



3 1293 01405 3379

This is to certify that the

dissertation entitled

TESTING SOCIOLINGUISTIC PROFICIENCY
WITH A MULTIPLE CHOICE TEST

presented by

Hisham Dzakaria

has been accepted towards fulfillment
of the requirements for

M.A. degree in Linguistics

Major professor

Date May 2, 1996

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

| DATE DUE | DATE DUE | DATE DUE |
|-----------------|-----------------|-----------------|
| MAY 18 2008 | _____ | _____ |
| _____ | _____ | _____ |
| _____ | _____ | _____ |
| _____ | _____ | _____ |
| _____ | _____ | _____ |
| _____ | _____ | _____ |
| _____ | _____ | _____ |

Testing Sociolinguistic Proficiency with a
Multiple-Choice Test (MCT)

by

HISHAM BIN DZAKIRIA

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the Degree of

MASTERS OF ARTS

Department of Linguistics

1996

ABSTRACT

This study is an attempt to validate the feasibility of a multiple-choice test (MCT) as a possible testing tool to test sociolinguistic proficiency, one aspect of communicative competence. Established techniques of test development and evaluation was used to develop a multiple-choice proficiency test of sociolinguistic knowledge - a desirable component of overall language ability. The results of the study show that such a test has a low but a usable degree of feasibility. Further research needs to be carried out to improve the reliability of such test.

For Ella, Amelia & Afiq

ACKNOWLEDGEMENTS

I wish to acknowledge the assistance and encouragement rendered to me by the faculty members of the Department of Linguistics at Michigan State University. Most importantly I would like to express my gratitude to my advisor Dr. Preston. His guidance during my studies here was always balanced; never too pushy but always willing to answer any questions I might have had. His great experience in sociolinguistics often allowed him to steer me away from dead ends. Certainly without his patience and expert direction, this thesis would not have been completed.

The other person I owe a great deal to is Dr. Susan Gass. Her immense knowledge of second language acquisition and interlanguage pragmatics has inspired me to learn more and pursue the accomplishment of this thesis.

The other members of my committee: Dr. Barbara Abbott and Dr. Grover Hudson also deserve a great deal of credit for providing support and pleasant working environment and always an open ear for discussions.

My appreciation also goes to the Northern University of Malaysia (UUM) for the financial support, and encouragement throughout the duration of this thesis.

Last, but certainly not least, I would like to thank my wife, Azilah Kasim for her patience and willingness to sacrifice her time and preferences in order to allow me to complete this thesis despite of her own busy schedule in completing her MSc. Through the years, she and my two wonderful kids Amelia Hanah and Afiq Danial have been my sources of motivation.

Also, I am greatly indebted to my parents although not being geographically close, for providing consistent encouragement and assurance that I could, and would complete the required tasks successfully.

All of the aforementioned can rightfully claim a share in the success of this thesis. However, any shortcomings or defects I willingly admit to being my own.

Contents

| | |
|----------------|------|
| LIST OF TABLES | viii |
|----------------|------|

| | |
|-----------------|------|
| LIST OF FIGURES | viii |
|-----------------|------|

| | | |
|---|---|----|
| 1 | STUDY INTRODUCTION..... | 1 |
| | 1.1 Research Proposal and Its background..... | 1 |
| | 1.2 Hypothesis..... | 2 |
| | 1.3 Terminology..... | 3 |
| | 1.3.a Communicative Competence..... | 3 |
| | 1.3.b Sociolinguistic Competence..... | 4 |
| | 1.3.c Linguistic Competence..... | 5 |
| | 1.3.d Discourse Competence..... | 5 |
| | 1.3.e Strategic Competence..... | 5 |
| | 1.3.f Speech Acts..... | 5 |
| | 1.4 Brief Details of the Following Chapters..... | 6 |
| 2 | LITERATURE REVIEW..... | 8 |
| | 2.1 General View of Communication..... | 8 |
| | 2.2 Dilemmas Confronting Second Language Learners..... | 10 |
| | 2.2.1 Performing Speech Acts:Various Plausible Forms (Formulas)..... | 17 |
| | 2.3 Language Testing:An Introduction..... | 18 |
| | 2.4 Issues in Communicative Testing..... | 20 |
| | 2.4.1 Testing and Communicative Competence..... | 20 |
| | 2.4.2 The Nature of Proficiency Tests..... | 21 |
| | 2.4.3 Methodological Issues..... | 23 |
| | 2.4.3.a Ethnographic Observation..... | 24 |
| | 2.4.3.b Role Play..... | 24 |
| | 2.4.3.c Acceptability Ratings..... | 25 |
| | 2.4.3.d Written Completion Task..... | 25 |
| | 2.4.3.e Multiple-choice Test..... | 26 |
| | 2.4.4 Norm Referenced Versus Criterion Referenced..... | 27 |

| | | |
|---------|--|-----------|
| 2.4.5 | The Notion of Authenticity..... | 28 |
| 2.5 | Previous Studies of Tests on Communication..... | 29 |
| 2.6 | Previous studies' Design and Methodology..... | 34 |
| 2.7 | Studies on Speech Acts:Revisited..... | 39 |
| 2.7.1 | The Speech Act of Apology..... | 39 |
| 2.7.2 | The Speech Act of Replying to Compliments.... | 41 |
| 2.8 | Theoretical and Practical Reasons for Choosing Speech Acts: Apology and Replying to Compliments for this study... .. | 42 |
| 2.8.1 | Apology and Its Rational..... | 42 |
| 2.8.2 | Compliments and Its Rational..... | 44 |
| 2.9 | Rational(s) for Measuring Sociolinguistic Competence With a Written Mode (MCT)..... | 45 |
| 2.9.1 | Scalar Behavior Versus Dichotomous Behavior.... | 50 |
| | | |
| 3 | METHODOLOGY..... | 52 |
| 3.1 | Purpose and Background:Revisited..... | 52 |
| 3.2 | Instrument..... | 53 |
| 3.2.1 | Development of the Instrument..... | 53 |
| 3.2.1.a | Instrument No.1 (Open-Ended Questions)..... | 54 |
| 3.2.1.b | Instrument No.2 (Multiple Choice Test)..... | 55 |
| 3.3. | Subjects..... | 56 |
| 3.3.a | Native Speakers..... | 57 |
| 3.3.b | Non-Native Speakers..... | 57 |
| | | |
| 4 | RESULTS..... | 59 |
| 4.1 | Performance of native Speakers..... | 59 |
| 4.2 | Performance of Non-Native Speakers..... | 59 |
| 4.3 | factors Affecting Variance Within the Non-Native Speakers..... | 60 |
| 4.3.1 | TOEFL Scores..... | 60 |
| 4.3.2 | Length of Stay..... | 61 |
| 4.4 | Native & Non Native Speakers:A Comparison..... | 62 |
| 4.5 | NNS Performance: Revisited-An Interpretation..... | 63 |
| | | |
| 5 | DISCUSSION & CONCLUSION..... | 67 |
| 5.1 | Validity..... | 67 |
| 5.2 | Reliability..... | 69 |
| 5.3 | Practicality..... | 71 |
| 5.4 | Evaluation..... | 73 |
| 5.5 | Suggestions..... | 74 |

| | |
|---------------------------|----|
| REFERENCES MATERIALS..... | 77 |
| The References..... | 77 |
| Appendices..... | 81 |
| Appendix A..... | 81 |
| Appendix B..... | 82 |
| Appendix C..... | 84 |

LIST OF TABLES

| | |
|--------------------------------------|----|
| 4.1 Findings Between NS and NNS..... | 60 |
| 4.2 TOEFL Scores..... | 61 |
| 4.3 Length of Stay..... | 61 |

LIST OF FIGURES

| | |
|---------------------------------|----|
| 4.1 NS and NNS Performance..... | 63 |
|---------------------------------|----|

Chapter One

Introduction

This study is an attempt to validate the feasibility of a multiple-choice test (MCT) as a possible testing tool to test sociolinguistic proficiency, one aspect of communicative competence. The study includes the stages of defining a body of knowledge or content, determining what native speakers norms are for knowledge of that content, and testing the collection of items on a representative sample of native speaker subjects and non-native speaker subjects to determine if such a test can distinguish between the two populations. The present study attempts to use established (see Appendix A flow chart describing the proposed research) techniques of test development and evaluation to develop a multiple-choice proficiency test of sociolinguistic knowledge. Through the identification of native speaker norms, and the development of a sufficient number of specific, multiple choice items testing knowledge of those norms, the study will determine whether such a procedure will allow us to distinguish non-native speakers' (NNS) knowledge from native speakers' (NS) knowledge, and to rank the knowledge of non-native speakers relative to that of native speakers, in a reliable way.

1.1 Research Proposal and Its Background

The major study on which this research is based is the study by Cohen and Olshtain (1981) on the speech act of apology (details of this study are elaborated in chapter two extensively). This study by Cohen and Olshtain was carried out in response to search for more empirical data on pragmatics in second language testing. This study can also be seen as

one of the most important studies which intended to create a tool to assess sociolinguistic competence in a second language. Their study set out to determine whether a "rating scale" for assessing sociocultural competence could be developed. They conclude that they have produced a "crude measure" of such competence. They feel that their testing instrument has certain inherent strengths. For example, the fact that the study was based on a contrastive analysis of sociocultural patterns, offers them a certain measure of "explanatory power" in dealing with the findings. By analyzing and then testing for a given speech act in terms of its semantic formulas, they were able to account for the specific formulas that NNS included in their responses (the formulas are dealt with in the literature review section).

Since Cohen and Olshtain (1981) have claimed that their design which uses a 'rating scale' is a possible tool to assess sociolinguistic competence, the present study is my attempt to create and validate a multiple choice test to assess sociolinguistic competence in English for second language speakers (specifically Malay learners of English). The test will focus on two different speech acts: *apologizing* and *replying to compliments* in English.

1.2 Hypotheses

It is the hypothesis of this research that a multiple choice test can be a possible tool to assess scalar behaviors such as sociolinguistic competence-one aspect of communicative competence. That is to say that the norms can be determined by the administration of a multiple choice test, and that the test is a feasible tool. A second hypothesis is that non-native speakers perform according to their levels

of proficiency. That is, lower proficiency subjects will score lower than those with higher proficiency levels (proficiency is to be measured by the subjects' TOEFL scores and/or placement test administered by the MSU English Language Center, and length of stay in the US). Meanwhile the native speakers will perform well on the multiple choice test, and thus establish the norms. To test the hypothesis, the multiple choice test is given to native speakers and non-native speakers, and their performances are compared.

1.3 Terminology

It would be useful before exploring previous research and studies in testing sociolinguistics competence and speech acts in the following chapter to review the major terms and concepts which are included in this study. However, this is only a brief review of each term for the purpose of understanding the succeeding chapters.

1.3.a *Communicative Competence*

Lets first look at a more general term - *communicative competence* which can be seen as encompassing other major terms (sociolinguistic competence, speech acts, etc) described in this research. Certainly, questions such as "What must be said, to whom, with what tone of voice and how the talk (or silence) of others is to be taken are some of the 'communicative competence' aspects of language ability" (Preston,1989:10). Thus, how is communicative competence best defined? Enormous amount of research and literature have been written on communicative competence, e.g. Hymes (1964); Davies,(1989), Spolsky (1989); etc. Nevertheless, the definition(s) given by different

researchers seems to be compatible with that of Stalker(1989: 182). His definition of the terms is:

... that part of our language knowledge which enables us to choose the communicative system we wish to use, and, when that selected system is language, to connect the goals and contexts of the situation with the structures which we have available in our linguistic repertoire through functional choices at the pragmatic level. In making this selections, language users accommodate linguistic features both consciously and unconsciously in order to adjust the social distance between the producer and the receiver.

The term encompasses both the speaker's knowledge of the linguistic rules as well as the sociocultural rules for appropriate use. Another version of the definition which is more basic in nature to communication is by Canale and Swain (1980). Their understanding of communicative competence is as "the underlying system of knowledge and skills required for communication." They further divide the term into four sub-competences: Linguistic competence, sociolinguistic competence, discourse competence and strategic competence.

1.3.b Sociolinguistic Competence

Sociolinguistic competence which is of primary interest to this research can be seen as know-how to use the knowledge of the target language in communicative situations. Specifically, the aspect of sociolinguistics that this research is concerned with here is the ability to use appropriate sociocultural rules of speaking, that is the ability to communicate in a culturally acceptable way in a context, and to use linguistically appropriate forms of that context (Hymes,1974, Canale and Swain,1980). The term sociolinguistic competence can be traced back to the above term 'communicative competence', which was introduced by Hymes (1964). This competence is furthermore dependent on contextual factors

such as status of participants, purpose of interaction, and norms of or conventions of interaction (Canale and Swain,1983).

1.3.c *Linguistic Competence*

Linguistic competence, also known as grammatical competence concerns the knowledge of the language code and system. It includes features and rules such as vocabulary, word formation, sentence formation, pronunciation, spelling, and some information on semantics. It focuses on the skills and knowledge required to understand and express the literal meaning of utterances accurately.

1.3.d *Discourse Competence*

Discourse competence, on the other hand, concerns the mastery of how to combine grammatical forms and meanings to achieve a unified spoken or written text, such as oral or written narrative, argumentative essay, business letter, speech and so forth.

1.3.e *Strategic Competence*

This term concern the mastery of verbal and non-verbal communication strategies which are used to compensate for breakdowns in communication and to enhance the effectiveness of communication.

1.3.f *Speech Acts*

The development of the theory on speech acts by Austin (1962), and then revised by Searle (1965,1969,) and other research has helped us to understand what a speaker needs to know in order to perform effectively and appropriately (in form and meaning) in communication. Speech acts,

the second most common term used in this research have been claimed to include speech situation, and speech event (Hymes,1972). This certainly has helped in providing a framework for researchers to study communicative competence. According to Searle (1969), speech acts, the minimal unit of discourse, are acts that we perform in our everyday communication and interaction. Much of the literature in interlanguage pragmatics has been conducted within the framework of speech acts. Speech acts can be defined in terms of discourse function. Thus, the minimal unit of communication is the performance of a linguistic act. A speech act can occur within various speech events. For example, one can compliment or apologize as part of a lecture or a conversation. However, different forms of utterances are required for different contexts. Most importantly, one needs to realize as posited by Gass and Selinker (1994: 183) that:

All languages have a means of performing speech acts and presumably speech acts themselves are universal, yet the 'form' used in specific speech acts varies from culture to culture. Thus, the study of second language speech acts is concerned with the linguistic possibilities available in languages for speech act realization and the effect of cross-cultural differences on second language performance and on interpretation by native speakers of second language speech acts.

Examples of speech acts include giving compliments, making statements, apologizing, making introductions, making requests, expressing gratitude, making refusals, etc.

1.4 Brief Details of the Following Chapters

Chapter two is intended to provide information on previous work done in the area of speech acts realization in interlanguage pragmatics, as well as informs readers about central issues pertaining to testing

and methodology. One obvious objective in chapter two is to exhibit how previous work on the area of sociolinguistic competence have greatly influenced this study as it pursues to testing sociolinguistic proficiency using a written mode 'multiple-choice test'.

Chapter three of this thesis is a section elaborating the present study with details and information on its methodology, and objectives. Detailed information on the study's instrumentation and the subjects are also included.

Chapter four is a chapter where the results of the two instrumentations employed in this study (Open-ended, and the multiple choice test) are revealed and analyzed.

Finally, chapter five will discuss the results obtained in this study. Most important of all, this chapter will discuss the results as to whether multiple-choice test is in fact a good test to test sociolinguistic competence or, vice-versa. Implication of this study and its limitation will also be a part of the final chapter.

Chapter Two

Literature Review

This chapter is a description and discussion of the methodologies and findings of previous studies conducted to create a tool for testing communicative competence in general, and specifically on sociolinguistic competence. The initial section of this chapter looks at communication in general and the dilemmas confronting second language learners when performing speech acts. The second section of this chapter introduces different testing methods used over the years by researchers, and presents some issues in testing communicative competence or performance in general. Next are the descriptions of previous studies on the speech acts that this study is most interested to investigate, namely the speech acts of *apology* and *replying to compliments*. This chapter ends by revealing the rationale for choosing a written-mode test (multiple-choice test) for testing sociolinguistic competence.

2.1 General View of Communication

Communication in general is a very complex process. Communication according to Wahlstrom (1992) is an 'ecosystem' composed of different interdependent systems. This includes culture, social institutions and language. Thus, because of its ecological nature, "figuring out how communication works means taking into account a host of variables. Communication is highly contextual, and without considering the other elements of the ecosystem, getting at the meaning of messages is impossible" (p.106). It is very much culturally related. Ways of communicating in talk according to Tannen (1987),

"are learned in the speech community, that is by talking to people with whom one identifies socially". Tannen further posits that social networks are always local, and not global. Thus, people in different cultures and communities have different ways of using linguistic means to communicative ends, and their ways of talking, like other cultural patterns, define them as 'community'. This can visibly be seen in a multicultural nation like Malaysia, Singapore, and many others including America which are composed of different ethnic groups, backgrounds, beliefs, cultures, religions, etc. In Malaysia, for example, the three major ethnic groups, Malays, Chinese and Indians are distinct from each other almost in every aspect of each culture including 'communication'. Therefore, an understanding of each community and culture is vital in order for one to have an effective and productive communication among the different groups.

Due to the complex nature of communicative aspects of language use and the poor state of our present knowledge of language use, it would appear that developing a measure of sociolinguistic competence in a second language is not an easy task. This would certainly explain why such measures have not been readily available. Nevertheless, there have been some research efforts and attempts to develop systematic means of measuring such competence. However, one should beware that the whole issue of measuring sociolinguistic competence is not yet conclusive. However, some of the researchers already involved in this area are:

Research done by Joel Walters (1979), Patricia Carrell (1981), Andrew Cohen and Elite Olshtain (1981;1983;1993), Nessa Wolfson (1980), Houck and Gass (in Gass & Neu,1995), Boxer (in Gass & Neu,1995) and others. Some of these studies will be described in this chapter.

2.2 Dilemmas confronting Second Language Learners

Due to the cultural differences in the use of language for communication, misunderstandings and misconceptions often occur when ESL speakers communicate with native speakers of English. The ways people use language to communicate can differ radically from society to society or from one culture to another. A knowledge of some of these cultural differences in the use of language will enable the language learner to avoid misunderstandings and enhance the effectiveness of communicating between interlocutors.

One should beware that almost everything in 'communication' is culturally related - from knowing what to say and how to say it, to other language features such as intonation, coherence, indirectness, etc. The issue of 'indirectness' for instance, can be one clear example to show differences between particularly Asian people (Non-native speakers (NNS) of English) and Westerners (i.e native speakers of English). To most Asians like the Chinese, Malays, Indians, Japanese, Koreans, etc. only a part of meaning resides in the words spoken. The largest part is communicated through different media and channels i.e by hints, assumptions, experience, etc. Native English speakers

like the British and Americans tend to ignore 'indirectness'. Their utterances of ideas, and words (both spoken and written) are concise and direct. They (native speakers of English) according to Tannen (1987) "...believe that words should say what they mean and people should be accountable only for what they say in words". Such perceptions and attitudes often put native speakers of English (i.e international businessmen) in an awkward position when dealing with Japanese, Mediterranean, etc. "for whom elaborate 'small talk' is big, essential, furnishing the foundation for any business dealings" (Tannen,1987). Quite often, business deals, the theme of a meeting, the main agenda; etc. are preceded by 'many' small talks or other issues which might or might not be relevant to the main issues at hand. Understanding and sensitivity to such behavior manifested by NNS of English is vital in achieving one's goal and objective in communication.

The fact is that, when entering into a different community with different language, culture, and background, one will face new rules, values, beliefs, etc. Similarly, non native speakers of English will encounter different and strange rules when they first communicate in English. Wolfson and Manes (1980) pointed out that a single speech act may vary greatly across cultures or speech communities. For example, an ESL speaker might offend a native speaker of English in making a request for assistance or a favor. He or she might use a wrong word or phrase. Perhaps this is due to the fact that the NNS of English is not well aware of the

felicity conditions that are necessary for a particular performance to occur. Thus, as a result the native speaker might take him or her as impolite or rude in making a request. The following examples illustrate a possible situation described above.

- 1.a **Can you help me with my books?** instead of,
- b **Could you help me with my books?**
2. **Can you pass me the salt?**

Example in 1. **Can you help me with my books?** instead of, **Could you help me with my books?**, is not a serious offense, but it is preferred and polite according to native speakers of English if "could" were used. In addition to the distinction of 'polite' versus 'impolite', there is also the distinction between 'formal' versus 'informal' as depicted in Preston (1989:4 adopted from Coffey, 1983:108; and Marquez & Bowen, 1983:72) which literally adds to confusions and problems for NNS. However, if the offense is serious, the native speaker might feel insulted, and a communication breakdown could occur. As briefly mentioned above, native speakers of English consider "could" more polite than "can". However, in this case, the ESL speaker may not know this fact because his lack of proficiency in English, or simply not remembering the rule when needed. In addition to this, Munn (1995) has stated that modals such as *can-could*, *will-would*, etc. do not 'clearly' attach to 'tense' in English. It is difficult for native speakers of English to differentiate and distinguish which of the two modals represent the 'present' tense or 'past' tense. It is even more difficult

for NNS to perform such an act. This too would contribute to another problem for NNS (although a different one) as to which of the two modals can be used and identified as the 'present' or 'past'.

The fact remains that speech acts are difficult and complex. Hence in example 1 & 2 above a NNS might ask whether this is a QUESTION or a REQUEST? What is the intent? Conversation is not structured and meaning is not derived from the locutionary but the deeper level, namely the illocutionary act. Thus, NNS has to be very attentive and careful in processing what they hear. They (NNS) have to look at the word order, punctuation, intonation, etc.. Searle (1969) has described their function as indicators, which can help to determine what or which illocutionary act is being performed. Failure to determine the 'intent' or illocutionary act performed may result in a communication breakdown.

In another example, an ESL speaker might be taken as rude when he or she does not open the window when a native speaker of English in the same room says that the room is warm, meaning to ask the addressee to do him a favor (open the window). In this case there is a clear cut case of communication breakdown between the addresser(NS) and the addressee (NNS).

Thus, clearly, it is easy to imagine how miscommunication and misunderstandings occur if the form of a speech act differs from culture to culture, or from one community to the next. Lets consider the response produced by a native speaker of Hebrew adopted from Cohen and Olshtain

(1993:54):

You promised to return a textbook to your classmate within a a day or two, after xeroxing a chapter. You held onto it for almost two weeks.

Classmate: *I'm really upset about the book because I needed it to prepare for last week's class.*

Response: *I have nothing to say.*

Without a doubt, the response such as the above sounds rude and impolite. It also suggests a lack of commitment or willingness to apologize. However, interestingly, what was meant was the translation of something like 'I have no excuses'. Thus, having different background and culture, one needs to realize that without shared knowledge and background, linguistic system, and values, Varonis and Gass (1985a.,1985b.) state that "when one interlocutor confidently [but inaccurately] interprets another's utterance, it is likely that participants will run into immediate problems because they do not share a common discourse space" (p.341).

Many more distinct differences can be observed among interlocutors from different background and cultures performing different speech acts. For example,with regards to the speech act of apology, many differences have been found between English and Hebrew (Wolfson and Manes,1980) in addition to the example adopted from Cohen and Olshtain (1993). They found that the Hebrew speakers treat apology differently than the English speakers. They suggested that there are great differences between English and other languages as well. An ESL speaker might not apologize appropriately in English when he or she does not know the

rule of apology in English. He or she might not apologize at all in certain offenses because he or she would not do so in the native language and culture. For example, it is a common practice in Malay culture to allow an additional thirty minutes (30) when meeting someone for an appointment, a meeting, luncheon, etc. Thus, if a Malay person 'A' arrives thirty minutes late for an appointment, he or she does not feel obligated or committed to apologize. Similarly, the other party or person 'B' should not be offended if he or she does not convey an apology. The situation in Hong Kong is much the same. Almost everybody can expect a one to one and a half hour delay to be at an appointment or a meeting due to the hectic and busy traffic in Hong Kong, and thus nobody would render or expect an apology, or feel offended.

Differences are also present in the speech act of replying to compliments. English speakers treat replying to a compliment differently from speakers of many other languages. Thus, ESL speakers are anticipated to encounter problems when they perform this speech act in English. Pomerantz (1978) for example, claims that 'compliment responses are subject to separate systems of constraints'. Firstly, the recipient of the compliment would either offer an agreement or disagreements (denials) to the compliment. That is, the person receiving the compliment agrees with the compliment given to him or her by showing gestures such as nodding his/her head, a smile, or being silence, or with utterances such as thank you; or disagree with the compliment given, although this disagreement may be lacking of its genuinity as in some

culture like the Malay, for example, where a disagreement in such a situation might represent a form of politeness, and being humble. The second constraint, the recipient of the compliment might accept or reject prior compliments. A recipient of the compliment might accept the compliment without doubts that he or she deserves to be complimented based on for example an achievement that he or she obtained recently. On the other hand, rejection of the compliment is also possible. That is, the recipient would not accept the compliment at all, and thus utterances such as "I don't deserve it", "It is not me that ought to be complimented", or "You have complimented the wrong person" , etc. could be plausible feedback from the recipient of the compliment. Thus, in a Malay community or Asian in general where indirectness is a norm, these non-native speakers of English would be anticipated to have problems replying to compliments. The problems are inevitable unless they are proficient and have lived in English speaking culture for some time and have learned all the rules of speaking. Certainly, this is not always the case. Even some ESL speakers who are very proficient in English are not aware of all the rules of speaking in English, and thus they too face the same problems. The fact of the matter is that when dealing with the production of speech acts, the immediate problem according to Cohen (in Gass & Neu,1995) is the evaluation of interlanguage speech act behavior. He further states that one needs to ask the following questions:

- a. "To what extent have learners acquired the sociocultural and sociolinguistic abilities needed to realize the particular speech act?"

- b. "To what extent is the learners' speech act behavior similar to or different from a native speaker's behavior under the same circumstances?"
- c. "What compensation strategies do learners use when their language is inadequate?"
- d. "What is the learners' selection route and decision making process with respect to strategy preference modification preference, content limitation, and illocutionary intent?"
(Cohen, 1993: 13)

Cohen stresses that there is a need to probe the actual decision making and selection process that learners at different levels of proficiency go through in order to identify strategies that lead to the successful production of speech acts in the target language.

2.2.1 Performing Speech Acts: Various Plausible Forms (Formulas)

It is a fact that any one communicative activity or function can be carried out in more than one way. This too constitutes a problem for non-native speakers of English. For an example, one apologizes by saying:

I am sorry or Please forgive me.

Many responses allow the speaker to achieve his communicative purposes, but it is impossible to predict which one will be used. Because of this it is not surprising too that native speakers of English could possibly make errors in their speech act production. The prediction is much harder with ESL speakers. The role of his/her native language, his/her overall proficiency in English and other factors will affect his or her communicative performance.

Sociolinguistic competence, as we know, is a scalar behavior. There is more than one particular utterance that can be used to accomplish each communicative task. For example, one can refuse an invitation to a dinner in different ways, and one can greet another in many ways too. The fact here is that sociolinguistic competence is not a dichotomous behavior. There is no right or wrong answer or response for each question or statement. There are however, levels or degrees of appropriateness for each communicative behavior.

2.3 Language Testing: An Introduction

At present, there are few reports, although they are increasing in numbers, on efforts to develop tools for testing sociolinguistic competence. However, language testing has been getting more attention in the field of second language. Literature in second language acquisition (SLA) for example, will show that various kinds of language testing have been carried out to test different aspects of SLA. However, as Kasper and Dahl (1991) have pointed out, until today, few efforts have been made to compare data collection techniques especially in the area of interlanguage pragmatic, and measuring 'sociolinguistic or sociocultural' competence in a second language (*brief introduction and comparison between different data collection techniques will be dealt with in section 2.4.3 of this chapter*). It is one area in which debate and discussion continues over the reliability and validity of various tests. For example, Rintell and Mitchell (1989) compared data collected with the *Discourse Completion Test* and *role play* and concluded the two basically

produce similar data. Bodman and Eisenstein (1988) compared data collected through *Discourse Completion Tests*, *open-ended role play*, and *field notes* on naturalistic data. Unfortunately, there has been no strong evidence to suggest that a particular collection procedure is superior to any other. In fact, the literature has shown that each approach has its own strengths and weaknesses; as Cohen (1993) stated "each technique has its own merit but it is the use of more than one that provides us with important triangulation." Most important of all, one needs to realize that speech act realization is complex. Thus, it is imperative that researchers in this area be cautious in developing a research methodology (testing tool) because presumably, both the methodology and the outcome of a research are interrelated. Certainly, issues pertaining to methodology are not at all to be ignored. Instead, researchers should learn from the issues surfaced in previous research, and continue to improvise and better their techniques and methodology for better results. As it is, methodological issues are still a central subject in relation to speech act realization in inter language pragmatic.

Regardless of the concern and problems on testing and methodology, many studies have been done on different speech acts since early 1980s. Speech acts is an area which according to Cohen (1993) "has been investigated and described from a variety of perspectives: Philosophical, social, linguistic and cultural". He further claim that "An

effort has been made to identify universal norms of speech behavior and to distinguish these from language specific norms in order to better understand and evaluate inter language behavior (Cohen,1993). However these studies cover only one speech act at a time, thus, only one aspect of sociolinguistic competence is studied. Studies which have been done to develop testing instruments are those done by Andrew Cohen and Elite Olshtain (1981). The details of this study will be elaborated in later section of this chapter.

2.4 Issues in Communicative Testing

2.4.1 *Testing and Communicative Competence*

Now, we turn to a discussion of some of the issues in testing communicative competence in general. This will then lead to the description of studies which are done to create tools for assessing communicative competence, and sociolinguistic competence in particular.

Current ideas about the role of language in communication have greatly influenced the development of tests for assessing communicative competence. The two important terms, "usage" and "use" are distinguished because the distinction is of great importance for teaching and testing. "Usage" is concerned mainly with formal language patterns, whereas "use" is concerned with communicative function. Instructional programs emphasizing "use" direct their pedagogies toward communicative competence. It is assumed that realistic communication will provide the learner with meaningful practice, and thus, he/she will conform

gradually to the norms of usage. However, the emphasis is on appropriacy rather than on formal linguistic correctness. The distinction between “usage” and “use” implies that the selection of the test items should not be chosen only from linguistic component, but also from the communicative one. Communicative competence states that one should have the linguistic knowledge of the language and also the ability to use that knowledge in various communicative situations. Then, communicative effectiveness can be taken as the criterion for success (Carroll,1980).

From the “usage” point of view, the mastery of the language patterns will teach the language user to cope with the situations or contexts he or she encounters. Therefore, a single test of the learner’s language proficiency (ie. grammar test) is adequate to measure his or her ability to cope with real situations. “Usage” views language as a unified entity of fixed grammatical patterns and a core of commonly used lexical items. On the other hand, “use” views the above view as “oversimplified”. From the “use” point of view, different patterns of communicative situations will require different configurations of language skills, and thus, different test contents (Carroll,1980). The regular classroom language proficiency test will not adequately measure how well an ESL student demonstrates his or her communicative skills. Such proficiency test will however, measure how much the student knows the grammatical patterns of the language he or she has learned.

2.4.2 *The Nature of Proficiency Testing*

Proficiency tests are measures of a subject’s overall knowledge of a subject relative to some norm. The results of a proficiency test rank subjects relative to that norm. The content of specific items on a

proficiency test can be evaluated only by the two criteria of a. how much and well they distinguish more from less knowledgeable subjects, and b. content validity. Proficiency tests are reliable and valid only as a collection of items that work together to achieve a single end, that of ranking subjects.

Proficiency tests need to be distinguished from diagnostic tests, in which each item represents a specific piece of knowledge that the tester wishes to measure the subject's knowledge. There are overall measure of quality of diagnostic tests since each item has specific reason for being included in the test. Diagnostic tests may be used to measure the achievement of students over a collection of materials they have been taught, or to determine what aspects of a given body of language knowledge students know and do not know, usually for the purpose of designing an instructional program.

Proficiency tests, therefore, are global measures of knowledge, ranking students against some external norm, and diagnostic tests are measures of specific aspects of knowledge and indicate only what from a list or collection of specific bits of knowledge is known and not known.

Proficiency test construction is a fairly well established process, at least as it applies to the development of multiple-choice tests. It consists of identifying a pool of prospective items. Pretesting them on a representative sample of subjects, item discrimination, and overall test reliability to determine which collection of items will reliably rank the subjects, indicating which subject's overall knowledge is greater or lesser than another's.

The content of proficiency test is frequently misunderstood. The content of specific items on a proficiency test is of no great interest

to the test designer, although they should represent a reasonable range of the type of knowledge being tested. Specific items appear on a proficiency test almost only because of what they can contribute to the overall goal of reliably ranking subjects.

2.4.3. Methodological Issues

According to Carroll (1980), there are problems to be considered in devising instruments for measuring communicative competence. One, is to discover how to carry out a rigorous measurement of language-based performance and still keep intact the essential features of communicative behaviors.

Since the departure of research pertaining to cross-cultural speech act, discussions and debate about the preferred way to collect data on speech acts have been a central issue. The fact that speech act realization and various possible strategies to meet one's objective in performing speech act are complex in nature, certainly a study in this area needs a careful development of research methodology. As Cohen (1993) points out that methodological issues cannot be ignored, for it is not clear to what extent differences in methodology yield differences in results.

A test should really measure what it is designed to measure and does not discard the criteria needed in such a test. That is, the test should not deviate from the objectives for which it was set. It should not become a test of something else. This section will first examine the literature in which issues pertaining to methodology issues are described. Particularly, different data elicitation measures are introduced, described and compared (i.e ethnographic observation,

acceptability ratings, role play, etc.) This section will also look at the distinction between “norm-referenced” and “criterion-referenced” tests, and the issues of authenticity.

2.4.3.a Ethnographic Observation

This approach involves the collection of naturally occurring data. Research on the speech acts of ‘complimenting’ pursued by Wolfson (1989) have proven that this method is effective. Unfortunately, ethnographic observation does not work effectively with other kinds of speech acts. It has been proven to be time consuming, perhaps expensive, and more over this approach may not be productive.

2.4.3.b Role Play

A situation can be called a ‘role play’ according to Cohen (1993) if the respondents or the subjects in the experiment are not aware that the event has been contrived for the purposes of collecting data, or, the subjects are aware and agrees to cooperate. Of the common data elicitation methods, open role play are the closest to what we might expect to reflect naturally occurring speech events. This approach have the advantage of allowing the researcher to set up situations in which the occurrence of a particular speech act is likely in circumstances in which the occurrence can be recorded using audio-visual tapping. Quite frequently, a role play may consist of a description of the situation that requires the subjects to respond. Subjects may also be required to provide one or more rejoinders to turn the role play into an interactive event. Audio and video-taping appeared to be the most popular form of recording data when a role play is used as a method. Inevitably, the

taping process itself may cause some problems depending how intrusively the research is done. Often, respondents would yield the feeling of uneasiness and uncomfortable when audio or video taping take place. Such situation according to Stubbs (1983) may engender certain reactive effects. For example, subjects might develop special verbal strategies for dealing with such an approach. Other than this, this approach is also cumbersome to administer and time-consuming. In addition, role play according to Houck and Gass (in Gass and Neu,1955) “ are just that, role plays, so again we are left with the question of the degree to which they mirror the linguistic behavior of individuals in the particular setting established by the researcher.”

2.4.3c *Acceptability Ratings*

This approach involves the obtaining of respondent's judgments as to how appropriate certain responses are for a given responses (Olshtain & Blum-Kulka,1985). When using this approach the respondents are required to select the most appropriate response from a series of possible responses given in a task. The responses given by the subjects would then be scaled on a continuum i.e could go from the least intensified to most intensified of a particular speech acts i.e an apology.

2.4.3.d *Written Completion Task*

In general, there are two types of written completion task commonly used. The first one is the '*open-ended elicitation*'. Normally, there is a written prompt followed by a space for the subjects to provide a written response. The second written completion task is known as the '*Discourse Completion Test*'. Unlike the open-ended elicitation, in the

latter approach the discourse is structured where some part is left open, and some part closed, providing both for the speech act and a rejoinder.

When comparing the written approach with role play, and ethnographic observation, the written approach seems to save time in data collection, and research done by Cohen, Olshtain and Rosenstein (1986) showed that it (the written approach) have been seen to provide a reasonable projective measure of the same kinds of data collected from oral role play. This approach have been seen as an effective means of gathering a large amount of data quickly. However, one visible short coming of this approach is its failure to include every detail required. For example, written completion tasks do not have the capacity to collect the prosodic and non-verbal features of oral interaction. The response given to possibly is shorter than would be the case in speaking.

2.4.3.e Multiple Choice Test

To begin with, MCT also falls under the category of a 'written mode test'. MCT is most suitable and practical for dichotomous behaviors. Scalar behaviors, however, could be assessed better and more reliably with tests such as an open-ended test or an oral interview. Thus, there is the problem of practicality of a MCT for scaling behaviors.

First of all, in order to use a MCT, behavioral criteria have to be determined before the test can be designed. However, this is not an easy task because there is a variety of utterances which can be generated. The best and most practical solution to this problem is to choose the most frequently occurring and the most typical patterns. In

this study, the responses will be generated by using an open-ended test which ask the subjects to write down what they are most likely to say in the situations described. The open-ended test is used because, like MCT, it is a written test, not an oral test. Therefore the responses in the open ended test will best suit the MCT because both tests generate responses by means of a written mode. If another method, such as an interview (which are produced orally) might not be suitable for the MCT because of the differences in mode. Thus, keeping the same mode (written) should maintain the suitability of the responses.

2.4.4 Norm-Referenced Versus Criterion Reference

Norm referenced test compares the performance of an individual with that of other individuals in a sample of comparable persons. According to Ingram (1977), the purpose of a norm-referenced test is "to compare the level of performance of an individual with the general standard of performance which is shown by the total group that he belongs to and can be compared with".

Criterion-referenced test, on the other hand, compares the performance of an individual with predetermined criteria. In his book, Carroll (1980) refers to "criterion-referenced" as criteria of specified communicative behavior, not "contributory linguistic skills". Therefore, the emphasis of criterion-referenced test is not on the level of a student with reference to his or her peers. It is on whether or not the individual student "knows something rather specific that he/she is supposed to know, or can perform something rather specific that he is supposed to be able to perform".

In short, Ingram states that norm-reference testing compares the

behavior of an individual with the behavior of others, while criterion-referenced testing describes the behavior of an individual with reference to "externally predetermined and specified objectives." Thus, a choice of using either one should be made very carefully so that the objectives of the test will be maintained. A consideration should be made of whether the results will be compared with another group, or with a predetermined norm. The two approaches are not mutually exclusive, and Carroll believes that both can help solve testing problems in that, the analysis of test items and the interpretation of test performance is the specified, communicative behavioral criteria.

2.4.5. The Notion of Authenticity

If this notion or principle is to be fully applied, Carroll (1980) emphasized that this would mean that all the tasks undertaken should be real life, interactive communication operations and not the typical routine examination responses to the tester's stimuli. The testing of the communicative performance should be based on real communicative activities, and not the regular class contents or textbook contents, which are usually and mainly grammatical patterns and rules of the language. That is, how language is used should be the major ingredient of the communicative testing. Therefore, communicative testing should be a different test altogether, and it should not include any material emphasizing the grammatical correctness of the forms but with the appropriateness of use in the real communicative situations.

According to Ramirez (1984), the theoretical base for the communicative competence construct are the anthropological, sociolinguistic and pragmatic views of language. The testing of

communicative competence tends to be associated with global ability involved in communicative acts such as giving directions, requesting permission, apologizing, etc. In most instances, the specific situations are described by the examiner, and rating scales are used to evaluate how effectively the examinee performed the communicative act. Thus, this leads to another problem in communicative act. Ramirez (1984) added that the measurement problems are compounded by the fact that the usual measurement of communicative competence is based on a global rating scale approach, while linguistic competence is generally measured in tests using a discrete item method. It was suggested that the presumed difference between measured linguistic and communicative competence may reflect a difference in method of assessment, and not the trait or construct being measured.

2.5 Previous Studies of Tests on Communicative Competence

The next section covers some previous studies on testing oral proficiency in the early stages of testing communicative competence in general and sociolinguistic competence in particular.

Oral proficiency testing received increasing attention until the term communicative competence emerged in early 1960s. From then on, 'communicative competence' has become the "buzz word" in foreign language proficiency testing. Before that, the oral interview was the instrument for measuring the so called communicative proficiency (now called communicative competence), even though some practitioners were not satisfied with it. Thus, the oral interview was the earliest instrument designed to measure such skills. Despite the shortcomings, it is still the most valid and the most reliable instrument available. The oral

interview described above is an oral testing of sociolinguistic competence. Even though it measures how well the subjects speak the languages tested, it is testing the appropriateness of the language used in the situations. It is therefore testing some aspects of sociolinguistic competence.

Another study that attempted to test communicative proficiency is the study by Palmer (1972). The purpose of this study was to search an answer to the two following questions: 1. Is there an oral communicative ability not being measured by foreign language proficiency test?, 2. Can this ability be objectively and reliably measured? For that purpose, Palmer (1972) conducted and administered two types of experimental tests. The first test, "COMTEST" was to measure two-way oral communication ability; the second, "PROTEST" was designed to measure oral production ability. The Michigan Test was selected as a representative foreign language ability test because TOEFL correlates highly with it. All test were administered to three groups of subjects: non-native students in the English language Institute (NNSLI), non native students in regular courses (NNSUM), and native speakers of English (NS) at the University of Michigan. The COMTEST asks for the students to obtain and communicate information in a dialogue with an examiner. The following scores were tabulated for each subject.

1. Elapsed time until the correct answer was reached.
2. Number of times the examinee could not understand the question and had to ask for clarification.
3. Number of times the examinee misunderstood the question and answered it, thus causing to receive incorrect information.
4. Number of times the examinee gave up.

5. Number of questions until the correct answer was obtained.

6. Number of guesses.

The test was given in three forms, with the second and third being the revised forms. However, the tasks remain the same, that is, the subject is shown a set of similar pictures, such as pictures of different TV sets. The examiner tells him that he is thinking about one of the pictures and the subject must ask him questions until he can identify which picture it is. The speed and the accuracy of the identification are graded and the subject was told about the criteria beforehand.

The PROTEST was prepared in one form. The subject was to describe orally one of four similar pictures so that the examiner could determine which one was described. The length of time of description before the examiner could determine correctly was scored.

An additional control test (ACTEST) was administered to test oral comprehension. The subject is shown two pictures. A recorded narration is played, which contains four clues, which identify only one of the two pictures. Subjects were scored on the elapsed time until they responded. Penalties were given for incorrect responses (Palmer, 1972).

Then, the Michigan Test with its three sub tests was administered. The scores of COMTEST, PROTEST and ACTEST were compared with scores of the Michigan Test. High correlations among the tests were recorded, which indicate that the experimental tests were measuring language related skills.

In conclusion, Palmer claimed that the tests were measuring factors not being measured in the traditional foreign language proficiency test battery. The reliability was higher than that obtainable in the oral interview of the same length. He thus, claimed

that the COMTEST was a useful and practical measure of oral communication ability (Palmer, 1972).

The next study is also an attempt to assess communicative ability. It is among the studies which followed Savignon's pioneering study in 1972. This study by Valdman and Moody (1979) was intended to test communicative ability at a relatively low level. It is called the Indiana University French Communicative Ability Test (IUFCAT). The authors set out in 1972 to develop a test which would provide valid and reliable evaluation and which could be administered effectively in a large multi-section foreign language program (Valdman and Moody, 1979).

In conclusion, the authors claimed that the test is relatively easy to administer to large groups of students if lab facilities are available. A related advantage is the relative ease of scoring (Valdman and Moody, 1979). Despite the inherent shortcomings in the area of scoring reliability, they claimed that the test has proven to be a generally valid, reliable, and practical measure of minimal communicative ability in French.

The final study, by Cohen And Olshtain (1981), is an important one and the study on which the present study is based. They have been doing extensive studies on the speech act of apology with an attempt to develop the right tool to assess sociolinguistic competence among ESL speakers. In their study eight apology situations were selected and designed to assess cultural competence among non-native speakers of English (specifically Hebrew speakers). The situations were specifically intended to assess "intensity of regret", "cultural competence", and "stylistic competence". The study was designed to measure how the NNS formulate their apologies in these situations, which range from formal

to informal. The variables are: status of participants and familiarity among them. Although the study was conducted to measure only the speech act of apology, the method adopted was interesting and practical. Furthermore, the method can be adapted to measure more than one speech act, which would count as more aspects of communicative competence. If more than one speech act can be measured in one test, then, it is possible that more aspects of the competence could be measured with one simple test. The speech act of apology was selected hoping to elicit more natural responses. Cohen and Olshtain selected eight apology situations; four were specifically intended to measure intensity of regret, and the other four, cultural competence and stylistic competence. The respondent was to read the description of the situation, and the tester was to role play the person deserving the apology. This approach was intended to avoid responses in indirect speech.

The subjects of the study were college students - native Hebrew speakers in their early twenties, enrolled in intermediate-level English classes at the Pre-Academic Center at the Hebrew University, and native English speakers from the US, enrolled in the Hebrew University's Overseas Program. The investigator, Cohen, first addressed the whole class, stating the purpose of the study. Then, they were tested one at a time in a separate room. The subjects were asked to read the situations (in randomized order). Then, the English speakers, and the Hebrew speakers serving as non-native speakers of English were asked to respond in English. The other Hebrew speakers were to respond in Hebrew.

In the analysis of the data, they first determined the range of patterns of apology among native speakers (both English and Hebrew), and then compared non-native responses to the native ones. A "rating scale"

was used as the tool of assessment. The frequency with which the three sets of subjects used one or more of the semantic formulas in each situation were calculated. A plus (one point) is given to non-native speakers each time they used a semantic formulae.

The results of the study were interesting, and significant differences were found in the way apology is treated in English and Hebrew. Both negative transfer and level of proficiency were found to influence the responses by each subject. In conclusion, Cohen and Olshtain stated that their study had produced a "crude measure" of sociocultural competence.

2.6 Previous Studies' Design and Methodology

Among the earliest and most common testing instrument developed was the 'oral interview'. This instrument was developed to meet the need for a testing procedure of the second language speaking skills. Now known as the American Council on the Teaching of Foreign Languages.Educational Testing Services (ACTFL/ETS) scale, the oral interview has shown a high degree of validity and reliability as a testing method. It is a procedure capable of measuring a wide range of speaking abilities from novice to native. In the oral interview, the testee converses with one or two trained testers on a variety of topics for approximately 10-40 minutes depending upon the student's proficiency level. The speech sample is then rated on a scale ranging from 0-5. The task requires the testee to perform to his functional ability in real-life situations. The administration of the test is generally made as normal as possible so that natural responses can be obtained. Overall, the oral interview has been praised for its high degree of validity and

reliability which makes it one of the best tool available in assessing communicative proficiency (Jones,1979).

The next study which consists of three tests was described by Spolsky et al (1972). The first test is the Spanish-English Language Dominance Assessment. The test was a 20 minute oral interview between a student and an interviewer and a recorder. The interviewer ask questions as naturally as possible while the recorder writes down the responses given in both English and Spanish. The interview contains three sections, each with a Spanish and an English component. Part A in Spanish and Part C, an English component. Part A in Spanish and Part C in English are series of questions about the child's language experience. The content of the answer and the evidence of how fluently he/she answers in each language are the two criteria being tested. Part B in Spanish and Part D in English each contains four-naming tasks, two calling for nouns and two for verbs. In part E, they are asked to describe what they see in some pictures, from which two to three minute tape recorded samples of their speech in each language are collected (Spolsky, et al,1972).

Next, is the Oral Placement Test for Adults which is oral test given to one subject at a time in English. It contains several sections. The first contains seven questions on personal details of the subject. The second section contains fourteen simple utterances, to be repeated with an understandable pronunciation. The next section is on phonology, and then short conversation. The conversation part ask the subject to convey understandable information within the social context suggested. The following section concerns the manipulation of grammatical structures. A success in each level will ensure that the subject will be

placed in the higher levels. These levels refer to the levels of instruction, which were divided into three levels: beginning, intermediate and advanced.

To conclude, all the above designs of tests use an oral mode as the instrument of assessment. Thus, more natural responses can be expected from the subjects. However, the test does not really ask for performance in more socially structured communicative situations. They do not ask for what or how the subject does in real communication, but only in a small part of the task. The tasks need to be more situational and realistic. The subjects were interviewed in an examination situation. In order to obtain more accurate responses, the interview should be administered in more natural settings. Thus, more accurate responses will be obtained from the situations or the communicative contexts.

The third study, by Palmer (1972) is another study to test communicative proficiency. The study experimented with two types of oral test, COMTEST and PROTEST; the former to measure two way oral communicative ability and the latter, oral production ability. In the COMTEST, the testee was required to ask questions which can identify the picture the tester had in mind. In the PROTEST, the testee was to describe one of the four pictures so that the tester can determine which one of the four pictures was described. Overall, the tasks required the testee to produce his oral communicative ability, and no other skills were asked for.

Indiana University French Communicative Ability Test is another study that was covered in this chapter. The study was conducted by Valdman and Moody (1979) to measure French communicative ability. The

test contains four sections: Part 1, pictorially cued responses; Part 2, personal questions; Part 3, question formulation; and Part 4, situational responses. The answers were evaluated among three parameters: semantic and pragmatic appropriateness, grammaticality and correct form of lexical items, and fluency and accuracy. In addition to these, the answers were judged on the basis of amount of information communicated.

Generally, the tasks of this test are mostly oral production of the testee. The tasks resemble real-life situations in some ways. The study on apology by Cohen and Olshtain (1981) was intended to measure sociolinguistic competence. It used the speech act of apology as the content of the test. For the purpose of their study, Cohen and Olshtain (1981) set up four semantic formulas for apology: 1. an expression of apology; 2. an acknowledgement of responsibility; 3. an offer of repair; and 4. a promise of forbearance. They make up eight apology situations to assess "stylistic appropriateness" and "intensity of regret." Their subjects were forty-four college students: twelve native English speakers and thirty-two native Hebrew speakers (twenty of them serve as nonnative speakers of English). The objective of the study is to identify and explain deviations of nonnative English in their formulas of apology in English. The testees were given eight specific apology situations and they were asked to apologize in the situation. After reading the eight situations silently, the subjects were asked to respond orally when cued by the investigator. The twelve native English speakers and twenty native Hebrew speakers were to respond in English and the twelve native Hebrew speakers in Hebrew. The subjects, English speakers and Hebrew speakers represent three groups. The results among

the groups were compared. In the results, they found that some situations where nonnative English speakers deviated from the cultural patterns of native English speakers are the result of "negative transfer" rather than a misconception of how to use the formulas in English. In these cases, the nonnative speakers did not "offer repair" as frequently as native speakers and were less likely to "express an apology" than were native English speakers. There are also situations where the problems are due to grammatical competence and not negative transfer. In these cases, the nonnatives were less likely to "offer repair" than both groups of native speakers and were also less likely to "acknowledge responsibility". Lastly, there were situations in which the non-native English speakers responded like native English Speakers even when Hebrew speakers responded quite differently in Hebrew. The nonnative English speakers also "expressed an apology" more frequently than Hebrew speakers, and they tended to "acknowledge responsibility" like native English speakers more than did speakers of Hebrew.

Cohen and Olshtain suspect that the deviation in the degree of intensity is due to grammatical competence; that is, the nonnatives may not have the necessary proficiency in English to know precisely how to use intensifiers. With respect to stylistic deviation, they suspect negligence on the nonnatives part. Also, stylistic variation among natives is slight and only in a few cases among nonnatives.

The test also requires oral production of the testees. However, the rating scale was used, and Cohen and Olshtain tried to validate if it can be used as the tool. This study now becomes the stepping stone for the present study. The suggestions by Cohen and Olshtain are applied in this study.

In conclusion, all the studies described in Chapter two used oral mode to assess communicative proficiency or competence. As claimed by the oral interview approach, the oral mode can provide a very reliable, valid and practical method of testing. The responses to the questions or tasks are more natural and thus, resemble real-life situations. The present study, however, will adopt a written mode, that is, the subjects will be asked to choose from responses given. The advantages and the disadvantages of both the oral and written modes of testing communicative competence will be presented in the final chapter of this thesis.

2.7 Studies on Speech Acts:Revisited

The following section covers the findings of the studies on the speech acts of apologizing, and complimenting in English.

2.7.1 *Speech Act of Apology*

Maeshiba, Yoshinaga, Kasper and Ross (in Gass & Neu,1995) in their research on 'Transfer and Proficiency in Inter language Apologizing' claim that apologies are the next best speech act studied other than the speech act of 'request' in descriptive, cross-cultural, and inter language pragmatic. Globally, in any culture and speech community, there must exist a system (i.e rule) where interlocutors are able to react in remedial action upon committing a wrong-doing or an offense. That is a person needs to apologize upon committing an offense. This according to Maeshiba et al. can "thus be regarded as a pragmatic universal." However, one needs to know that the conditions which require an apology

is not universal. We know that one culture differs from the other in many aspects and situations. Thus, it can also be generalized that what counts as an offense in a Malay culture for example, is different than in Japanese culture, and is different than in American culture, and vice versa. Thus, it is clear that a non-native speaker of particular target language (i.e English) has to learn the conditions for apology in that target culture in order to function well in that culture. Specifically, he or she has to learn the ideal strategies, and linguistic means by which an apology can occur and be implemented.

Truly, there have been a large number of research examining both the NS and NNS' patterns of apologizing which was first proposed by Olshtain and Cohen (1981) and empirically followed by other studies by the same authors (Olshtain 1983; 1989; Olshtain-Cohen 1989). These studies have found that generally, there are many strategies that one can utilize to convey regret, or proposing remedy. Other plausible strategies described by Olshtain (1989) and Berman & Kasper 1993) are: upgrading apologetic force; downgrading the severity of the offense; offering repair; etc. As one would expect, all these strategies are visible cross-culturally. However, the use of a particular strategy is highly sensitive to contextual conditions and situations. It is also subject to variation that exist cross-culturally.

Clearly, the obligation to perform an apology on the offender affects the choice of apologetic formula and intensification of apologetic force (Cohen & Olshtain,1981; Volmer and Olshtain,1989; Bergman & Kasper,1993). It seems that the severity of the infraction have the strongest effect on the speech act of apology.

2.7.2 *The Speech Act of Complimenting*

Complimenting may vary from one culture to the other (Holmes & Brown, 1987; Creese, 1991; Han, 1992; Nelson, El Bakary, Al Batal (in Gass & Neu, 1995)). Wolfson and Manes (1981) claim that complimenting is different across cultures. A compliment in American English could easily seem very insulting to other speakers. In addition to understanding the words, one must also know the rules for interpreting them. Nonnative speakers are often unsure of the meanings of compliments because they could have different meanings than what they think. Often the result is misunderstandings.

The study of complimenting by Wolfson and Manes (1980) was aimed to elicit the forms (linguistic realization) of compliments in English, that is what constitutes a compliment. What linguistic forms are used in formulating a compliment in English? Another focus of this study was on the appropriateness of complimenting, that is, what is appropriate to compliment and what needs to be complimented.

Wolfson found that many languages use proverbs and other "ritualized phrases" in compliments. For example, she noted cultural differences in complimenting among Iranian and Arabic speakers and posited that Iranian and Arabic speakers tend to use proverbs and other precoded ritualized expressions when complimenting. One example that Wolfson (1981: 120) described was a compliment given by an Arabic speaker to a friend's child: English equivalent translation: "She [the child] is like the moon and she has beautiful eyes". However, Americans use a very restricted set of lexical and syntactic structures. Thus, a learner of American English must become familiar with the appropriate syntactic structures and lexical items used in compliments. The most

important finding in this study (Wolfson and Manes, 1981) is that compliments in American English “totally lack of originality”. Both the subjects of compliments and the lexical items are formulas. They found that eighty percents of all compliments are adjectival in form. Of the seventy-two positive adjectives occur, five: “nice, good, beautiful, pretty and great” were used frequently. The two most common adjectives found in the compliments were “nice” (22.9%) and “good” (19.6%).

The proceeding section describes the brief outline of the previous studies done on the two speech acts which are the components of the test. The analysis of previous work done on the speech acts of apology and replying to compliments, and studies on communicative competence in general, and a comparison of different testing tools used for testing sociolinguistic competence have greatly influenced the objective of this present study. In addition to that, previous work on the subject matter have given this paper specific reasons for choosing each of these speech acts. As stated earlier in terminology section in chapter one, “speech act” is the minimal term in the study of communicative competence. It is the act we perform when we speak, and is defined in terms of discourse function. Thus, the following section looks at: (1.) the rationale for choosing the speech acts of apology and replying to compliments as a primary content of this present study, and (2.) rationale for using a written mode (MCT) to conduct this study.

2.8 Theoretical and Practical Reasons for Choosing Speech Acts: Apology and Replying to Compliments for This Study.

2.8.1 *Apology, and Its Rationale*

The first speech act ‘apology’ was chosen because of the many components

it contains in its semantic formulas. The components are very clear cut and can easily be identified in each apology statement. Cohen and Olshtain have done an extensive study on apology which they believe could provide the basis for a study regarding the testing of sociolinguistic competence. There are five "semantic formulas" present when the offender is positively inclined to apologize (Cohen and Olshtain, 1981). They are: 1. expression of apology; 2. explanation or account of the situation; 3. acknowledgement of responsibility; 4. offer of repair; and 5. promise of forbearance. Cohen and Olshtain claim that in most cases, one of the formulas is sufficient to perform an apology, but often two or more are combined and thus create an apology of higher intensity. A small offense would probably need only a weak expression of apology, for example, "*I am sorry*"; however, a serious offense might need two or three formulas combined, as in "*I am very sorry, are you alright? Let me help.*"

Different cultures use different semantic formulas in making an apology, and thus an ESL speaker is susceptible to having problems in apologizing in English. Some cultures do not even apologize if the offense is small and insignificant; the speaker does not feel that he or she needs to apologize and the addressee does not feel offended and thus does not expect an apology from the speaker. In such cultures, both the speaker and the addressee usually keep silent without mentioning anything about it. On the other hand, some cultures feel that an apology is necessary for every offense regardless of its intensity. For example, people from Asia or the Orient i.e Japanese and Indians treat apology very seriously. Usually, the apology is accompanied with body movements such as taking a bow, or kneeling at the feet of the person apologize

to.

Thus, the many formulas in English apologies and the differences across cultures are some of the reasons the speech act of apology was selected as one of the components of the test. Again, as in other speech acts, in apology too, many factors affect the performance of the speaker. A more proficient ESL speaker might perform better than the less proficient ones. The native language of the speaker might sometimes influence the speaker and thus keep him/her from using the formula in English. In this case, the speaker might use, his native language formula. An ESL speaker might not use the formula in the target language because his proficiency level is low, thus, he/she does not have the knowledge of the formula or the necessary grammatical skills.

2.8.2 *Compliments and Its Rational*

The second and last speech act chosen in this present study is replying to compliments in English. This speech act was selected due to the following reason: Complimenting in English has been described by Holmes & Brown (1987) tend to be troublesome aspect of English for learners of English that come from different communities and cultural backgrounds. For instance, the high frequency of complimenting in western culture like Americans in general can be an embarrassing experience for NNS (Holmes & Brown, 1987; Wolfson, 1981). Thus the unsureness to react or to reply to complement may be a serious constrain faced by NNS of English.

Complimenting has been studied extensively by Wolfson and Manes. They both indicated that the major function of compliments is the "establishment or reinforcement of solidarity between the speaker and

the addressee" (Wolfson and Manes, 1980). They focused mainly on the forms of the compliments, that is, how compliments are formulated and what constitutes a compliment. They also focused on the appropriateness of complimenting - what is appropriate to compliment and what needs to be complimented.

Up till this point, review of previous studies and familiarity on the existence of different testing techniques used to assess sociolinguistic competence have given us a solid ground on choosing a multiple choice test as the tool in this study. Following is a section explaining the rationale and logic for using MCT.

2.9. Rationale(s) for Measuring Sociolinguistic Competence With a Written Mode (MCT)

The following is a brief discussion of why this thesis attempts to measure sociolinguistic competence with a written, multiple-choice test. Even though proponents of oral interviews claim that the oral mode is a very reliable, valid and practical method of testing, Harris(1969) claimed that oral interviews tend to have low reliability. He claims that no two interviews are conducted alike, even by the same interviewer. Therefore, it is clear that the test reliability will be affected. Similarly, no interviewer can rate consistently throughout a large number of interviews. Thus, this inevitable shift will lower the rater reliability of the measure.

On the other hand, it is not so with the written mode, i.e the

multiple choice test. Since there is one correct or preferred answer for each question, both the test reliability and the rater reliability are fairly high. The scoring is also consistent throughout a large number of tests.

Another weakness of oral interview is the validity of the interview itself. Does it really assess what it is supposed to assess? Even though the oral interview emphasizes ease and comfort during the interview, the interviewee will not be able to represent himself/herself as natural and calm as he or she could. The subject is confronted with an interviewer or more who ask(s) him or her questions. Nervous and uncomfortable feelings are present, and these feelings affect the subject's performance. The performance of the interviewee in a particular interview does not necessarily accurately reflect one's true ability. Even though proponents of oral interview claim that it offers a realistic means of assessment in a "natural" speech situations, the opponents argue that the interview is still artificial and unrealistic (Heaton,1975). The students are placed not in natural, real-life speech situations, but in examination situations. Thus, they are susceptible to psychological and emotional tensions which will affect their performance.

However, the validity of a multiple choice test is fairly high. It measures what it is supposed to measure because the test writer can easily choose any item or material that he wants to be tested. The points are discrete. Each point to be tested can clearly be presented in each question. The environment of the multiple choice testing also eliminates the feeling of uneasiness and nervous on the testee's part.

Even though it does not provide a natural speech situation, it does provide a better environment for taking a test, free of psychological tensions and emotional constraints.

With regards to scoring, the oral interview also presents another weakness. The scoring is not clear cut, and sometimes can be inaccurate. Usually the scoring is done during the interview. Heaton (1975) states that the scoring can range from an impression mark to a mark arrived at on the basis of a fairly detailed marking scheme. The interviewee will feel uncomfortable when the interviewer writes down some notes after he/she responds to the questions. Therefore, the inconsistency in scoring affects the reliability of the measure. In contrast, the scoring of the multiple choice test is faster and easier than of the oral interview. The problem of inter-rater reliability does not exist at all in a multiple-choice test. The raters do not have to judge any of the responses. Thus, opinion and inaccurate judgment are eliminated in the scoring. Other weakness of the oral interview includes the administration of the interview itself. An oral interview is not easy to administer; the fact that it is expensive prohibits immediate retesting. Furthermore, it is very time consuming and the interviewers have to be fairly experienced and well trained.

Unlike the oral interview, a written mode, a MCT presents fewer criticism with respects to the points discussed above. It presents fewer problems than the oral interview. The scoring is clear cut. That is, there is always one correct or preferred answer for each question. Thus, the test reliability and the inter-rater reliability are fairly high and consistent. Furthermore, the responses are easily scored and the scoring is less time-consuming. The ease of administration makes it less

expensive, and the retesting can easily be arranged. The testers do not have to be experienced and well trained.

It has been established that the oral interview can be a valid, reliable, and practical tool for assessing oral proficiency and some aspects of sociolinguistic competence. Under some conditions, it can also be a practical tool. The discrete point multiple-choice test, on the other hand, can be a valid, reliable, and practical tool for measuring certain aspects of language competence, especially that competence that is relatively dichotomous, such as grammatical knowledge. What has been adequately investigated is the degree to which a practical instrument such as a discrete-point multiple-choice test can also be used to reliably and validly measure less dichotomous knowledge such as sociolinguistic competence.

It would be useful at this point to clarify the nature of the knowledge we are attempting to measure. Grammatical knowledge, the knowledge that makes up the content of many traditional discrete point multiple-choice tests, is relatively dichotomous in nature. That is, speakers of a language, either native or non-native, tend to be able to make yes/no acceptable/unacceptable decisions about a particular grammatical phenomenon. For example, on the following item from a TOEFL test:

Refrigerating meats.....the spread of bacteria.

- A. retards***
- B. retarding***
- C. to retard***
- D. is retarded***

Virtually all native speakers of English will choose A. as the correct

answer. Their internal grammatical knowledge makes a clear distinction of possibility or acceptability between A. and the distracters on this item. An internal grammar that allows one of the distracters, differ from that of essentially all native and competent speakers of English.

On the other hand, sociolinguistic knowledge is less dichotomous in nature. For a given speech act, even native speakers may accept a variety of forms or realizations. It is, in general, much more difficult to find aspects of sociolinguistic knowledge on which native or other competent speakers will reliably and consistently make a single judgment of correctness or acceptability. For example, taking one item from the test used in this study:

Dropping a pen while handing it back to your friend.

A. Gee. I'm sorry about that!

B. (You say nothing/ being silent)

C. Ooops sorry

D. Excuse me!

While most native speakers will accept C. as the form most probably to occur, they might accept some of the others also, and other native speakers might judge one of the other distracters as most probable. In this sense, sociolinguistic knowledge tends to be more scalar than grammatical knowledge.

This presents a challenge for a test writer. If proficiency tests are to measure the knowledge of some population of language learners, and if we wish to include sociolinguistic knowledge in the overall knowledge to be measured, we must find a way to evaluate the knowledge of less competent speakers against that of more competent speakers, even though that knowledge may be somewhat scalar in nature.

The oral interview, as claimed by its proponents, is a reliable and practical tool of assessing scalar behaviors such as sociolinguistic knowledge. On the other hand, the multiple-choice test is also a reliable and practical tool for testing, but not for scalar behaviors. It is more practical for dichotomous behavior such as grammatical knowledge or scientific facts.

2.9.1. *Scalar Behavior Versus Dichotomous Behavior*

First, what is scalar behavior, and what is dichotomous behavior? A scalar behavior is one that is describable by a number that can be represented by a point on a scale. That is to say, a scalar behavior has no definite point of preference; it is represented along a scale. It is also a matter of opinion, not a fact, and thus, allows no yes/no judgment. It involves personal judgment and opinion. On the other hand, a dichotomous behavior is one that is readily divided into two parts or groups. It is the type of behavior that can be judged on a yes/no or correct/ incorrect basis. It does not involve judgment or opinion because the preferred part has been decided, and that part is a fact. Thus, in the case of sociolinguistic knowledge, there is no fixed or correct choice. When one behavior (response) is given, it cannot be labeled correct or incorrect because there is no correct behavior. The behavior however, will be judged as to which point on the scale it represents. However, in the case of grammatical knowledge, there is one fixed or correct choice, and that choice has been decided. If other than that choice is selected, the choice is incorrect.

Despite this fact, this present study attempts to measure scalar

behavior with a multiple-choice test. It attempts to discover the feasibility of testing sociolinguistic competence with a multiple-choice test. Although this is no contradiction to the facts above, it is predicted here that a multiple-choice test can be a feasible in testing sociolinguistic competence, a scalar behavior. A preferred utterance or pattern of utterance needs to be chosen so that any other patterns that occur will be deviations from the chosen one. Thus, one preferred utterance or pattern of utterance will become the correct answer for each item in the multiple choice test.

As a summary, the section on issues in testing elaborated in this chapter was intended to discuss issues which may also be encountered in this study. Lastly, the sections on previous studies on testing communicative competence and speech acts were presented so that the present study can relate its content and implications to those studies. The studies on oral interviews were covered so that comparison can be made as to the practicality of the multiple-choice test in assessing sociolinguistic competence. The findings of studies on speech acts were included so that assumptions can be made of what could be expected from this study.

The details of the present study are found in the following chapter. Specifics of the instrument, subjects, task of the study are included.

Chapter Three

Methodology

This chapter provides details on the methodology and the design of the present study. The study will use the speech acts of 'apology' and 'replying to compliments' as the content of the test.

3.1 Purpose and Background: Revisited

This present study is ask non-native speakers to make a yes/no answer or judgment about knowledge that is, to some degree, a matter of opinion, not fact. Thus, this becomes a norm-referenced test; that is, it measures what non-native speakers do against a pattern of what native speakers have decided. As such, the solution to the problem is two folds. First a norm has to be established. This is accomplished by giving the multiple choice test to native speakers, and then determining whether native speakers have strong preferences for one or another choice (distracters). A sufficient number of items on which native speakers demonstrate a high degree of agreement will be identified, thus establishing a reliable norm for this type of knowledge among a comparable population of native speakers.

Then, the second step is to compare the performance of non-native speakers on the total number of items with the performance of native speakers on the same item. The degree of similarity (high score) or difference (low score) will then be a measure of the similarity of the knowledge of the test-takers to that of the norming populations. This test is also a proficiency test; that is, it is measuring a globally defined element of sociolinguistic competence. Thus, it is only the

aggregate or total score on the test that will be of significance. It is presumed that the performance on individual items is not very significant. It is the score on the overall test rather than on individual items that will determine the level of NNS performance relative to that of native speakers. It will be possible for native speakers to make the same choices on some of the questions as non-native speakers do, but is unlikely to happen on the test as a whole. If the reliability is high, then it can be asserted that the test, overall, and as a unit, is a good measure of sociolinguistic competence. The differences between native speakers and non-native speakers can be seen as a difference in their overall sociolinguistic competence. Performance on individual items may vary, but if the test is a reliable one, performance on the overall test will still be a valid and reliable measure of the degree of similarity of the sociolinguistic knowledge of the two populations.

The following are the details of the design of the present study.

3.2 Instrument

Unlike other studies, this study uses a multiple-choice test to measure sociolinguistic competence among ESL speakers. It is a criterion-referenced test of which the criteria are predetermined by the administration of the multiple-choice test to native speakers, thus establishing a norm or set of criteria against which the performance of the non-native speakers can be assessed.

3.2.1 *Development of the Instrument*

The two speech acts: apology and reply to compliments were selected because of the work done on them previously. The development of the instrument for the present study proceeded in several stages.

3.2.1.a *Instrumentation No. 1 (Open-Ended Questions)*

First, an open-ended test consisting of six situations each calling for the production of each of the speech acts was prepared. The purpose of this test was solely to generate responses for each situation of the speech act. The subjects were to write down the responses or utterances that they are most likely to say or what they would say in the situations given. The responses or utterances should be those which the subjects would naturally say or respond to when given the situations. This test was then given to twenty-five native speakers of English to gather possible forms they might use in each situations. The test was also given to twenty-five non-native speakers to determine if they would use different forms.

Then the responses from the open-ended test were analyzed. The responses or patterns of response (semantic formulas) which were supplied most frequently in each situation was taken as the preferred answer of the multiple-choice test. One dispreferred response from the native speakers and two from the non-native speakers were selected as the three distracters for each situations. These responses were those with the lowest percent of occurrences. Other criteria that were used in selecting the dispreferred responses as distracters were relevancy, source and appearance (grammatical versus not grammatical).

First of all, the distracters were chosen because of their

relevance to the situations. Even though they were dispreferred responses, each response is relevant to the speech act situations described in the test. Also, they were generated from the same test administration. If they were taken from somewhere else, they might not have the same level of relevancy. Furthermore, all the responses were generated from the same subjects, college students, and by using the same mode of testing (written). Because they were generated from the same source, by using the same mode and from the same test administration, the responses sound more natural than if they were made up. Even though some changes were made in the wordings, they still sound natural.

3.2.1.b Instrument No.2 (Multiple Choice Test)

Based on the information gathered from the open-ended test, a multiple-choice test was created using the same situations as the open-ended test. This is the second stage in the development of the instrument. This version serves as the multiple choice test of the study. The test contains twelve items, and asks the subjects to choose one preferred choice from the four provided.

All the situations in the test are related to or relevant to college students at Michigan State University, thus addressing the content validity of the test. They are designed so that each subject will respond differently according to his or her knowledge of sociocultural rules. For example, a low proficiency student might only say "I'm sorry", after an intense offense such as bumping into an old lady and knock her down, because he or she does not know the most appropriate response. However, a more proficient student might use "I'm

very sorry" and " Let me help you," because he or she knows more the appropriate response and feels necessary to offer help. In the speech act of apology, the intensity of the offenses varies from one another. The addressees of the situations range from close friends to total strangers. The situations are selected to require subjects to respond by including different components mentioned above for each situation.

In the speech act of replying to compliments, the situations range from less formal to more formal. The addressees range from close friend to stranger. The situations are intended to elicit the general semantic formulas of replying to compliments in English. One can reply to compliments by staying silent, complimenting, denying the fact, or most generally by thanking.

Overall, the situations in the test are intended to elicit the general semantic formulas of responses used by the native speakers and also by the non-native speakers, particularly Malay learners of English. The results will then be compared.

The following are the details of the subjects for each administration of the instruments.

3.3 Subjects

For the open-ended test, the subjects were twenty-five native English undergraduate and graduate students majoring in various disciplines at Michigan State University, and twenty-five non-native English students (specifically they are Malay learners of English) at Michigan State University.

3.3.a Native Speakers

As mentioned briefly above, they come from various disciplines. However, none of the 25 native speakers of English who participated are Linguistics, or Teaching of English as a Second Language (TESL) or Communication majors. They are presumed to have little knowledge about linguistics particularly on the subject matter that this study is interested in. Thus, their production in the open ended test can be expected to be more natural (according to their best intuition and judgment). The open-ended questionnaires were distributed at random and have no formal settings. For example, most of the open-ended questionnaires were handed to subjects sitting at the lobby in the 'Union' building at MSU. They were first questioned if they were at all interested to participate in the research. The consent letter was also given for them to get a general idea of the research. Upon completion, subjects were asked to submit the questionnaires to the researcher of this study who was located at the far end of the Union lobby.

The second instrumentation (Multiple Choice Test) also took the same course of distribution. However, none of the 25 subjects in the instrumentation #2 (MCT) are the ones participated earlier.

3.3.b Non-Native Speakers

Their ages ranges from nineteen to forty-three. All the Malay students have been in the US for at least one year, and are currently enrolled in regular university courses. Similar to native speakers, none of the Malay students too are associated with linguistics, TESL or communication. They too know very little about linguistics and the

subject matter of this study. They are linguistically naive subjects. Thus, their involvement and production in the open-ended are expected to provide various responses and problems to be used as possible distractors in the multiple-choice test. No formal settings were involved in gathering the data from the Malay subjects. The subjects were obtained from volunteers who wanted to participate. The subjects were told that if they agreed to participate, there are 2 instrumentations involved, and they were advised that they follow the course. However, they have the final say of whether to proceed or discontinue the research at any time. The open-ended questionnaire was then distributed, and the subjects were asked to return them upon completing the task. The subjects were also advised to take their time and feel relaxed when filling out the questionnaire. Instruction such as 'try to answer how you would normally say in the given situations' was also stressed when distributing the questionnaires.

Instrumentation 2 or the multiple-choice test was distributed to the same subjects approximately two and the half weeks after the open-ended questionnaires were collected.

The following chapter describes the analyses of the test results. Descriptive statistics includes mode, median, mean, range and variance. The reliability, and t-test are also included.

Chapter Four

Results

This Chapter reports the overall results of the administration of the multiple-choice test. First, the performances of both the native and the non-native (Malay learners of English) are presented separately. Then, the performance of the native speakers and non-native speakers are compared.

4.1 Performance of Native Speakers

As expected, the native speakers performed very well on the multiple choice test. Of the 25 subjects, nineteen of them scored between 10 and 12. Thus, approximately 76% of them scored very highly. The mean for the 12 items was 10.640. The highest score by the native speakers was 12, and the lowest score was 7. Thus, the range of their score for the 12 items was 05, with a standard deviation of 1.524. It can be concluded that most of the items in the MCT were fairly easy for the native speakers.

4.2 Performance of Non-native Speakers

The non-native speakers also performed well on the test, but not as well as the native speakers of English. Scores ranged from 5 to 9 for the twelve items MCT. The mean for the twelve items was 6.520, with a standard deviation of 1.418. The Malay subjects seem to have more difficulty with the test. The following is a summary of the basic statistics for both the native speakers and non-native speakers involved in this study.

Table 4.1: Descriptive Statistics on the findings between the 02 groups (NS versus NNS) tested in this study.

| Speakers | Native Speakers | Non-Native |
|--------------------|-----------------|------------|
| NUMBER OF CASES | 25 | 25 |
| MODE | 12 | 5 |
| MEDIAN | 9.5 | 7 |
| MEAN | 10.640 | 6.520 |
| RANGE | 12-7= 5 | 9-5= 4 |
| STANDARD DEVIATION | 1.524 | 1.418 |
| VARIANCE | 2.323 | 2.010 |
| STANDARD ERROR | 0.305 | 0.284 |

4.3 Factors Affecting Variance within the Non-Native Speakers

Two factors existing within the non-native speaker population were not investigated directly but could have had an effect on their performance. They were length of stay in the US and TOEFL scores.

4.3.1 TOEFL Scores

Twenty three of the twenty five non-native speakers reported their TOEFL scores. The TOEFL scores are divided into three groups: between 451 and 500; between 501 and 550, and 551 and higher. Those who have TOEFL scores between 501 and 550 top the other three groups in terms of the mean scores for the multiple-choice test, the group with TOEFL scores of

551 and above place second, followed by the group with TOEFL scores between 451 to 500. Below are the mean scores for the above variable.

Table 4.2: NNS's score on the 12 items MCT based on TOEFL scores

| TOEFL Scores | 12 Items MCT MEAN | STANDARD DEVIATION |
|----------------|----------------------|--------------------|
| 501 - 550 | 7.700 | 1.160 |
| 550 and higher | 6.400 | 0.894 |
| 451 - 500 | 5.000 | 0.000 |

4.3.2 Length of Stay

The variable length of stay is divided into three groups: one year or less, between one year and three years, and three years and more. The second group (between one year and three years) has the highest mean score for the multiple-choice test, followed by the group with three years or more, and the group with one year or less place last. Both length of stay and TOEFL scores do not seem to correlate in a simple linear fashion with the means on the multiple choice test. Below are the mean scores for the above variable.

Table 4.3: NNS score on the 12 items MCT based on their length of stay in US.

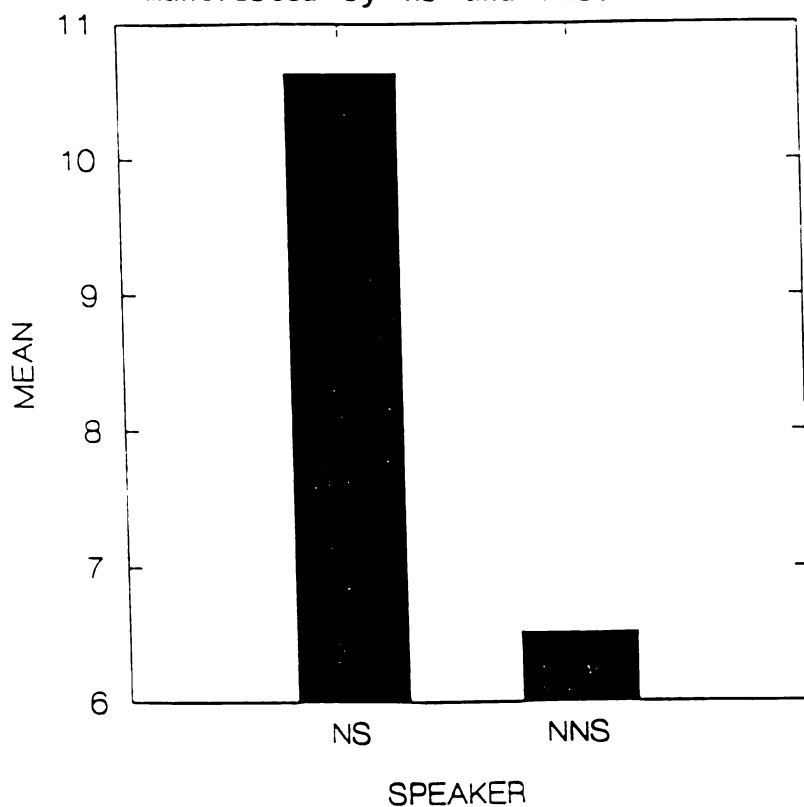
| Number of years | 12 Items MCT MEAN | STANDARD DEVIATION |
|----------------------|----------------------|--------------------|
| One to Three years | 7.500 | 1.690 |
| Three years and more | 6.333 | 1.225 |
| One year and less | 5.625 | 0.744 |

4.4 Native & Non Native Speakers: A Comparison

Overall, the native speakers performed better than their non-native counterpart, thus adhering to the norms defined for native speakers by the test construction process. The modes for native speakers are also higher than their Malay counterparts. Generally, based on the range obtained for the two groups, it seems that the non-native speakers perform very consistently; that is, everyone seemed to perform at about the same level. There were only a few who did not score highly on the multiple-choice test. On the other hand, the performance of the NS subjects seemed to be slightly more variable (range of 5 as compared to 4 obtained by the NNS).

A t-test (paired samples T-Test) was calculated to determine the significance of the difference in performance between the two population of native speakers and of the non-native speakers. The results of the t-test indicate that the probability is less than 0.001 (>0.001). This means that the probability that the hypothesis of this study is rejected is less than 0.001. That is the findings or results showing the differences between group 01 and group 02 (Ns versus NNS) are significant. In addition, the significant differences in performance between the two groups suggest that 'overlap' between the two groups is not likely to occur (standard error for NS: $X + SE = 10.6 + 0.305$, and for NNS: $6.52 + 0.284$ This can be further illustrated by the following graph:

Figure 4.1: Bar graph illustrating the performance means manifested by NS and NNS.



As a conclusion, the results indicate that the performance of the two population differed significantly and that the standard error did not overlap.

4.5 NNS Performance Revisited-An Interpretation

Generally one would expect that subjects with the highest TOEFL scores and subjects that have lived the longest in a native English speaking nation would have done better in a test such as the MCT used in this study. This is not the case revealed by the NNS subjects in this study. Interestingly, the results of the scores on the MCT do not show any

correlation in a simple linear fashion. For example, in the TOEFL score, the subjects that scored the highest on the MCT were the ones who have a TOEFL score ranging from 501-550. Subjects with TOEFL scores ranging from 550 and higher place second best compared to the group that performed well on the MCT. Subjects with TOEFL score 451-500 place last within the three groups. Further analysis revealed that the differences between the groups are significant with a probability of 0.000 (< 0.0001). However, which group is more significant than the other was not known. Thus, a 'Tukey HSD' multiple comparisons was performed, and it appears that the differences in all three groups are significant (< 0.05). For example, comparison between group 1 and 3 reveal 0.034; group 2 and 1 give 0.028 reading; and group 2 and 3 yield 0.000.

Similarly, subjects who have lived in USA the most (three years and more) do not performed better than subjects who have lived in USA ranging from one year to three years. Their mean scores as reported in table 4.3 are 6.333 and 7.500 respectively. Further analysis show that the differences between the three groups are significant with a probability of 0.025 (< 0.05). However, the Tukey HSD multiple comparison reveals that group 1 and 3 is the only comparison that is significant (0.020- < 0.05).

Thus, the one question that need to be raised here is how can the findings pertaining to the subjects' TOEFL score and length of stay be interpreted? This is certainly not an easy task. One can only speculate. However, two factors that could possibly influence the results manifested by the NNS are the 'attitude' and 'motivation' of the NNS themselves. It might not be surprising if the subjects with TOEFL score

ranging from 501-550 and the subjects who have lived here one to three years scored the highest on the MCT. Perhaps these NNS subjects are 'highly motivated' and have the right attitude towards the target language. In addition, it might also be their desire or primary objective to improve their second language. Thus, with the right motivation and the right attitude they would pursue every opportunity to use the target language. 'Practice makes perfect' - although the subjects might not achieve native like accent and quality, certainly these subjects would improve their mastery in using the target language and this might well include the performance of speech acts .

The failure to have the right attitude and low level of motivation to improve their knowledge and mastery of the target language could be reflected in their performance. This could be the case with the subjects with the highest TOEFL score and the subjects who have lived the longest in USA who performed collectively as second best in this study. Having a low level of motivation and the wrong or negative attitude towards the target language certainly would inhibit one from pursuing every given opportunity to improve his or her second language. They might not be socializing as much. They may not have a string of native English speakers as friends that they can talk to or discuss. Most unfortunate if they limit or choose only to socialize with speakers who speak the same native language (Malay). Certainly, spending a lot of time in the latter circle of friends would not add or provide new experiences, new vocabularies, mastery to use proverbs, and certainly knowledge to use different speech acts effectively would also be hurt. It is important to note here that this is only a speculation. It is difficult to be sure as to what factors or variables could have influenced the findings

manifested by the NNS subjects.

Chapter Five

Discussion

As shown by the results, this study has shown that a multiple-choice test of sociolinguistic competence is feasible, at least in principle. The criteria that indicate the feasibility of this instrument are its validity, its reliability, and its practicality.

5.1 Validity

The validity of a test is the extent to which it measures what it is supposed to measure. It must aim to provide a true measure of the particular skills or body of knowledge which it is intended to measure. There are a number of types of validation applied to tests; however, the most important type in this study is content validity. This type concerns the content of the test itself. The test should be constructed so that it contains a representative sample of the course or the skill being tested. Thus, the objectives of the course should be the guidelines in the selection of the test items. For example, if the test is designed to measure mastery of tenses in English, the test should be expected to provide enough representative samples of the phenomenon being tested, and nothing else. There is, however, no statistical procedure used to measure content validity. A rigorous analysis by the test writer will determine the content validity of the test.

The particular speech acts used in this study are not of great importance to the test. They were selected because they represent demonstrated aspects of sociolinguistic knowledge, and thus, they are part of that body of knowledge. Since they represent part of that body

of knowledge, the speech acts of apology and reply to compliments are indications of the content validity of the test.

For the test used in this study, there is also an indirect positive evidence of content validity. The results of the test administration showed that there was no simple linear fashion correlation between the test scores and the TOEFL scores of the non-native speakers. Therefore, it can be claimed that the test measured something other than the linguistic proficiency levels of the subjects, or what is measured by the TOEFL. If high correlations were found between the test scores and TOEFL scores however, then, it would be possible that the test was measuring linguistic proficiency, and not sociolinguistic competence.

Given the procedures used to construct this test, the type of information included on it, and the relatively low correlation with known measures of linguistic competence, it can be concluded with some assurance that the knowledge the test is measuring is in fact, sociolinguistic knowledge and that the test has content validity.

Indirectly, two factors which will be discussed later in this chapter relate to the content validity of the test. They are the length of test and the homogeneity of the items. They relate to the adequacy of the sample and the appearance of the test respectively. An adequate number of items will ensure that adequate representative samples of the content are included in the test. Also, the content of whatever is aimed to be measured must be carefully defined, or the test will be measuring something else.

Most important among the results of this study, and a measure that contributes to the validity of the test, was the t-test, in which two

means (of native speakers and non-native speakers) were compared. The results of the test showed very high discrimination between native speakers and non-native speakers. This showed clearly that the test was effective in discriminating between the knowledge of the native speakers and the non-native speakers.

5.2 Reliability

In general, reliability refers to the stability or consistency of the test scores. Reliability is a necessary characteristic of any good test. A test cannot measure anything well unless it measures it consistently. It is the extent to which how well and consistently the test measures what it does measure.

Reliability can be measured in several ways. For the purpose of this study, a measure of the internal reliability of the test was used. Internal reliability differs slightly from other reliability measures such as stability, test-re-test, or split-half. According to Hatch and Farhardy (1982), there are four factors affecting the reliability of a test. The factors are:

1. length of test
2. homogeneity of test
3. variability of group ability
4. sufficient test-taking time

With regard to the present study, the first three factors from the above list seemed to have affected the reliability of the multiple-choice test. The reliability of the test could have been higher if any or all of the four factors were increased.

First, the length of the test affected the reliability of the test

quite significantly. If the test were longer, it is possible that it could be more reliable. What is of concern here is the number of items on the test. The test for this study contained twelve items, which is a relatively small number. If the items were doubled in number, the reliability might be increased significantly. That is, all other things being equal, the more items on the test, the more reliable the test is.

Thus, the reliability of the MCT in this study would probably be higher if there were greater number of items included in the test. The instrument would have been a more reliable one if the number of items were increased.

The second factor that affected the reliability of the test was the homogeneity of the test items. If the test items were testing the same trait, the reliability would be higher. Although it is true that the MCT in this study was testing sociolinguistic competence, several different sub-traits were included. Two different speech acts were used as the content of the test; thus, there were two different sub-traits being tested, even though all of them were called sociolinguistic competence. If the traits were more or less similar or homogeneous, the reliability of the test could have been higher and thus, the test would have been a more reliable instrument. If the traits were made homogeneous, then the objectives of the test to measure some aspects of sociolinguistic competence (two speech acts) would have been changed, although it is possible that the reliability would be increased. The test would then measure only one trait, which is in contrast with the objective of the test - to measure the two traits or sub-traits.

The final factor that affected the reliability of the multiple-choice test in this study is the variability of performance within the

group. If the subjects of the test had a wider range of ability, the test reliability could be increased. The subjects of the present study have a rather narrow range of overall ability.

Thus, the above factors could have affected the reliability of the test, but not significantly. If however, the abilities of the subjects had a wider range, the reliability of the test could have been increased. The third criterion is the practicality of the test itself.

5.3 Practicality

This study has shown that it is possible to measure sociolinguistic competence with a multiple-choice test. This means that the knowledge can be measured more easily, faster and more reliably than with other instruments such as the oral interview.

However, there are other basic practical questions that need to be answered. For examples, what is the purpose of testing sociolinguistic competence?; Why should such knowledge be tested, or should it be tested at all? Is it necessary to test sociolinguistic competence? These are some of the questions which should address some practical implications of the test.

A good knowledge of sociolinguistic rules would enhance the communication one is engaged in. It is especially useful to non-native speakers of English, who possess such knowledge but different from English. A study by Carpenter (1983) on "foreigner talk" indicated that foreign students have difficulties in communication during appointments with their professors. Carpenter reported that the appointments showed lack of coordination in turn taking and in topic flow. The students seemed to include more "backtracking and attempt at clarification,

moving forward choppily, chiefly through the professors' initiation, in question-answer sequences or in professors' explicit instruction" (in Wolfson and Judd (eds):187)

The appointments between professors and foreign students seemed to contain little discussion of "extended topics and few of the comments, digressions, negotiation, and references to shared beliefs which tended to characterize the native speaker interactions". Thus, it indicates that the social, cultural and linguistic resources, which many foreign students probably do not have are essential components of language ability. They need this information in order to be better participants in communication.

A recent example of this kind of deficiency is the difficulties faced by foreign teaching assistants in many American colleges including Michigan State University (as revealed by the State News in 1995). These teaching assistants do not have adequate social or cultural knowledge of English which would allow them to become better assistants. Added to the deficiency in linguistic knowledge, these teaching assistants often have difficulties explaining to the students what they are trying to convey and understanding what the students are asking from them. The students in turn, are faced with the problem of understanding what the teaching assistants are trying to say or explain mainly because of the linguistic deficiency on the teaching assistant's part.

Therefore, there is a practical need for sociolinguistic knowledge. It helps people to understand one another better. This would decrease the possibility of any misunderstanding. Sociolinguistic knowledge is not a necessary but a desirable component of language ability in higher academic settings. It should be tested especially with

foreign students who intent to study in American colleges. However, it should not be used as a major instrument nor should it be used as a major determinant of any ability of the students. However, a test such as this could be used in research or study on sociolinguistic knowledge. It may also be useful or practical to use it with other test batteries. But, before this can be confirmed, further research needs to be done to find further practical use for such a test.

One further practical issue is the matter of the linguistic content of a test such as this one. A test of sociolinguistic knowledge necessarily involves using some complex linguistic forms, i.e. vocabulary and structures. The language forms will, at some point, become problems to some test takers. Such test cannot be used with lower level students because at some point of the test, the reading and the vocabulary are going to make it a linguistic proficiency test, rather than a sociolinguistic test. The students will encounter comprehension problems, which will then contribute to low scores on the test. The vocabulary also will hinder the students from scoring well on the test because they do not understand some of the words in the test. Thus, the students would be taking test of reading and vocabulary, and not test of sociolinguistic competence.

5.4 Evaluation

After analyzing the results of the test, I believe that this study has succeeded, to a certain degree, in demonstrating that a MCT of sociolinguistic competence is a feasible instrument. The reliability measures showed that the instrument developed in this study is a usable one even though the measures were rather low. However, there are very

good reasons for the low reliability (as discussed in the previous section). Thus, scalar behavior can, after all, be measured with a MCT.

Despite all the shortcomings, the test was successful at measuring the sociolinguistic competence of the subjects. The test did show content validity, that it was measuring sociolinguistic competence, and not other traits, such as reading or vocabulary. However, the instrument could be improved, which could then reveal more accurate and reliable results.

5.5 Suggestions

Due to the shortcomings mentioned in the previous sections, any attempt to further extend the present study or to replicate its should consider the following suggestions. These suggestions, if applied, would change some statistical results of the study, and possibly decrease the number of shortcomings and drawback that may exist with this type of instrument.

The first item to be considered seriously is the number of items in the test. The reliability of a test depends heavily on the number of test items. The greater the number of test items, the more reliable the test will be. The test in this study had only twelve items, which could hardly generate a high reliability. Any attempt to replicate or further this study should at least double the number of items. This should increase the reliability significantly to a degree that the test can be strongly claimed as a reliable instrument.

Even though not as important as the numbers of items, the number of subjects taking the test should also be considered. A large number of subjects would increase the representatives of the population being

tested. If a smaller number of subjects are involved, the number might not adequately represent the population being measured. Thus, the measure would not be as reliable as it is expected. However, a large number of subjects would represent the population more accurately.

Another item for consideration is the variability of group ability. This is one of the factors affected the reliability measure. A wider range of ability would probably increase the reliability measure. A narrow range of ability will not ensure high reliability because the test will be scored at a relatively similar level. However, if the subjects had a wider range of ability, the test scores would be widely spread, and not concentrate to a particular level.

In their study, Cohen and Olshtain (1983) presented the test items in random order to avoid any "response set effect". However, the present study did not randomized the items to keep the directions simple. The items were not randomized because they represented two speech acts groups. Each group was introduced with a different instruction. If the items were randomized, a different direction would have to be given before every item; which is very wordy and meticulous. Thus, randomization of test items was not possible for the present study, in both instrumentation: open-ended and the MCT. The "response set effect" was not expected to be very significant in the MCT because the choices for each item were not similar.

However, the effect was more significant in the open-ended test. The responses given by the subjects seemed to show the effect. For example, in the speech act of reply to compliments, the majority of the responses by the native speakers were "thank-you." Even though different situations of reply to compliments were presented in each item, the

response for the first item seemed to have effects on the remaining items in that group.

Thus, anyone who attempts to utilize such test should consider this suggestion if a more accurate and detailed outcomes are to be expected.

In conclusion, sociolinguistic knowledge is not a necessary but a desirable component of overall language ability. However, further research needs to be carried out to improve the reliability of such test. Also, further study should be done as to whom should be tested and at what level should be tested.

What this study has done is to investigate the feasibility of measuring sociolinguistic competence with a short and practical MCT. The results of the study show that such a test has a low but a usable degree of feasibility.



References

- Austin, J.L. (1962). *How to do things with words*. Oxford: Oxford University Press.
- Bergman, M.L. & Kasper, G. (1993). *Perception and Performance in Native and Non-native Apology*, in Kasper, G. & Blum-Kulka, S. (eds.), p.82-107.
- Blum-Kulka, S.; Shoshana; Juliana, H.; Kasper, G. (eds) (1989). *Cross-Cultural Pragmatic: Request and Apologies*. Norwood, NJ: Ablex.
- Bodman, J.; Eisenstein, M (1988). May God Increase Your Bounty: The Expression of Gratitude in English by Native and Non-Native Speakers. *Cross Currents*. 15:p.1-21.
- Bongaerts, T. & Poulisse, N. (1989). Communication Strategies in L1 and L2: Same or Different. *Applied Linguistics*, 10,3, p.253-267, Oxford University Press.
- Canale, M. & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing, *Applied Linguistic*, i,1, p.1-47.
- Canale, M. & Swain, M. (1983). From Communicative Competence to Communicative Language Pedagogy, in Richards, S. & Schmidt, R. (eds.) *Language and Communication*, p.3-27.
- Carpenter, C. (1983). Foreigner Talk in University Office-Hour Appointments, in Wolfson, N. and Judd, E., (eds.) *Sociolinguistic and Language Acquisition*, pp.184-194
- Carroll, C. (1980). *Testing Communicative Performance*. New York, Pergamon Press.
- Cohen, A.D. et.al (1981). Developing a Measure of Sociocultural Competence: The Case of Apology, *Language Learning*, 31, p.113-134.
- Cohen, A.D. & Olshtain, E. (1993). The Production of Speech Acts by EFL Learners. *TESOL Quarterly*, 27, p.33-56.
- Creese, A. (1991). Speech Act Variation in British and American English, *PENN Working Papers*, 7,2, p.37-58.
- Davies, A. (1989). Communicative Competence as Language Use. *Applied Linguistics*, 10, 2, p.157-169, Oxford Uni.Press

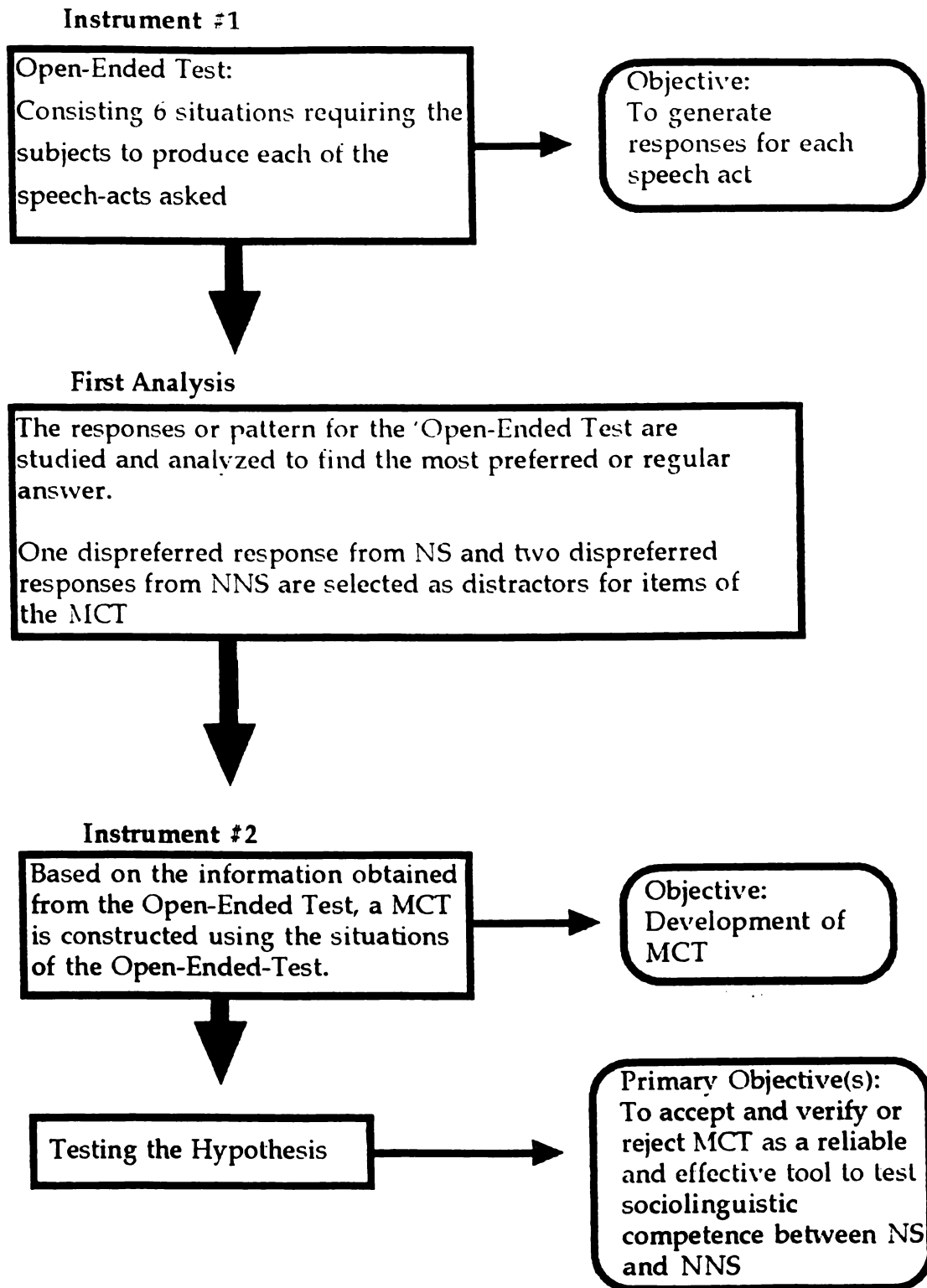
- Deborah, T. (1990). The Pragmatic of Cross-Cultural Communication. *Applied Linguistics*, 5,3, p.189-125
- Edmonson, W.; et.al (1990). Learning the Pragmatic of Discourse: A Project Report. *Applied Linguistics*, 5,2, p.113-127.
- Gass, S. & Selinker, L. (1994). *Second Language Acquisition: An Introductory Course*. Hillsdale, NJ.: Lawrence Erlbaum Associates, Publishers.
- Gass, S. & Neu, J. (1995). *Challenges to Communication in Second Language Acquisition (SLA)*. Berlin: Mouton de Gruyter.
- Han, C.H. (1992). A Comparative Study of Compliment Responses: Korean Females in Korean Interactions and In English Interactions. *Educational Linguistics*, 8,2, p.17-31.
- Harris, D.P. (1969). *Testing English as a Second Language*. New York: McGraw-Hill Books Co.
- Heaton, J.B. (1975). *Writing English Language Test*. London: Longman.
- Holmes, J. & Brown, D.F. (1987). Teachers and Students Learning About Compliments. *TESOL-Quarterly*, 21,3, p.523-46.
- Hymes, D. (1964). Directions in Ethnolinguistic Theory. *American Anthropologist*, 6,3(2), p.6-56.
- Hymes, D. (1974). *Foundation in Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.
- Ingram, E. (1977). Basic Concepts in Testing. In Allen, J.P.B, & Davies, A. (eds). *The Edinburgh Course in Applied Linguistics*. Oxford; Oxford University Press.
- Jones, R.L. (1979). *The Oral Interview of the FSI*, in Spolsky, B. (eds.) *Some Major Tests*. Arlington, VA: Center for Applied Linguistic, p.104-116.
- Kasper, G. & Blum-Kulka, S. (1993). *Inter language Pragmatic*. New York: Oxford University Press.
- Kasper, G.; Dahl, M. (1991). Research Methods in Inter language Pragmatic. *Studies in Second Language Acquisition*. 13, p.215-247.
- Munn, A. (1995). *Series of Lectures in Syntax in Linguistics 434*, Fall Semester 1995. Room C112, Wells hall, Michigan State

University.

- Olshtain, E. & Cohen, A.D. (1985). Cross cultural Pragmatic and the Testing of Communicative Competence, *Language Testing*, 2, p.16-30.
- Olshtain, E. & Cohen, A.D. (1991). The Learning of Complex Speech Act Behavior, *TESL Canada Journal*, 7, p.45-65.
- Petrey, S. (1990). *Speech Acts and Literary Theory*. Routledge, Chapman and Hall, Inc.: New York.
- Palmer, A.S. (1972). Testing Communication. *IRAL*, 10, p.35-45.
- Pomerantz, A. (1978). Compliment Responses; Notes on the Cooperation of Multiple Constraints. In J.Schenkein (Ed.). *Studies in the Organization of Conversational Interaction*. New York: Academic Press, p.79-112.
- Preston, D. (1989). *Sociolinguistics and Second Language Acquisition*. Oxford: Basil Blackwell.
- Rintell, E. & Mitchell, C.J. (1989). *Studying Requests and Apologies: An Inquiry into Method* in Blum-Kulka, S.; House, J. & Kasper (eds), p.248-272.
- Searle, J. (1965). "What is a speech act?" In M.Black, ed., *Philosophy in America*. Ithaca: Cornell University Press.
- Searle, J. (1969). *Speech Act*. Cambridge: Cambridge University Press.
- Stalker, J. (1989). Communicative Competence, Pragmatic, Functions, and Accommodations. *Applied Linguistics*, 10, 2, p.182-193.
- Spolsky, B. (1989). Communicative Competence, Language Proficiency, and Beyond. *Applied Linguistics*, 10, 2, p.138-155, Oxford Uni.Press.
- Thomas, J. (1990). Cross-Cultural Discourse as 'Unequal Encounter Towards a Pragmatic Analysis. *Applied Linguistics*, 5, 3, p.223-235.
- Valdman, A. & Moody, M. (1979). Testing Communicative Ability. *The French Review*, 52, 4, p.221-234.
- Varonis, E. & Gass, S. (1985a). Miscommunication in Native/Non-native Conversation. *Language in Society*, 14, p.327-343.

- Varonis, E. & Gass, S. (1985b). Non-Native/Non-Native Conversation: A Model for Negotiation of Meaning. *Applied Linguistics*, 6, p.71-90.
- Vollmer, H.J., & Olshtain, E. (1989). *The Language of Apologies in German*, in Blum Kulka, S., House, J. & Kasper, G. (eds.), p.197-218.
- Wahlstrom, B.J. (1992). *Perspectives on Human Communication*. University of Minnesota: Wm.C.Brown Publishers.
- Widdowson, H.G. Knowledge of Language and Ability for Use. *Applied Linguistics*, 10,2, p.128-137, Oxford Uni.Press.
- Wolfson, N. (1989). *Perspectives: Sociolinguistics and TESOL*. Cambridge: Newbury House.
- Wolfson, N. & Elliot, D. (eds.) (1983). *Sociolinguistics and Language Acquisition*. MA: Newbury House.
- Wolfson, N. (1981). Compliments in Cross-Cultural Perspectives, *TESOL Quarterly*, 15,2, p.117-124.

APPENDIX A



APPENDIX B
Questionnaire

- Please circle: a. I am a native speaker of English
b. I am not a native speaker of English

You are asked to reply to the description of the different situations given in this questionnaire. Write what you think you are most likely to say. Remember that there is no CORRECT and WRONG or POLITE or IMPOLITE answers here.

A. REPLYING TO COMPLIMENTS

Directions: Given different situations below, how would you reply to compliments in each situation:

1. Your roommate says that you look nice today.

2. You invite a friend to have dinner at your place and after the dinner, he or she says it was delicious.

3. Your English instructor says you did an excellent job on your paper.

4. You just had a hair-cut and your friend says you look better with the new hair cut.

5. Your friend is impressed with your apartment and says that you are a good decorator.

6. Your next door neighbor (looking at your back yard) says that your rose garden is beautiful.

B.TO APOLOGIZE

Directions: Given different situations below, what would you say?

1. Dropping an eraser while handling it back to a friend.

2. Forgetting a study group meeting with your classmates.

3. Stepping on someone's foot while standing in line to get into a theater at the Meridian Mall.

4. Damaging your roommate's expensive camera.

5. Submitting an assignment late to an instructor for the second time.

6. While backing up your vehicle, you knock over your neighbor's garbage cans.

Thank you for participating. Have a good day.

APPENDIX C
Questionnaire

Age:.....

- Please circle:
- a. I am a native speaker of English
 - b. I am not a native speaker of English
if you are not a native speaker of English,
please continue responding to the following
questions:
 - b.1. Length of stay in US.....
 - b.2. TOEFL/MSU Placement Test.....

I am interested in what you might say in various social situations. The following are some brief descriptions of situations. Choose one of the choices which is what you are MOST LIKELY to say. Do not spend too much time reading them; just circle your choice.

A. REPLYING TO COMPLIMENTS

Directions: Circle the most appropriate apology statement in each situation below.

1. Your roommate says that you look nice today.
 - A.You think so?
 - B.You don't look so bad yourself.
 - C.No. I'm not.
 - D.Thank you
2. You invite a friend to have dinner at your place and after the dinner,he or she says it was delicious.
 - A.Thank you.
 - B.Oh, it's just something I threw together.
 - C.Thank you, I'm glad you liked it.
 - D.Thank you, but all the credit should go to my wife.
3. Your English instructor says you did an excellent job on your paper.
 - A.Thank you.
 - B.Really?
 - C.Thank you, I'm glad you enjoyed reading it.
 - D.Thank you, I really put a lot of work on it.
4. You just had a hair-cut and your friend says you look better with the new hair hair-cut.
 - A.Thank you.
 - B.Hey, you look nice too.
 - C.It's in the genes
 - D.The compliment should go to my barber.

5. Your friend is impressed with your apartment and says that you are a good decorator.
 - A.Thank you, I'm glad you like it too.
 - B.You think so?
 - C.I just want to make it fit my personality.
 - D.That is what I do best.

6. Your next door neighbor (looking at your back yard) says that your rose garden is beautiful.
 - A.Would you like one?
 - B.Yes, the rain helps out a lot.
 - C.Thank you.
 - D.You can do it too with the new fertilizer.

B.TO APOLOGIZE

Directions: Circle the most appropriate apology statement in each situation below.

1. Dropping an eraser while handling it back to a friend.
 - A.Gee, I'm sorry about that!
 - B.(You say nothing or silent)
 - C.Oops, sorry!
 - D.Excuse me!

2. Forgetting a study group meeting with your classmates.
 - A.I'm sorry, I completely forgot.
 - B.Why didn't you call me?
 - C.Sorry!
 - D.Please accept my apologies.

3. Stepping on someone's foot while standing in line to get into a theater at the Meridian Mall.
 - A.Are you OK?
 - B.Pardon me!
 - C.Excuse me.
 - D,I'm very sorry,how clumsy of me!

4. Damaging your roommate's (John) expensive camera.
 - A.I'll pay for the damage.
 - B.Sorry John.
 - C.John, I broke your camera, but I don't know how it happened.
 - D.I'm sorry about the camera John, I'll pay for it somehow.

5. Submitting an assignment late to an instructor (Dr.Mason) for the second time.
 - A.Dr.Mason, this report is late again, and I am really sorry.
 - B.Sorry!
 - C.I'm sorry this is late, but it won't happen again.
 - D.I don't believe it, but my report is late again.

6. While backing up your vehicle, you knock over your neighbor's garbage cans (Please assume that the neighbor witnessed the situation).
- A. I'm very sorry.
 - B. I'm very sorry, I just didn't see them.
 - C. I'm very sorry, how clumsy I was.
 - D. I'm sorry, I'll pay for the damage.

Thank you for participating. Have a good day.

MICHIGAN STATE UNIV. LIBRARIES



31293014053379