CHAPTER THREE

METHODOLOGY

3.0 INTRODUCTION

The methodology used to conduct this research is a grammatical assessment applied to sentences to test the degree of certainty, focus, perspective and time along the lines of a framework of a previous research. The grammatical test consists of four dimensions which test the speech act of five speakers, in an interview, for their commitment to the truth proposition using modality which is a linguistic device.

3.1 THE ANALYSIS FRAMEWORK

There will be two types of analysis; the first one is a linguistic analysis conducted manually using a Categorization Model Framework and the other is a frequency count output of epistemic modals and expressions. The research will adapt the framework of three authors Rubin, Liddy and Kando (2006) in their article Computing Attitude and Affect .

(http://www.cnlp.org/publications/04rubinkandoliddy2004_certainty.pdf). Their research was based on identifying explicit certainty information in 32 newspaper articles. This report is a qualitative and quantitative analysis examining the sentences as truth propositions uttered with conviction by the speakers and it will use the categorization model and analytical framework for certainty identification.

3.2 DATA COLLECTION PROCEDURES

The data is based on only one television interview, selected at random focusing on discussions about The War in Iraq from the archives of MSNBC. NBC is a broadcasting station in Washington D.C. in The United States of America which airs

news, interviews and opinions. The television interview in discussion is a regular program which has the late Tim Russert as the anchorman interviewing five American politicians and asking them their views of the current affairs of the country.

The five speakers are listed below:

- Senator Bob Graham, a member of the Democrat and a Senator from the state of Florida. He was a presidential candidate for the 2004 US elections.
- James Carville, a Democrat strategist who is a political contributor to CNN as well as a political consultant.
- Newt Gingrich, a Former House Speaker of the Republican Party.
- Patrick Buchanan, presidential candidate for the 1992, 1996, and 2000 US elections as well as a White House speech writer and political advisor.
- Mary Jo Matalin, political consultant and advisor to President George W.Bush. (wife of James Carville listed above.)

This research will only concentrate on *The War in Iraq*, considered here as the variable and the use of modality in the utterances of the speakers. The anchorman's questions for the show will not be included in the analysis. The interview was conducted on September 5, 2004 and broadcasted on September 5 which was a Sunday. The programme was aired in Malaysia on CNBC between 5am and 6am early Monday morning. The transcript is available from the archives of MSNBC. Several other current issues were discussed but the sections discussing war was extracted from the transcript to be analysed in this research.

A word count was done on each speaker discussing the topic, *War in Iraq*, is listed below:

- Sen.Graham -852 words
- James Carville-1500 words
- Newt Gingrich -1300 words

- Patrick Buchanan- 1160 words
- Mary Jo Matalin-850 words

The research is carried out by comparing the distribution of modal verbs as well as by using the computer to extract phrases and sentences which are epistemic in nature. Compleat Lexical Tutor is the software which will be used to extract the typological modals while manual grammatical analysis will be used to identify epistemic expressions. The created corpora will reflect the type of language used, and here, only a grammatical analysis of modals and modalized statements expressing truth propositions will be carried out. Lexical predicates which behave as modality devices throughout the interview will also be identified and analysed. The main transcript of the interview is analyzed and a selection of excerpts is done according to each speaker respectively (Refer to Appendix 2 3,4,5,6 and 7).

The interview is separated into five separate discourse texts, each being an individual compilation of excerpts of the five speakers and each is analyzed separately for the presence of linguistic devices discussed in the research (Refer to Appendix 3.4.5.6 and 7). The presence of epistemic modals and lexical predicates operating as epistemic modals are highlighted in a different color from the main text for each speaker (Refer to Appendix 3 to Appendix 7). The number of modals and lexical predicates will be tabulated after a concordance is done (Refer to Appendix 10). The hypothesis will be tested with the results obtained. The researcher then processes the text file for epistemic expressions and their extent of truth or certainty levels.

3.3. DATA PROCESSING CRITERION 1-4

The manual grammatical analysis of the data will be at the sentence level. The analysis covers four areas which are certainty levels, perspective or view, certainty focus as well as time. This research will adapt the framework of a research done by three authors, Rubin, Liddy and Kando (2006) who designed the Categorization Model by that name. A review of the analysis method is presented in Chapter Two. Each time a modalized statement is analyzed the questions to be answered are as below:

1. What is the certainty level? Is it absolute, high, moderate or low?

2. Whose perspective is being presented? Is it the speaker's view or some professional's view?

3. What are the details about? Factual information, opinions, emotions , judgments, or hypothetical situations.

4. What time frame is indicated? Is it in the past, the present or the future?

According to the Categorization Model, an analysis is only done when the sentence is a marked sentence. A marked sentence is a sentence which differs from an unmarked sentence by the presence of a modal or lexical predicate acting as a modal. The focus of this research will be modalized sentences so unmarked sentences will not be analyzed. The total number of sentences in the transcript is numbered from 1 until 357, with the marked sentences highlighted with a different colour (Refer to Appendix 2 for the marked sentences).

3.3.1 DEFINITIONS OF EACH CRITERION.

There are four criteria by which a sentence is analysed for i) its certainty level ii) the focus of the certainty iii) the perspective and iv) time

A sentence is only certain when it only identifies concrete facts, events and states which are considered factual information. Opinions, judgments, attitudes,

beliefs, emotions, assessments, and predictions are considered abstract information according to the Categorization Model of Certainty Identification. This criterion is called *certainty focus*. *Certainty levels* are divided into four groups which are absolute certainty, high certainty, moderate certainty and low certainty. This correlates to the strength of the linguistic devices being investigated which in the research is ***modality***. For an understanding of the precise meanings of a few lexical predicates used in an epistemic context, their dictionary definitions have been used. In Chapter Two, several tables of taxonomies classifying levels of certainty according to researches on modality will be referred to in the grammatical analysis in Chapter Four. This will help to determine which level of certainty the sentence belongs to.

Perspective is another criterion of this model, which implies that the notions in the sentences are analyzed according to whom the text was written or spoken by. The "Perspective" section in the table analyses the information as to whether it is the speaker or writer's personal view or the view of a reliable source, for example government experts' report on issues or professional's comments based on findings.

To further explain all the above criteria, let's look at an analogy below. An example is used here for further clarification. Results on a Richter Scale stating the strength of an earthquake is considered as a reliable source because it is empirical. Seismic activity noted and analyzed and commented on is classified as a professional view when it is quoted by a seismologist or a volcanologist. This section also indicates involvement of individuals or groups either as victims, witnesses or as experts and authority. The comments based on the experiences of the victims of an earthquake are considered involvement of individuals or groups.

Being subjective about some issue refers to whether emotions are related to the opinion or views. Dictionaries define subjectivity as the product of the mind or a part particular state of mind, caused by the feelings, or temperament of the person's

thinking. A subjective utterance is classified as an expression which is personal and opinionated and does not need proof of occurrence while an objective expression is based on evidence and facts.

The opposite of a subjective opinion is an objective opinion which is not influenced by personal feelings, interpretations or prejudice. It is based on facts. For example, Science is objective and objective judgments based on Science are absolutely true as the results are empirical in nature. So the third section of the analysis which is *perspective* is based on these definitions.

The last section of the analysis is the *Time* factor which takes into account the grammatical temporal element of linguistics which is the past tense, present tense or the future tense. The past and the future tenses are considered as non factual tenses while the present tense is considered to be factual (Palmer 1990).The conditional IF makes the statement unreal. However, past actions which have been documented are factual because they have occurred. It is important to note that when epistemic devices within sentences start with the IF conditional then the whole sentence becomes an abstract sentence.

After the manual analysis of sentences using the Certainty Categorization Framework, a frequency count is done to find out what are the most common epistemic expressions, in the discourse text of the four American politicians.

3.4. LANGUAGE SOFTWARE AND THE RECORDING METHOD.

Compleat Lexical Tutor, a language software will be used to identify the linguistic tools used as modality throughout the interview. Each person's discourse will be treated separately to pick out the language devices used as research material in the study. Each speaker will have his own series of analysis tables presented and analysed sentence by sentence.

The sentence to be analysed will be on top of the text box while the criteria will occupy the text box created and the analysis will be recorded in detail in the text box. Refer to Table 3.1.Cross referencing of the epistemic content of the utterance is done with details in Chapter Two. A conclusion will be derived each time an utterance is made based on its certainty level and this will be done for every speaker. The analysis will be written inside the table. A Frequency count will be done on the number of auxiliary modals and lexical predicates behaving as modals used throughout the discourse text of the five speakers. Refer to Table 3.2 and Table 3.3. The details are summed up in Chapter Five.

Four-Dimensional Certainty Categorization Model

Sentence with the epistemic expression

A Detailed Linguistic Analysis

Certainty Levels	Analysis
Perspective	Analysis
Certainty Focus	Analysis
Time	Analysis

Table 3.1: Linguistic Analysis of Epistemic Expressions

Frequency Count Output:

An example of a tabulated analysis of modal auxiliaries

Modal	Frequency
Can	
Could	
Will	
Would	
Shall	
Should	
May	
Might	
Must	

Table 3.2 Frequency count output of auxiliary modals

More examples of epistemic expressions

Epistemic Modals and Expressions	Frequency	Percentage
Modal Auxiliaries		
Modal Lexical predicates		
Modal Adjectives		
Modal Adverbials		
Multiword Epistemic Expressions		
Semi- modal auxiliaries		

Table 3.3: Frequency Count of lexical predicates

3.5   STAGES IN DATA ANALYSIS

Firstly in Step 1**,** the modalized statements are extracted from the concordance text and the number of auxiliary modals and lexical predicates acting as modals are counted. In Step 2**,** a framework is decided in analyzing the epistemic devices. The researcher will use the Four-Dimensional Certainty Categorization Model to analyze each sentence in the interview. This is a framework which the researcher

will use as a guide to analyze the article. (The framework is presented as in Table 3.1, Table 3.2 and Table 3.3 under heading 3.4 on page 46). Step 3 then proceeds with an analysis involving the four dimensions which are certainty levels, perspective or stance, the focus of the epistemic propositions and the time which is represented with the use of the past, present and future tense. The analysis is done at the sentence level. In Step 4 the analysis is tested and compared with the hypothesis. When the sentences are linked cataphorically or anaphorically, it is noted, however paragraph level analysis involving relationships between sentences is not the focus of this research. This research applies Ferdinand de Saussure's Sign Based Theory principles to investigate the subjectivity of a sentence and the sign being that of a modality notion. In Step 5 the frequency of the criterion in the linguistic test by each speaker is tabulated and interpreted. In Step 6 a conclusion is made based on the analysis.

SUMMARY

Chapter Three is about the method used to extract information and how the results are tabulated. It is a systematic technique of logical thinking and it involves four phases. It is mainly about links between the spoken text and the speaker as well as the source of information. It is about categorizing and creating a divide between what is abstract like an opinion and what is a fact so that an analysis can be made once the focus of the utterance as a truth propositions is decided.

Firstly, the sentences to be analysed which are refered to as marked sentences are placed above each table and the criteria for analysis and the findings are recorded inside the table. Each marked sentence is analysed according to its certainty level depending on the modality used. The certainty levels are along a continuum from low to absolute or high certainty. Next, the perspective or actual source of the

utterance is identified by asking the question "whose view is it" which then leads to the study of objectivity or subjectivity of the utterance and finally the time factor of the utterance which is considered by looking for parts of grammar and their clues within the sentences indicating the actual event taking place somewhere in the past, present or future. The analysis is only done on marked sentences which means, sentences which carry epistemic signs or notions are separated from the rest of the text for analysis from the very beginning. Meanwhile unmarked sentences are not analysed at all because they do not have epistemic notions within them. Then this is followed by a frequency count which is done on the modals and lexical predicates used in the discourse texts of the five speakers and they are recorded in a frequency table. Cross referencing is done on sentences identified for analysis with the literature review in Chapter Two and criteria definitions in Chapter Three to ascertain the level of truth in each expression in the discourses. A language computer software known as Compleat Lexical tutor will be used to extract marked sentences with the notions and signs mentioned for analysis as it will help in the frequency count. The methodology outlined in this chapter will be used as a guideline to the analysis of data in Chapter 4. The chapter uses the framework created by three language researchers, Rubin, Liddy and Kando (2006) whose project was analyzing certainty and the extent of truth reported in newspaper articles. Their research framework and its four criteria have been adapted in this research and it studies instead truth propositions of speakers in a life television interview.