages in this region. Both the instability of word-classes and the tendency of clauses to be verb-only are central features of these languages which cannot be properly accounted for in essentialist syntactic theories developed on Indo-European languages.
5.4 Main verbs are assumed

The concept of a clause centered main verb is deeply engrained in linguistic theory and when serial verbs are brought up as a topic that could make this concept problematic, many linguists seem to dismiss arguments that this could be the case. The logical form of a sentence assumes the existence of a subject and a predication. Clause boundaries are marked by main verbs and Western trained linguists interpret a series of verbs as syntactic coordination or subordination. A single clause is comprised a subject and one finite verb. How then can a series of what semantically appear to equally ranked verbs be analyzed?

The central importance of the idea of a main verb is simply assumed in syntactical theory. This point is noted by Goldberg, who argues that Construction Grammar avoids the problem inherent in the “widespread claim in current linguistic theories that syntax is a projection of lexical requirements. This claim is explicit in the Projection Principle of Government and Binding Theory (GB) (Chomsky, 1981), the Bijection Principle of Lexical Functional Grammar (Bresnan, 1982), and in all current accounts which attempt to predict overt syntax from semantic roles or theta role arrays. In all of these frameworks, it is the verb which is taken to be of central importance. That is, it is assumed that the verb determines how many and which kinds of complements will co-occur with it. In this way, the verb is analogized to the predicate of formal logic, which has an inherent number of distinct arguments. The verb is taken to be an n-place relation "waiting" for the exactly correct type and number of arguments” (Goldberg,1995). If an acceptable sentence is Thai can be eight verbs in sequence as shown in (63) in Chapter 3, then this view of sentence structure needs some “corrective knowledge” to accommodate languages where the strict separation of lexical versus grammatical meaning seems problematic.

An example of the misunderstanding of what serial verbs can do comes from (Newmeyer, 2004) who brings up the issue that in English one can say “go get the book” which shows that serializing verbs is possible in English as well. He does not mention the fact that in English this construction is an imperative sentence with “go” understood as a reduced clause. Are serial verbs then simply reduced clauses as examples in English have been analyzed to be? Examples such as “he pulled the door open” deriving from “he pulled the door open and it opened” (Bybee, Perkins, and Pagliuca, 1994) might be comparable syntactical constructions.
But there is abundant evidence that the use of serial verbs as predicates involve much more than this. As Matisoff (1969: 71) puts it, “SVCs 'serve to provide in a uniform way the sort of information that in the surface grammar of languages like English is handled by a formally disparate array of subordinating devices: complementary infinitives, -ing complements, modal auxiliaries, adverbs, prepositional phrases, even whole subordinate clauses.” This observation was made more than 40 years ago and it seems to have opened few minds. Essentialist assumptions about plausible linguistic descriptions are interfering with properly investigating this productively rich area of Thai languages.

5.5 Contrasting numbers of function words

In this section a randomly chosen sentence from a Thai short story will be compared with its free English translation to show the great contrast in the ratio of grammatical words with content words between the two languages.

This sentence below (75) is a perfectly ordinary sentence taken from probably the most famous story in Thai literature. It is not a context dependent short simplex clause with elided arguments but rather a fairly normal Thai sentence of moderately high register. It has eight verbs versus three verbs in the English free translation. But what is more different and what is a general characteristic of Thai at all registers is that the ratio of grammatical words versus content words is much fewer than English.

(75)

Mom  man  kāō  kraʔdōot  nāam  wāaj  paj  khāap  ?ōw
N  IT  LINKER  V  N  V  V  V  V
Mom  it  then  jump  water  swim  go  hold.in.mouth  take

kīgmāj  klāp  ma  hāj  naaj
N  V  V  V  N
stick  return  come  give  master

Mom then jumped into the water and swam back to take the stick back to the master.
The grammatical morphemes are underlined above. English twelve grammatical morphemes while Thai has two grammatical morphemes (if we take ma to be a directional preposition we can gloss ma as “back” thus reach a total of three grammatical morphemes. Thus we have a ratio of comparison of at least 4 to 1. I have done this with several dozen Thai sentences and generally the ratio is at least 2 to 1 in every case. If we accept that the grammatical morphemes make a type of syntactic grid that open class words can slot into, this grid looks a lot different in Thai than it does in English. What seems to be different is not just the fact that there are more verbs but that there are less functors or grammatical morphemes to syntactically structure or limit the possible meanings of each word. Essentialist descriptions which often prioritize bleached grammatical morphemes are not well equipped to handle languages like Thai which express the same concepts but always include significant additional semantic information.

5.6 Pre-essentialist description of Ahom

What follows is a description of a Thai language that has no more living speakers but much written text. It was written nearly a century ago and appears in the Linguistic Atlas of India where Grierson (1928) describes the Tai Kadai language of Ahom. The description of the language was written decades before the current universalist view language structure emerged as orthodoxy in linguistics. I think that it is an insightful, jargon-free description of current low register conversational Thai. The concept of a word as primarily a semantic entity rather than a syntactic entity emerges as the key point in the description of Ahom and, I believe, in Thai. In particular the quote “each particle affects the whole sentence” will be examined in detail in this chapter.

Ahom was the language of the Tai conquerors who invaded Assam in 1228 and ruled until the 18th century. By 1880 there were few living speakers but much literature. Every word, without exception denotes, primarily, the idea of some thing action or condition, such as a man, a tree, striking, going, life, death, distance, propinquity, goodness. Thus if in Ahom we wish to express the idea "slept" we say "sleep-completion", if we wish to express "sleeps" we say "sleep-existence". It will be seen that the processes of what we call declension and conjugation do not properly occur in Ahom, nor can we divide the vocabulary into parts of speech. The relation which in Aryan languages by these two
processes of inflexion are in Ahom indicated partly by the position of the two words in the sentence and partly by compounding words together. We cannot properly talk of nouns and verbs. We can only talk of words performing the functions of nouns and verbs. Particles that give the idea of tense have their own meaning. Thus "uu", the particle of present time means "existence". As each particle affects the whole sentence, Tai languages can afford to be economical of their use. If many words are performing the functions of verbs one tense particle is supplied for all. (Grierson G. A., 1928: 73)

Grierson’s description was not bound by the strong essentialist theories of the late 20th century and seemingly as a consequence he easily observed that word-classes are a function of syntactic roles and not an immutable property of lexical items.

5.7 Semantic analysis of verbs used in the Malee bridge sentences

Now that the point has been made that verbs in Thai are more semantic entities than syntactic entities a detailed semantic analysis of the words involved in the Malee sentences in Chapter 3 is presented as a prologue to the analysis which follows in Sec 5.7.

The words that were glossed as verbs in the Malee sentences will be examined and contrasted after some word meaning issues are discussed.

$jôn = \text{(dictionary lists as verb) to come back, to answer back, to retort, to be retroactive, to go against.}$
Table 7: Common collocations of jón

<table>
<thead>
<tr>
<th>jón kláp</th>
<th>to turn back</th>
</tr>
</thead>
<tbody>
<tr>
<td>jón tòp</td>
<td>answer back</td>
</tr>
<tr>
<td>jón néua mài</td>
<td>to go against wood grain</td>
</tr>
<tr>
<td>jón roi</td>
<td>retrace steps, get back at someone</td>
</tr>
<tr>
<td>jón sêañ</td>
<td>take a picture with the lens facing the sun</td>
</tr>
<tr>
<td>jón lâŋ</td>
<td>retrospective, retroactive</td>
</tr>
<tr>
<td>neuk jón lâŋ</td>
<td>to recall, to think back to</td>
</tr>
<tr>
<td>mái mee pôn jón lâŋ</td>
<td>(legal term) not retroactive</td>
</tr>
</tbody>
</table>

Asking informants and from personal knowledge this word can be paired with the verbs used in the sentences to mean:
- jón khâm = return across
- jón paj = turn around
- jón kláp = return back, do a U-turn
- jón ṭɔɔk = return the way out the exit you came

kláp = (dictionary lists as verb, adjective) to return, to turn back/round, to change, to reverse; instead, unexpectedly

Table 8: Common collocations of kláp: (of 28 listed)

<table>
<thead>
<tr>
<th>kláp kẽun</th>
<th>return, replace, give back</th>
</tr>
</thead>
<tbody>
<tr>
<td>kláp kan</td>
<td>vice versa, on the contrary</td>
</tr>
<tr>
<td>kláp paj kláp ma</td>
<td>back and forth</td>
</tr>
<tr>
<td>kláp kham</td>
<td>go back on your word</td>
</tr>
<tr>
<td>sôn kláp</td>
<td>to send back</td>
</tr>
<tr>
<td>han kláp</td>
<td>to turn around</td>
</tr>
</tbody>
</table>

troy = (dictionary lists as adjective, preposition) straight, upright, honest, accurate, direct, same, at, in
Table 9: Common collocations of トル: (of 29 listed)

<table>
<thead>
<tr>
<th>トル kan</th>
<th>identical</th>
</tr>
</thead>
<tbody>
<tr>
<td>トル (kan) ชนะ</td>
<td>opposite, contrary, on the contrary</td>
</tr>
<tr>
<td>トル นัน</td>
<td>right there at that spot</td>
</tr>
<tr>
<td>トル นิ</td>
<td>right here</td>
</tr>
<tr>
<td>トル นี้</td>
<td>where, at which place</td>
</tr>
<tr>
<td>トル นิ</td>
<td>in front</td>
</tr>
<tr>
<td>トル ตรง</td>
<td>direct route</td>
</tr>
<tr>
<td>トル ผจญ</td>
<td>straightforward, forthright, frank</td>
</tr>
<tr>
<td>บอก トル ตรอ</td>
<td>to tell the truth, speak frankly</td>
</tr>
</tbody>
</table>

ชนะ = (dictionary lists as verb) to cross, to go over, to skip

Table 10: Common collocations of ชนะ (of 16 listed)

<table>
<thead>
<tr>
<th>ชนะ ทานน์</th>
<th>cross the street</th>
</tr>
</thead>
<tbody>
<tr>
<td>ชนะ 合击</td>
<td>cross a border</td>
</tr>
<tr>
<td>ชนะ รูา</td>
<td>to cross a stretch of water by boat</td>
</tr>
<tr>
<td>ชนะ ชนะ</td>
<td>(of pupils) to skip a grade</td>
</tr>
<tr>
<td>ชนะ น้า</td>
<td>to pass over, overlook, bypass</td>
</tr>
<tr>
<td>ชนะ น้ามชนะ ทะนัย</td>
<td>to endure hardships before achieving success</td>
</tr>
<tr>
<td>トル ชนะ</td>
<td>opposite, contrary</td>
</tr>
</tbody>
</table>

ผจญ = (dictionary lists as verb, adjective, adverb) to be out, to leave, to issue, to emit, to start, to show, to bud; out, off (one of the most common words in Thai).

Common collocations (105 listed)

ผจญ pay = to go out; get out!

No collocations with the four verbs used above in Sentence #1-#5.
So if we look at the sentence #2 again aware of these semantic meanings and common collocations of these five words certain facts emerge.

(76) Survey Sentence #2

Maalii  wîŋ  jóng  kláp  troŋ  khâam  saphaan  paj
Malee  run  reverse  return  go.straight  cross  bridge  go

“Mali ran back straight, crossing the bridge, away from the speaker”

Primarily the fact that troŋ does not mean “go straight” but rather has adverbial or stative verb properities as well as being part of the locative collocation with khâam meaning “opposite to” or “opposite of”. Secondarily, jóng kláp is commonly paired together to mean “turn around” or “do a U-turn”. This sentence had many possible construals but the two interpretations that emerged in my interviews as is seen in the translations can be explained by the grouping and glossing below in (77). In a sense this sentence could be considered a semantic garden path sentence.

(77) Maalii  wîŋ  (jóng  kláp)  (troŋ  khâam)  saphaan  paj
Malee  run  U-turn  opposite from  bridge  go

The theoretical assumption that these words are simply tokens of a class or category of words that are semantically stable outside of a clause usage is simply not a safe assumption. What happened in this sentence is that the English glosses along with word category errors affected the analysis of these structures. Words glossed as verbs can also function in Thai as prepositions would in English. Words assumed to be separate may be compounds.

The free translations given along with the English glossing hardly began to narrow down the possibility space of these sequenced verbs to one agreed upon motion event scenario. A few detailed questions asking about details of the motion event showed that most versions of the speakers who claimed to understand the meaning were inconsistent and very loosely construed. Event verbs are simply the actions that in any particular language have been tagged with a name. But all languages can be shown to greatly under-determine the range of possible actions. Words are prompts for the imagination and any one word can involve many possible conceptualizations. Personal experience as a technical writer has proven that a picture really is worth a thousand words.
Attempt a description to tell a friend how to tie a bowline knot over the telephone and one quickly realizes the limit of words to explicitly describe a certain series of explicit actions.

Testing the Malee sentences was by turns frustrating, hilarious and extremely informative linguistically speaking. The interpretations of the sentences proved most revealing since the versions offered by many respondents could not withstand scrutiny of some basic questions such as those below:

a. Is Malee running away from the speaker?
b. Did Malee cross the bridge?
c. Did Malee come back across the bridge after she crossed it?
d. Did Malee turn back before she reached the bridge?

I created several more sentences to narrow the possibilities of various interpretations of this motion event scenario. I gave pen and paper and asked the informants to imagine three friends, one whose name was Malii and draw a picture locating where these three people were with respect to a bridge.

(78)
Maalii jùu troŋ khâam saphaan
Malee is straight cross bridge

(79)
Maalii wîŋ troŋ khâam saphaan
Malee run straight cross bridge

For (78) the two friends are talking and Malii is on the opposite side of the bridge. The words troŋ khâam are obviously being interpreted as prepositions, meaning “opposite” just as the dictionary definition also defines for this collocation. The verb jùu here keys that meaning. For (79) Malii is running straight crossing the bridge since wîŋ troŋ is interpreted as “running straight”. The claim that the re-ordering for sentences #3-#5 in 3.5.3 was possible was shown by the survey results to be an invalid claim. The interpretations offered by informants varied and seemed more attempts at possible construals of a confusing sentence. The meanings, if at all retrievable, were certainly not the same.
In many cases, a semantic understanding must precede a successful syntactic parse—a fact which gives lie to the essentialist attempt to divorce syntax from semantics.

5.8 Grammaticality as a normative judgment

Thai has styles or registers of usage and the idea of grammar being the “rules of proper speech” is widely accepted. As far as grammatical judgments are concerned, the Thai informants that I have tested these sentences with believe that their judgments are either correct or not. They assume that their knowledge about something is being tested and this knowledge is normative knowledge. They do not understand the concept that they can have an opinion that a sentence is correct or not, much less that they have an inner competence that comes with being a native speaker that would allow them to automatically access this source of correctness. This is a very foreign concept to most Thais and probably to the large majority of the world’s population. It is not the focus of this thesis to research the history of the concept of “grammar” in Thailand but a few remarks are necessary to understand and interpret the results of my survey.

Diller makes a remark regarding Thai linguistic research “any attempt to do linguistics becomes sociolinguistics quite rapidly either overtly or covertly” (Diller, 1985). Speech marks one’s social status here as it does everywhere and Thais call low register speech “phasas talaat” (market speech). Diller (1989) goes on to say that “modern Thai has its roots in a traditional form of diglossia, into which has been grafted normative-prescriptivist material, partly assimilated from Western sources”.

By “a traditional form of diglossia” he refers to the high register of speech appropriate to addressing royalty. All Thais are aware of this whether they are capable of using this speech register fluently or not and educated Thais have to learn this register as a matter of course. The reference to the prescriptivist Western material refers to models of grammar that Thais are taught in schools that have been heavily influenced by Eurocentric models of what constitutes a proper sentence. In this sense Diller claims that looking for SVO sentences in Thai is a simple matter because Thais have been taught this model in school and can readily produce examples of this form when asked to produce well formed sentences. Being aware of this historical background is also crucial to an understanding of the significance of current Thai judgments as to whether or not particular sentences are “grammatical”.

80
The sociolinguistic situation is Thailand as regards speech register alone is extremely complicated. Other complicating factors exist, such as the fact that dialects which differ to the point of mutual unintelligibility are spoken by a majority of Thais and for most Thais, Standard Thai is learned at school (Smalley, 1994). This creates a situation where asking for a grammaticality judgment can put a speaker at some risk of being judged themselves.

I believe it was the influence of speech register that was involved in the different judgments for sentence #19 shown in (82) below. A bare majority considered this sentence acceptable and most of the translations caught the gist of Smythe’s translation. But many considered this an unacceptable sentence probably because it seemed “phasaa talaat” and not proper. This is a common L register example of speech only without the provided context clues for help in interpretation.

(80) Sentence #19

Ø  tõŋ  rûp  klâp  pâj  râk  hâj  Ø  maa  bûk  Ø
(you) must hurry return go call cause (him) come tell (me)
“(you) must hurry back (and) summon (him) to come and tell (me).”

Attempting to collect simplistic grammaticality judgments in the Thai linguistic environment as suggested by essentialist linguistic theories is impossible as the intermingling of intuition and multiple normative pressures is inevitable.
Chapter 6
Conclusion

6.1 Conclusion
In this section, I take a look back at some assumptions and some perspectives gained.

Jackendoff’s tree diagram showing no words at the bottom of his tree diagram that he used to illustrate the point that syntax concerns parts of speech not words, is important to bring back into discussion. Jackendoff’s point was that words are “pieces of phonology” and that the syntax essentially was the structure of the tree. We have seen Muansuwan’s tree structure which was hypothesized to be the syntactical structure of her example sentences. When the informants were asked how they would change these sentences to make them more “correct” most said to do that they would have to know what the sentences meant.

This simple fact shows, to my mind, that meaning dictates syntax rather than the reverse. If Malee had in fact crossed the bridge then troj khaam would be two words not one. In this simple case we see illustrated an example of where semantics can explain syntax but then syntax can hardly be autonomous. In Thai orthography generally there are no spaces between words to mark a word boundary. Verb clusters can be interpreted as successive actions or can cohere and be interpreted as referring to a single event. These semantic facts of meaning and usage cannot then be ignored in building abstract theoretical models since the quality of a theory is based on the quality of the data, and semantic and pragmatic data simply is the primary data.

The Essentialist practice focusing on context free sentences as the unit of analysis can create pseudo problems assuming that disambiguation has to be done through syntax. Thai speakers make inferences without the obligatory grammatical structures that force one strict interpretation of a clause or sentence. A general feature of languages in this region is that speaker/listeners have to keep open more interpretational options and pay more attention to context in order to interpret meaning at the clause level.
Collecting large data bases and empirically finding patterns of distribution for sets of verbs is definitely a worthy research goal. As Steedman (2011) notes, grammars of natural languages are complex—more like almanacs than algebraic formulae. His remark that “the real world need is for grammars that directly support low-complexity derivation of a considerable variety of surface constructions” is true for English but also true for languages with radically pragmatic grammars as found in SE Asia.

As shown by my word scramble and the sentence from the story Mom, Thai has few grammatical morphemes that syntactically structure the sentence and mark relations between the words. A syntactical grid with slots for large open classes of words does not exist but rather many collocated verb patterns some of which cohere into larger potential event descriptions. The selection restrictions are semantic and pragmatically governed and context is important for reconstructing meaning. What is seen in linguistics today as a problem, that is how to fit serial verb constructions into the greater conceptual scheme of description and classification, is evident by these quotes from Newmeyer. He quotes Aikhenvald’s (2006) list of characteristics of serial verbs then remarks, “the problem is that the components of her definition, namely notions like ‘acting as a single predicate’, ‘monoclusal’, ‘dependency of any sort’, ‘conceptualizing as a single event’, and so on are themselves either so vague or so controversial that her definition does not help us very much in separating the class of SVs from the class of non-SVs. The problem is that continua and prototypes explain nothing. At best they provide an ordered set data in search of a theory” (Newmeyer, 2004: 4).

Newmeyer then offers 13 pages of examples from the literature showing that current theories, both formal and functional, do not solve the puzzle of serial verbs. In Newmeyer’s view the suggestion serial verb constructions represent a fundamentally different view of viewing reality is “drastic and unwanted”. He concludes: “Suffice it to say in summary that the general confusion about what characterizes serial verbs at a descriptive level is matched within the generative community by uncertainty as to how to characterize them at a theoretical level” (Newmeyer, 2004: 17).

The admission of confusion is refreshing and spells opportunity for Thai speaking linguists to step up and explain why Thai and other isolating languages in this region manage well without structures thought to be necessary in the Eurocentric approaches
to syntactical analysis. Saussure originated the analogy between the grammar of a language and the rules of chess, ‘the respective value of the pieces depends on their position on the chessboard just as each linguistic term derives its value from its opposition to all the other terms’ (Saussure, 1913: 88). This model is at the heart of the Essentialist view of grammar. The question remains however, can language be modeled with a formal system of rules?

Jackendoff’s paper, The State of the Art (2007), upholds the concepts of formal algebraic structures for syntax but chides Chomsky for his “syntactocentrism”. The title to his paper is obviously taken from a book by Hockett written forty years ago where Hockett developed a summary of 19 points that had to be accepted to agree with Chomsky’s formal model of language. Hockett disagreed with Chomsky on 17 of those points. In the book Hockett created another analogy which I believe more accurately describes a natural language. “In sand lot chess inhabitants of an isolated village play chess their whole life only the rules are never written down and all learning is via observation and participation. Furthermore, the “rules” can be changed. They are whatever you can get your partner to accept. To our way of thinking, sand lot chess is not nearly so desirable a game as real chess. It is not very much like a language. But it is much more like a language than is real chess” (Hockett, 1968: 85).

On a personal note, I confess that my age and background had led me to assume that linguists were interested in learning natural languages and that most if not all linguists would speak several languages. I worked at the National Security Agency during my four years in the US Air Force as an intelligence analyst translating Russian and Spanish between 1970 and 1974. My exposure to academic linguists during that time involved very early efforts at machine translation and I was extremely impressed with the few professional linguists that I met and who all spoke several languages fluently and also seemed to have a profound respect for data.

That this is not the case now and that I was profoundly wrong about this is the only reason I can think of for this quote displayed proudly on a linguist’s website. “Asking a linguist how many languages they speak is like asking a doctor how many diseases they have”. The fact that this quip is considered humorous enough to post on a linguistic website for professional linguists speaks volumes about the modern attitude about what the subject should be construed to be by professional linguists. I would like to conclude with some comments as to why an intimate knowledge of a language under
study might affect one's approach to linguistic theory. In Chapter 5 I discussed at length the descriptions of Thai and languages in this region of the world by Western trained linguists in terms of what the languages lacked. I would like to end with an example that bears on this issue.

Wittgenstein famously remarked, “Es gibt nur die Beispiele” (There are only examples). I understand this to mean that the object/data/referent under study has a greater truth claim than an analysis or interpretation of that example. In other words, postulated underlying principles are suspect and prone to be personal and unexaminable. They are further from “reality” than “some-thing” that can be observed equally by all interested observers. I offer as a final example, a word for word glossing of a Thai book title recently shown to me.

(60) Sentence #20

\[ \text{tāaj lēw paj naj} \]

die already go where

One possible English free translation might be, “Where do we go when we die” or perhaps there are better interpretations. For English speakers, the above gloss of just the four English words seems cryptic and open to different interpretations. We think we need to add referents for the verbs to make the phrase or clause understandable. Linguists with firm convictions about language structure, and this seems to be the huge majority of linguists, seem inclined to confidently leap into an analysis using abstract words that have theoretical underpinnings in present day linguistic theory. Let us try to examine this simple specific example in detail with open minds and see how translation and theory are intertwined.

For an English speaker it does seem that there are words missing. But are the words missing or is it that the English glosses of the Thai words have allowed the easy inference that the Thai words do have just these simple meanings shown by their English glosses. In other words, is the act of glossing in English hiding something? In my view the Thai words have differently delimited meaning per word than do English words. They have extremely general semantic senses as my quote from Grierson in Chapter 5 illustrates concerning Ahom. This fact about Thai is what confuses, frustrates and bewilders so many who attempt to learn the language. It is the opposite “complaint” of Thais who are frustrated with English and continually wonder why a
language needs so many extra words that do not need to be there. The genius of Thai is that often less can be more. I have to learn this lesson often when my translations of English sentences into Thai have baffled my listeners.

Knowing what one does not have to say to be understood by others, can be a most difficult lesson to learn. English speakers believe that these unsaid words simply must be there in the language- in some “deeper structure”. Just as linguists all believe that “main verbs” have to be there in all clauses since the form of a clause is modeled on the form of a logical proposition which clauses are assumed to be.

My experience with translating Thai into English is that verbs are often better thought of as gerunds. Perhaps “dying, then going where” is a better English translation of the book’s title. Linguists might say, this is a translation issue not a matter of theoretical linguistics but I maintain that it is a matter of utmost importance in the methodology of theoretical linguistics. How can implicit assumptions be challenged and brought out into the open if linguistic analysis is carried out in one language only? If linguists draw all their conclusions from English glosses, their perspective is seriously skewed and the almost ideological conviction that we all speak basically the same language becomes a truism. This issue may help explain the Gordian Knot of serial verb description discussed in detail in Chapter 5.

I believe that we will make progress in understanding why language is the way it is only by studying language as it is. Our methodology must acknowledge the limitations of our imaginations, our biases, and our proven tendencies as human beings to seek to maintain unexamined presumptions that serve to support our current notions of truth. I think the field of theoretical linguistics needs to embrace a more humble and appropriate attitude toward the still enormous challenge of natural language study where mysteries abound and where far more remains to be discovered than is currently understood.
Bibliography


Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. LSA Presidential Address, delivered in Atlanta, January 2003.


Steedman, Mark. 2011. Romantics and Revolutionaries: What Theoretical and Computational Linguists need to Know about Each Other* (*But were Afraid to Ask Linguistic Issues in Language Technology, 6 unpaged.


RESUME

Name: Terry Adams

Place of Birth: USA

Institutions Attended: University of South Florida, BA 1970

Built 37ft Steel Sailboat, cruised aboard 1975-1980
AP Physics teacher 1981-2004