

Discourse Markers as Predictors of Success for  
the TOEFL

by

Theda Nunn

A Thesis Presented in Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts

Approved April 2014 by the  
Graduate Supervisory Committee:

Dr. Claire Renaud, Chair  
Dr. Matthew Prior  
Dr. Mark James

ARIZONA STATE UNIVERSITY

May 2014

## ABSTRACT

The teaching of formulaic sequences (FSs) to improve speech fluency is a time honored tradition in the field of English as a Second Language (ESL). However, recent research seems to indicate that certain discourse markers, specifically transition and personal stance markers, are more useful than other FSs. This study is an attempt to partially replicate (on a very small scale) one of these studies to see if the findings are similar when the standardized test materials are from the Test Of English as a Foreign Language (TOEFL) rather than the International English Language Testing System (IELTS). The hope is that teacher researchers could have access to readily available, standardized assessment materials with which to create their own research studies consisting of a standardized pretest and posttest. Four students of various levels in an Intensive English Program (IEP) were given a practice listening and speaking exam utilizing TOEFL preparation materials found online. The results were analyzed to see if there was a noticeable correlation between the use of the specified discourse markers on the speech portion of the test and the performance of the students on the listening portion of the test. The findings show some discrepancy between the two studies' results. It appears possible to have a high perceived fluency rate and still have a lower overall speaking fluency when taking into account listening comprehension and various other measures.

## DEDICATION

“In the beginning was the Word,  
and the Word was with God,  
and the Word was God.”

John 1:1 KJV

This thesis is dedicated to the Word.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	v
CHAPTER	
1 INTRODUCTION .....	1
A Tale of Two Studies .....	2
Background and Need.....	3
Purpose of the Study .....	3
2 LITERATURE REVIEW .....	5
Action Research .....	5
Formulaic Sequences .....	8
Speech Proficiency vs. Speech Fluency vs. Speaking Ability.....	11
Explicit vs. Implicit Teaching.....	13
Questions of Cognitive Development .....	15
Two Studies in Detail.....	20
Replication Research.....	24
3 METHODS .....	26
Introduction .....	26
Profile of Participants.....	26
Materials .....	28
Measurement Instruments.....	31
Procedures .....	31

CHAPTER	Page
4 DATA ANALYSIS AND RESULTS .....	34
Data Analysis .....	34
Results .....	35
5 DISCUSSION AND CONCLUSIONS .....	38
Limitations.....	40
Implications for future research.....	41
REFERENCES.....	44
APPENDIX	
A INTERVIEW INSTRUMENT.....	47
B TOEFL LISTENING AND SPEAKING TEST TRANSCRIPTS.....	49
C COMPLETE LIST OF FORMULAIC SEQUENCES TALLIED .....	63
D IELTS 2005 PREPARATORY MATERIALS .....	69

## LIST OF TABLES

Table		Page
1.	Various Labels of FSs .....	10
2.	Participant Characteristics.....	27
3.	Listening .....	29
4.	Speaking.....	30
5.	Listening Portion of TOEFL Preparatory Materials.....	35
6.	Speaking Portion of TOEFL Preparatory Materials.....	36

## CHAPTER 1

### INTRODUCTION

Within the field of second language acquisition (SLA), teacher action research is maligned and neglected as a research methodology on a recurring basis. At the same time, there is an oft cited need for further research into the efficacy of implicit versus explicit teaching of formulaic sequences (FSs). Most research done in this area is conducted by researchers with few subjects leading to a paucity of data for statistical analysis purposes. Those stakeholders in such research that have the greatest access to subjects are either teachers of a second language (L2) or researchers working with a standardized test company. When teachers do research, it is usually identified as *action research* because the *research* is an *action* in response to a problem within the classroom that the teacher wishes to resolve and because the *research* will, hopefully, provide direction as to what *action* to take to attain resolution of the issue. When researchers work in collaboration with teachers to do research, that can also be referred to as action research. Crookes (1993) indicated that action research is different from other research done by teachers because it is not generalizable beyond the teacher researcher's classroom or institution. However, there are broader definitions (Elliott, 1991; Dörnyei, 2010) that, under ideal conditions, would allow the value of the results to go beyond the classroom and/or specific institution of the teacher researcher and allows for applicability within the specific area of enquiry. When and how to teach FSs is important in every classroom and every institution in which such teaching occurs, and many teachers would welcome the opportunity to share in the discovery process of how best to teach FSs if it were not for two common problems: lack of time and lack of money for creating and

implementing research projects in the classroom. This study was created in response to the need for an inexpensive and less time consuming approach to involving teacher researchers in the process of action research in order to facilitate the study of the efficacy of implicit versus explicit teaching of FSs, specifically within the topic area of English as a second language (ESL).

### A Tale of Two Studies

If a researcher has the opportunity to work with a standardized test manufacturer, as in the case of Ohlrogge (2009), it can be the best of times for creating a study of some significance. However, if you are a teacher who is also a researcher in need of resources, as in the case of Khodadady and Shamsaee (2012), it can be, maybe not the worst of times, but maybe, a good time to use some ingenuity to create a meaningful study. In 2009, Ohlrogge identified eight types of formulaic language used by a diverse group of intermediate level learners of English when taking the written portion of the Examination for the Certificate of Competency in English (ECCE), a high-stakes English as a Foreign Language (EFL) examination. Ohlrogge was involved with the development of the ECCE composition assessment rubric and had access to materials that were used during the development of the rubric.

In a follow-up study in 2012, Khodadady and Shamsaee attempted to identify these same types of formulaic language as used by intermediate level Iranian students of English when participating in a short practice for the Listening/Speaking portion of the International English Language Testing System (IELTS). They found that two types of the eight types of FSs identified by Ohlrogge had a high correlation with student success on a speaking proficiency test. The study by Khodadady and Shamsaee is of interest



because of their creative use of readily available materials (free, online, standardized test preparation materials) to make their study with a minimum of outside cost and a maximum of standardization and replicability.

### Background and Need

It is, almost by definition (Block, 2000; Crookes, 1993), difficult for teacher researchers to create studies of any generalizable value (with results that are applicable outside of their own classrooms) due to various financial and logistical constraints. Time and money are always in short supply. If teacher researchers had ready access to fairly standardized materials when creating tests for their studies, this would solve some of the logistical problems by reducing (or even eliminating) the time required to design a study. This standardization of approach would allow for greater ease (and therefore a lower cost) in aggregating data from various sources in order to create a much larger study, called a meta-analysis, out of many small studies.

### Purpose of the Study

In order to do action research on the impact of explicit versus implicit teaching of FSs to students of English as a second language (L2), the ideal study(ies) should use recognized standards of research of a pretest, a posttest, a number of subjects and a control group (Mackey & Gass, 2005). To create such a study is close to impossible for any full time teacher due to the money and time constraints involved in creating, administering, and analyzing the results of test materials. Thanks to the study by Khodadady and Shamsaee (2012), it is clear that one test, freely available online, could be used for either a pretest or a posttest, and there is another standardized test with preparation materials freely available online. That is the TOEFL. Because the IELTS is

used with less frequency than the TOEFL at Arizona State University (the site of this study), it seems probable that there would be a high interest among potential subjects for participation in a replication study of the Khodadady and Shamsaee (2012) study using the TOEFL instead of the IELTS. Both tests elicit speech production from their examinees, and the data elicited is then used to create correlations. Therefore, if the correlations produced by the use of TOEFL materials are similar to those produced by Khodadady and Shamsaee (2012) using IELTS, it might be possible to create the ideal situation of having a pretest that uses preparation materials from one study, and a posttest that uses preparation materials from the other test (IELTS or TOEFL), creating virtually free and almost universally accessible research study tools. The purpose of this study is to determine if the TOEFL online preparatory materials can be used in a similar way to obtain similar results as the IELTS materials that were used in Khodadady and Shamsaee (2012). A secondary purpose is to note the development of the use of discourse markers (defined on page 21) at intermediate and advanced levels of language proficiency.

## CHAPTER 2

### LITERATURE REVIEW

#### Action Research

In the field of teacher action research, there is a frequently quoted definition attributed to Elliott (1991). He says it is an excellent summation of his description of action research on pages 49-56 of the referenced work, but he does not think it is a direct quote. He says it “may well be a direct quote from another article of mine” (personal correspondence). The definition is as follows:

Action Research is the process through which teachers collaborate in evaluating their practice jointly; raise awareness of their personal theory; articulate a shared conception of values; try out new strategies to render the values expressed in their practice more consistent with the educational values they espouse; record their work in a form which is readily available to and understandable by other teachers; and thus develop a shared theory of teaching by researching practice.

In this paper, I would like to suggest that such a definition allows for the collaborative efforts of teachers around the world who share an interest in determining the best practices for teaching FSs to their students, specifically in the area of EFL/ESL.

According to Dörnyei (2010), the most important principle of action research is that the teacher and the researcher are either, one and the same, or the two are working together very closely. Dörnyei goes on to explain that, although the idea sounds great, there have been some very serious problems with this concept.

Other researchers seem to agree with his analysis (Block, 2000; Crookes & Chandler, 2001). These problems were caused primarily by the fact that teachers lack the time, the incentives, and the expertise or professional support necessary to do meaningful research. Dörnyei (2010) goes on to state that this is not a viable form of research at this time (in his opinion) due to “the current unsupportive teaching climate” (p. 193), and he goes on to imply that administrative support for teachers to be able to do such research would be of tremendous help toward making this great idea actually feasible. Although Block (2000) seems to have come to the same conclusion as Dörnyei (2010), he does it through a process of looking at the perceived gap between researchers and teachers. Block (2000) states that teachers have a lower professional status than researchers, and he opines that in our current culture, teachers are viewed no differently than sales clerks and bus drivers. To him, this means that when teachers and researchers work together, it is less of a collaborative situation and more of a situation where the researcher directs the teacher as to what needs to be done. He also points out that researchers use a different vocabulary from teachers. This is also confirmed by Erlam (2008), who describes how the research project she reports on gave teachers a new vocabulary, thereby improving their ability to collaborate with others in meaningful research endeavors. Block (2000) also notes that teachers and researchers have different views of how “the growth of knowledge” (Block, 2000, p. 135) occurs. According to Block, researchers believe that knowledge is grown in a structured and predictable manner, while teachers see knowledge growth as more like a conversation with stops, starts, and relationships involved. And while Block (2000) briefly considers whether action research could possibly bridge the gap between

researchers and teachers, he promptly dismisses it as an impossibility. Instead, he holds out hope for the idea of teachers and students collaborating in the classroom to solve “puzzles” about acquiring knowledge, which is actually a form of action research, but it is only hopeful to help improve teaching and learning inside the four walls of the classroom in which the “research” occurs. The way he describes it, it does nothing to bridge the gap between researchers and teachers.

While it is true that teachers have a very difficult time doing research activities, it is also true that researchers have a very difficult time gaining access to potential subjects for their studies. Erlam (2008) states that the researchers in her study were surprised that they “encountered resistance from some teachers when it came to negotiating access to their classrooms for the purpose of research” (p. 258). One of the researchers (who had a current teacher registration) was asked, “And what do you researchers know about language teaching?” (p. 259). Erlam’s study is exceptional in that her researchers made an effort to recognize the teachers as professionals, and she did seem to obtain quite a bit of cooperation and collaboration with teachers. However, her researchers were, still, only able to obtain access to two thirds of the classrooms she had targeted. She, too, notes the gap between researchers and teachers, but her study makes use of several techniques to bridge that gap. Providing accessibility to the language of researchers by glossing unfamiliar terms in her materials, led to the teachers having a shared language to discuss the research with each other and with the researchers. This led to a perceived relevance to the teachers, and, perhaps the most important aspect of how her study bridged the gap was that it engaged the teachers instead of it being imposed on them.

The study by Schmitt, Dörnyei, Adolphs, and Durow (2004) could also be defined as action research in the broadest sense. It was conducted in classrooms, and Schmitt, et al. (2004) note that “in order to secure the cooperation” of the teachers, the target FSs needed “to be seen as useful to students and worthwhile to teach” (p. 56). Therefore, after an initial review of various corpora and classroom materials, the researchers presented a relatively short list of FSs from which the teachers were to choose 20 as targeted items. The researchers also made it a point to provide teachers a great deal of autonomy in how they approached the teaching of the target FSs. (This study is discussed in greater detail in the section Formulaic Sequences.)

Both Erlam (2008) and Schmitt, et al. (2004) seem to believe the way to encourage higher rates of participation in studies that require assessments would be through authenticity of the assessment experience. Authenticity is defined as “the degree of correspondence of the characteristics of a given language test task to the features of a target language task” (Brown & Abeywickrama, 2010, p. 346). In order for testing materials to seem authentic to participants, the materials should relate somehow to the goals of the participants. For instance, Khodadady and Shamsaee utilized the free online preparatory materials provided by the IELTS organization for their study. Authenticity is guaranteed in this case because future English teachers would most likely be very interested in preparing for the IELTS.

### Formulaic Sequences

One of the most iconic works on the subject of FSs is that of Nattinger and DeCarrico (1992), which was an attempt to summarize the history and research of lexical phrases and to apply what was known at that time to the teaching of language.

It is a magnificent work that explains the thinking behind many of the terms used for formulaic language, and it provides some possible applications from the research, such as the need to teach commonly used macro organizers (words or phrases that mark the overall direction in discourse) to students of American English for academic purposes. However, in their conclusion they summarize the situation for the many areas of study within formulaic language by stating, “A general observation to be made about all of these areas is that they require additional empirical fieldwork” (p. 174).

Another perhaps equally iconic work on formulaic language is the definition put forward in Wray (2002). As defined by Wray, a FS is:

*a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (p. 9)*

According to Wray (2013), in 1874, a doctor named John Hughlings Jackson, who was working with patients who had incurred brain damage, “formed the view that some language was ‘automatic’ and ‘non-propositional’, and was processed by the right hemisphere rather than the linguistically dominant left” (p. 320). By 1964, Firth had made his famous comment, “You shall know a word by the company it keeps” (Firth, 1964, as cited in Wray, 2013, p. 320), a statement that reflects the idea that certain words are frequently found together, as in FSs. And now, with the advent of relatively cheap, high speed computational processes, the creation and analysis of numerous corpora of remarkable size—in the millions and even billions of words—have led to a marked

increase in the study of such language and to the proliferation of terms applied to it. Those terms often vary in their meaning depending on the study they are found in. It is because of this variability that authors and speakers need to define what they mean when using any of those terms. Table 1 contains some examples of the various labels applied to FSs.

Table 1  
*Various Labels of FSs*

Chunk	Collocation	Concgram	Conventionalized Form
Fixed Expression	Formula	Formulaic Language	Formulaic Speech
Formulaic Utterance	Holophrase	Idiom	Lexical Bundle
Multiword Expression	Multiword Unit	Prefabricated Routine	Ready-made Utterance

Whether we are speaking our native language or a L2, we all use FSs to help us speak cogently and fluidly. FSs allow us to formulate our future expressions even while we are speaking our previous thoughts, thus eliminating many long pauses while we search our memories for just the right word. Instead, because of our knowledge of formulaicity, we automatically know that certain words belong together, and this automaticity gives us the appearance (if not the reality) of proficient speech (Erman & Warren, 2000; Pawley & Syder, 1983). They enable us to communicate complex concepts with a minimum amount of effort (Sinclair, 1991), and they comprise a rather large portion of our speech. Studies using different definitions for the FSs observed (and



using different methodologies, as well) have obtained estimates ranging from 32.3% (Foster, 2001) to 58.3% (Erman & Warren, 2000). It seems probable that using Wray's (2002) very broad definition of FSs could lead to even higher estimates than that of Erman and Warren. In summation, FSs are essential to speech. In particular, they are essential to speech fluency and therefore to speech proficiency, as well.

#### Speech Proficiency vs. Speech Fluency vs. Speaking Ability

In the field of SLA, it seems that speech proficiency is often conflated with speech fluency. Speech fluency has been researched extensively and is fairly closely defined. "Across all of the studies of spoken fluency and its development, there has been a remarkable degree of agreement on the types of temporal variables to be tracked. These are rate of speech, measured as syllables uttered per minute, amount of pauses and the length of runs, measured as number of syllables uttered between pauses" (Wood, 2009, p. 40-41).

De Jong, Steinel, Florijn, Schoonen, and Hulstijn (2013) mention three types of fluency: cognitive fluency, utterance fluency, and perceived fluency. While utterance fluency is an objective measurement of the words produced per minute (much like Wood's definition of spoken fluency), cognitive fluency and perceived speech fluency are actually subjective measurements. Cognitive fluency describes the speaker's "ability to efficiently plan and execute his speech" (De Jong et al., 2013, p. 894). This is something quite difficult to measure in an objective way, as it requires the ability to see what is happening in the speaker's brain. Perceived fluency is the perception that listeners have of a speaker's fluency based on a sample of their speech. It seems possible that an overreliance on perceived fluency could lead to the

overuse and misuse of FSs through explicit teaching in an attempt to create a shortcut to proficiency, and thus only giving the impression of someone with a strong grasp of the language, while there are still considerable pragmatic, grammatical, and syntactical issues that will need to be overcome to reach an academically communicative level of speech. Therefore, although we have an objective definition of at least one type of fluency (utterance fluency), there are other forms of fluency (cognitive and perceived) that are far more subjective in their determination.

However, “there is nothing even approaching a reasonable and unified theory of proficiency” according to Lantolf and Frawley (1988, p. 186). In fact, in their same work, they state, “Vollmer (1981) points out that after years of investigation the only thing we are able to say about proficiency is that proficiency is what the proficiency tests measure” (p. 185). The work of Vollmer (1981) and Lantolf and Frawley (1988) is around three decades old now, but even today attempts to find any contemporary research (or even discussion) about what constitutes speech proficiency failed to produce an objective definition.

All discussions of speech proficiency assume either the use of a standardized test (which is a type of circular logic, we are defining tests by the thing the tests are supposed to define) or a comparison against the subjective concept of a “native speaker” as the defining standard. But which native speaker is the defining standard: British, American, or one of the many regional variants of each? So Vollmer’s (1981) comment still holds true. There is no truly objective definition of speech proficiency, and therefore the definition used for this paper will be the same as the definition given for speaking ability in Khodadady and Shamsaee (2012). It is as follows: “the

mean score obtained on speech fluency as well as the knowledge of grammar, knowledge of vocabulary, pronunciation, and semantics and pragmatics of any given language” (Khodadady & Shamsaee, 2012, p. 45). Although this definition has an objective approach to speech fluency, any test of knowledge is bound to be subjective. It is also worth noting that regional expectations in each of the subjective areas can vary wildly.

One other term that is often used interchangeably with speech fluency and speech proficiency is speech (or speaking) ability. This term is most often used as a synonym for speech (or oral) proficiency, but it also seems to be synonymous with speech fluency. In this paper, these terms will be used only as synonyms for speech proficiency.

#### Explicit vs. Implicit Teaching

The perceived importance of FSs to the development of a native-like proficiency in an L2 has led to numerous studies of the use of these items by both native speakers and L2 learners. Schmitt, et al. (2004) examined the acquisition of a target set of FSs and the potential influence of various individual differences on the process of acquiring FSs. Their study utilized 62 students participating in a university’s two month English for Academic Purposes (EAP) program and an additional 32 students participating in the same university’s three month EAP course. There was a pretest within the first week of the program, and the study required that during the course of the program teachers would present the target FSs at least once to their students. Teachers were also required by the study to draw attention to each of the targeted FSs at some point during the course of the session without indicating that those FSs were being targeted. Other than those two

requirements, teachers were free to introduce the FSs in any way or order they wished. The results indicated that the students already “knew a considerable number of these FSs, and that they enhanced this knowledge over the course of the 2-3 month” program (p. 69). Schmitt et al. (2004) also found that the individual differences that they looked at (aptitude/attitude/motivation factors) did not influence this enhancement. They suggested investigating whether explicit teaching was superior to implicit teaching.

In Wood’s (2009) oft cited study, he came to the conclusion that there was a need for further study as to how best to implement explicit teaching of FSs, but his own study suffered from a lack of a control subject. He had only one subject, and that subject seemed to benefit from a workshop explicitly teaching FSs. However, the subject, Sachie, was an English learner enrolled in an intensive study abroad program. Although Wood states “it is unlikely that these changes can be attributed solely to other aspects of her English language experience over six weeks” (p.53), it is not clear that the intervention is what made for the strong increase in fluency, either. For comparison purposes, it would have been very helpful to have data on even just one other student of similar level, sex, age, etcetera, who chose not to do the workshop.

Webb, et al.’s study (2013) was conducted at three universities in Taiwan. The participants were 161 first and second year students at the universities. They had all been studying English as a foreign language for at least 6 years. The purpose of the study was to examine the incidental learning of collocations. There were experimental groups and a control group. There was a pretest and four immediate posttests. Overall, this study seems to demonstrate all of the requirements of careful methodological investigation. Nevertheless, the results were deemed somewhat uncertain due to the advanced

knowledge of the participants coming into the study and because the test may have provided an excessive amount of informational cues that may have inflated the results of the interventions. This study is of interest because it is approaching the question of implicit teaching from the view of how collocations are learned. It demonstrates the fact that, while researchers may in fact be looking at the same phenomenon, they may label that phenomenon in different ways. It also demonstrates the lack of standardization in the testing procedures and data elicitation processes currently in play in the study of FSs.

All three studies demonstrate the wide variety of studies that fall under the term action research. This wild variability in numbers of participants, assessment tools and procedures only adds to the difficulty of identifying studies that are similar enough for any meaningful meta-analyses. It seems logical to expect that standardization of assessment tools and procedures could both improve the overall data collection practices for all concerned, and it could encourage a higher rate of participation by teachers and their classes if a focus was made on preparing students for the high stakes tests they face while gathering useful data for understanding the teaching of FSs.

### Questions of Cognitive Development

It is a widely accepted fact that the appropriate and fluent use of FSs is very much correlated with the more advanced levels of L2 development (Nattinger & DeCarrico, 1992; Pawley & Syder, 1983; Sinclair, 1991, to just name a few). However, it also remains a matter for further research as to whether or not FSs should be taught explicitly, implicitly, or some combination of the two (Boers & Lindstromberg, 2012; Meunier, 2012; Schmitt et al., 2004). Humans learn different things at different times. Human development studies have shown us that teaching some things explicitly can actually be a

waste of time and energy for all parties involved, and it may even cause irreparable damage to the subject's development. A case in point is Dr. Arnold Gessell's famous twins study.

Arnold Gesell, M.D., conducted a famous experiment (with identical twin babies) that seems apropos to consider here:

In 1927, in collaboration with Dr. Helen Thompson, he undertook a comparative study in which two highly identical twin girls were observed from early infancy to determine, first, their developmental correspondence and, second, their developmental divergences, as these might be affected by training confined to one twin. These identical twins, T and C, showed a great degree of similarity which was established by elaborate and repeated examinations later documented in publications. At the age of 46 weeks twin T was trained daily in climbing a stair that had five treads and after seven weeks was able to perform the coordination complex in 26 seconds. Twin C, at the age of 53 weeks, without any previous training or experience, climbed the stair unaided in 45 seconds. Then twin C was trained for two weeks and at the age of 55 weeks required only ten seconds to accomplish this task. Twin C was at the age of 55 weeks far superior to twin T at 52 weeks, even though T had been trained for seven weeks in the beginning. (Miles, 1964, p. 70)

Thus was proved the age old grandmothers' adage "they'll do it when they're ready." The unfortunate result of the researchers' explicit interventions was that the baby, who was trained to climb the stairs, missed out on the free play time that her sibling enjoyed. It seems possible, based on the information above, that this may have actually delayed or

permanently damaged some form(s) of development for the baby receiving the intervention.

Clearly the subjects of Dr. Gesell's experiment were infants, and the area being explicitly taught was a kinesthetic learning activity, whereas the subjects in the study conducted by Khodadady and Shamsaee (2012) were university students performing oral and aural tasks: two very different groups developmentally and two very different skill sets being developed. Nevertheless, it does seem wise to consider that human learning often, maybe even always, includes cognitive development issues.

As Khodadady and Shamsaee (2012) mention, the learning and use of FSs affects how our brains function. There are also indications in the linguistic research literature that teaching FSs too early in an L2 learner's development could lead to serious misuse of the formulas learned. It is important that researchers (whether they are teachers or not) make every effort to produce not only sound data, but also sound analyses of that data. Conclusions drawn based on faulty data or faulty analyses could have unintended consequences as seen in Dr. Gessell's twins study. And it is important to note the warning in Crookes (1993), "the stronger the claims for general applicability that a study makes and the more damage such claims, if wrong, could do, the greater the demand that should be made for reliability, validity, and trustworthiness" (p. 136). Krashen's (1985) Input Hypothesis states that, to be of value to a learner of a new language, input must be comprehensible to the learner. He argues that learners who are forced to utilize grammatical structures that they do not understand will often substitute a grammatical rule from their L1 that they do understand. This rule may or may not be similar to the rule being taught in the L2. Because of this, the learner will produce errors that are deemed to

be interference from the L1 when, according to Krashen, the real problem is that the learner did not receive comprehensible input for the L2. FSs can be viewed as grammatical structures of a type. If they are not fully comprehensible to the learner, they will tend to be misused or abused when explicitly taught. For instance, the use of certain grammatical forms have been shown to be acquired in a predictable sequence as stated in the Natural Order Hypothesis (Krashen, 1985; but also see Luk & Shirai, 2009), if one were to introduce a FS that uses a grammatical structure that the student is not ready to use, one could cause the student to learn the sequence incorrectly. An example might be, *he also serves who only stands and waits*, meaning it is sometimes okay not to do anything other than wait. Because the grammatical “s” on “serves”, “stands”, and “waits” is one of the last structures for English language learners (ELLs) to acquire (Luk & Shirai, 2009), the explicit teaching of this saying could come out, *he also serve who only stand and wait*.

According to Ellis (2006), implicit knowledge of language grammar is the knowledge that allows us to speak fluently, and the most common approach to acquiring grammar seems to be the present-practice-produce (present the item to be learned, practice the item, then produce the item in novel statements or situations) approach to obtain implicit knowledge. This is an explicit teaching process. While there seems to be some controversy (according to Ellis) regarding which of various approaches to use, it is clear that his work has led him to believe that practice (repetition) is how implicit knowledge is attained. Please note, the difference between implicit knowledge (knowledge that allows fluency of speech because of familiarity with the grammar) and implicit teaching (teaching that is providing a rich context for students to develop their



language skills without explicit instruction in the grammatical forms). Implicit instruction guarantees that the student is ready for the grammatical concept learned because they have discovered the rule on their own. Explicit instruction is given by the instructor when the student may or may not be ready to learn.

Wray and Fitzpatrick (2010) conducted a study in which they had ESL students memorize FSs that the participants would then use in an authentic situation within a relatively short period of time. This exercise did allow the participants in their study to speak more fluently and feel greater confidence when speaking with native English speakers in an authentic situation (i.e., visiting the doctor, or paying a bill in person), but Wray and Fitzpatrick note that many times the students did not use the forms as rehearsed. Because explicit instruction requires much repetition to cause the phrase to move from short term memory to long term memory (and it is in long term memory that implicit knowledge resides), it is possible the student will always have this saying in their mind incorrectly if they cannot reproduce it accurately when they do their repetitions.

There is also the view that, from an educational standpoint, it is necessary for students to receive multiple exposures to new information before they will actually be able to acquire the information and make use of it. There is an argument that a combination of implicit exposure to FSs in the basic and intermediate levels of study and more explicit exposure in the advanced levels may be the best combination of explicit and implicit teaching of FSs (Ellis, 2006; Meunier, 2012; Schmitt et al., 2004). Boers and Lindstromberg (2012) highlight the fact that ESL students are able to recognize FSs before they are able to reproduce those FSs reliably. In their discussion of a study by Webb, Newton, and Chang (2013), they state, “Rather disconcertingly, however, as many

as 15 encounters in such a short span of time still did not guarantee full scores on any of the posttests” (p. 91). It is a well-known fact among educators that it is necessary to create multiple exposures over long periods of time for the human brain to reliably move data from short-term memory to long-term memory. Dr. Gessell’s over-trained and eventually underperforming twin, the Input Hypothesis from Krashen (1985), along with the comments by Ellis (2006) regarding the importance of practice, and the difficulties in retaining memorized FSs experienced in the study by Wray and Fitzpatrick (2010) are all examples of difficulties that may arise from the teaching of FSs, and they help to underscore the urgent need for more and better studies on the teaching of these vitally important parts of speech.

#### Two Studies in Detail

In 2009, Ohlrogge created a study in which “a small corpus of 170 compositions” (p. 378) were selected to answer two research questions:

1. What types of formulaic language are used by intermediate-level learners in a high-stakes writing examination?
2. Do high-scoring and low-scoring writers use particular formulaic sequences with the same frequency? (p. 371)

Ohlrogge’s goal was to see if “formulaic sequences can be a useful criterion in discriminating between different levels of proficiency—the primary purpose of most language tests” (p. 377).

The candidates who wrote the compositions used in Ohlrogge’s study came from nine different first language (L1) backgrounds: Greek (85), Spanish (37), Portuguese

(25), Vietnamese (10), Arabic (5), Italian (3), Romanian (2), Catalan (2), and Macedonian (1).

During his analysis, Ohlrogge identified eight categories of FSs: collocations, idioms, phrasal verbs, personal stance markers, transitions, generic rhetoric, and irrelevant biographical information. Ohlrogge was careful to define each of these FSs for his study. However, for the purposes of this paper, only transitions and personal stance markers will be discussed. His definitions for these two FSs are as follows:

Personal Stance Markers: Expressions that signal a writer's personal view or opinion. These can be regarded as an example of what Pawley & Syder (1983) call "sentence frames". Such expressions generally occur at sentence beginnings and are common in argumentative and expository writing. Examples include *in my opinion*, *I strongly believe* and *without a doubt*.

Transitions: Sequences used to signify the relationship between sections of a text. Examples include *on the one hand*, *first of all* and *in conclusion*. (p. 380)

Transitions and personal stance markers are each a type of discourse marker. Müller (2005) says, "discourse markers contribute to the pragmatic meaning of utterances and thus play an important role in the pragmatic competence of the speaker" (p. 1). Schiffrin (1987) defines discourse markers as "sequentially dependent elements which bracket units of talk" (p. 31). A very simple explanation might be to say that discourse markers connect parts of discourse and indicate where the discourse is going. Ohlrogge's results showed that transitions occurred with the greatest frequency overall, and that only idioms

and personal stance markers were used with increasing frequency as the composition grade increased.

For their study, Khodadady and Shamsaee (2012) utilized 40, female, Iranian university students who were majoring in Teaching English as a Foreign Language and Translation. The students' ages ranged from 18 to 35, and they were enrolled in a listening and speaking course. Students were not identified by level. As future English teachers, these students will, most probably, be taking the official IELTS at some point in the future. They will certainly have students, someday, who will wish to prepare for the IELTS. It seems safe to assume a fairly high level of motivation for these students to participate in a study that will allow them to practice the necessary skills for taking the IELTS successfully.

Khodadady and Shamsaee (2012) were not interested in replicating Ohlrogge's work exactly. In fact, they chose to focus on Listening and Speaking skills instead of writing skills. Because they did not provide a copy of the IELTS speaking specimen (2005) utilized by them, nor the IELTS Listening specimen also utilized in their study, I contacted Dr. Khodadady, and he confirmed that they were the online preparatory materials provided on the IELTS website, in 2005 (see Appendix D), to help future test takers prepare for the IELTS examination (personal correspondence). I have used those materials for comparison with the TOEFL materials.

Using Ohlrogge's definitions for categories of FSs, Khodadady and Shamsaee (2012) evaluated the data they collected and found a correlation between the use of transitions and personal stance markers and speech fluency, which they seem to refer to interchangeably with speaking ability. In their conclusion, they define speaking ability as

“the mean score obtained on speech fluency as well as the knowledge of grammar, knowledge of vocabulary, pronunciation” (p. 45). Khodadady and Shamsaee also state, “Personal Stance Markers and Transitions are the only types of FS that have a significant relationship with learners’ speech fluency” and “can predict overall speaking ability better than speech fluency” (p. 45). They then come to the conclusion that the data resulting from their study can be used to determine an order of acquisition for FSs that can be used for explicit instruction. They also express the belief that the use of these FSs helps to cause the cognitive development necessary for proficient language development. In other words, they seem to be advocating a focus on the teaching of transitions and personal stance markers as a priority over other forms of FSs because they believe it will improve the overall speaking ability of L2 students.

It appears that they may have fallen prey to the post hoc fallacy or post hoc ergo propter hoc (“after this; therefore, because of this”) fallacy. So the question is, does the correlation discovered by Khodadady and Shamsaee between the two FS types and student scores actually indicate that students who receive explicit teaching of these two FSs will obtain higher scores than if they receive only implicit instruction, or is it actually an indicator that the developmental level of students who are able to successfully use those types of FSs is such that they are also better able to handle all other aspects of the L2 as well? In other words, which comes first, proficiency or explicit teaching? This is a question that still needs answering, even though Khodadady and Shamsaee seem to feel it is already resolved.

## Replication Research

Replication research, as its name implies, is research that repeats a previous study, but it is not necessary to repeat the previous study exactly. In fact, an exact repetition of a previous study is called a reproduction. “Other variants can address the robustness or generalizability of a study by the introduction of further variables or contexts alongside those used originally and which might be thought to provide further knowledge about, for example, whether different approaches produce different results” (Porte, 2013, p. 6). Action research is often a cyclical process also (Crookes & Chandler, 2001). This makes it easy to meld the two research methods.

This study is an attempt at a small scale replication of the study by Khodadady and Shamsaee (2012) that was inspired by Ohlrogge’s study in 2009. There are enough differences between the two previous studies to make it a borderline case as to whether Khodadady and Shamsaee were actually attempting a replication. It may be more correct to say that they were inspired by Ohlrogge’s study to see if they could find similar results in spoken data as those found by Ohlrogge in the written data used for his study. But where Ohlrogge had access to data that he had already elicited while working on a rubric for a standardized test maker, Khodadady and Shamsaee needed to find a way to elicit their own data unassisted. Ohlrogge’s data was written and from the Examination for the Certificate of Competency in English; Khodadady and Shamsaee’s data was spoken and was from their creative use of the IELTS preparatory materials found online. And finally, Ohlrogge’s study was one of discovery. He set out to determine what types of FS discourse markers are most commonly used by intermediate-level learners in a high stakes writing exam and if

those FSs are used with the same frequency by both high and low scoring writers. However, Khodadady and Shamsaee, following after Ohlrogge, used his definitions to see if they found the same usage for FSs in speech as that found by Ohlrogge in writing. They also sought to know which of the eight different FSs defined by Ohlrogge best predict overall speaking ability among intermediate level English learners. Khodadady and Shamsaee seem to have successfully completed what could be termed an action research project with their study that made use of free authentic assessment materials and their efforts to benefit their students by helping them prepare for a high stakes exam. This raises the question, could their method work in other settings and with other standardized tests? This could be useful to know as it could provide a way to standardize test methodologies for many small, action research projects with little to no funds available for them to purchase or create tests.

## CHAPTER 3

### METHODS

#### Introduction

This quantitative, exploratory study was conducted to attempt to answer the following questions:

1. Which transition markers and personal stance markers are used by intermediate and advanced level nonnative speakers when practicing for a high stakes listening and speaking test (TOEFL)?
2. What is the variation by level in frequency and variety of use?
3. Do these variations have the same import on the TOEFL as other research (Khodadady & Shamsaee, 2012; Ohlrogge, 2009) claims they do on the IELTS?
4. Can a comparison of the data obtained above provide useful insights for future research and pedagogical activities?

#### Profile of Participants

A convenience sampling procedure was utilized. Participants were volunteers from conversation groups which were part of an Intensive English Program (IEP) conducted by ASU on the Tempe campus. Four people participated (three females, one male) who were between the ages of 19 and 35 years old. As can be seen in Table 2, they represented a variety of L1s as well as educational backgrounds. English was the only L2 of all participants except for Participant 2 who also spoke five other languages (identified in Table 2) in addition to her L1 and English.



Table 2  
*Participant Characteristics*

Participant	Sex	Age	Level*	Educational Level	L1	L2(s)	Previous TOEFL Experience
1	M	35	I1	Completed Master's Degree	Arabic	English	May have taken a practice test 3 years ago
2	F	26	I2	Completed Bachelor's degree	Pashtun	Dari, Arabic, English, Balochi, Russian, Urdu	Took the TOEFL once with no practice
3	F	21	I2	3rd year Undergraduate	Spanish	English	Self-practice on reading and writing portions only
4	F	19	A2	Entering Undergraduate	Swedish	English, French, Mandarin	None

*Note:* \* The various levels of the participants were initially determined by the IEP through an in-house test given after arrival and registration. After the entrance exam there are no further placement tests, and level is determined strictly on the basis of course work. The levels provided correspond roughly to intermediate and advanced levels of proficiency.

## Materials

Instead of the preparatory materials for the IELTS used by Khodadady and Shamsaee (2012), I used the online preparatory materials for the TOEFL exam. The TOEFL exam is more commonly used at ASU, and the speaking and listening portions of the two tests are quite different from each other. As can be seen in Tables 3 and 4, the TOEFL actually provides much less time than the IELTS for the examinee to speak, and the process is completely computerized. And, whereas there is no human interaction in the TOEFL, the IELTS Speaking Test involves human interaction throughout. This difference inspired me to wonder if there would be any noticeable difference between the levels of proficiency in participants' usage of transitions and personal stance markers under these conditions.

For the purposes of this study, a brief interview with each participant was also included to provide the necessary personal information for my subjects with the spoken data they produced. This interview contained thirteen questions, and covered topics such as age, sex, natal country, languages spoken, time spent in the United States, and previous experiences with both English and the TOEFL. The interview took between five and ten minutes to complete. See Appendix A for the full instrument.

The IELTS is a four skills English proficiency exam that assesses participants' reading, writing, listening, and speaking skills. It is jointly owned by British Council, IDP: IELTS Australia and Cambridge English Language Assessment. According to their website, the IELTS is administered in over 130 countries. See Tables 3 and 4 for a comparative overview of the Listening and Speaking portions of the two tests.

The TOEFL is also a four skills English proficiency exam. The TOEFL is owned and administered by Educational Testing Service (ETS), the private non-profit test and assessment company also responsible for creating and administering numerous other standardized tests such as the SAT (once known as the Scholastic Aptitude Test and then, later, as the Scholastic Assessment Test) and the Graduate Record Examinations (GRE). According to the TOEFL website, the TOEFL, like the IELTS, is also administered in over 130 countries.

A description of the Listening and Speaking portions of the TOEFL and IELTS tests is provided in Tables 3 and 4 respectively. I chose to focus on the Listening and Speaking portions of the TOEFL in order to allow for a replication of Khodadady and Shamsaee (2012).

Table 3  
*Listening*

TOEFL	IELTS
60-90 minutes long	30 minutes
34-51 questions	40 questions
Listen to lectures, classroom discussions, and conversations spoken in a fairly standard American form of English.	Listen to four monologues by different speakers whose English may be accented according to geographic origin of the speaker.
Multiple choice answers	Combination of multiple choice and short answers including fill in the blank, matching and labeling.

Table 3 shows that the TOEFL provides a greater variety of settings of listening experiences than the IELTS, and it is delivered in a fairly standard American accent, while the IELTS speakers may have various national or regional accents (e.g., Australian, American, British). The TOEFL also requires only one type of answer format, whereas the IELTS has various answer formats. On average, both tests have a similar number of questions. The IELTS always has 40 questions for this section, but the TOEFL number can vary from 34 to 51 questions. The TOEFL allows much more time for this portion of the exam than the IELTS does.

Table 4  
*Speaking*

TOEFL	IELTS
20 minutes long computer delivery	14 minutes long human interaction
6 questions	3 part interview format
The first two questions allow 15 seconds to prepare and 45 seconds to answer	Part 1 is an interactive discussion about the examinee and his/her family
The second two questions provide 30 seconds to prepare and 60 seconds to answer	Part 2 is prompted by a task card. Examinee reads card, prepares and then speaks on topic
The third two questions allow 20 seconds prep time and 60 seconds for the answer	Part 3 is a one to one conversation on the topic provided in Part 2
Examinee spends approximately 5.5 minutes responding. The remaining time is spent listening and preparing.	The examinee talks most of the time while being prompted by the interviewer as necessary.

As can be seen in Table 4, the speaking portion of the TOEFL is completely computerized, while the IELTS is based entirely on human interaction. And although the IELTS is of a shorter duration, it allows more time for the examinee to actually speak because it does not allocate a set amount of time for each answer.

#### Measurement Instruments

The Listening and Speaking tests were the TOEFL exam preparation materials obtained from the ETS website and can be found here:

[http://www.ets.org/toefl/ibt/prepare/sample\\_questions/](http://www.ets.org/toefl/ibt/prepare/sample_questions/)

Transcripts of the tests used are provided in Appendix B. For the preparation materials, ETS only provides two listening situations and a total of only 12 questions. For the Speaking Section there are only 6 questions, which are progressively more difficult to comprehend and to answer.

#### Procedures

Research was conducted on the ASU campus. A dedicated research room equipped with computers was used. Two participants were tested individually, and, due to schedule constraints, two participants were tested in a somewhat simultaneous fashion. No more than four people were ever in the room at the same time. Timers were necessary in order to replicate an actual TOEFL exam setting, and this did become an issue (see Discussion section for details).

All activities were individual. Using practice questions provided by the TOEFL (from the ETS TOEFL website) for test preparation purposes, participants were asked to perform a listening exercise and a speaking exercise, both of which were provided by computer and recorded on a separate digital sound recording device. (No recording was

made on the computer. It was only a means of conveying the questions.) Please see Appendix B, “Modified TOEFL sample questions,” which includes the official TOEFL Listening and Speaking exam scripts (including exam procedures). Please see Appendix A for the Interview this researcher created and inserted between them.

The Listening portion was administered first, and an answer sheet (please see Appendix C, “Answer Sheet”) was provided to be used by the respondent to record their answers. Approximately 15 minutes were allocated for this portion of the project. Respondents were required to listen to a conversation between two people and then answer five questions about that conversation. They then listened to a lecture and answered another seven questions. This portion of the test was used as the determinant of speech proficiency. Thus I was able to compare the number of discourse markers in question and the total word count (which is being used as the measure of fluency) at “predicting” (in hindsight) the listening score (overall speech proficiency). The Listening portion was followed by the Interview and the Speaking portion, both of which were electronically recorded as audio only.

The Interview took approximately five minutes and served to insure that the spoken data collected corresponded with the demographic of the speaker. Participants then began the speaking portion of the test that is completely timed by computer on the actual TOEFL with a total time of only 20 minutes to complete all tasks. Each participant spent approximately one hour on the entire process (for further detail, please see the Discussion section). All data were collected within a two week time period.

At the end of the procedure, each participant received a 3 x 5 card to write their name on and to deposit in a container by the door. The card was their entry into a raffle

for a \$5 Starbucks gift card. The raffle drawing was held after the completion of data collection.

## CHAPTER 4

### DATA ANALYSIS AND RESULTS

#### Data Analysis

In the listening portion of the TOEFL, participants received one point for each correct answer for a maximum of 12 points. Question number 3 requires two correct answers to be complete and is worth two points. The remaining questions are each worth only one point. In the speaking portion of the TOEFL, the total number of words produced by each participant was tallied. Because the focus is on transition and personal stance markers, all instances of the use of either of these FSs were also tallied for each participant.

I used the same definition for personal stance markers as Ohlrogge (2009) and Khodadady and Shamsaee (2012). However, it seemed that the definition for transitions was not as well defined by them. Longman (1999) identifies transitions as a type of linking adverb that serves “to make semantic connections between spans of discourse of varying length, thus contributing to the cohesion of discourse” (p. 558). Although Longman does not provide an actual definition of transitions, two examples of transitions are provided as part of the section on linking adverbials. The two examples provided by Longman of linking adverbs used as transitions, are the words, “incidentally” and “now”. These two transitions are both provided as single words (not part of an FS) providing a conjunction between rather long stretches of news media and academic text. Because the requirements of the exam used for this study were for more casual speech forms than the written examples in Longman (1999), and because they must be multi-word, I have chosen to further clarify the definition of transitions as follows:



Sequences used to signify the relationship between sections of a spoken text whether within a sentence or a paragraph.

Some examples of transitions identified in the study are: “and any other”, “for example”, and “the first time is”.

## Results

Table 5 and Table 6 show the results of the Listening and Speaking tests, respectively.

Table 5  
*Listening Portion of TOEFL Preparatory Materials*

Participant	Level	Raw Scores *	Score as a percentage
1	I-1	6/12	50%
2	I-2	8/12	67%
3	I-2	6/12	50%
4	A-2	12/12	100%

*Note:* \* Raw scores are composed of the number right over the total number of questions.

As shown in Table 5, Participant 4 (who is at the highest Advanced level in the IEP) had the highest score, Participant 1 (who is in the lowest Intermediate level) had the lowest score, and one of the I-2 level participants (Participant 2) had a score between the I-1 and the A-2. Even though I-2 is the highest Intermediate level, Participant 3 has the same score as Participant 1, who is the lowest level participant in the study. However, Participant 3 has mentioned a tendency to become very stressed during tests—to the point of it affecting her performance. So in spite of the apparent anomaly of Participants 1 and 3 having the same scores, these results do seem to reflect the level of the participants quite well.

Table 6  
*Speaking Portion of Exam*

Participant	Level	Number of Transitions	Number of Personal Stance Markers	Total Word Count
1	I-1	16	5	512
2	I-2	10	6	446
3	I-2	9	3	330
4	A-2	29	11	762

During the Speaking portion of the test, the number of transitions used appears to be directly related to the total word count, with the greatest frequency of use of transitions, and the highest word count, belonging to Participant 4, while the lowest number of transitions used and the smallest word count belong to Participant 3. Level is clearly not as important as word count, as shown by the fact that the second highest word count and the second highest usage of transitions belong to Participant 1 (level I-1—the lowest level), while the two participants at the I-2 level (Participants 2 and 3) both had lower word counts and fewer transitions used. The apparent anomaly of Participant 3 having lower scores than Participant 1 is probably due to her tendency to experience some panic when taking oral tests.

Some examples of the transition markers used by participants in the study are:

- Participant 3: For example, if it is raining or the sun is very hot, additionally, you can do the same activities that you can do in a restaurant or a café like eat something or play.

- Participant 4: ...felt stressed about it because of the finals coming up
- Participant 1: ...he mentioned that if there are any other activities on campus like recreation activities and any other activities that will find to improve student skills.
- Participant 2: This was the first reason.

Some examples of the personal stance markers used by study participants are:

- Participant 4: I remember the first day when I started gymnasium.
- Participant 2: I think it is more fun to spend time with friends...
- Participant 1: My advice, my recommendation in this case, there are two ways
- Participant 3: ...or the best thing is that he should assist to some study groups

From an observational standpoint, these sequences all seemed “to be prefabricated” and not “subject to generation or analysis by the language grammar” (Wray, 2002, p. 9). They came easily to the speakers’ lips and helped to maintain the flow of their speech. However, they do seem subject to some misuse or abuse as in the example of the transition marker, “Additionally, you can do”. Clearly the speaker did not mean to make a modification to her statement, but more likely, she meant to provide a compare and contrast statement like, “Instead, you can do”. It is important to note that the error here may have nothing to do with how she had been taught these clauses in the past. There are many reasons such errors occur, and in this case, nervousness may have been the cause.

## CHAPTER 5

### DISCUSSION AND CONCLUSIONS

The purpose of this study was to replicate a study conducted by Khodadady and Shamsaee (2012) and to ascertain if similar results would be obtained when using materials from the TOEFL online preparation site instead of those from the IELTS preparation site. Four questions were asked:

1. Which transition markers and personal stance markers are used by intermediate and advanced level NNSs when practicing for a high stakes listening and speaking test (TOEFL)?
2. What is the variation by level in frequency and variety of use?
3. Do these variations have the same import on the TOEFL as other research (Khodadady & Shamsaee, 2012; Ohlrogge, 2009) claims they do on the IELTS?
4. Can a comparison of the data obtained above provide useful insights for future research and pedagogical activities?

For question #1, the results show a variety of sequences used by participants. However, the strings of words show a fairly limited vocabulary for all participants through repetitious use of various words and phrases. Nevertheless, each participant used their limited vocabulary in multiple formulaic ways (see Appendix C for complete list of FSs used by participants).

Certain sequences occurred with much greater frequency than others. An example would be “for example”, which was used in three different formulations by Participant 1:

The first one for example, which were used by chimpanzees, for example, when they use sticks to move and jump from place to other place and were too used by

insects. I think that. The other example which was the broad definition. Which were, for example, the same stick used by elephant to itching his back.

Another example would be “because it” which was used three times by

Participant 1:

The opinion of the man was that it was a bad decision because it will waste money. The man mentioned to the student association that it was a bad decision because it, he he—mean that it will waste money because that building is 200 years old...., and the man was has difficulties with the exam because it contains many functions and equations.

For question #2, the data seem to show that level does correspond with increased frequency and variety of FSs used in the case of personal stance markers. However, total word count seems to be more closely related to the overall frequency and variety of FSs than proficiency level alone.

For question #3, the results seem to indicate mixed outcomes with the TOEFL preparation materials when compared with the IELTS preparation materials. The similarity is that personal stance markers do seem to relate closely to the proficiency scores (Raw Scores in Table 5). The difference is that transition markers and fluency seem to be equally predictive of raw (proficiency) scores with the TOEFL materials, as opposed to the results Khodadady and Shamsaee recorded for the IELTS materials.

For question #4, it appears that, because of the mixed results, it may be possible to use either IELTS or TOEFL for a pretest and the other for the posttest in a study of ESL or EFL use of FSs. Because both of these exams are being used as elicitation tools, and because the results must be correlated, it could be possible to use two such very different

tests for a pretest and a posttest. However, it would be much better if the measurements were the same for both transitions and personal stance markers. The differences in the results may have been from the differences in the populations used. However, the study by Mackey and Gass (2005) states that, “Findings in second language research are highly dependent on the data collection (often known as *data elicitation*) measures used” (p. 55). Because of the many technical difficulties involved in adapting the computerized TOEFL materials to something usable in a study, it seems one could create a higher level of validity and reliability by using almost the same amount of effort to modify the IELTS materials to create a second test.

#### Limitations

This study had many limitations, not the least of which was the small number of participants. There is no statistical significance possible with such a small sample group. Fortunately, the materials used (official preparation materials for a commonly taken high stakes test) were of high interest to the participants. Otherwise, there may have been no participants. It is important to note that authenticity of the test materials played a part in there being any participants at all. The people participating in the study did it for the opportunity to practice for the TOEFL.

Another limitation was the inability of the researcher to create a true to life testing environment. One problem was the impersonal nature of the TOEFL does not work as well as the interview format of the IELTS in the relatively intimate environment created by a very small, or one to one testing situation. This led to far more verbal interactions between the test takers and the researcher than would be normal in a typical TOEFL testing situation. A second environmental problem was that the researcher had no one to

use for a dry run of the test setup other than herself. This led to a miscalculation on the difficulty of timing the various parts of the test. This difficulty with timers may or may not have affected the overall results. Having no collaborator also led to the need to use the Listening portion of the TOEFL as the measure of overall speech proficiency. This provided a far more objective measure for speech proficiency than if the lone researcher had graded the spoken portion of the assessment for grammar and syntagmatic accuracy. With no one to check the analysis of the participants' performance, the results could have been highly subjective.

A third limitation was the fact that with only one researcher, there is no inter-rater reliability for this assessment. This becomes an issue when deciding which groups of words actually qualify as the target FSs. There is no standard list of every FS of any type. Although it was pretty simple to tell where a FS began, it was often difficult to decide where it should stop. Perhaps further research on the prosodic nature of FSs (Lin, 2012) will lead to greater clarification on how to define them. For now, it is up to the researcher to clearly state the definition used and then to conscientiously apply that definition to the best of their ability.

#### Implications for future research

In 2002, Wood lamented the fact that there were no commercially available materials for teaching that recognized the significance of formulaic language in acquisition and production of an L2. In 2012, Meunier decided to do a review of the current state of the curricula available to see if the situation had changed any since Wood's complaint. In the conclusion of her review, she stated that, "though L2 teaching no longer ignores the formulaic nature of language, the exact paths to follow to better

teach it are still insufficiently lit” (p. 123). So there is much work to be done in the study of how best to teach FSs, but there seems to be little in the way of funds available for such study. Researchers often create studies that have only one or a few subjects. It is difficult to make much headway in researching such a large field with such small studies. However, if one were able to obtain hundreds of such studies, produced by teacher researchers (for instance), in a variety of locations, and aggregate them into one large study via meta-analysis, it should be possible to obtain meaningful and actionable information for the field in a relatively short period of time. One way to do this is to create a ready-made study methodology involving a pretest, posttest, and control group format that is easily adapted to various situations. If one used authentic materials (e.g., standardized test preparation materials) to create the assessment tools, one may get a higher level of volunteerism from potential subjects under circumstances that might preclude many volunteers otherwise.

Although this current study is also not able to answer the question asked in the Literature Review—what came first, proficiency or explicit teaching—it could pave the way for additional Action Research using Replication to answer that question. The cost of testing materials is prohibitive for some studies, Khodadady and Shamsaee (2012) may have found a useful approach to solving this problem. One could design a study that uses either the TOEFL preparatory materials or the IELTS preparatory materials as a pretest and the other as a post treatment test, but it may be better to simply adapt the IELTS materials. If one includes a control group each time that opts out of the explicit teaching but receives the same lessons otherwise, one could have a solid study for almost no cost. If one could do this numerous times, one should be able to obtain statistically significant



numbers. With crowd sourcing being used in many new ways lately, it seems possible one could even use an online journal to obtain help from other teacher researchers interested in the question. Under these circumstances, this small study could be the beginning of a trend.

## REFERENCES

- Biber, D. (1999). *Longman grammar of spoken and written English*. Harlow, England; New York: Longman.
- Block, D. (2000). Revisiting the gap between SLA researchers and language teachers. *Links & Letters*, 7, 129-143.
- Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, 32(83), 83-110. doi:10.1017/S0267190512000050
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. Pearson Education, Inc., White Plains, NY.
- Crookes, G. (1993). Action research for second language teachers: Going beyond teacher research. *Applied Linguistics*, 14(2), 130-144.
- Crookes, G., & Chandler, P. M. (2001). Introducing action research into the education of postsecondary foreign language teachers. *Foreign Language Annals*, 34(2), 131-140. doi:10.1111/j.1944-9720.2001.tb02818.x
- De Jong, N. H., Steinel, P., Florijn, A., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34(5), 893-916. doi:10.1017/S0142716412000069
- Dörnyei, Z. (2010). *Research methods in applied linguistics*. Oxford University Press, Oxford; New York
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83-107.
- Erlam, R. (2008). What do you researchers know about language teaching? Bridging the gap between SLA research and language pedagogy. *Innovation in Language Learning and Teaching*, 2(3), 253-267.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Interdisciplinary Journal for the Study of Discourse*, 20(1), 29-62. doi:10.1515/text.1.2000.20.1.29
- Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In Bygate, M., Skehan, P., & Swain, M. (Eds.). *Researching pedagogic tasks: second language, learning, teaching and testing* (pp. 75-93). Harlow: Longman.

- Khodadady, E. (2012). Formulaic sequences and their relationship with speaking and listening abilities. *English Language Teaching (Toronto)*, 5(2), 39-49.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London; New York: Longman.
- Lantolf, J. P. & Frawley, W. (1988). Proficiency: Understanding the construct. *Studies in Second Language Acquisition*, 10(2), 181-195.
- Lin, P. (2012). Sound evidence: The missing piece of the jigsaw in formulaic language research. *Applied Linguistics*, 33(3), 342-347.
- Luk, Z. & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning*, 59(4), 721-754.
- Mackey, A., & Gass, S. (2005). *Second language research: Methodology and design*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111-129. doi:10.1017/S0267190512000128
- Miles, W. (1964). *Arnold Lucius Gessell, 1880 – 1961: A biographical memoir by Walter R. Miles*. National Academy of Sciences retrieved from: <http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/gesell-arnold.pdf>
- Müller, S. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam; Philadelphia: John Benjamins Pub.
- Nattinger, J. R. & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford, England. Oxford University Press.
- Ohlrogge, A. (2009). Formulaic expressions in intermediate EFL writing assessment. In Corrigan, E. A. Moravcsik & H. Ouali (Eds.), *Formulaic language, volume 2 : Acquisition, loss, psychological reality, and functional explanations* (pp. 375-389). Amsterdam, NLD: John Benjamins Publishing Company.
- Pawley, A. & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds), *Language and communication*, (pp. 191-225). London: Longman.
- Porte, G. (2012). Introduction. In Porte, G. (Ed.), *Replication research in applied linguistics* (pp. 1-17). Cambridge, England: Cambridge University Press.

- Schiffrin, D. (1987). *Discourse markers*. Cambridge; New York: Cambridge University Press.
- Schmitt, N., Dörnyei, Z., Adolphs, S. & Durow, V. (2004). *Formulaic sequence: Acquisition, processing, and use*. Amsterdam; Philadelphia: John Benjamins Pub.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- Vollmer, H. J. (1981). Why are we interested in "general language proficiency"? In J. C. Alderson & A. Hughes (Eds.), *Issues in language testing* (pp. 152-175). London: The British Council.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91-120. doi:10.1111/j.1467-9922.2012.00729.x
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, 12(1), 39-57.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, U.K.; New York: Cambridge University Press.
- Wray, A. (2013). Formulaic language. *Language Teaching*, 46, 316-334. doi:10.1017/S0261444813000013

APPENDIX A

INTERVIEW INSTRUMENT

## **Interview**

This portion of our procedure is not a part of the TOEFL exam, but it will help you to relax and prepare a little for the speaking portion of the exam that we will do. It will also provide us with some necessary personal information. Please speak as clearly as possible.

1. What is the identification number assigned to you for this study?
2. Are you male or female?
3. What year were you born?
4. What is your first language?
5. Do you speak any other languages? (If yes) Which languages do you speak?
6. What country were you born in?
7. How old were you when you first began to learn English?
8. What is your level at AECP?
9. How long have you been at AECP?
10. What will you do after you finish your course work at AECP?
11. Have you ever lived in an English speaking country prior to this? (If yes) Which one(s) and for how long?
12. Have you taken the TOEFL before? (If yes) How many times?
13. Have you used the TOEFL online practice materials before? (If yes) How often, and when was the last time you did?

## APPENDIX B

### TOEFL LISTENING AND SPEAKING TEST TRANSCRIPTS

## Listening Section

**Directions:** The Listening section measures your ability to understand conversations and lectures in English. In this sample, you will read one conversation and one lecture and answer questions after each conversation or lecture. The questions typically ask about the main idea and supporting details. Some questions ask about a speaker's purpose or attitude. Answer the questions based on what is stated or implied by the speakers. Most questions are worth one point. If a question is worth more than one point, it will have special directions that indicate how many points you can receive.

- In an actual test, you will be able to take notes while you listen and use your notes to help you answer the questions. Your notes will not be scored.

### CONVERSATION TRANSCRIPT

**(Narrator)** Listen to a conversation between a student and her basketball coach and then answer the questions.

**(Male coach)** Hi, Elizabeth.

**(Female student)** Hey, Coach. I just thought I'd stop by to see what I missed while I was gone.

**(Male coach)** Well, we've been working real hard on our plan for the next game . . . I've asked Susan to go over it with you before practice this afternoon, so you'll know what we're doing.

**(Female student)** Okay.

**(Male coach)** By the way, how did your brother's wedding go?

**(Female student)** Oh, it was beautiful. And the whole family was there. I saw aunts and uncles and cousins I hadn't seen in years.

**(Male coach)** So it was worth the trip.

**(Female student)** Oh definitely. I'm sorry I had to miss practice, though. I feel bad about that.

**(Male coach)** Family's very important.

**(Female student)** Yep. Okay, I guess I'll see you this afternoon at practice, then.

**(Male coach)** Just a minute. There are a couple of other things I need to tell you.



**(Female student)** Oh, okay.

**(Male coach)** Uh . . . First, everybody's getting a new team jacket.

**(Female student)** Wow. How did that happen?

**(Male coach)** A woman who played here about 20, 25 years ago came through town a few weeks ago and saw a game, and said she wanted to do something for the team, so . . .

**(Female student)** So she's buying us new jackets?

**(Male coach)** Yep.

**(Female student)** Wow, that's really nice of her.

**(Male coach)** Yes, it is. It's great that former players still care so much about our school and our basketball program . . . Anyway you need to fill out an order form. I'll give it to you now, and you can bring it back this afternoon. I've got the forms from the other players, so as soon as I get yours we can order. Maybe we'll have the jackets by the next game.

**(Female student)** OK.

**(Male coach)** Great. And the next thing is, you know Mary's transferring to another college next week, so we'll need someone to take over her role as captain for the second half of the season. And the other players unanimously picked you to take over as captain when Mary leaves.

**(Female student)** Wow. I saw everybody this morning, and nobody said a word.

**(Male coach)** They wanted me to tell you. So, do you accept?

**(Female student)** Of course! But Susan's a much better player than I am. I'm really surprised they didn't pick her.

**(Male coach)** They think you're the right one. You'll have to ask them their thoughts.

**(Female student)** Okay . . . I guess one of the first things I'll have to do as captain is make sure we get a thank-you card out to the lady who's buying us the jackets.

**(Male coach)** Good idea. I have her address here somewhere.

**(Female student)** And I'll make sure the whole team signs it.

**(Male coach)** Good. That's all the news there is. I think that's it for now. Oh, let me get you that order form.

1. What are the speakers mainly discussing?
  - a. How the woman should prepare for the next game
  - b. The woman's responsibilities as team captain
  - c. Things that happened while the woman was away
  - d. The style of the new team uniforms

2. Who is buying new jackets for the team?
  - a. The coach
  - b. The captain of the team
  - c. A former player
  - d. A group of basketball fans

3. There are two answers for the next question. Mark two answers.

Why is the woman surprised to learn that she has been chosen as the new team captain?

- a. She is not the best player on the team.
- b. Her teammates did not tell her about the decision.
- c. She does not have many friends on the team.
- d. She has missed a lot of practices.

4. Read part of the conversation again. Then answer the question.

**(Female student)** I'm sorry I had to miss practice, though. I feel bad about that.

**(Male coach)** Family's very important.

What does the man mean when he says: "Family's very important."

- a. He hopes the woman's family is doing well.
  - b. He would like to meet the woman's family.
  - c. The woman should spend more time with her family.
  - d. The woman had a good reason for missing practice.
5. Why does the coach say: "Good. That's all the news there is. I think that's it for now."
- a. He wants to know if the woman understood his point.
  - b. He wants the woman to act immediately.
  - c. He is preparing to change the topic.
  - d. He is ready to end the conversation.

## LECTURE TRANSCRIPT

**(Narrator)** Listen to part of a lecture in a literature class.

**(Male professor)** Today I'd like to introduce you to a novel that some critics consider the finest detective novel ever written. It was also the first. We're talking about *The Moonstone* by Wilkie Collins. Now, there are other detective stories that preceded *The Moonstone* historically—Um, notably the work of Poe . . . Edgar Allen Poe's stories, such as "The Murders in the Rue Morgue" and . . . "The Purloined Letter." Now these were short stories that featured a detective . . . uh, probably the first to do that. But *The Moonstone*, which follows them by about twenty years—it was published in 1868—this is the first full-length detective novel ever written.

Now, in *The Moonstone*—if you read it as . . . uh, come to it as a contemporary reader—what's interesting is that most of the features you find in almost any detective novel are in fact already present. Uh, its hard at this juncture to read this novel and realize that no one had ever done that before, because it all seems so strikingly familiar. It's, it's really a wonderful novel and I recommend it, even just as a fun book to read, if you've never read it. Um, so in *The Moonstone*, as I said, Collins did much to establish the conventions of the detective genre. I'm not gonna go into the plot at length, but, you know, the basic set-up is . . . there's this diamond of great . . . of great value, a country house, the diamond mysteriously disappears in the middle of the night, uh, the local police are brought in, in an attempt to solve the crime, and they mess it up completely, and then the true hero of the book arrives. That's Sergeant Cuff.

Now, Cuff, this extraordinarily important character . . . well, let me try to give you a sense of who Sergeant Cuff is, by first describing the regular police. And this is the dynamic that you're going to see throughout the history of the detective novel, where you have the regular cops—who are well-meaning, but officious and bumblingly inept—and they are countered by a figure who's eccentric, analytical, brilliant, and . . . and able to solve the crime. So, first the regular police get called in to solve the mystery—Um, in this case, detective, uh, Superintendent Seegrave. When Superintendent Seegrave comes in, he orders his minions around, they bumble, and they actually make a mess of the investigation, which you'll see repeated—um, you'll see this pattern repeated, particularly in the Sherlock Holmes stories of a few years later where, uh, Inspector Lestrade, this well-meaning idiot, is always countered, uh, by Sherlock Holmes, who's a genius.

So, now Cuff arrives. Cuff is the man who's coming to solve the mystery, and again he has a lot of the characteristics that future detectives throughout the history of this genre will have. He's eccentric. He has a hobby that he's obsessive about—in this . . . in his case, it's the love of roses. He's a fanatic about the breeding of roses; and here think of Nero Wolfe and his orchids, Sherlock Holmes and his violin, a lot of those later classic detective heroes have this kind of outside interest that they . . . they go to as a kind of antidote to the evil and misery they encounter in their daily lives. At one point, Cuff says he likes his roses because they offer solace, uh, an escape, from the world of crime he typically operates in.

Now, these detective heroes . . . they have this characteristic of being smart, incredibly smart, but of not appearing to be smart. And most importantly, from a kind of existential point of view, these detectives see things that other people do not see. And that's why the detective is such an important figure, I think, in our modern imagination. In the case of *The Moonstone*—I don't want to say too much here and spoil it for you—but the clue that's key to . . . the solving of the crime is a smeared bit of paint in a doorway. Of course, the regular police have missed this paint smear or made some sort of unwarranted assumption about it. Cuff sees this smear of paint—this paint, the place where the paint is smeared—and realizes that from this one smear of paint you can actually deduce the whole situation . . . the whole world. And that's what the hero in a detective novel like this . . . brings to it that the other characters don't—it's this ability to, uh, see meaning where others see no meaning and to bring order . . . to where it seems there is no order.

6. What is the lecture mainly about?

- a. A comparison of two types of detective novels
- b. Ways in which detective novels have changed over time
- c. *The Moonstone* as a model for later detective novels
- d. Flaws that can be found in the plot of *The Moonstone*

7. In what way is *The Moonstone* different from earlier works featuring a detective?

- a. In its unusual ending
- b. In its unique characters
- c. In its focus on a serious crime
- d. In its greater length

8. According to the professor, what do roses in *The Moonstone* represent?
- a. A key clue that leads to the solving of the mystery
  - b. A relief and comfort to the detective
  - c. Romance between the main characters
  - d. Brilliant ideas that occur to the detective
9. Why does the professor mention a smeared bit of paint in a doorway in *The Moonstone*?
- a. To describe a mistake that Sergeant Cuff has made
  - b. To show how realistically the author describes the crime scene
  - c. To exemplify a pattern repeated in many other detective stories
  - d. To illustrate the superior techniques used by the police
10. What can be inferred about the professor when he says this: “Uh, it’s hard at this juncture to read this novel and realize that no one had ever done that before, because it all seems so strikingly familiar.”
- a. He is impressed by the novel’s originality.
  - b. He is concerned that students may find the novel difficult to read.
  - c. He is bored by the novel’s descriptions of ordinary events.
  - d. He is eager to write a book about a less familiar subject.

11. What does the professor imply when he says this: “. . . well, let me try to give you a sense of who Sergeant Cuff is, by first describing the regular police.”
- a. Sergeant Cuff is unlike other characters in *The Moonstone*.
  - b. The author’s description of Sergeant Cuff is very realistic.
  - c. Sergeant Cuff learned to solve crimes by observing the regular police.
  - d. Differences between Sergeant Cuff and Sherlock Holmes are hard to describe.

Key to Listening section:

- 1. c
- 2. c
- 3. a, b
- 4. d
- 5. d
- 6. c
- 7. d
- 8. b
- 9. c
- 10. a
- 11. a

## Speaking Section

**Directions:** The Speaking section in the test measures your ability to speak about a variety of topics.

- In questions 1 and 2, in an actual test, your response will be scored on your ability to speak clearly and coherently about familiar topics.
- In questions 3 and 4, in an actual test, you will first read a short text and then listen to a talk on the same topic. You will have to combine appropriate information from the text and the talk to provide a complete answer. Your response will be scored on your ability to accurately convey information, and to speak clearly and coherently. In this sampler, you will **read** both the text and the talk.
- In questions 5 and 6, in an actual test, you will listen to part of a conversation or lecture. Then, you will be asked a question about what you have heard. Your response will be scored on your ability to accurately convey information, and to speak clearly and coherently. In this sampler, you will **read** the conversation.
- In an actual test, you will be able to take notes while you read and while you listen to the conversations and talks. You may use your notes to help prepare your responses.
- Preparation and response times for an actual test are noted in this text. Candidates with disabilities may request time extensions.
- Sample candidate responses and score explanations can be found in the online version of the sampler. The scoring rubric used to score actual responses can be found on the TOEFL website's "Download Library" page.

1. Talk about a pleasant and memorable event that happened while you were in school. Explain why this event brings back fond memories.

**Preparation Time: 15 seconds**

**Response Time: 45 seconds**



2. Some people think it is more fun to spend time with friends in restaurants or cafés. Others think it is more fun to spend time with friends at home. Which do you think is better? Explain why.

**Preparation Time: 15 seconds**

**Response Time: 45 seconds**

3. Read the following text and the conversation that follows it. Then, answer the question.

The Northfield College Student Association recently decided to make a new purchase. Read the following announcement in the college newspaper about the decision. (Reading time in an actual test would be 45-50 seconds.)

### **Good News for Movie Fans**

The Student Association has just purchased a new sound system for the Old Lincoln Hall auditorium, the place where movies on campus are currently shown. By installing the new sound system, the Student Association hopes to attract more students to the movies and increase ticket sales. Before making the purchase of the new equipment, the Student Association conducted a survey on campus to see what kind of entertainment students liked best. Going to the movies ranked number one. “Students at Northfield College love going to the movies” said the president of the Student Association, “so we decided to make what they already love even better. We’re confident that the investment into the sound system will translate into increased ticket sales.”

**(Male student)** I really think the Student Association made a bad decision.

**(Female student)** Really? Why? Don’t you like going to the movies?

**(Male student)** Sure I do. But this new purchase is just a waste of money.

**(Female student)** What do you mean? It’s supposed to sound really good.

**(Male student)** Yeah, well, I’m sure it does, but, in Old Lincoln Hall? I mean that building must be 200 years old! It used to be the college gym! The acoustics are terrible.

**(Female student)** So you’re saying there’ll be no improvement?

**(Male student)** That’s right. And also, I seriously doubt that going to the movies is the number one social activity for most students.

**(Female student)** Yeah, but that’s what students said.

**(Male student)** Well, of course that's what they said. What else is there to do on campus?

**(Female student)** What do you mean?

**(Male student)** I mean, there isn't much to do on campus besides go to the movies. If there were other forms of, uh recreation, or other social activities, you know, I don't think most students would have said that going to the movies was their first choice.

**Question:** The man expresses his opinion of the Student Association's recent purchase. State his opinion and explain the reasons he gives for holding that opinion.

**Preparation Time: 30 seconds**

**Response Time: 60 seconds**

4. Read a passage from a psychology textbook and the lecture that follows it. Then answer the question. (Reading time in an actual test would be 45-50 seconds.)

### **Flow**

In psychology, the feeling of complete and energized focus in an activity is called flow. People who enter a state of flow lose their sense of time and have a feeling of great satisfaction. They become completely involved in an activity for its own sake rather than for what may result from the activity, such as money or prestige. Contrary to expectation, flow usually happens not during relaxing moments of leisure and entertainment, but when we are actively involved in a difficult enterprise, in a task that stretches our mental or physical abilities.

**(Male professor)** I think this will help you get a picture of what your textbook is describing. I had a friend who taught in the physics department, Professor Jones, he retired last year. . . . Anyway, I remember . . . this was a few years ago . . . I remember passing by a classroom early one morning just as he was leaving, and he looked terrible: his clothes were all rumpled, and he looked like he hadn't slept all night. And I asked if he was OK. I was surprised when he said that he never felt better, that he was totally happy. He had spent the entire night in the classroom working on a mathematics puzzle. He didn't stop to eat dinner; he didn't stop to sleep . . . or even rest. He was that involved in solving the puzzle. And it didn't even have anything to do with his teaching or research; he had just come across this puzzle accidentally, I think in a mathematics journal, and it just really interested him, so he worked furiously all night and covered the blackboards in the classroom with equations and numbers and never realized that time was passing by.

**Question:** Explain *flow* and how the example used by the professor illustrates the concept.

**Preparation Time: 30 seconds**

**Response Time: 60 seconds**

5. Read the following conversation between two students and then answer the question.

**(Female student)** How's the calculus class going? You're doing better?

**(Male student)** Not really. I just can't get the hang of it. There're so many functions and formulas to memorize, you know? And the final . . . It's only a few weeks away. I'm really worried about doing well.

**(Female student)** Oh . . . You know, you should go to the tutoring program and ask for help.

**(Male student)** You mean, in the Mathematics building?

**(Female student)** Ya. Get a tutor there. Most tutors are doctoral students in the math program. They know what they're talking about, and for the final test, you know, they'd tell you what to study, how to prepare, all of that.

**(Male student)** I know about that program . . . but doesn't it cost money?

**(Female student)** Of course. You have to register and pay by the hour . . . But they've got all the answers.

**(Male student)** Hmm . . .

**(Female student)** Another option, I guess, is to form a study group with other students. That won't cost you any money.

**(Male student)** That's a thought . . . although once I was in a study group, and it was a big waste of time. We usually ended up talking about other stuff like what we did over the weekend.

**(Female student)** But that was for a different class, right? I've actually had some pretty good experiences with study groups. Usually students in the same class have different strengths and weaknesses with the material . . . if they're serious about studying, they can really help each other out. Think about it.

**Question:** Briefly summarize the problem the speakers are discussing. Then state which solution you would recommend. Explain the reasons for your recommendation.

**Preparation Time: 20 seconds**

**Response Time: 60 seconds**

6. Read part of a lecture in a biology course and then answer the question.

**(Female professor)** Human beings aren't the only animals that use tools. It's generally recognized that other animals use tools as well . . . use them naturally, in the wild, without any human instruction. But when can we say that an object is a tool? Well, it depends on your definition of a tool. And in fact, there are two competing definitions—a narrow definition and a broad one. The narrow definition says that a tool is an object that's used to perform a specific task . . . but not just any object. To be a tool, according to the narrow definition, the object's gotta be purposefully changed or shaped by the animal, or human, so that it can be used that way. It's an object that's made. Wild chimpanzees use sticks to dig insects out of their nests . . . but most sticks lying around won't do the job . . . they might be too thick, for example. So the sticks have to be sharpened so they'll fit into the hole in an ant hill or the insect nest. The chimp pulls off the leaves and chews the stick and trims it down that way until it's the right size. The chimp doesn't just find the stick . . . it . . . you could say it makes it in a way.

But the broad definition says an object doesn't have to be modified to be considered a tool. The broad definition says a tool is any object that's used to perform a specific task. For example, an elephant will sometimes use a stick to scratch its back . . . it just picks up a stick from the ground and scratches its back with it . . . It doesn't modify the stick, it uses it just as it's found. And it's a tool, under the broad definition, but under the narrow definition it's not because, well, the elephant doesn't change it in any way.

**Question:** Using points and examples from the talk, describe the two different definitions of tools given by the professor.

**Preparation Time: 20 seconds**

**Response Time: 60 seconds**

## APPENDIX C

### COMPLETE LIST OF FORMULAIC SEQUENCES TALLIED

**Participant 1 I-1**

# of personal stance markers: 5

I remember that

I think it's better to

My advice

My recommendation in this case

I think that

# of transitions: 16

because there will be

because it will

because it

because that

and he mentioned that if there are any other

and any other

because it

the first way is

There is another

which will be

which were the first

and the second

the first one for example

for example

the other example which was

which were for example

**Participant 2 I-2**

# of personal stance markers: 6

I think it is more

I prefer to

That will give me

I will enjoy

I think the

I am sure, I cannot do this

# of transitions: 10

because that will

rather than

This was the first

The second

because the... may have

only because the...are not

The first was to

and the second was to

So the problem for

but the rest

**Participant 3 I-2**

# of personal stance markers: 3

I think it is better

you can do another thing...like

the best thing is that

# of transitions: 9

because you can

or you can

For example, if it is

additionally, you can do

And another thing is that maybe

because he can

and he can

the first time is

And another...is like



**Participant 4 A-2**

# of personal stance markers: 11

I remember the first

I prefer to

you don't really have to

But I still thinks

So he thinks it's a bad idea

The only reason why

It would be better to

You don't really care about

He didn't really have to

I would recommend the second option

it could be good

# of transitions: 29

but it was

but as soon as

So it was really

'cuz you can

But at the same time

you can do some

so he

And then he also says that

because there's nothing else to do

And this man

and he was

And this

and this man

and he said

and he was

And he started

which he

and he

because of the...coming up

then another option is

because it's

instead of

'cuz they have

first it's the

which means that

one example is

because otherwise it won't be




And then there is the

And one example is

## APPENDIX D

### IELTS 2005 PREPARATORY MATERIALS

# Listening Module

Candidate Name: _____	Candidate Number: _____								
<b>INTERNATIONAL ENGLISH LANGUAGE TESTING SYSTEM</b>									
<b>LISTENING</b>									
<b>SPECIMEN MATERIALS</b>									
Approx. 30 minutes									
<b>TIME</b> Approx. 30 minutes									
<b>INSTRUCTIONS TO CANDIDATES</b>									
Do not open this booklet until you are told to do so.									
Write your name and candidate number in the spaces at the top of this page.									
You should answer all questions.									
All the recordings will be played ONCE only.									
Write all your answers on the Question Paper.									
At the end of the test you will be given ten minutes to transfer your answers to an answer sheet.									
Do not remove this booklet from the examination room.									
<b>INFORMATION FOR CANDIDATES</b>									
There are 40 questions on this question paper.									
The test is divided as follows:									
<table border="1" style="width: 100%; border-collapse: collapse;"><tr><td style="width: 50%; padding: 5px;">Section 1</td><td style="width: 50%; padding: 5px;">Questions 1 – 10</td></tr><tr><td style="padding: 5px;">Section 2</td><td style="padding: 5px;">Questions 11 – 20</td></tr><tr><td style="padding: 5px;">Section 3</td><td style="padding: 5px;">Questions 21 – 30</td></tr><tr><td style="padding: 5px;">Section 4</td><td style="padding: 5px;">Questions 31 – 40</td></tr></table>		Section 1	Questions 1 – 10	Section 2	Questions 11 – 20	Section 3	Questions 21 – 30	Section 4	Questions 31 – 40
Section 1	Questions 1 – 10								
Section 2	Questions 11 – 20								
Section 3	Questions 21 – 30								
Section 4	Questions 31 – 40								
<div style="display: flex; justify-content: space-around; align-items: center;"><div style="text-align: center;"><b>BRITISH COUNCIL</b></div><div style="text-align: center;"><b>IELTS AUSTRALIA</b></div><div style="text-align: center;"><b>UNIVERSITY of CAMBRIDGE ESOL Examinations</b></div></div>									

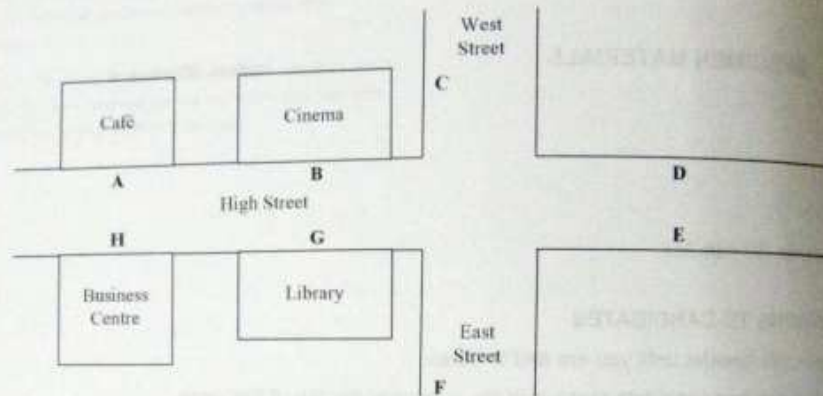
## SECTION 1

### Questions 1 – 10

Questions 1 and 2

Label the map below.

Write the correct letter A-H next to questions 1 and 2.



Example

Answer

Mrs Reynolds:

**B**

1 Elderly woman:

2 Thieves' car:

Questions 3 – 5

Choose the correct letter, A, B or C.

- 3 When Mrs Reynolds saw the thieves, she
- A ran after them.
  - B telephoned the police.
  - C went to help the elderly woman.
- 4 The elderly woman was
- A badly hurt.
  - B unhurt.
  - C very upset.
- 5 Which woman had once had her bag stolen in the past?
- A the elderly woman
  - B Mrs Reynolds
  - C Mrs Reynolds' friend

Question 6

Choose **TWO** letters, A-F.

- 6 The bag contained
- A a purse.
  - B £50.
  - C a cheque book.
  - D a cheque card.
  - E a bus pass.
  - F a door key.

Questions 7 – 10

Complete the table below.

Write **NO MORE THAN TWO WORDS AND/OR A NUMBER** for each answer.

	age	build	hair colour	distinguishing marks
younger man	about 17	7 .....	8 .....	none
older man	about 9 .....	medium	brown	a 10 ..... on chin

**SECTION 2**

**Questions 11 – 20**

*Questions 11 – 13*

*Choose the correct letter, A, B or C.*

- 11 The Bridge Hotel is located in
- A the city centre.
  - B the country.
  - C the suburbs.
- 12 The newest sports facility in the hotel is
- A a swimming pool.
  - B a fitness centre.
  - C a tennis court.
- 13 The hotel restaurant specialises in
- A healthy food.
  - B local food.
  - C international food.

*Questions 14 and 15*

*Choose TWO letters, A-E.*

Which **TWO** business facilities are mentioned?

- A internet access
- B mobile phone hire
- C audio-visual facilities
- D airport transport
- E translation services

Questions 16 – 20

Complete the table below.

Write **NO MORE THAN THREE WORDS AND/OR A NUMBER** for each answer.

SHORT BREAK PACKAGES		
Length of stay	Cost (per person per night)	Special features
2 days	16 £ .....	Full cooked breakfast Entertainment in the 17 .....
3 days	£60	As above, plus — a 18 .....
5 days	19 £ .....	As above, plus — 2 days' free beauty therapy — full-day membership of a 20 .....



**SECTION 3**

**Questions 21 – 30**

Questions 21 and 22

Complete the sentences below.

Write **NO MORE THAN TWO WORDS AND/OR A NUMBER** for each answer.

21 Last year, Dina got a grade ..... for the Theory and Practice option.

22 Dina has some free time because her ..... has been cancelled.

Question 23

Choose **ONE** letter, **A-E**.

Which book does Dina advise against?

A Brown: *Observing Theory in Practice*

B Jespersen: *Theory's Crucible*

C Piresi: *On Giants' Shoulders*

D Willard: *Practical Theories in the Social Sciences*

E Williams: *Knowledge Theory*

*Questions 24 – 30*

*Complete the notes below.*

*Write NO MORE THAN TWO WORDS AND/OR A NUMBER for each answer.*

**How to use the Recall System**

- Take a **24** ..... from librarian's desk.
- Complete the details of the book. Write your **25** ..... address on back. Hand it in at the **26** .....
- Check mail in your department twice a day to see if book is ready to collect. It normally takes 3 days.
- Cost: **27** ..... per book.

**Dina's advice on organising a study group to work on an assignment**

- Find two or three people on the course who live near you.
- Divide up the reading load.
- Take it in turns to **28** ..... what you have read for the others.
- Explain your **29** ..... to each other.
- Write first draft of essay.
- Exchange drafts and **30** .....
- Write final version of essay.

**SECTION 4**

**Questions 31 – 40**

Questions 31 – 33

Complete the sentences below.

Write **NO MORE THAN THREE WORDS** for each answer.

**PEREGRINE FALCONS**

- 31 The Peregrine falcons found in ..... are not migratory birds.
- 32 There is disagreement about their maximum .....
- 33 When the female is guarding the nest, the male spends most of his time .....

Questions 34 – 37

Complete the table below.

Write **NO MORE THAN THREE WORDS AND/OR A NUMBER** for each answer.

Age of falcons	What occurs
20 days old	The falcons 34 .....
28 days old	The falcons are 35 .....
2 months old	The falcons 36 ..... permanently
1-12 months old	More than half of falcons 37 .....

Questions 38 – 40

Complete the notes below.

Write **NO MORE THAN THREE WORDS** for each answer.

**Procedures used for field research on Peregrine falcon chicks**

First: catch chicks

Second: **38** ..... to legs

Third: **39** ..... of chicks

Fourth: take blood sample to assess level of pesticide

Fifth: check the **40** ..... of the birds

## Speaking Module

The structure of the module is summarised below.

Part	Nature of interaction	Timing
<b>Part 1</b> Introduction and interview	Examiner introduces him/herself and confirms candidate's identity. Examiner interviews candidate using verbal questions selected from familiar topic frames	4-5 minutes
<b>Part 2</b> Individual long turn	Examiner asks candidate to speak for 1-2 minutes on a particular topic based on written input in the form of a candidate task card with content-focused prompts. Examiner asks one or two questions to round off the long turn.	3-4 minutes (including 1 minute preparation time)
<b>Part 3</b> Two-way discussion	Examiner invites candidate to participate in discussion of a more abstract nature, based on questions thematically linked to Part 2 topic.	4-5 minutes

### EXAMPLE PART ONE

Let's talk about your home town or village

- What kind of place is it?
- What's the most interesting part of your town/village?
- What kind of jobs do the people in your town/village do?
- Would you say it's a good place to live? [Why?]

**EXAMPLE PART TWO: Candidate task card**

Describe something you own which is very important to you.

You should say:

where you got it from  
how long you have had it  
what you use it for

and explain why it is important to you.

You will have to talk about the topic for 1 to 2 minutes.  
You have one minute to think about what you're going to say.  
You can make some notes to help you if you wish.

**EXAMPLE PART TWO: Rounding-off questions**

- Is it valuable in terms of money?
- Would it be easy to replace?

**EXAMPLE PART THREE**

Let's consider first of all how people's values have changed

- What kind of things give status to people in your country?
- Have things changed since your parents' time?

Finally, let's talk about the role of advertising

- Do you think advertising influences what people buy?