

International Language Testing Washback – standing the monster on its head

Thor May
Brisbane 2013



International language testing, and particularly English language testing, is a multi-billion dollar business. For language students, both employment chances and international working or study visas are mediated by such tests. The tests engage the ingenuity of millions of students in a struggle to master or outwit purpose designed language obstacle courses. In this life transition institutional language learners are assisted by an army of specialists who might or might not be skilled in the process, rather as corporations are assisted by armies of lawyers and para-legals, or hospitals assisted by an entourage of medical professionals.

At the top of the assessment pyramid are multinational testing corporations, best known by the names of their standardized tests, such as IELTS, TOEIC, TOEFL, BULATS, TKT, Cambridge ESOL main suite, or G-TELP (there are many other aspirants). In some ways these testing companies can be thought of as the Big Pharma corporations (i.e. drug companies) of the educational world. Like Big Pharma they are subject to constant challenges to their ethics and reliability from within and without, and like Big Pharma they are rather prone to corrupt the issues which they were designed to assist with.

The possible corruption of language learning by the requirements of testing is known as wash-back. Wash-back is not always malignant. Language learning is typically a long, complex and discouraging experience, especially in mass educational institutions. The clear objective of a high-stakes exam at the end of study can be powerfully motivating. The explicit requirements of external language tests can organize the multitudes of teachers and institutions who have only a tenuous grasp of productive language teaching processes.

Those countries committed to the mass-skilling of their populations in foreign languages have also become aware that home-grown diplomas of language achievement often convey little reliable information about the performative ability of their graduates in foreign languages. International language tests claim to overcome this handicap. However, experience has repeatedly shown that the international tests are also elusive guarantors. For example, a level 6.5 score in IELTS may be a poor predictor of success for students venturing into English language tertiary institutions overseas. Both the learned methodologies for mastering such tests, and the actual language skills which they teach may scarcely prepare the new graduate for real life success in a foreign language environment. There are often good arguments that test wash-back into schools has actively inhibited the acquisition of real language ability.

The dilemmas of learning and testing just outlined have no easy solutions. They engage difficult to control collisions of politics, profits, aspirations and realized skills. The mixture is different in every country, it varies with social strata, and it varies with the culture and structure of every local college. That is, there is unlikely to be any simple fix for wash-back in mass educational institutions.

However there can be stratagems to bend the process in local ways for maximum language learning benefit.

The analysis which follows is a tentative attempt to manipulate the wash-back from an international test in a manner which actually assists genuine language acquisition. The test chosen for manipulation is G-TELP, which is less well known than major players like IELTS. Nevertheless the discussion and principles engaged can apply anywhere.

The present analysis is in fact an assembly of four Masters level degree assignments in a single sequence. This work was completed in 2005 at the University of Newcastle, NSW. I thought it would be useful to collect the assignments here as a spur to creative thinking on a very important topic (not as a crib for lazy degree candidates!). I have retained the original cover pages with the assignment questions since this seems a useful device to provoke a response from readers themselves. The opinions expressed are of course my own, and are open to fault or amendment.

Part I - Test Specification

- Goal Evaluation From A Proficiency Test

TABLE OF CONTENTS

Statement of the Problem.....	11
Problems common to large scale standardized tests.....	13
Test Content.....	16
Grammar.....	16
Listening	17
Speaking.....	18
Reading & Vocabulary	19
Writing.....	20
Types of text.....	22
Listening	22
Speaking.....	22
Reading	22
Writing.....	22
Addressees of texts	23
Lengths of texts	23
Topics	23
Readability	23
Structural Range.....	23
Vocabulary Range	24
Dialect and style	24
Speed of Processing.....	24
Task Procedure, Timing and Medium.....	24
Task Procedure	24
Number of Items	25
Timing.....	26
Medium	26
Organization	26
Techniques.....	27

Section 1 Grammar.....	27
Section 2 Listening	27
Section 3 Speaking.....	28
Section 4 Reading	28
Section 5 Writing.....	28
Criteria levels of performance and Marking Scales	28
Scoring Context	28
Scoring for the diagnostic purposes of the current test cycle	31
Grammar.....	31
Listening	31
Speaking.....	31
Reading	32
Writing.....	33
Test Construction Procedures.....	33
Sampling.....	33
Item writing and Moderation.....	34
Informal Trialling	34
Trialling and Analysis	34
Validation	35
Speaking - concurrent validity.....	35
Candidate Instructions.....	35
Teacher’s Handbook	36
Rationale:.....	36
Development and Validation:.....	37
Description of the test:	37
Sample items:	38
Section 1 Grammar	38
Section 2 Listening	38
Section 3 Speaking	39
Section 4 Reading.....	39
Section 5 Writing	39
Advice on preparing for the test	39
Interpretation of test scores	40
Training Materials:	40

Test administration:	41
Handbook for Candidates.....	41
How to prepare for the test	42
What do your test scores mean?	42
What is in the test ?	43
Sample test items:	44
Section 1 Grammar	44
Section 2 Listening	44
Section 3 Speaking	44
Section 4 Reading	45
Section 5 Writing	45
APPENDIX 1 G-TELP Speaking Score Criteria	45
G-TELP Proficiency Levels for Speaking	46
APPENDIX 2 G-TELP Writing Score Criteria	48
APPENDIX 3 G-TELP CONTENT SPECIFICATIONS FOR LEVEL 3.....	50
- Grammar	50
- Listening	51
- Reading and Vocabulary	51
- Writing	52
=> Task Contents	52
- Speaking.....	53
Test Specification References.....	136

Part II Test Construction

TABLE OF CONTENTS

Assignment topic & comments: Writing Test Items for listening, reading, writing, speaking, grammar / vocabulary	56
Listening Test.....	58
Reading Test.....	62
Writing Test.....	67
Speaking Test	73
Grammar Test.....	76
Appendix 1 Listening Skills.....	83
Appendix 2 - Reading Test.....	84
Appendix 3 - Speaking and Writing Tests - Picture Panel Cues for text construction	86
Appendix 4 - G-TELP Specifications for Levels 3 to 5, Grammar Tests	87

Part III - PRACTICE ENGLISH LANGUAGE TEST

Level : Intermediate

(based on the G-TELP international language test)

TABLE OF CONTENTS

Part III - PRACTICE ENGLISH LANGUAGE TEST	91
Test Instructions	91
1. Speaking Test	92
Instructions	92
Hints For Talking About The Pictures	92
Picture Panels Showing The History of Smallpox	93
2. LISTENING TEST	93
General Instructions.....	93
2. LISTENING TEST - SECTION 1	94
Recording : Telephone Duties for a Secretary	94

Part A : Multiple Choice Questions	94
Listening - Section 1, Part B - Dialogue Choice	95
LISTENING TEST - SECTION B.....	96
Recording : “Haven’t We Met Before?”	96
Listening Part A - Dialogue Completion	96
Listening - Section 2, Part B : Multiple Choice Questions	96
3. Reading Test.....	97
General Instructions	97
Reading Section 1 - A Tourist Brochure	98
Reading Section 1 - Multiple Choice Questions (Ideas	99
Reading Section 1 - Multiple Choice Questions (Vocabulary).....	100
Reading Section 2 - An Interview.....	101
Reading Section 2 - Multiple Choice Questions (Ideas).....	102
Reading Section 2 - Multiple Choice Questions (Vocabulary).....	102
4. Writing Test	103
Hints For Writing A Business E-mail	103
5. Grammar Test	105
Grammar - Section A (Multiple Choice Questions)	105
Grammar - Section B (Phrase Completion)	108
Grammar - Section C (Sentence Correction)	109

Part IV – Practical Aspects of Language Test Construction

TABLE OF CONTENTS

1. The notion of testing, and practical consequences	112
2. Test Construction.....	113
4. The Purpose of the Practice Test	115
5. The Application of the Practice Test applied to its purpose	116
6. The Content of the Test	116
7. Test Specifications.....	117
8. Instructions to Candidates	118
9. The Quality of Individual Test Items.....	119
10. The Representativeness of the Sample.....	119
11. Reliability and Validity.....	120
12. Scoring Procedures	120

13. Aspects of Moderation	121
14. Summary Evaluation of the Test.....	121
Appendix 1 : Error Summary of Kim Joon-woo’s Practice Test	123
1. Listening Test.....	123
2. Reading Test.....	123
Tester comment.....	123
3. Grammar Test.....	124
Tester comment :	126
4. Writing Test.....	127
Tester comments :	128
a) Formal grammatical errors :	129
b) Register (the language of social marking).....	129
i) The erratic use of register	129
ii) Terms of address	130
iii) Discourse etiquette.....	130
c) Information Content.....	131
5. Speaking Test	132
Testee transcript	132
Tester comment.....	133
Dialect.....	133
Fluency	133
Formal grammatical errors	134
Lexical choice	134
Appendix 2 Post-test e-mail from the Candidate	135
Appendix 3 - References	137

University of Newcastle, NSW

Assignment 1 : Language Testing And Evaluation

Topic : Write a set of test specifications for a comprehensive ESL proficiency OR achievement test

Provide : 1) Test Specifications
 2) Instructions to Candidates
 3) Criterial Levels of Performance

Note on Topic adaptation : The set topic has been adapted in a non-conventional way. A proficiency test widely used in South Korea has been used as a model for initiating a teacher-assisted, student process of self-diagnosis and planning.

The open purpose of this test framework developed here is to provide English Major students in a Korean university with information about their own skills relative to an external proficiency test commonly used by large employers in South Korea. The test design in this project is therefore that of a proficiency test, but its function is diagnostic and ultimately motivational.

The underlying teacher-purpose of this test is to harness the backwash of external test motivation by addressing it directly, giving students clear instrumental goals relative to their own strengths and weaknesses, then build on (subvert?) this foundation into genuine communicative language abilities.

Test Specification - Goal Evaluation From A Proficiency Test

Statement of the Problem

The open purpose of this test is to provide English Major students in a Korean university with information about their own skills relative to an external proficiency test commonly used by large employers in South Korea. The test design in this project is therefore that of a proficiency test, but its function is diagnostic and ultimately motivational. Because this test is expected to be integral in a teaching program, the sample presented here will only be the first of a series in a cycle, with each successive cycle increasing in complexity and length. Each completed test cycle will be subject to detailed post-mortem and critique by both the students and the teacher.

The external test in question is the American G-TELP . The G-TELP claims to have a functional, task-oriented approach and to be criterion referenced. Hence its attraction to employers. In fact the G-TELP, like most of the major international language tests, has some quite striking cultural biases and sometimes dubious linguistic content, at least from the viewpoint of a Korean learner.

Nevertheless, most English learning in South Korea is instrumentally driven, and the pressure above all is to excel in tests like G-TELP which will lead to employment (regardless of actual L2 fluency).

Note that the real usage of English in Korean business frequently involves regional varieties of English rather than American English. Accepting such variation is particularly important in spoken language, although G-TELP makes no such allowance.

Tertiary foreign language teachers in Korea and their employers commonly ignore the international test mastery requirements which are so important to their students, unless they are specifically employed to coach these external test skills. In contrast, the underlying teacher-purpose of this test is to harness the backwash of external test motivation by addressing it directly, giving students clear instrumental goals relative to their own strengths and weaknesses, then build on (subvert?) this foundation into genuine communicative language abilities.

Tests are normally seen as tools to be employed by teachers, administrations, institutions and employers to categorize students. Text books on testing may ask rhetorically "why do we have tests?" (Moskovsky 2005) , but the "we" implied typically excludes students. In effect tests of all kinds (not only language tests) are instruments of social control. Much of the stigma attaching to tests and examinations comes precisely from this power function. The test specification discussed in this project stands the normal power equation on its head. It aims to help students master the test culture on their own terms.

In order to measure G-TELP ability, the current test must mirror the general structure of the G-TELP specifications. However, the class test will go beyond those specifications, especially in later cycles but not so dramatically in this sample, to bring students to an implicit awareness of how G-TELP study can be extended into non-American communicative environments. The specifications of the test outlined in this program assumes a level of approximately 3 to 4 on the G-TELP scale (that is, roughly an intermediate level of competency in L2).

Problems common to large scale standardized tests

The brief discussion below forms a necessary context for the specification of the current test since the design undertaken will be influenced by following realities.

Large scale tests such as G-TELP suffer from many deficiencies, which are widely discussed in the literature (e.g. Spolsky 1995) and may even be a hot political issue in places like South Korea. Indeed, G-TELP itself, by aiming to be criterion rather than norm referenced, tries to circumvent some of the most serious limitations of norm referencing.

The speaking/listening division of G-TELP is a credible attempt to elicit measurable holistic performance, while the writing division is somewhat less convincing. Both depend to a degree upon holistic rather than item-based marking, with the accompanying penalty of lower inter-marker reliability. The process is advertised as cost-effective because local country markers are employed. That is non-native speaker markers are employed. G-TELP does not give reliability statistics for this outcome or any other on its web site. Although descriptors are given for reading levels, no tests appear to be available in this division. The grammar tests are exactly the same kind of multiple choice types and categories found in large scale norm referenced tests such as TOEFL and TOEIC.

From a language teaching/learning viewpoint, some of the most damaging limitations of current large scale standardized tests are :

a) It is in the nature of the testing process to emphasize discrete point accuracy over fluency or communicative success. This is true

even where evaluation appears to be holistic, since markers often respond implicitly to discrete point elements.

b) The large international tests take no serious account of the shifting importance of fluency Vs accuracy in the language acquisition process, or in different communicative contexts. This balance has been subject to much professional debate. The (non-dogmatic) position of this writer is that up to around intermediate level, fluency is much more important than accuracy, though the balance will vary with the activity. That is, learners need to experience communicative success and acquire traction in using the language. Native speaker speech is, after all, littered with violations of formal 'rules'. Depending upon learner goals, more attention may be paid to accuracy in the target language after intermediate level. With regard to context, there is obviously extreme variation in the need or desirability for real-life accuracy. Vagueness, imprecision or even poor language control can be beneficial (as many a politician, many a Romeo, and many a foreigner getting into scrapes knows very well).

c) The actual analyses of linguistic accuracy implicit in most language tests are naive (I believe). Taking syntax for example, the easiest errors to measure are local violations such as tense inflections or preposition choice. The most difficult errors both to analyse and to test are global violations (e.g. see Birt & Kiparsky 1972). Global violations (such as scope problems, and confusions of assignment in thematic relations) lead to communicative breakdown because the listener can't extract the message reliably. Local violations rarely lead to communicative breakdown. Some local errors persist for years in the speech of many L2 users without causing any serious handicap. Nevertheless, the large published

tests almost invariably test for local syntactic errors, which are easily amenable to multiple choice testing.

d) To the extent that speaking is tested analytically at all, testing analysis nearly always focuses on segmental rather than non-segmental phonology. For example, tests for minimal pair discrimination are fairly common. In real communicative practice, non-segmental problems are probably more likely than minor phoneme confusion to lead to communicative breakdown and/or social rejection. This is especially the case with learners from syllable timed languages trying to make the transition to a stress timed language such as English. The G-TELP does make reference to " a foreign intonation and rhythm", but gives no guidance as to what this might actually mean, or how testers might evaluate it. The widespread omissions of testers in this area may well have a lot to do with their general ignorance of the whole field (as well as, of course, the ignorance of many teachers).

Since the open focus of the current diagnostic test is on external test mastery, the early cycles of the diagnostic process (shown here) will reflect the known limitations of the external tests. Those limitations will only be made explicit to participating students at a later stage.

Test Content

The actual test content specifications of G-TELP (from the G-TELP website) are given at the end of this section. The content specifications for the diagnostic cycle undertaken here are a notional extraction from the more comprehensive Level 3 G-TELP specification.

Grammar

"An examinee at [Level 3] is required to demonstrate understanding of basic grammatical structures in two-to three-sentence contexts". This may be the least useful part of the G-TELP test, especially given the multiple choice format. However, it is also the part least likely to be misunderstood by non-native English markers and falls within student expectations.. This first test training cycle will follow G-TELP multiple choice format. Later cycles will test the forms in a more genuinely contextual way by posing communicative problems; (e.g. I am personally impressed by the way that Pimsleur audiolingual materials pose such problems at lower levels in L1 for solution in L2. This is extremely effective).

A sample will be taken from the fairly arbitrary selection of structures chosen by G-TELP for Level 3, which is :

*Verbs

- past progressive
- future possible conditional with will, would, and would have
- present perfect

*Coordinators and Subordinators

- and, but, or, nor, so
- after, because, etc.

Word forms

- adjectives: equatives, superlatives

Listening

" An examinee at [Level 3] can understand native speakers of English who are speaking at a slower than normal rate with pauses, repetitions, and paraphrasing intended for nonnative speakers". The sample chosen for modelling here will be :

".. directions for a procedure"

Procedural directions have the advantage of being easily adapted to many different levels of complexity, so that the listening task can be graduated in difficulty to discriminate amongst learner abilities.

Note that the actual test format for listening in G-TELP poses difficulties which are out of all proportion to the language being evaluated. That is, six questions are presented orally prior to a spoken passage, the passage is then given, then the six questions are asked again. Responses are in multiple choice format, but the questions are nowhere written down. The questions relate to quite specific information. This is more of a brute memory test than a listening test.

In a self-test with the Level 3 G-TELP website sample, I found it almost impossible to keep six explicit oral questions about a meeting in mind while I listened to the passage, then respond to

queries about them. Even as a native speaker I would have needed to take notes. (Long ago, as a dispatch officer announcing departures at Sydney airport, I learned that the listening public has great difficulty processing even single, simple information items). The short term language memory of an L2 speaker is generally much shorter than that of a native speaker. This differential needs to be calibrated carefully in a listening test. The G-TELP material shows little evidence that this has been done.

Speaking

The G-TELP Level 3 descriptor states that "The Speaker is generally intelligible and grammatically accurate in performing the tasks. Speech may be interrupted by occasional pauses, false starts, rephrasing, etc. Speech evidences some consistent pronunciation and stress errors and a foreign intonation and rhythm, which only occasionally interfere with communication."

Assessment criteria : "GST assesses six different areas such as grammar, vocabulary, pronunciation, content, fluency, and interaction". The actual balance amongst these criteria is not specified. One might anticipate rather different weightings of evaluation criteria to be applied by native and non-native speaking testers.

The full G-TELP offers a series of 11 speaking tasks which are said to be graduated challenges from Level 5 (elementary), to Level 1 (near native speaker proficiency). In fact, most of the challenges could be handled at several levels of proficiency with communicative success. One intractable problem for all language tests which elicit open-ended student performance is that some individuals are naturally economical in expression while some are verbose. This can

easily lead to false assessment, since the taciturn individual may follow instructions, but not reveal his full capacities to the examiner. Also, it is not clear how the G-TELP concept of "a foreign intonation and rhythm" is to be evaluated.

The sample chosen for modelling here will be :

"Narrating a story from pictures. The speaker is able to relate a simple narrative illustrated in a series of pictures. The speaker appropriately describes the setting, people, and/or objects depicted, and coherently relates the chronology of events".

This task is actually rated by G-TELP as the last in Basic performance. Though it is true that it can be attempted at a marginally basic level, clearly the response can be elaborated to any level. It is a good choice for the initial diagnostic cycle of the current project.

Reading & Vocabulary

G-TELP : "An examinee at this level can read simplified and/or simple authentic texts of the type found in textbooks written for learners of English as a second or foreign language. A person at this level is also required to demonstrate understanding of selected vocabulary in the context of the reading tasks described".

In fact, the tasks nominated by G-TELP descriptors are not those normally found in ESL textbooks, though it is true in Korea anyway, that little else is read or written in L2. Indeed, as a trainer of Korean graduate teachers of English I found that few read or wrote anything even in Korean (L1) beyond the bare minimum required by

their employment. This makes the authentic selection of textual material beyond classroom content difficult, except in ESP contexts.

The question of vocabulary recognition is also complex, and from a testing point of view, problematic. Lists for word recognition are rather pointless. Extracting meaning from context is a skill or aptitude which in any instance might or might not depend upon prior acquaintance with the word. Parry (1991) "found unexpectedly that after even a short delay the words had often been forgotten. Moreover, it was often the best guessers who were the worst forgetters. Good guessers were able to perceive the main lines of a text very quickly, fill in semantic gaps left by unknown words, and then show no vocabulary gain between pretest and posttest."

Given the preceding limitations, for this diagnostic cycle a task has been selected from the G-TELP descriptors which will hopefully strike some chord with most of the target students :

"..descriptions of landmarks and places of interest based on texts from tourist guidebooks, travel magazines, and brochures".

Writing

G-TELP : "Grammar, vocabulary, organization, logicity, and substance are assessed in the following 11 tasks in the range of three levels, Basic, Intermediate, and Advanced, at five levels".

As with speaking, the writing division of G-TELP is graded into eleven task sections, supposedly at escalating levels of sophistication :

Basic - Task 1 Sentence construction(syntax) / Task 2 Use of appropriate synonyms & pronouns / Task 3 Sentence length consistency / Task 4 A suitable example //

Intermediate - Task 5 Organization of the body / Task 6 Variety of expressions / Task 7 Validity of the subject /

Advanced - Task 8 Clarity of the conclusion / Task 9 Logical coherence / Task 10 Originality of the essay / Task 11 Development of related ideas.

These task sections are no doubt a useful checklist for a writer. An employer for certain occupations would certainly be interested in a candidate's management of the categories. Unfortunately, they have nothing to do with the mastery of L2 per se, excepting perhaps Tasks 1 and 2; (it is not clear to me that Task 3 is relevant to writing skills at all). A student who performs poorly by these criteria in L2 will almost certainly show the same weaknesses in L1. This problem is generic to all the elements of L2 testing, but becomes stark in writing skills. On the other hand, from a teaching & learning (not a testing) perspective, the 'problem' may be an opportunity. Learning to write well in L1 may be the shortest route to skilled writing in L2, if account is taken of discourse style differences.

In practice, the Writing Grade Scale used by G-TELP is very broad and very vague : Poor, Fair, Fairly Good, Good. Intermediate writing skills are said to : "Address the writing task generally, but neglects some aspects. Some grammatical errors, but the writing remains intelligible. Logic is unclear at some points. Lacks variety in vocabulary and sentence structure".

For the purpose of first cycle diagnostic testing, a writing topic that highlights the so-called intermediate skills will be chosen. However, 'basic' skills mastery will obviously form part of the assessment, while performance on the supposedly advanced skills will be de-emphasized.

Types of text

Listening

- Practical directions to follow a set of procedures in a new employment situation

Speaking

- A student conducted narrative of events in a series of picture panels

Reading

- Travel brochures prepared for English speaking tourists in a Korean travel bureau

Writing

- Write an extended e-mail to a foreign businessman who is visiting your company and South Korea for the first time. Give clear directions about when and where he will be met. Give a brief itinerary for his two day stay in Korea. Explain his accommodation. Give a little cultural background about your company and the country.

Addressees of texts

- Young tertiary educated South Korean adults in an employment context. The writing segment requires them to address another professional in a cross-cultural context.

Lengths of texts

- Texts or dialogues, up to 500 words. The grammar segment deals with multiple choice sentence fragments. The written response formats for listening and reading tasks are also multiple choice.

Topics

- Situations, events and communicative demands relevant to young tertiary educated South Korean adults in an employment context

Readability

- Intermediate, as conceived by the Level 3 scale of G-TELP

Structural Range

- The test constructor considers this to be fairly arbitrary, but for the purpose of mirroring the G-TELP test, evaluated structures will conform to G-TELP specifications for Level 3 and below (see the Content section)

Vocabulary Range

- As found in the G-TELP test series, but with a particular emphasis on Korean business employment contexts and local-to-visiting-foreigner interaction

Dialect and style

- Constrained by the heavily American orientation of the G-TELP test, but tending to relaxed formal, as found in a congenial business environment when Koreans deal with foreigners locally. American idiom preferred (for G-TELP purposes), but some allowance for other international varieties of English where meaning is not impaired.

Speed of Processing

- It is not considered that this can or should be validly assessed at intermediate level, unless delays seriously inhibit communication. Short term memory demands in the listening tasks need to be carefully calibrated.

Task Procedure, Timing and Medium

Task Procedure

Five major sections - grammar, speaking, listening, reading, writing.

Section 1 - Grammar - Complete a set of multiple choice items on selected structures

Section 2 - Listening - One text; listen to the passage on employment procedures, then respond to spoken questions via a written multiple choice format. (Although G-TELP rates procedural language at Level 2 - more advanced than Level 3 - I feel that its strong contextual frame can actually make some procedural instructions simpler than the G-TELP nominated Level 3 tasks).

Section 3 - Speaking - generate a coherent narrative of events from a picture panel story. Apply appropriate vocabulary, including linking words to give the text proper cohesion. Speak within an intonation and phonemic range that will be comprehensible to native English speakers not familiar with 'Konglish' (Korean English).

Section 4 - Reading - read extracts from an English language Korean travel brochure. Extract specific information, follow reasoning, and recognize some inferences implicit in the text.

Section 5 - Writing - an extended email with explicit information for a foreigner visiting Korea. Convey time, dates, places, and an itinerary. Give a brief description of the local situation as it affects a visiting businessman. Demonstrate a level of textual organization adequate to convey the information without ambiguity. Avoid formal errors that will seriously confuse the message. Length : adequate to do the job.

Number of Items

Section 1 - Grammar - 10 items

Section 2 - Listening - 5 items

Section 3 - Speaking - 5 to 15 minutes

Section 4 - Reading - 10 items

Section 5 - Writing - roughly 300 to 500 words

Timing

Section 1 - 20 minutes

Section 2 - 15 minutes

Section 3 - 15 minutes

Section 4 - 30 minutes

Section 5 - 30 minutes

Total - 1 hour, 40 minutes, with the speaking and listening sections done at a different time to the other sections.

Medium

Pen and supplied paper, plus a centrally controlled language laboratory cassette and cassette recorder to play the listening text and record the picture panel narrative. There will be a separate answer sheet and question booklet.

Organization

- Students undertake the test under supervision.
- The grammar, reading and writing sections will be done in a group environment.
- The speaking and listening sections will be done in a language laboratory, also in a group environment, at a different time to the grammar-reading-writing test.
- The listening test will be done before the speaking test. The listening task with spoken questions will be available on the tape consoles for ten minutes.

- The speaking test will be done after the writing test. The picture panel story will be distributed with instructions. Students will be allowed up to fifteen minutes to orally compose and record a narrative related to the picture panel story.

Techniques

Section 1 Grammar

1. Multiple choice with one correct answer and three distractors (as per G-TELP; other formats will be introduced in later test cycles).

Example:

1. The KAL flight was delayed _____ the pilot became ill in Singapore
a) but; b) so; c) because

Section 2 Listening

Listen to the job instructions. Answer each of the five spoken questions. Select the right answer for each question from three choices.

Example :

[Spoken] When are invoices collected from the warehouse?

[Written prompt] a) immediately after lunch ; b) immediately you arrive at work; c) immediately after yesterday's orders are finished

Section 3 Speaking

Make one story from the cartoon pictures you see here. Describe what you see, what happened, and why it happened. Say what happened next. You can talk for five to ten minutes.

Section 4 Reading

Read the travel brochure, then answer the questions. Select the answer which is most true from what the brochure says. Each question has three answer choices.

Example :

- a) Chungju dam is the main water supply for Seoul
- b) Chungju dam is the main dam in South Korea
- c) Chungju dam is the main reason visitors come to Chungju

Section 5 Writing

Write an email to a foreign businessman visiting your company. Tell him how he will be met at Incheon Airport. State times, dates, places, and give him an itinerary. Give a brief description of Chungju and local information about your company.

Criteria levels of performance and Marking Scales

Scoring Context

Criteria levels in this diagnostic test are calibrated approximately to the Level 3 descriptors of the G-TELP test (see the Content section). Since the real purpose of the test is to bring students to an understanding of G-TELP test criteria, and the limitations of those

criteria, actual performance will be evaluated in post-test discussion and analysis with the students. Further, the sample of diagnostic tasks is not extensive enough to derive a valid evaluation of overall student ability.

The final score of individual students is not significant in this context. The important thing is for students to arrive at their own understanding of the techniques that G-TELP uses to judge them, and to develop strategies for managing G-TELP evaluations while improving their real mastery of English; (this implies a process which goes beyond the scope of this presentation exercise).

The G-TELP test itself offers this profile of achievement for the speaking task (stated as 'mastery') :

The level of mastery

Separate level scores for each of the three levels

Percentage profiles summarizing the performances on each task

Evaluations of the performances in the five Skill Areas across all 11 tasks: Content, Grammar, Vocabulary, Pronunciation, Fluency, and Interaction

Descriptions of the Skill Area Performance at the examinee's level of mastery.

Level Mastery Score

The Mastery Score is determined by the number of skill areas in which the examinee has accumulated a Skill Area Score of 75% or more. The Mastery Score is based on the Task/Structure Scores and the Skill Area Scores.

At Levels Five, Four, Three, and Two, three Mastery scores are possible:

Mastery – the examinee achieves Skill Area Scores of 75% or greater in all three skill areas
Near Mastery – the examinee achieves Skill Area Scores of 75% or greater in two of the three skill areas
No Mastery – the examinee achieves Skill Area Scores of 75% or greater in less than two skill areas

At Level One, two Mastery Scores are possible:

Mastery – the examinee achieves Skill Area Scores of 75% or greater in the two areas
No Mastery – the examinee achieves Skill Area Scores of 75% or greater in less than two skill areas

G-TELP
GENERAL TESTS OF ENGLISH LANGUAGE PROFICIENCY

Level: **2**
Mastery

PROFILE A - Skill and Task Structure

Skill Area	Sub-Skill	Score	75%	70%	65%	60%
Listening	NAVIGATION	57%				
	PERCUSSION	83%				
	NEGOTIATION	86%				
Reading and Writing	PROCESSES	100%				
	TEXT ACCURACY	100%				
	NON-TECH ARTICLE	100%				
Speaking	ONCEAS ENTRY	100%				
	DISPENSE LETTERS	80%				
	CHANGES	100%				
Grammar	FIRST UNPAID CORO	80%				
	DISBANCTION	50%				
	NOVEM CHANGES	71%				

PROFILE B - General Information Type

Score	Correct Total	Expression of Question Type
10	20	Literal information questions ask the examinee about information which is explicitly stated in the passage.
5	11	Inferential information questions ask the examinee to deduce information which is not specifically stated, but which is implied by the passage.
0	8	Vocabulary questions ask the examinee to select synonyms for words appearing in detail questions or the correct of the vocabulary.

Total Score: 255
You have answered 85% of all the questions on this test correctly.
DATE: 190-880113-915-712

The Meaning of 'Mastery' in G-TELP

Thus, a score of 75% is taken to represent mastery at any given level. That score may be comprised of discrete point items (as in the multiple choice tests), or be based on impressionistic marking (as for speaking and writing tasks). The balance of criterial weighting is probably different for each marker in impressionistic marking, but G-TELP claims that "the performance of the examinees is evaluated by local experts in English as a foreign language, who have been trained and certified by G-TELP Services to assign objective and consistent scores to in-country nonnative speakers of English". Inter-rater reliability statistics are not given.

Comment : Nobody can assign a subjective score objectively. Even discrete point criteria are, in the end, assigned significance subjectively. The best that any language test can do is to arrive at some kind of social consensus about criteria, and strive for consistency.

Scoring for the diagnostic purposes of the current test cycle

Grammar

The grammar test checks the ability to use several grammatical constructions, extracted from the Level 3 G-TELP specification. This is clearly only a sample, not a comprehensive test of grammatical knowledge applying to any level of English mastery. There are **ten multiple choice questions**.

Each multiple choice question will be worth 1 mark, with an overall score of 7 indicating 'mastery' level. No half marks will be awarded.

Listening

The listening test involves a single task with responses to **five multiple choice questions**. This is clearly only a sample, not a comprehensive test of listening skills.

Each multiple choice question will be worth 2 marks, with an overall score of 8 indicating 'mastery' level. No single marks will be awarded.

Speaking

The speaking test requires the construction of a narrative based on a series of picture panels. This is clearly only a sample, not a comprehensive test of speaking skills.

The G-TELP mastery criteria for the picture narrative task is :

Narrating a story from pictures. The speaker is able to relate a simple narrative illustrated in a series of pictures. The speaker appropriately describes the setting, people, and/or objects depicted, and coherently relates the chronology of events.

The G-TELP assesses speaking in six categories : Grammar / Vocabulary / Pronunciation / Fluency / Content / Interaction (overall comprehensibility). Descriptors for these categories are given in Appendix 1 and the end of this section. For this diagnostic test, Grammar & Vocabulary will be combined into a single category, yielding a total of five categories.

The speaking task will be scored out of 10 marks. Two marks will be awarded for each of the 5 categories above. That is , 2 marks will be awarded in each category when the student performs according to the Level 3 Speaking criteria in G-TELP (see Appendix 2 below). An overall score of 7 will indicate Level 3 'mastery' for the purposes of the test.

Note that the scoring here is a formality. It does not imply that the equal weighting of categories necessarily applies in true communicative situations.

The student diagnostic purpose in this context is to highlight areas that require attention.

Reading

The reading test involves a single text, a travel brochure, with responses to 10 multiple choice items. This is clearly only a sample, not a comprehensive test of reading skills.

Each multiple choice question will be worth 1 mark, with an overall score of 7 indicating 'mastery' level. No half marks will be awarded.

Writing

The writing test requires the construction of a formal e-mail, incorporating some nominated information. This is clearly only a sample, not a comprehensive test of writing skills.

The G-TELP assesses writing in **four grades within five categories**. The grades are Poor / Fair / Fairly Good / Good. The categories are Grammar / Vocabulary / Organization / Logic / Substance. Descriptors for these grades and categories are given in Appendix 2.

The writing task will be awarded an overall score of 20, comprising four marks for each category. Poor = 1 mark; Fair = 2 marks; Fairly Good = 3 marks; Good = 4 marks. An overall score of 15 will indicate 'mastery' for the purposes of this diagnostic test.

Note that the scoring here is a formality. It does not imply that the equal weighting of categories necessarily applies in true communicative situations.

The student diagnostic purpose in this context is to highlight areas that require attention.

Test Construction Procedures

Sampling

Texts, structures and pictures for the diagnostic test will be based on the G-TELP framework which this test is seeking to clarify for students. However, this content will be given a local Korean

reference where possible. The underlying message of content choice will be that English is an international language first, and a national language only incidentally.

Item writing and Moderation

Items will be written according to modified G-TELP specifications. A key will be provided for multiple choice answers, and a list of marking criteria for holistically marked elements. Using a checklist, items will be moderated by a fellow native speaker teacher of English.

Informal Trialling

The test will be trialled by native speaking American co-workers of the tester (if their cooperation can be secured).

Trialling and Analysis

Trialling of two versions of texts and items will be carried out with a colleague's English major students at Chungju National University (if their cooperation can be secured). Half the class will complete a different version from the other half. Any administration and scoring problems will be noted. If possible, the tested students will be involved in a post-test analysis of the test content and design, as well as their own performance. This procedure is central to the notional purpose of the test.

Basic statistical analysis will be used to select the 'best' options for inclusion future versions of the test. A qualitative analysis will also

be done, drawing extensively upon a post-test analysis with test-takers.

Validation

Sections : Grammar, Speaking, Listening, Writing, Reading

Items constructed and selected for the final version of the test will be checked against the modified G-TELP specifications, as well as comparable items from the actual G-TELP test.

Students' scores will be checked against their previous performances in other tests and courses that involved an evaluation of grammar, listening, speaking, reading or writing in English.

Speaking - concurrent validity

To check concurrent criterion-related validity a random sample of students will complete a 10 minute interview and discussion, with two native-speaking teachers scoring. The scores achieved in this performance will be compared with those from the full diagnostic trial.

The attainment of a particular formal validity coefficient between informal trialling and full diagnostic application of the test is not important at this stage for meeting the central purpose of the test as a focus for student discussion and self-improvement.

Candidate Instructions

Unless understanding testing instructions is formally assessed, it seems difficult to justify giving instructions in L2, especially at lower

levels. To do otherwise is simply to insert another unmeasurable variable into the assessment process. That is, the validity of the test may be compromised. Nevertheless, it is common language testing practice to give instructions in L2. The G-TELP sample test site gives instructions partly in English and partly in Korean.

The sample instructions in this diagnostic presentation are in English, but will be translated where the actual application of the test is required.

Teacher's Handbook

One handbook will be prepared as a summary for the current tester, and for anyone who seeks to develop the material for further application. A separate handbook will be prepared for use by students to help in preparation for the diagnostic test.

The test is a diagnostic and analytic tool for students who want to to enhance their ability to pass international proficiency tests such as G-TELP. It is intended for post-test discussion by tested-students. Later application of the material will show how it can form a foundation for true communicative language learning.

Rationale:

Internationally standardized tests such as G-TELP have a huge backwash effect on the learning behaviour of language students in South Korea, even quite young students. This backwash is often not beneficial, and the formal syllabi of schools and universities are often not coordinated in any productive way with this underlying engine of testing motivation. The aim of the current diagnostic test

is to make these relationships explicit, particularly to the students concerned, and involve them in a process which can harness the testing content of international standardized tests as an instrument for genuine language learning.

Development and Validation:

The test was developed by carefully abstracting selected elements of the international G-TELP test. That is, it is an informal subset of the G-TELP test, and written according to slightly modified G-TELP specifications. The general content and design of the G-TELP and comparable tests is familiar to most South Korean tertiary students who study English.

Description of the test:

The test comes with five divisions : grammar, speaking, listening, reading and writing.

The grammar, reading and writing tests will be done together and take 1 hour, 20 minutes. The speaking and listening tests will be done together at a different time in a language laboratory, and take 30 minutes.

The grammar test contains ten multiple choice questions. Students are allowed 20 minutes to complete this.

The speaking test contains one task, constructing a narrative from pictures. The student narrative is recorded on tape for later marking and review. Students are allowed 15 minutes to complete this.

The listening test involves listening to a short passage of instructions and answering five multiple choice questions. Students are allowed 15 minutes to complete this.

The reading test involves reading a text and answering and answering ten multiple choice questions. Students are allowed 30 minutes to complete this.

The writing test involves writing a formal e-mail containing certain given information. Students are allowed 30 minutes to complete this.

Sample items:

Section 1 Grammar

Example:

1. The KAL flight was delayed _____ the pilot became ill in Singapore
a) but; b) so; c) because

Section 2 Listening

Listen to the job instructions. Answer each of the five spoken questions. Select the right answer for each question from three choices.

Example :

[Spoken] When are invoices collected from the warehouse?

[Written prompt] a) immediately after lunch ; b) immediately you arrive at work; c) immediately after yesterday's orders are finished

Section 3 Speaking

Make one story from the cartoon pictures you see here. Describe what you see, what happened, and why it happened. Say what happened next. You can talk for five to ten minutes.

Section 4 Reading

Read the travel brochure, then answer the questions. Select the answer which is most true from what the brochure says. Each question has three answer choices.

Example :

- a) Chungju dam is the main water supply for Seoul
- b) Chungju dam is the main dam in South Korea
- c) Chungju dam is the main reason visitors come to Chungju

Section 5 Writing

Write an email to a foreign businessman visiting your company. Tell him how he will be met at Incheon Airport. State times, dates, places, and give him an itinerary. Give a brief description of Chungju and local information about your company.

Advice on preparing for the test

Students should make themselves familiar with the online version of the G-TELP test at http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp04.asp .

They are likely to already be familiar with the testing pattern of multiple choice grammar items.

- They should review the grammar items for G-TELP level 3.
- They can practice constructing a narrative using picture cues, and listening for specific information from a recorded text.
- The class can study informational brochures from businesses and government departments, as well as office procedure manuals.
- They can practice writing formal e-mails, particularly to foreign businessmen who may not be familiar with conditions in South Korea.

Interpretation of test scores

This is a diagnostic test, intended for review and discussion by the tested students themselves. The actual scores are not very significant, except as a guide to areas needing attention. The G-TELP test takes a score of 75% in any area to indicate 'mastery' at a particular level from 1 to 5. Level 5 is near native speaker standard. The diagnostic test is calibrated approximately for students at about G-TELP Level 3.

Training Materials:

This diagnostic test does not require a native speaker, or even a trained teacher for administration. However, the follow-up marking, analysis and guided student discussion of the content and outcome does require trained teachers who are thoroughly familiar with standardized tests in South Korea, with Korean educational norms, and with Korean cultural sensitivities.

Test administration:

The question booklet and answer sheets are separate. Students are not to write in the question booklet. Question booklets should be collected from all students at the end of the test. However, since the question booklets themselves will be analysed by the students in a post-test analysis, they are not externally confidential in the same way as normal standardized test documents.

The test should be administered within the specified time limits and under standard test conditions. Answers should be written in blue or black ink for legibility, not pencil. The test will be done in normal class time, either in a classroom with separated desks, or in the case of the speaking and listening tests, a language laboratory. The testing atmosphere should be cooperative, but disciplined. Students should have a prior understanding of the purpose of the test, and be aware of their later role in analysing the results.

Scoring of the tests should be done with reference to the key and rating scales provided. Where possible, writing and recorded speaking tests should be cross-checked by another teacher. All results will ultimately be discussed and peer moderated by the students themselves.

Handbook for Candidates

1. This is a test to help you with external proficiency examinations like G-TELP. You will learn from this test what your strengths and weaknesses are in English. The teacher can then design a program to help you.

2. After the teacher has marked the test, he will discuss it with you step by step. In this way you will not only learn to do the G-TELP test successfully, but also how to use test study to really become a better English speaker, reader and writer.

How to prepare for the test

Check the online version of the G-TELP test at http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp04.asp .

- Especially review the grammar items for G-TELP Level 3.
- Practice speaking by making up a story. Use pictures or cartoons for ideas.
- Practice listening to any recorded speaking that gives information like instructions or announcements. Can you remember what the speaker said?
- Study informational brochures from businesses and government departments in English. Look at some instruction manuals. Can you understand this information?
- Practice writing formal e-mails like you might have to working in a company. Write e-mails to foreign businessmen who may not know about conditions in South Korea.

What do your test scores mean?

Don't worry if you find some parts of the test a bit hard. The marks don't matter too much. The whole idea is to help you do better in a real test, and to become a better English speaker.

The G-TELP test itself takes a score of 75% to mean 'mastery' at some level from 1 to 5. Level 5 is near native speaker standard.

This test you are doing now is set for students at about G-TELP Level 3.

What is in the test ?

The test comes with five divisions : grammar, speaking, listening, reading and writing.

The grammar, reading and writing tests will be done together and take 1 hour, 20 minutes. The speaking and listening tests will be done together at a different time in a language laboratory, and take 30 minutes.

The grammar test contains ten multiple choice questions. Students are allowed 20 minutes to complete this.

The speaking test contains one task, constructing a narrative from pictures. The student narrative is recorded on tape for later marking and review. Students are allowed 15 minutes to complete this.

The listening test involves listening to a short passage of instructions and answering five multiple choice questions. Students are allowed 15 minutes to complete this.

The reading test involves reading a text and answering and answering ten multiple choice questions. Students are allowed 30 minutes to complete this.

The writing test involves writing a formal e-mail containing certain given information. Students are allowed 30 minutes to complete this.

Sample test items:

Section 1 Grammar

Example:

1. The KAL flight was delayed _____ the pilot became ill in Singapore

a) but; b) so; c) because

Section 2 Listening

Listen to the job instructions. Answer each of the five spoken questions. Select the right answer for each question from three choices.

Example :

[Spoken] When are invoices collected from the warehouse?

[Written prompt] a) immediately after lunch ; b) immediately you arrive at work; c) immediately after yesterday's orders are finished

Section 3 Speaking

Make one story from the cartoon pictures you see here. Describe what you see, what happened, and why it happened. Say what happened next. You can talk for five to ten minutes.

Section 4 Reading

Read the travel brochure, then answer the questions. Select the answer which is most true from what the brochure says. Each question has three answer choices.

Example :

- a) Chungju dam is the main water supply for Seoul
- b) Chungju dam is the main dam in South Korea
- c) Chungju dam is the main reason visitors come to Chungju

Section 5 Writing

Write an email to a foreign businessman visiting your company. Tell him how he will be met at Incheon Airport. State times, dates, places, and give him an itinerary. Give a brief description of Chungju and local information about your company.

APPENDIX 1 G-TELP Speaking Score Criteria

[Material in these appendices is taken from the G-TELP website

http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp04.asp]

Grammar

The speaker's control of basic sentence structure

Vocabulary

The speaker's available vocabulary in the performance of certain required speaking tasks

Pronunciation

The speaker's control over the sound system

Fluency

The speaker's flow of speech and degree of spontaneity as well as facility in the performance of the required tasks

Content

The topics/contexts about which the speaker can speak

Interaction

The Overall Evaluation, which confirms the examinee's Mastery Level, is based on comprehensibility, which is a more general assessment of the speaker's communicative ability, influenced to some extent by the other, separately evaluated, skills.

G-TELP Proficiency Levels for Speaking

No Mastery

The speaker cannot perform or understand the tasks.

Level

5 Beginning

Speaking Skills

The speaker is not comprehensible even in simple speech due to frequent pauses, rephrasing, pronunciation errors, limited vocabulary, and lack of grammatical control.

Level 4 Basic

Speaking Skills

The speaker can successfully complete, and be understood overall, in performing the tasks. Speech is slow and labored, and is characterized by a strong accent and frequent grammatical errors. Transmission of the message may be impeded frequently by phonemic, stress, and intonation problems.

Level

3 Intermediate

Speaking Skills

The Speaker is generally intelligible and grammatically accurate in performing the tasks. Speech may be interrupted by occasional pauses, false starts, rephrasing, etc. Speech evidences some consistent pronunciation and stress errors and a foreign intonation and rhythm, which only occasionally interfere with communication.

Level 2 Advanced

Speaking Skills

The speaker is almost always intelligible in performing the tasks. Speech is fairly smooth and spontaneous although having some mispronunciation of individual sounds. The speaker's vocabulary is extensive and varied, exhibiting good control of the structure of the language.

Level1 Authentic

Speaking Skills

Level One is not evaluated by the Speaking Test, although mastery of Level Two indicates that the speaker is approaching the acquisition of oral proficiency skills at the native speaker level.

APPENDIX 2 G-TELP Writing Score Criteria

Grammar

1 - Poor: The composition shows very little grammatical or structural control. It has errors that interfere with meaning.

2 - Fair: The composition has some grammatical errors, but these do not interfere with meaning.

3 - Fairly good: The composition shows control over basic structures, with attempts to use some complex structures. There are some grammatical errors, but these do not interfere with meaning.

4 - Good: The composition has control over structures. Overall, the writing is grammatically correct. The few errors rarely interfere th meaning.

Vocabulary

1 - Poor: The composition shows limited vocabulary. Many words are used incorrectly.

2 - Fair: The composition shows limited vocabulary. Word choice is faulty at times. However, overall meaning is still conveyed.

3 - Fairly good: The composition uses vocabulary appropriately and accurately, with occasional misuse. It shows the writer's ability to paraphrase.

4 - Good: The vocabulary is varied and extensive. The writer rarely uses words incorrectly, but when he/she does, the intended message is communicated.

Organization

1 - Poor: The composition has no outward forms, introduction, main discourse (body), or conclusion. In content, the composition falls short in each process.

2 - Fair: The composition has outward forms but, in content, it does not have a definite form in each process.

3 - Fairly good: The composition has outward forms, introduction, main discourse (body), and conclusion, and it has a generally realized content.

4 - Good: The composition has a well-developed introduction, main discourse (body), and conclusion. It is written in the complete essay form.

Logic

1 - Poor: The composition has no generally developed logic. Only pieces of ideas have been presented.

2 - Fair: The composition has logic, but ideas often go astray and miss the logic.

3 - Fairly good: The composition has a consistent flow of logic in general but sometimes goes astray and misses the logic.

4 - Good: The composition is consistently logical and presents suitable logical examples.

Substance

1 - Poor: The content is not appropriate for the task. The composition has many irrelevant information or incomprehensible expressions.

2 - Fair: The composition addresses some aspects of the writing topic. It has some irrelevant information.

3 - Fairly good: The composition adequately addresses the writing task. It provides relevant information, but lacks focus.

4 - Good: It effectively addresses the writing task. The composition clearly and persuasively communicates to the reader.

APPENDIX 3 G-TELP CONTENT SPECIFICATIONS FOR LEVEL 3

(http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp05.asp)

- Grammar

An examinee at this level [Level 3] is required to demonstrate understanding of basic grammatical structures in two-to three-sentence contexts. Structures tested at this level include:

*Verbs

- past progressive
- future possible conditional with will, would, and would have
- present perfect

*Coordinators and

Subordinators

- and, but, or, nor, so
- after, because, etc.

Word forms

- adjectives: equatives,
superlatives, comparatives

- Listening

An examinee at this level can understand native speakers of English who are speaking at a slower than normal rate with pauses, repetitions, and paraphrasing intended for nonnative speakers.

Tasks at this level include:

- *public announcements which give simple information
- *explanations of the physical functions or attributes of familiar objects
- *personal accounts in which a personal event and its chronology are described
- *directions for a procedure

- Reading and Vocabulary

An examinee at this level can read simplified and/or simple authentic texts of the type found in textbooks written for learners of English as a second or foreign language. Tasks at this level include:

- *descriptions of landmarks and places of interest based on texts from tourist guidebooks, travel magazines, and brochures.
- *historical biographies based on texts from travel literature and history texts

*informal narratives in personal correspondence between friends
*announcements concerning the activities, meetings, and accomplishments of organizations such as clubs and professional groups
A person at this level is also required to demonstrate understanding of selected vocabulary in the context of the reading tasks described above.

- Writing

Grammar, vocabulary, organization, logicity, and substance are assessed in the following 11 tasks in the range of three levels, Basic, Intermediate, and Advanced, at five levels.

=> Task Contents

Basic

- Task 1 Sentence construction(syntax)
- Task 2 Use of appropriate synonyms & pronouns
- Task 3 Sentence length consistency
- Task 4 A suitable example

Intermediate

- Task 5 Organization of the body
- Task 6 Variety of expressions
- Task 7 Validity of the subject

Advanced

- Task 8 Clarity of the conclusion
- Task 9 Logical coherence
- Task 10 Originality of the essay
- Task 11 Development of related ideas

- Speaking

The content of the G-TELP Speaking Tests reflects expectations in specific performance tasks in which the examinee is expected to demonstrate comprehensible oral proficiency while progressing through the three levels. The task content consists of the following everyday situations:

Level 4 - Basic

Task 1

Giving personal information. The speaker is able to answer simple questions about self and family, responding in words, phrases, or simple sentences.

Task 2

Describing a familiar setting/objects. The speaker is able to name and describe a familiar setting and the objects within that setting. The speaker is able to locate and describe the spatial relationships among the objects in the setting. He/She may go beyond accurate naming of what is illustrated to describing some physical characteristics such as size, shape, etc.

Task 3

Describing habitual activities. The speaker is able to name and describe personal activities so that the description is understood as a coherent sequence of events. The speaker may go beyond a mere listing of events or activities to include some elaboration of them.

Task 4

Narrating a story from pictures. The speaker is able to relate a simple narrative illustrated in a series of pictures. The speaker appropriately describes the setting, people, and/or objects depicted, and coherently relates the chronology of events.

Level 3 - Intermediate

Task 5

Expressing and supporting an opinion.

The speaker is able to present a personal opinion and express personal preferences, and to substantiate these opinions or preferences with at least two simple supporting details.

Task 6

Giving autobiographical detail about a place or event.

The speaker is able to describe a place or event in adequate detail.

Task 7

Responding to requests for information about places of interest.

The speaker is able to supply appropriate and well-formed questions and responses in a dialogue centering on requests for information.

Task 8

Discussing advantages/disadvantages of two related objects.

The speaker is able to compare and contrast the objects by giving at least two advantages and/or disadvantages of each.

Level 2 - Advanced

Task 9

Giving directions from a map.

The speaker is able to give directions for reaching a location with clarity and thoroughness so that the route for reaching the destination can be understood.

Task 10

Presenting a solution to a specific problem.

The speaker is able to clearly state and express a position on a topic and present convincing arguments supporting such position. The speaker is able to relate an abstract issue to the specifics of the situation.

Task 11

Presenting a solution to a complex hypothetical problem.

The speaker is able to explain, with relevant and convincing detail, how he/she would solve the problem(s) associated with a hypothetical situation.

Assessed Areas

GST assesses six different areas such as grammar, vocabulary, pronunciation, content, fluency, and interaction.

Part II Writing Test Items for listening, reading, writing, speaking, grammar / vocabulary

University of Newcastle, NSW

Assignment 2 : Language Testing And Evaluation

Topic : *Write a set of between 20 and 30 test items for each of the following language skills :*

- *listening, reading, writing, speaking, grammar / vocabulary . For each sample test prepare a list of 'DOs' and 'DON'Ts'*

Notes on Assignment Content :

1. I have followed the intention stated in the first assignment of roughly modeling test content and structure on the G-TELP test, as it is used in South Korea. However, in several sub-sections I have used a more innovative testing format.
2. Genuine test material of the kind dealt with here would be preceded by an analysis of sample language produced by target candidates for the test. This has not been done in the present case since the target candidates are more or less hypothetical. However, the general level is pitched around an intermediate level of English ability as conceived by the G-TELP test (Level 3 G-TELP).
3. I have preceded each set of test questions by a short commentary on the limitations of formal testing in the particular skill area. I feel this is necessary because although test constructors and writers such as Arthur Hughes (course text) do genuflect towards the limitations of testing, they generally

accept that the limitations they mention can be lived with, and their analysis overlooks a great deal. I am rather more skeptical of the whole process.

4. The purpose of a request for a list of DO's and DON'Ts attaching to each skill item is not quite clear to me, (unless it is to check that I have read the text!). I do understand that there are design problems associated with test questions. The Hughes text is largely about spelling some of these out, and at various points lists of advice are given on choosing and constructing questions. Thus page 83 on gives advice on writing test construction (some of it poorly thought through in my view); page 124 gives advice on speaking test construction, page 154 for reading test construction, page 164 for listening tests, page 174 for grammar, and page 180 for vocabulary.

Since I have chosen to model the assignment tests on the G-TELP test as used in South Korea (see Assignment 1) , the general question formats have been largely selected for me. The assignment request for 30 questions on writing and speaking does not really fit the pattern employed by G-TELP. For those skills I have therefore constructed a set of cue questions to supplement the less explicit stimuli of a writing/ speaking topic and picture panels.

5. A recording for the testing content of the listening segment can be provided on request. A transcript of the recording is included in Appendix 3.

Listening Test

Tester's note : In real life contexts, listening is usually supported by the discourse context and well as the physical situation. The listener can immediately clarify confusions, repeat elements to assist in recall, use body language and engage a variety of other stratagems and resources. the measure of listening success is in the listener's responses (immediate or later behaviour). Information which is extensive, precise and complex is rarely delivered in a single monologue. In such contexts, speech is often used to clarify and support written material. This would almost certainly be the case with the "Telephone Duties for a Secretary" used in the recording transcribed in Appendix . For these reasons amongst others, genuine listening ability is difficult to measure reliably within the parameters of formal controlled test. It is envisaged that the testees in the context below would have access to the recording for replay for a fixed amount of time (say, 15 minutes).

Extracting the information required for correct answers below would be difficult for some native speakers of English, depending upon their intelligence and education. Such variables will contaminate most second language tests also. I feel that the demands upon memory and focus required by the multiple choice style of the G-TELP test may be excessive in a listening test. However, since the test for this assignment is modeled on G-TELP I have mostly followed their format. Part B is my own innovation, but again may not be a very easy option for many testees. An easier testing option for lower level L2 testees would be to give them a transcript with missing elements for cloze completion.

Marker's note : The correct answer has been italicized in each question for the purposes of this assignment presentation.

Recording : Telephone Duties for a Secretary [Appendix 1]

Part A : Multiple Choice Questions

Instructions : Choose the best answer for each question. *The best answer is what you hear on the recording. The best answer is NOT what you guess from experience. Circle a, b or c in each of the following questions.*

1. When you answer the telephone, you should

- a. speak kindly
- b. speak quietly
- c. *speak clearly*

2. The rules apply to

- a. all employees
- b. *all secretaries*
- c. all office staff

3. You give the caller an option

- a. *of waiting or being called back*
- b. of waiting or calling back
- c. of waiting or calling another number

4. Give progress reports

- a. every 30 minutes
- b. every 3 minutes
- c. *every 30 seconds*

5. When you return to the phone

- a. tell the caller to ring later
- b. *thank the caller for waiting*
- c. thank the caller for their call

6. You should take accurate messages
 - a. *to avoid confusion later*
 - b. to avoid criticism later
 - c. to avoid calls later

7. Your record of a telephone message should include
 - a. the time to call back
 - b. *the time of the call*
 - c. the time the call lasted

8. Tell the caller
 - a. when the person they want is at work
 - b. when the person they want is leaving
 - c. *when the person they want is expected*

9. Offer the caller information
 - a. if it is inappropriate
 - b. *if it is appropriate*
 - c. if it is accurate

10. Do not give the caller promises about
 - a. *when the call will be returned*
 - b. when the call will be reviewed
 - c. when the call will be rerouted

Part B - Dialogue Choice

Instructions :

- a) *Imagine that you are a manager's secretary. You have to follow the procedures on the recording.*
- b) *From the list below, choose the ten best statements you can make.*

1. "This is Jane Smith. How can I help you?"

2. *"Cooperative Extension, Cook County Office, Jane speaking."*
3. "Hello. Cooperative Extension Office, extension 594."
4. *"Will you hold on, or will I call you back?"*
5. "Hold the line please."
6. "The manager is busy now. Please call back later."
7. "You can give details when you talk to the manager."
8. *"May I check that again? You want me to tell the manager you got a post card from someone called Sadie?"*
9. "I don't know anything about Sadie, sorry. However I will tell the manager you called."
10. "This morning the manager has been called out on an emergency. You can reach him at 137 5557 8352."
11. "This morning the manager is visiting his mother in hospital. You can reach him at 137 5528 1935."
12. *"This morning you can reach the manager at our Orange County branch. The number is 137 5521 1689."*
13. *"Is there anyone else here who can help you?"*
14. "I'm sorry. I can't transfer you. You will have to ring the sales supervisor directly."
15. "That other person you mentioned works in the warehouse. I don't know his number. Sorry."
16. "I will tell the manager to call you at 11 a.m. I'm sure he can help you."
17. *"The manager should be here about 11 a.m. I will tell him about your call."*
18. "The manager is sure to be in by 11 a.m. He will call you then."
19. *"The person you need is Mr Jones in Sales. Would you like me to transfer you?"*
20. "The person you need is in Sales. I will transfer you now."
21. "The person you need is in another department. Hang on a moment and I

will transfer you."

22. "You have the wrong extension sir. I'll put you back to the switchboard."

23. " I don't know why the call center directed you here sir. We can't help you."

24. " *I'm afraid Mr White can't help you sir. If you can wait a few moments, I will find you the right person to speak to.*"

25. "If you like, I will transfer you to Mr Walters, who knows about this problem. His extension is 480."

26. "Just hang on while I put you through to Mr Walters."

27. " Call the main number again sir, and press 7 on the second beep."

28. "Hello; I want to make a booking for my boss."

29. "Hi, I'm Sarah. My boss needs an air ticket."

30. "Hello, my name is Sarah Winter from Agricultural Services. Can I make a booking please?"

Reading Test

[see Appendix B for the text of the passage to be read]

Tester's note on reading tests in a second language :

Most reading tests, at bottom, are comprehension tests. However "comprehension" is only partially a learned skill, and only partially related to functional language knowledge. It is intimately related to the underlying aptitudes of the readers, to their learned or native capacity for inference, to their cultural presuppositions and the patterns of thinking supported by their foundation culture. In other words, any reading test is only incidentally a test of those language skills which are a necessary but not sufficient component for success. Of course, such a diffuse instrument may be suitable for the purposes of various end users. For example, both universities and employers require a certain amalgam of abilities. Nevertheless, the outcomes of such

tests may give only ambiguous information to both language learners and language teachers.

Traditionally reading, or comprehension tests have been layered in sophistication by the use of questions which are purely factual, which require simple inference, or which require deduction from fairly complex arguments. A test which is trying to evaluate purely second language knowledge faces extreme difficulties here. Both habits and patterns of inference are greatly influenced by cultural background. Thus the safest and easiest questions are those which are fairly literally tied to the print on the page (although ability with these will not necessarily predict success in university courses etc. requiring critical evaluation).

Texts can have other purely linguistic uses in language classes. For example students can be asked to identify certain word classes, or certain kinds of syntactic constructions, or certain kinds of language functions. Although this kind of exercise can be structured as a test, it is not obvious that the results of any test along these lines would convey useful information about the subject's ability to decode textual information in a practical manner.

The questions used in this assignment are mostly comprehension questions of a fairly literal type. Some vocabulary query items are added at the end. For the purposes of the assignment, the correct answers are italicized.

Instructions

Read the material from a KNTO tourist brochure.

Choose the **best answer** in each multiple choice question. The best answer here is what the article says or implies, not what you think personally. In each question, circle a, b or c.

1. The quality tag on clothing tells you

- a) that the item was made in South Korea
- b) *that the item has been checked for faults*
- c) that the item is a good buy

2. The tourist brochure advises you to check clothing for

- a) *careful stitching*
- b) whether it fits you
- c) whether it is fashionable

3. A good place to buy clothing is

- a) In front of Korea University
- b) *In Itaewon shopping district*
- c) Around Chongno-sam Ga

4. You should keep your receipt

- a) to prove that you bought the item if challenged
- b) to claim a tax refund
- c) *to make it easy to return or exchange goods*

5. Before you buy leather, you should

- a) consider if you really need it
- b) *consider the price*
- c) consider how long it will last

6. A good place to find leather goods is

- a) hotel shops
- b) duty free shops
- c) *department stores*

7. Tourists prefer Itaewon because

- a) *there are high quality goods at reasonable prices*
- b) the shopkeepers speak English
- c) it is recommended by the tourist office

8. Custom goods can be ordered at

- a) leather factories
- b) some high class shops
- c) *some shops owned by leather factories*

9. When you are ordering custom items

- a) you are shown many designs to choose from
- b) *it is best to take pictures of a design you like*
- c) it is best to take professional advice

10. You should research the model number of electronic items

- a) *before you buy anything*
- b) after you see something you like
- c) before you come to Korea

11. With notebook computers

- a) the good models will suit everyone's needs
- b) *different computers are best for different kinds of people*
- c) price is a good guide to quality

12. A five megapixel digital camera

- a) is necessary for professional standard photography
- b) is probably best for professional standard photography
- c) *is recommended by the Tourist Office for professional standard photography*

13. For digital still photography

- a) *some camcorders can be used*
- b) most camcorders can be used
- c) all camcorders can be used

14. Technomart sells

- a) only digital cameras
- b) goods of all kinds

c) *all kinds of electronic goods*

15. Youngsan Electronics Mart

- a) mostly sells used goods
- b) is famous for used goods
- c) *contains both new and used goods*

16. Shops show that they prefer cash sales

- a) by refusing most credit cards
- b) *by offering discounts on cash purchases*
- c) by offering gift vouchers with cash purchases

17. Cell phone chains are popular gifts because

- a) they are beautifully designed
- b) they are traditionally given
- c) *they have a practical use*

18. Myeongdong is less popular than Insadong to buy

- a) *traditional cell phone chains*
- b) trendy cell phone chains
- c) expensive cell phone chains

19. Insadong has a lot of

- a) art galleries
- b) *tourists*
- c) museums

20. Postcards are suitable souvenirs because

- a) they are easy to buy
- b) *they are easy to carry*
- c) they show many Korean themes

21. In this reading passage [paragraph 1], "on the fabric" means

- a) printed on the garment you are buying
- b) *about the material in the garment you are buying*
- c) on the information tag about the garment you are buying.

22. In this reading passage [paragraph 1], "outlets" means

- a) places where water or other liquids come out of a pipe.
- b) ways people use to express their talents, or relieve tension
- c) *places where things are sold*

23. In this reading passage [paragraph 1], "case" means

- a) a kind of solid bag for carrying things in
- b) a job involving one patient or client, worked on by a professional person
- c) *a circumstance, a situation*

24. In this reading passage [paragraph 2], "custom" means

- a) the way people usually do things
- b) a kind of tax
- c) *goods made for the special individual needs of a buyer*

25. In this reading passage [paragraph 2], "sample" means

- a. *an example of a particular product or design*
- b. to try something out
- c. a small container of a patient's urine taken by a doctor

Writing Test

[see Appendix 3 for the picture story panel used to cue testees in the second testing option for this component]

Tester's Note : The elements which contribute to successful writing in any language may be partially catalogued, but their individual importance within certain tolerances is not constant. The total process includes a writer, a text, and readers. Each act of reading activates this tripolar relationship in a unique

way. Thus, one writer may write with apparent technical mastery but fail to engage readers, while another may seem clumsy but superbly effective. With the preceding in mind, it is not surprising that markers of academic writing show huge inconsistencies both in their own marking and compared to colleagues. Equally, 'testers' of writing show very low inter-tester reliability unless the agreed criteria are so narrow and mechanical that the overall product of writing is not really evaluated at all.

There are genres of writing where certain criteria must be met to qualify a piece as a candidate for acceptance. The poet writing sonnets must conform to the accepted pattern of sonnets. In a more contemporary context, industrial societies have given rise to a profession of 'technical writers'. Technical writers work within very tight specifications of style and content. They will be assessed accordingly. It may be no accident that such technical literature is widely considered to be almost unreadable, and only consulted under duress; (most people with industrial experience are aware of this, including technical writers themselves). The peer assessment of academic literature also leads to a conformity of style with covert criteria which are less concerned with genuine communication than prestige; (the public pretense is objectivity). The result is an infestation of literature which is scarcely read even by specialists in particular fields. Wherever material is subject to editing, whether by school teachers, news editors, fashion magazines or romance novel publishers, then the pattern of acceptable language is delimited.

The restrictions on acceptable writing just discussed are not irrelevant to the 'testing' of writing of second language learners. In addition to all the usual elements of style and content, the L2 writer is open to judgement on his or her technical ability to employ appropriate syntax and vocabulary, as well as manage register and the expected discourse organization. These may be assessed, as they are in speaking, listening and reading. Speaking and writing are so-called active skills in that memory is less prompted by an external text or discourse. On the other hand, the speaker or writer is free to employ vocabulary and structure suitable to their L2 mastery, while they can be overwhelmed by complexity in listening and reading. Unfortunately such

*speaker/writer choice of complexity is not a reliable guide to mastery.
Extremely able people may choose great economy of expression.*

There are two alternative sets of test content outlined below.

*The **first alternative** requires candidates to construct a business e-mail containing considerable detail. This is a practical task likely to be understood by most tertiary South Korean students. An extended set of cue questions is provided below. In practice most business e-mails are likely to be much shorter than the cue questions suggest. Most South Koreans do not do a lot of reading or writing, even in Korean, beyond the briefest employment or academic requirements. They prefer the telephone. E-mails I have seen (including some from employers) which should have detailed complex information, tended to be cryptic to the point of enigma, and often lacked even a signature.*

*The **second alternative** writing test is lifted directly from the speaking test. It heavily cues the production of particular elements in the writing of candidates. The cues take the form of a series of picture panels, as well as a long list of elements which need to be communicated. Although such a process does indeed test whether the student has sufficient mastery to satisfy the cues (hence making marking easy), it adds up to a poor test of overall writing ability. This is because in a genuine writing situation the writer must identify and marshal such cues, or key points, himself.*

The picture panels used for cueing are taken from an excellent ESL book, "Storylines ... picture sequences for language practice" by Mark Fletcher and David Birt, page 36; published by Longman 1983.

1. First Writing Test

Instructions

Write an email to a foreign businessman who will visit your company. Tell him how he will be met at Incheon Airport. State times, dates, places, and give him an itinerary. This is his first visit to South Korea. Give a brief description of Chungju and local information about your company.

Here is a list of questions to suggest some information which your e-mail could contain. You will need to organize your information. The questions below are not in any special order.

1. What is the proper greeting format to a stranger in a business e-mail?
2. Who are you and what is your role in the company?
3. What does the foreign businessman already know about his itinerary?
4. Are you going to TELL him his itinerary, or make suggestions?
5. Who will meet him at Incheon Airport?
6. What time will his flight be arriving?
7. What is the weather like in South Korea at this time of year, and what clothing should he bring?
8. How long should he expect passport and customs formalities to take?
9. Is it alright for him to bring product samples in his hand luggage into South Korea, or should they be packed separately?
10. Where will he be met in Incheon Airport?
11. Is there anything he needs to know about the layout and services of the airport if he is not met for some reason?
12. What is the cell phone number of the person who will meet him?
13. What does the person who will meet him look like?

14. How long will it take him to get from Incheon to Chungju?
15. What transport will be used for that trip?
16. Where will he be staying in Chungju?
17. What are company and accommodation contact numbers in Chungju?
18. How many days will his itinerary last?
19. What will he be doing on each day of his itinerary?
20. Will he be provided with a car, or a driver?
21. What seminars or other presentations (if any) will he be asked to give in South Korea?
22. Who will he be meeting during his stay?
23. Who speaks English in your company?
24. What are some useful facts he should know about your company?
25. What are some useful facts he should know about Chungju?
26. How far is Chungju from Seoul and what are the transport connections?
27. What are some special attractions of the Chungju area?
28. What educational and research facilities, and what libraries are available near Chungju?
29. What internet, telephone or other electronic services are available to visitors?
30. What banking services are available to visitors?

2. Alternative Writing Test [based on material used in the Speaking Test]

Part A Instructions

Study the panel of eight pictures and diagrams. Use the panel to help you write a short history of smallpox vaccine. The marker must also find the information below in your writing. However, just answering the questions is not enough to make your essay. The order of the questions has been mixed up.

1. Where are reserve vaccination supplies of smallpox vaccine kept?
2. Who was the discoverer of smallpox vaccine?
3. Why was 1965 an important date in the history of smallpox eradication?
4. Where was smallpox most common before vaccinations began?
5. When was smallpox finally eradicated?
6. What are possible reasons that some countries were more effective than others in the early control of smallpox?
7. What actually is vaccination?
8. Which countries continued to be seriously affected by smallpox in 1900?
9. Where do we find early records of smallpox?
10. What kind of instruments are used for vaccination?
11. What do you think the WHO is?
12. What would make a vaccination program difficult to carry out in some countries?
13. When did the discoverer of smallpox vaccine live?
14. How much do you think villagers in remote African countries knew about the worldwide smallpox eradication campaign?
15. Which countries had most successfully controlled smallpox by 1900?
16. When were the last outbreaks of smallpox in South America, Asia and Africa?
17. Why was smallpox a problem?
18. Which countries do you think might take part in WHO decisions?
19. How many reserve vaccinations are kept in stock?
20. How might the successful smallpox eradication program have affected public opinion about medical science?

Part B Instructions**

Imagine that you are a journalist interviewing a doctor. The doctor is working on a village vaccination program in East Africa. Make interview questions from the five statements below. Then make an extra five interview questions of your own.

1. Ask how long she has worked in East Africa.
2. Ask about living conditions in the villages.
3. Ask her what is the main difficulty of her work.
4. Ask her what she likes about the job.
5. Ask her how the vaccination program is going.

+ questions 6, 7, 8, 9, 10

** Editorial note : the exercises in Part B are taken directly from Fletcher & Birt 1983, p. 36, as referenced below..

Speaking Test

[see Appendix 3 for the picture story panel used to cue testees in this component]

Tester notes : Except for the very special genres of speech making and private inner speech, speaking is normally done in a discourse context, assisted by both turn-taking cues and body language (unless it is a phone conversation). Intonation plays a central role, and there is an instant opportunity for requesting clarification and repetition. Thus although the test below is stimulated by a picture panel, it is a thoroughly unreal test of discourse performance. Candidates may be realistically more, or less,

articulate than the test suggests. As every university tutor knows, getting even an average native speaker to perform coherently in the the formal context of a tutorial is often extremely difficult. Probably the only genuine test of "speaking ability" is to tag someone electronically twenty-four hours a day for a week. That is scarcely amenable to a controlled testing format, even ignoring other practicalities.

The questions below are provided essentially for the requirements of this assignment. Most speaking tests of the picture stimulation kind do not provide the additional cues of pre-formatted questions (the G-TELP test does not). It would be possible to create an interview format where the tester did provide some substitute for normal discourse exchange in the form of questions. Unless the questioning process was strictly structured, this would introduce new variables and hence reduce inter-rater reliability (probably one reason that large standardized tests like G-TELP tend to forego the interview format). Of course, asking a testee to speak without question cueing is a greater test of creativity, which is only partially related to the L2 situation.

PART A Instructions :

Study the panel of eight pictures and diagrams. Use the panel to help describe the history of smallpox vaccine. Here is some of the information the tester will be listening for. The questions below are only suggestions

1. When was smallpox finally eradicated?
2. Why was smallpox a problem?
3. Where do we find early records of smallpox?
4. Who was the discoverer of smallpox vaccine?
5. When did he live?
6. What actually is vaccination?

7. What kind of instruments are used for vaccination?
8. Where was smallpox most common before vaccinations began?
9. Which countries had most successfully controlled smallpox by 1900?
10. Which countries continued to be seriously affected by smallpox in 1900?
11. What are possible reasons that some countries were more effective than others in the early control of smallpox?
12. What would make a vaccination program difficult to carry out in some countries?
13. How might the successful smallpox eradication program have affected public opinion about medical science?
14. Why was 1965 an important date in the history of smallpox eradication?
15. What do you think the WHO is?
16. Which countries do you think might take part in WHO decisions?
17. How much do you think villagers in remote African countries knew about the worldwide smallpox eradication campaign?
18. When were the last outbreaks of smallpox in South America, Asia and Africa?
19. Where are reserve vaccination supplies of smallpox vaccine kept?
20. How many reserve vaccinations are kept in stock?

Part B Instructions**

Imagine that you are a journalist interviewing a doctor. The doctor is working on a village vaccination program in East Africa. Make interview questions from the five statements below. Then make an extra five interview questions of your own.

1. Ask how long she has worked in East Africa.
2. Ask about living conditions in the villages.
3. Ask her what is the main difficulty of her work.
4. Ask her what she likes about the job.
5. Ask her how the vaccination program is going.

+ questions 6, 7, 8, 9, 10

** Editorial note : the exercises in Part B are taken directly from Fletcher & Birt 1983, p. 36, as referenced above.

Grammar Test

[see Appendix 4 for a set of G-TELP specifications relating to this segment]

Tester's note : The underlying template for testing in this series of assignments is the G-TELP test (discussed in detail in Assignment 1). The G-TELP test at intermediate level chooses to evaluate control of particular syntactic structures, in addition to those required at lower levels of the test. The relationship of these structures to any actual order of language acquisition is open to debate. Further, G-TELP testers give no indication that they distinguish between command of local elements (which are easy to test), and global factors (which influence meaning critically, but are rather more difficult to test reliably than the local variety). The G-TELP grammar questions are actually a rather unimaginative collection of multiple choice questions of the kind familiar from other large standardized tests like TOEIC. Nevertheless the questions below draw on the structures required by intermediate level G-TELP, with an emphasis on those at Level 3.

The tester is aware that a concentration of syntactically related structures, used contrastively in a single test can in itself cue and tutor the second language testee to reconsider answers. Although this might seem to detract from the validity of an achievement test, in the context of this test design, which is ultimately intended to coach students (see Assignment One) it is actually an advantage.

Note that there is more than one possible answer to some questions in Part B. However the number of possible answers is small enough for markers competent in English to expect maximum inter-marker reliability.

Questions, Part A

Instructions

The following items need a word or words to complete the sentence.

From the four choices which follow each item, choose the best answer. Circle the letter of the correct choice.

1. In 1796 Edward Jenner _____ a milkmaid when he realized that cowpox could immunize against smallpox.

- a. treated
- b. treating
- c. is treating
- d. was treating

2. If we take care, smallpox _____ not kill large numbers of people again.

- a. would
- b. will have
- c. will
- d. have

3. If smallpox does not break out again, scientific cooperation _____ successfully eliminated the disease.

- a. will
- b. will have

- c. would have
- d. would

4. If the WHO had not existed, there _____ been no easy way to coordinate the anti-smallpox campaign.

- a. will
- b. will have
- c. would have
- d. had

5. Smallpox has now ceased to be a major health problem.

- a. is
- b. _____ [nothing]
- c. was
- d. has

6. Smallpox has been eliminated _____ doctors must continue to be alert.

- a. and
- b. or
- c. but
- d. so

7. Political organization _____ science are both needed to fight epidemics.

- a. and
- b. but
- c. or
- d. nor

8. Neither money _____ goodwill are enough to overcome major health problems.

- a. and
- b. but
- c. nor
- d. with

9. Either a lot of luck _____ intense research will be needed to beat the AIDS virus.

- a. and
- b. but
- c. nor
- d. or

10. Logical inquiry together _____ advanced technology is needed to solve many modern problems.

- a. through
- b. by
- c. with
- d. and

11. _____ smallpox was eradicated, emergency supplies of vaccine were stored in Geneva and New Delhi.

- a. because
- b. after
- c. while
- d. however

12. Edward Jenner is famous _____ he devised a vaccine to immunize against smallpox.

- a. although

- b. when
- c. because
- d. after

13. Smallpox has apparently been eradicated _____ we can never be sure there won't be another outbreak somewhere.

- a. although
- b. when
- c. while
- d. after

14. Modern medical science has performed miracles. _____, there are still many major medical challenges to overcome.

- a. and
- b. although
- c. while
- d. however

15. _____ the anti-small pox campaign was underway, many city doctors went to remote regions to help.

- a. during
- b. although
- c. while
- d. because of

16. Smallpox caused many fatal epidemics _____ it was eradicated.

- a. because
- b. before
- c. or
- d. during

17. Governments can always gamble that there will not be epidemics _____ not prepare for emergencies.

- a. but
- b. or
- c. and
- d. because of

18. Doctors can only treat the most urgent cases first, _____ try to give quick help to all patients.

- a. and
- b. but
- c. or
- d. before

19. Helping remote communities is emotionally rewarding _____ often costs doctors in lost income.

- a. although
- b. however
- c. and
- d. but

20. Mass vaccinations cannot be conducted often _____ the high costs involved.

- a. when
- b. because
- c. because of
- d. whether

Part BInstructions

Join the words in columns A and C by writing a suitable word in column B.

	A	B	C
21	In medicine, prevention is		than trying to cure sick people
22	Careful planning eventually costs		in suffering than ignoring medical risks
23	Smallpox was one of the		successes of modern medicine
24	In 1965, the anti-smallpox campaign was the		vaccination campaign ever attempted
25	Persuading people to cooperate		always very difficult
26	The smallpox vaccination campaign has often		used as an example of international cooperation
27	A virus		a kind of living organism
28	Defeating epidemics is		difficult than curing a single patient
29	The defeat of epidemics is a/an		experience for doctors
30	Not many diseases can be		and easily defeated by doctors

Appendix 1 Listening Skills

Content : General Secretarial Duties - Management of the Telephone

Source :

<http://www.ca.uky.edu/internal/Office%20Manual/13%20secretarial%20duties/13.htm> - University of Kentucky County Extensions Office

Transcript :

Speak clearly, slowly and distinctly.

How to answer and what to say - This applies to all secretaries and agents.

"Cooperative Extension, Cook County Office, Jane speaking."

Leaving the line - When you must leave the line to obtain information, it is courteous to give the caller the option of waiting or of being called back. You might say "Will you hold or shall I call you back?" If the caller waits, use the hold button. Give progress reports every 30 seconds and say "I'm still looking for ____." When you return to the line, thank the caller for waiting.

May I take a message? - Take accurate messages to avoid misunderstandings later. Include all the information you have knowledge of: what may seem insignificant to you may be extremely important to the person receiving the message. Include date, time, your initials, the caller's name, telephone number, the business or service represented and/or the information desired. Tell the caller when the person they want is expected or whether he/she can be reached at another location IF APPROPRIATE. Offer whatever information you can. Inquire whether anyone else can help them. Do not make promises to the caller regarding when the call will be returned. Do advise the caller when the person is expected to return or call in for messages. Say "I'll give the message" or "I'll tell _____ you called."

Transferring calls - Tact, courtesy, and efficiency are essential when calls are

being transferred. Before transferring a call, explain to the caller that he/she is being transferred and ask him/her if he/she is are willing to be transferred. You might say, "_____ is responsible for touring information. May I transfer you to his/her extension?" Transferring tips: Provide the caller with number of the called party in case you are disconnected during the transfer process. If you are not sure about how to transfer a call, check instructions before taking action. Do not guess when transferring calls. If you are not certain to whom the call should be transferred, ask the calling party to wait a moment while you check. Consult with co-workers and find out all you can to avoid run around!

Placing calls - Check the phone directory, number finder, or other lists before calling. Identify yourself and your organization immediately and state the purpose of your call.

Remember time zones and avoid calling during lunch hours. Allow time for the party to answer. Leave clear, concise messages on answering machines or with answering services. At the minimum, leave your name, your organization, the date, the time, a phone number, and a brief message regarding the purpose of the call.

Appendix 2 - Reading Test

Source : Korea National Tourist Office - 'Things to Buy in Korea'

Clothing When purchasing clothing, consider the material, quality of the stitching, design, and price. Make sure to check for the quality tag (a tag with information on the fabric and washing instructions) on the inside of the garment to ensure that it has undergone quality control. Clothing is usually purchased in department stores, brand shops, outlets, Dongdaemun Market, Namdaemun Market, Myeongdong, Itaewon, and in front of Ehwa Women's University. In the case of brand name goods, it is a good idea to wait for sales in January, April, July, and October (may be subject to change). Make sure to

keep your receipt in case you need to exchange or return something.

Leather Goods Leather goods come in a wide variety- clothing, shoes, and bags. Before purchasing, consider the type of leather, quality of the dye, stitching, inner lining, and price. Good places to find leather goods are department stores, Namdaemun Market, Dongdaemun Market, Itaewon, and more. Tourists prefer markets at Itaewon because high quality goods are offered at reasonable prices. Many of the shops have their own factories, so custom goods can be ordered at some. When ordering a custom made item, you're shown many samples of materials before choosing one- and it's a good idea to take along pictures from magazines as well with designs you like.

Electronics : Among items sought out by tourists, the most popular are notebook computers, digital cameras, camcorders, mp3 players, PDAs, and more. When making a purchase, you should research the name and model number of the product beforehand, and consider the following: For a notebook computer, consider the uses of the computer and whether or not it fits your lifestyle. Will you watch movies? Play games? Design Graphics? With your answers, select the processor and monitor accordingly. In the case of digital cameras, check for the megapixels and other functions. For the casual photographer, 3.0 megapixels is plenty, while 5.0 megapixels or more are recommended for professional use. With camcorders, you must also consider resolution and monitor size- some can even double as a digital camera, so check before you buy. Good places to purchase electronics are department stores, the Yongsan Electronics Market, discount marts, Technomart, and more. Of these, the Yongsan Electronics Market is the most famous- a good selection of goods for great prices. Computers, household appliances, and game software is available- there's even a market for used goods. Keep in mind that most stores offer discounts for cash purchases.

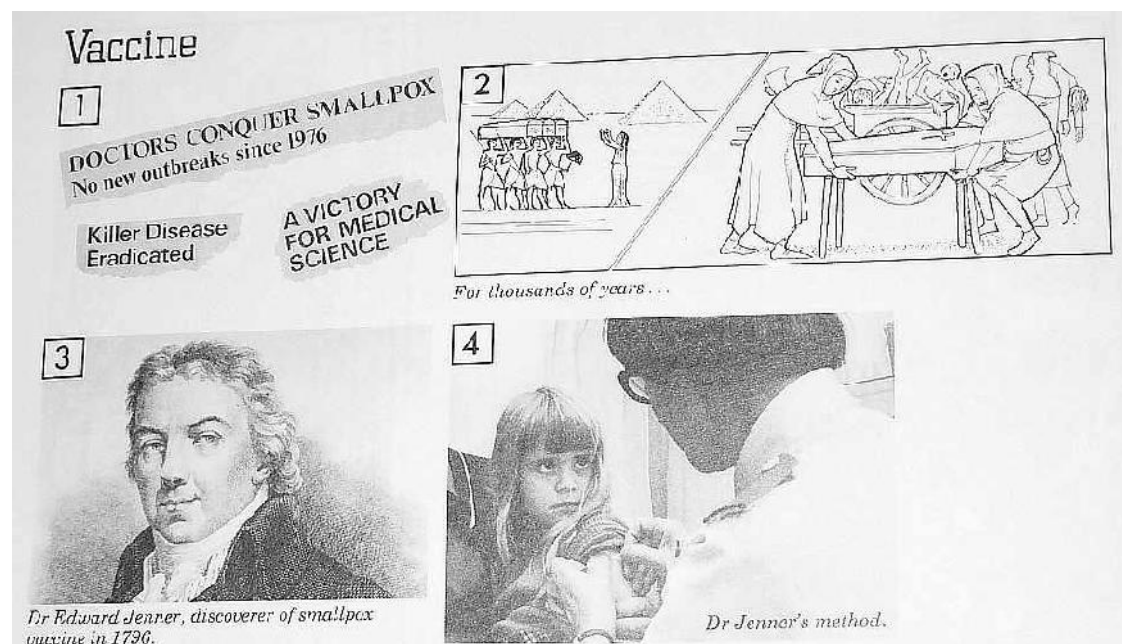
Cell Phone Chains Appreciated as a gift with actual use, cell phone chains range from cute and cheap plastic to delightful and more expensive handcrafted offerings. When purchasing as a gift, traditional Korean chains

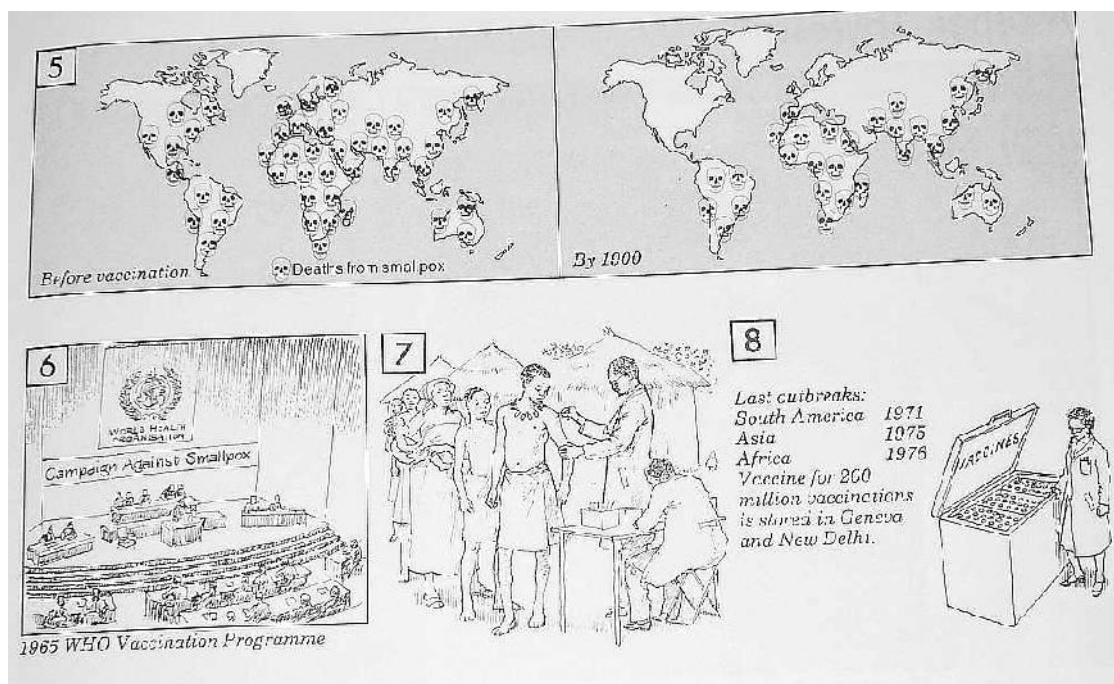
are popular- they cost about 2~3 times more than regular chains, but many are beautifully designed with great workmanship. Trendy hand phone chains can usually be purchased in Myeongdong, Jongno, Apgujeong, Ehwa Women's University, and more, while traditional chains can be found in Insadong.

Postcards Postcards are economical and easy to transport, a perfectly convenient and appropriate souvenir. The most popular post cards have traditional Korean themed photographs or artwork, and can be purchased at sites with high tourist traffic such as Insadong, palaces, art galleries, and museums. Prices are from 500 won to 1000 won.

Appendix 3 - Speaking and Writing Tests - Picture Panel Cues for text construction

Source : "Storylines ... picture sequences for language practice" by Mark Fletcher and David Birt, page 36; published by Longman 1983.





Appendix 4 - G-TELP Specifications for Levels 3 to 5, Grammar Tests

Source : G-TELP website http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp04.asp

This is the G-TELP focus for structure (content extracted directly from G-TELP literature) :

Level 5 (lowest level)

General Description This level assesses the ability of the test taker to produce and recognize minimal amounts of English. The test taker has had limited exposure to English in the classroom, and knows only basic elements of the language.

- Grammar A test taker at this level is exposed to a variety of basic grammatical structures, such as frequently used irregular verbs and

present tense. He/she is required to demonstrate understanding of basic grammatical structures in one-and two-sentence contexts.

Structures tested at this level include:

*Verbs

- present progressive
- simple present
- simple past
- future tense with will and be going to
- constructions with there is/there are
- constructions with there was/there were

*Nouns

- plurals
- singular

*Pronouns

- personal pronouns: subject, possessive

*Formation of Questions

- yes/no questions with to be
- information questions with to be
- short answers with to be

Level 4 descriptors

General Description This level assesses the abilities of a test taker to handle basic communicative functions. The test taker is capable of basic formulaic expressions such as those used in greetings and introductions, and is successful with basic communicative transactions. Although a test taker at this level has had only limited exposure to English outside of the classroom, he/she is able to use English in simple communication with native speakers in a narrow range of tasks.

- Grammar An examinee at this level is required to demonstrate understanding of basic grammatical structures in one-and- two sentence contexts. Structures tested at this level include:

*Verbs

- be
- commands
- negatives
- constructions with there is/there are

*Pronouns

- personal pronouns: subject, object, possessive

*Formation of Questions

- yes/no question with be, was, were, and do, does, did
- short answers to yes/no questions with be, was, were, and do, does, did
- question-word questions (Wh-words)

Level 3 descriptors

General Description

This level assesses the abilities of a test taker to communicate in English on an elementary level. This person has been exposed to English mainly in the classroom with little extended contact with native speakers. It should be noted, however, that while a person at this level is not an independent user of the language and is not expected to cope with unmodified English, he/she would be able to function within a limited range of authentic English language situations.

- Grammar

An examinee at this level is required to demonstrate understanding

of basic grammatical structures in two-to three-sentence contexts.

Structures tested at this level include:

*Verbs

- past progressive
- future possible conditional with will, would, and would have
- present perfect

*Coordinators and Subordinators

- and, but, or, nor, so
- after, because, etc.

Word forms

- adjectives: equatives, superlatives, comparatives
-

Part III - PRACTICE ENGLISH LANGUAGE TEST

Level : Intermediate

(based on the G-TELP international language test)

Test Instructions

1. This test is designed to give you practice in doing international language tests like G-TELP. The G-TELP test itself (<http://www.gtelp.co.kr>) is an American test used in job hiring by many big companies in South Korea.
2. This test has five parts : Speaking, Listening, Reading, Writing and Grammar.
3. Because this is a practice test, we are not too worried about strict time limits. Take your time. Do your best. Be honest with yourself and the tester. The idea is to coach you to do better next time. "Cheating" won't help you.
4. Don't use a dictionary, don't ask for help, and don't do anything you couldn't do in a real job testing situation.

*Tester : Thor May
thormay@yahoo.com
Department of English
Chungju National University
South Korea*

1. Speaking Test

Your Name : _____ Your e-mail : _____ Date : _____

Instructions

1. Study the panel of eight pictures and diagrams.
2. Use the picture panel and hints to help describe the history of smallpox vaccine.
3. The short hints below are only ideas; (they are not proper English sentences!). Make your own story.
4. Try to speak for several minutes (no more than 10 minutes !)
5. Record your story and send it to the tester**

How to record your story : You can use a tape recorder, or record on an MP3 player, or on computer.

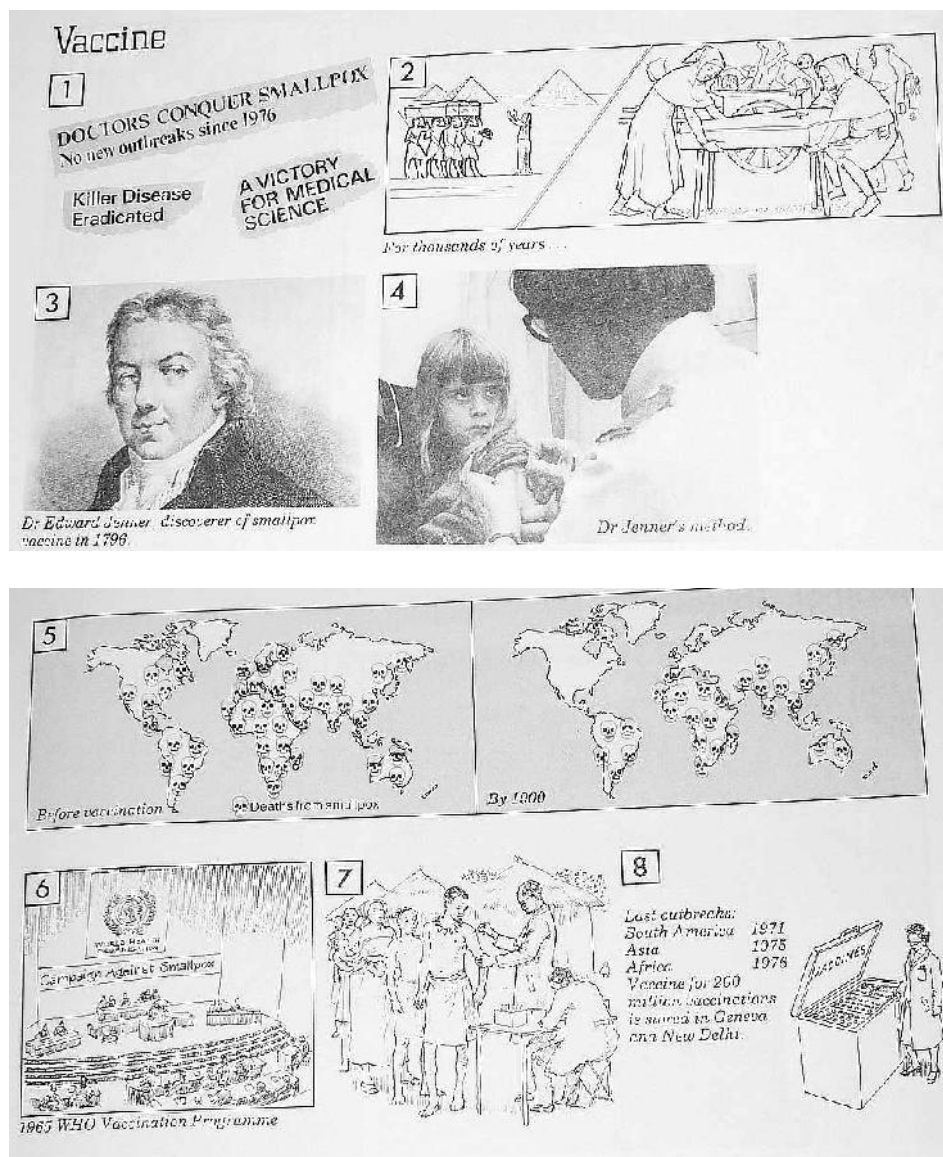
**The tester : Thor May; e-mail : thormay@yahoo.com ; Department of English, Chungju National University, 123 Gyomdon-ri, Iryu-myon, Chungju, Chungbuk 380702, South Korea

Hints For Talking About The Pictures

1. When eradicated?
2. Why a problem?
3. Where early records?
4. Who discoverer ?
5. When live Jenner ?
6. What vaccination?
7. What used for vaccination?
8. Where most common before program?
9. Which countries most successful 1900?
10. Which countries affected 1900?
11. What reasons some countries more effective?
12. What make vaccination difficult ?
13. How might success program affect public opinion?
14. Why 1965 important ?
15. What WHO is?
16. Which countries WHO decisions?
17. How much villagers Africa knew about campaign?
18. When last outbreaks South America, Asia and Africa?
19. Where reserve vaccination supplies ?

20. How many vaccinations kept ?

Picture Panels Showing The History of Smallpox



2. LISTENING TEST

Your Name : _____ Your e-mail : _____ Date : _____

General Instructions :

The listening test is in two sections. Each section has two parts.

Section 1 The recording tells a secretary in an American office about how to use the telephone.

- a) Listen carefully to the recording. You may re-play the recording to listen several times.
- b) Part A1 : Answer multiple choice questions about telephone duties.
- c) Part B1 : Choose the best answers a secretary can give to someone calling her office.

Section 2 This is a conversation when a young man meets a woman in America

- a) Listen carefully to the recording. You may re-play the recording to listen several times.
 - b) Part A2 : Fill in the gaps in the sentences. Only use words you hear in the recording.
 - b) Part B2 : Answer multiple choice questions about the conversation.
-

2. LISTENING TEST - SECTION 1

Recording : Telephone Duties for a Secretary

Part A : Multiple Choice Questions

Instructions :

1. Choose the best answer for each question.
 2. The best answer is what you hear on the recording. The best answer is NOT what you guess from experience.
 3. Circle a, b or c in each of the following questions.
-

1. When you answer the telephone, you should
 - a. speak kindly
 - b. speak quietly
 - c. speak clearly
2. The rules apply to
 - a. all employees
 - b. all secretaries
 - c. all office staff
3. You give the caller an option
 - a. of waiting or being called back
 - b. of waiting or calling back
 - c. of waiting or calling another number

4. Give progress reports
 - a. every 30 minutes
 - b. every 3 minutes
 - c. every 30 seconds

 5. When you return to the phone
 - a. tell the caller to ring later
 - b. thank the caller for waiting
 - c. thank the caller for their call
-

Listening - Section 1, Part B - Dialogue Choice

Instructions :

- a) Imagine that you are a manager's secretary. You have to answer the telephone in a certain way.
 - b) Listen carefully to the "Telephone Duties for Secretaries".
 - c) Now, what are the best answers you can give to a caller from the questions below ?
 - d) Circle a, b or c in questions 1 to 5 below.
-

- 1a. "This is Jane Smith. How can I help you?"
- 1b. "Cooperative Extension, Cook County Office, Jane speaking."
- 1c. "Hello. Cooperative Extension Office, extension 594."

- 2a. "Will you hold on, or will I call you back?"
- 2b. "Hold the line please."
- 2c. "The manager is busy now. Please call back later."

- 3a. "You can give details when you talk to the manager."
- 3b. "May I check that again? You want me to tell the manager you got a post card from someone called Sadie?"
- 3c. "I don't know anything about Sadie, sorry. However I will tell the manager you called."

- 4a. "This morning the manager has been called out on an emergency. You can reach him at 137 5557 8352."
- 4b. "This morning the manager is visiting his mother in hospital. You can reach him at 137 5528 1935."
- 4c. "This morning you can reach the manager at our Orange County branch. The number is 137 5521 1689."

- 5a. "Is there anyone else here who can help you?"
- 5b. "I'm sorry. I can't transfer you. You will have to ring the sales supervisor directly."
- 5c. "That other person you mentioned works in the warehouse. I don't know

his number. Sorry."

LISTENING TEST - SECTION B

Recording : "Haven't We Met Before?" Listening Part A - Dialogue Completion

Instructions :

- a) Listen to the recording carefully
 - b) Fill in the gaps in the following sentences. Only use words you hear in the recording.
-

1. We _____ on campus last week
 2. Let me guess. It's Sherry, _____?
 3. No, but you got the _____ letter
 4. You got it . . . and _____ on the fourth try.
 5. Not bad. _____ what was your name?
-

Listening - Section 2, Part B : Multiple Choice Questions

Instructions :

- a) Choose the best answer for each question.
 - b) The best answer is what you hear on the recording. The best answer is NOT what you guess from experience.
 - c) Circle a, b or c in each question to show your choice.
-

1. The man
 - a) thinks that he has the right qualifications to teach English as a second language
 - b) wants to study to get the right qualifications to teach English as a second language
 - b) isn't sure if he has the right qualifications to teach English as a second language

2. What does the man want to study most at the moment?

- a) Marketing
- b) International business
- c) Accounting

3. The woman says

- a) she will miss her class
- b) she must hurry to get to her class
- c) her class began ten minutes ago

4. The man

- a) invites the woman to a dance
- b) thinks the woman is interested in dancing
- c) was wondering if there was a dance tonight

5. The woman

- a) is very interested in seeing the man again
- b) might be interested in seeing the man again
- c) doesn't really want to see the man again

3. Reading Test

Your Name : _____ Your e-mail : _____ Date : _____

General Instructions

1. The reading test is in two sections
 2. Section 1 is advice taken from a South Korean tourist brochure
 3. Section 2 is taken from an interview with a South Korean student who has moved to the United States
 4. Both Section 1 and Section 2 contain multiple choice questions to check your understanding of ideas
 5. Both Section 1 and Section 2 contain multiple choice vocabulary questions to check your understanding of words and phrases
-

Reading Section 1 - A Tourist Brochure

[Source : Korea National Tourist Office - 'Things to Buy in Korea']

Instructions :

1. Read the material from a KNTTO tourist brochure.
 2. Choose the **best answer** in each multiple choice question. The best answer here is what the article says or implies, not what you think personally.
 3. In each question, circle a, b or c.
 4. The underlined words are for vocabulary questions
-

Things To Buy In South Korea

Clothing When purchasing clothing, consider the material, quality of the stitching, design, and price. Make sure to check for the quality tag (a tag with information on the fabric and washing instructions) on the inside of the garment to ensure that it has undergone quality control. Clothing is usually purchased in department stores, brand shops, outlets, Dongdaemun Market, Namdaemun Market, Myeongdong, Itaewon, and in front of Ehwa Women's University. In the case of brand name goods, it is a good idea to wait for sales in January, April, July, and October (may be subject to change). Make sure to keep your receipt in case you need to exchange or return something.

Leather Goods Leather goods come in a wide variety- clothing, shoes, and bags. Before purchasing, consider the type of leather, quality of the dye, stitching, inner lining, and price. Good places to find leather goods are department stores, Namdaemun Market, Dongdaemun Market, Itaewon, and more. Tourists prefer markets at Itaewon because high quality goods are offered at reasonable prices. Many of the shops have their own factories, so custom goods can be ordered at some. When ordering a custom made item, you're shown many samples of materials before choosing one- and it's a good idea to take along pictures from magazines as well with designs you like.

Electronics : Among items sought out by tourists, the most popular are notebook computers, digital cameras, camcorders, mp3 players, PDAs, and more. When making a purchase, you should research the name and model number of the product beforehand, and consider the following: For a notebook computer for example, consider the uses of the computer and whether or not it fits your lifestyle.

Reading Section 1 - Multiple Choice Questions (Ideas)

1. The quality tag on clothing tells you
 - a) that the item was made in South Korea
 - b) that the item has been checked for faults
 - c) that the item is a good buy

2. The tourist brochure advises you to check clothing for
 - a) careful stitching
 - b) whether it fits you
 - c) whether it is fashionable

3. A good place to buy clothing is
 - a) In front of Korea University
 - b) In Itaewon shopping district
 - c) Around Chongno-sam Ga

4. You should keep your receipt
 - a) to prove that you bought the item if challenged
 - b) to claim a tax refund
 - c) to make it easy to return or exchange goods

5. Before you buy leather, you should
 - a) consider if you really need it
 - b) consider the price
 - c) consider how long it will last

6. A good place to find leather goods is
 - a) hotel shops
 - b) duty free shops
 - c) department stores

7. Tourists prefer Itaewon because
 - a) there are high quality goods at reasonable prices
 - b) the shopkeepers speak English
 - c) it is recommended by the tourist office

8. Custom goods can be ordered at
 - a) leather factories
 - b) some high class shops
 - c) some shops owned by leather factories

9. When you are ordering custom items

- a) you are shown many designs to choose from
- b) it is best to take pictures of a design you like
- c) it is best to take professional advice

10. You should research the model number of electronic items

- a) before you buy anything
 - b) after you see something you like
 - c) before you come to Korea
-

Reading Section 1 - Multiple Choice Questions (Vocabulary)

11. In this reading passage [paragraph 1], "on the fabric" means

- a) printed on the garment you are buying
- b) about the material in the garment you are buying
- c) on the information tag about the garment you are buying.

12. In this reading passage [paragraph 1], "outlets" means

- a) places where water or other liquids come out of a pipe.
- b) ways people use to express their talents, or relieve tension
- c) places where things are sold

13. In this reading passage [paragraph 1], "case" means

- a) a kind of solid bag for carrying things in
- b) a job involving one patient or client, worked on by a professional person
- c) a circumstance, a situation

14. In this reading passage [paragraph 2], "custom" means

- a) the way people usually do things
- b) a kind of tax
- c) goods made for the special individual needs of a buyer

15. In this reading passage [paragraph 2], "sample" means

- a. an example of a particular product or design
 - b. to try something out
 - c. a small container of a patient's urine taken by a doctor
-

Reading Section 2 - An Interview

Instructions :

1. Read the Pacific University interview
 2. Choose the best answer in each multiple choice question. The best answer here is what the interview says or implies, not what you think personally.
 3. In each question, circle a, b or c.
 4. The underlined words are for vocabulary questions
-

Pacific University Interview

[<http://mcel.pacificu.edu/as/students/korea/culture-interview.html>]

In this interview Ben Bagley asks Theresa Han about the difference between Korean and American culture. Theresa is a teenager who recently moved to the United States so she has an excellent perspective for understanding the differences and similarities between these countries.

[BAGLEY] This is Ben Bagley, and I'm going to interview Theresa Han about Korea. Could you introduce yourself?

[HAN] My name is Theresa Han, I'm from South Korea, I'm 18 years old, and I'm a freshman in College.

[BAGLEY] How long have you lived in America?

[HAN] I think a little bit less than 3 years.

[BAGLEY] Where did you live in Korea?

[HAN] I lived in Pyoung Tek, It's right below Oosan, where the American Air force is located.

[BAGLEY] What were the people like where you lived?

[HAN] They're really busy. Fathers go to their work. Mothers if they have a job go to their work, and students go to school, so they don't have enough time to communicate with each other, like time to spend together, because mostly students come home like 10:00pm-11:00pm.

[BAGLEY] What did you do with your friends?

[HAN] We mostly go to each other's house, rent a movie or something, watch it, and do homework usually, because we have a whole bunch of homework. On the weekends we would go downtown; it's kind of like a shopping mall. It's a street. There are small restaurants, small clothes shops and all that stuff. It was kind of fun, but nothing special I think.

Reading Section 2 - Multiple Choice Questions (Ideas)

1. Which of these statements is true ?

- a) Theresa spent most of her life in Korea
- b) Theresa spent most of her life in America
- c) Theresa spent about equal time in Korea and America

2. Which of these statements is true ?

- a) The American Air Force base is in Pyoung Tek
- b) The American Air Force base is below Oosan
- c) Theresa lived near The American Air Force base

3. Which of these statements is true ?

- a) Students in Pyoung Tek see a lot of their mother and father
- b) Students in Pyoung Tek are not good communicators
- c) Students in Pyoung Tek are too busy to communicate much with their families

4. Which of these statements is true ?

- a) Students visit each other's houses both to watch movies and to do homework
- b) Students visit each other's houses mostly to do homework
- c) Students visit each other's houses mostly to watch movies

5. Which of these statements is true ?

- a) Downtown Pyoung Tek is a very exciting place
 - b) Downtown Pyoung Tek has streets of small shops
 - c) Downtown Pyoung Tek has a single street of small shops
-

Reading Section 2 - Multiple Choice Questions (Vocabulary)

6. In this reading passage, "perspective" means

- a) skill
- b) reason
- c) viewpoint

7. In this reading passage, "similarities" means

- a) problems
- b) things rather like each other
- c) things the same

8. In this reading passage, "right below" means
- a) just a short distance south on the map
 - b) a little less important
 - c) a part of a town
9. In this reading passage, "have a whole bunch of" means
- a) have very important
 - b) have a lot of
 - c) have to do
10. In this reading passage, "all that stuff" means
- a) things like that
 - b) things to do
 - c) things we liked
-

4. Writing Test

Your Name : _____ Your e-mail : _____ Date : _____

Instructions

1. Write a detailed email to a foreign businessman who will visit your company.
 2. Tell him how he will be met at Incheon Airport**. State times, dates, places, and give him an itinerary.
 3. This is his first visit to South Korea.
 4. Give a brief description of Chungju and local information about your company.
 5. ** If you don't know South Korea well, or don't know Chungju, you can use another country or city in your e-mail instead.
 6. Below is a list of short hints. The hints suggest some information which your e-mail could contain. You don't have to use all of them; (the hints are not proper English sentences !).
-

Hints For Writing A Business E-mail

1. Proper greeting format to stranger in business e-mail?
2. Who you ? What your role ?
3. What foreign businessman already know => itinerary?
4. You TELL him itinerary, or just suggestions?
5. Who meet him Incheon ?
6. What time flight ?
7. What weather like South Korea this time ? Clothing ?
8. How long passport and customs ?
9. Alright product samples in hand luggage ? Packed separately?
10. Where met in Incheon ?
11. Airport layout, services if not met ?
12. What cell phone number person meet him?
13. What person meet him look like?
14. How long from Incheon => Chungju?
15. What transport that trip?
16. Where staying Chungju?
17. Company, accommodation contact numbers Chungju?
18. How many days itinerary ?
19. What doing each day ?
20. Provided with car, driver?
21. Seminars, presentations to give ?
22. Who meeting during stay?
23. Who speak English your company?
24. Useful facts about company?
25. Useful facts about Chungju?
26. How far Chungju => Seoul; what transport connections?
27. Special attractions Chungju area?
28. Educational, research facilities, libraries near Chungju?
29. Internet, telephone, other electronic services fo visitors?
30. Banking services for visitors?
31. Signature format

5. Grammar Test

Your Name : _____ Your e-mail : _____ Date : _____

1. The grammar test has 40 questions in three sections.
 2. Section A asks you to complete multiple choice questions.
 3. Section B asks you to join two phrases or sentences with a correct word.
 4. Section C asks you to correct some sentences which may have grammatical errors.
-

Grammar - Section A (Multiple Choice Questions)

Instructions

1. Each sentence below needs a word or words to finish it.
 2. Choose the best answer for the missing word in each sentence.
 3. Circle the letter of the correct choice.
-

1. In 1796 Edward Jenner _____ a milkmaid when he realized that cowpox could immunize against smallpox.
 - a. treated
 - b. treating
 - c. is treating
 - d. was treating
2. If we take care, smallpox _____ not kill large numbers of people again.
 - a. would
 - b. will have
 - c. will
 - d. have
3. If smallpox does not break out again, scientific cooperation _____ successfully eliminated the disease.
 - a. will
 - b. will have
 - c. would have
 - d. would

4. If the WHO had not existed, there _____ been no easy way to coordinate the anti-smallpox campaign.
- will
 - will have
 - would have
 - had
5. Smallpox has now ceased to be a major health problem.
- is
 - _____ [nothing]
 - was
 - has
6. Smallpox has been eliminated _____ doctors must continue to be alert.
- and
 - or
 - but
 - so
7. Political organization _____ science are both needed to fight epidemics.
- and
 - but
 - or
 - nor
8. Neither money _____ goodwill are enough to overcome major health problems.
- and
 - but
 - nor
 - with
9. Either a lot of luck _____ intense research will be needed to beat the AIDS virus.
- and
 - but
 - nor
 - or

10. Logical inquiry together _____ advanced technology is needed to solve many modern problems.

- a. through
- b. by
- c. with
- d. and

11. _____ smallpox was eradicated, emergency supplies of vaccine were stored in Geneva and New Delhi.

- a. because
- b. after
- c. while
- d. however

12. Edward Jenner is famous _____ he devised a vaccine to immunize against smallpox.

- a. although
- b. when
- c. because
- d. after

13. Smallpox has apparently been eradicated _____ we can never be sure there won't be another outbreak somewhere.

- a. although
- b. when
- c. while
- d. after

14. Modern medical science has performed miracles. _____, there are still many major medical challenges to overcome.

- a. and
- b. although
- c. while
- d. however

15. _____ the anti-small pox campaign was underway, many city doctors went to remote regions to help.

- a. during
- b. although
- c. while
- d. because of

16. Smallpox caused many fatal epidemics _____ it was eradicated.

- a. because
- b. before
- c. or
- d. during

17. Governments can always gamble that there will not be epidemics _____ not prepare for emergencies.

- a. but
- b. or
- c. and
- d. because of

18. Doctors can only treat the most urgent cases first, _____ try to give quick help to all patients.

- a. and
- b. but
- c. or
- d. before

19. Helping remote communities is emotionally rewarding _____ often costs doctors in lost income.

- a. although
- b. however
- c. and
- d. but

20. Mass vaccinations cannot be conducted often _____ the high costs involved.

- a. when
- b. because
- c. because of
- d. whether

Grammar - Section B (Phrase Completion)

Instructions

1. Join the words in columns A and C by writing a suitable word in column B.

	A	B	C
21	In medicine, prevention is		than trying to cure sick people

22 Careful planning eventually costs	in suffering than ignoring medical risks
23 Smallpox was one of the	successes of modern medicine
24 In 1965, the anti-smallpox campaign was the	vaccination campaign ever attempted
25 Persuading people to cooperate	always very difficult
26 The smallpox vaccination campaign has often	used as an example of international cooperation
27 A virus	a kind of living organism
28 Defeating epidemics is	difficult than curing a single patient
29 The defeat of epidemics is a/an	experience for doctors
30 Not many diseases can be	and easily defeated by doctors

Grammar - Section C (Sentence Correction)

Instructions

1. Some of the sentences in Part C have incorrect grammar.
 2. Write out the sentence in each question again. If there is a grammar problem, fix it. If there is no problem, don't change anything.
-

31. An active president has chosen our country.

Answer : _____

32. Escaped the professor from prison.

Answer : _____

33. The neighbours say both children he teases.

Answer : _____

34. Surprises me that he wears a wig.

Answer : _____

35. There was a very tall man here for you.

Answer : _____

36. She has been smoking, isn't it ?

Answer : _____

37. I must can catch this train.

Answer : _____

38. Three packages sent yesterday.

Answer : _____

39. The secretary won't be here today.

Answer : _____

40. Why we bow to each other ?

Answer : _____

This is the end of the test. Thank you very much for trying.

University of Newcastle, NSW

Assignment 3 :

Topic : *Do a trial test administration, collect the data and analyse the results.*

In about 1000 - 1200 words discuss practical aspects of language test construction, then critically analyse your own test considering the following :

- *purpose of the test*
- *content of the test*
- *test specifications*
- *instructions for candidates*
- *quality of individual test items (e.g. ambiguous or not)*
- *representativeness of the sample*
- *reliability and validity*
- *scoring procedures*
- *aspects of moderation*
- *a summary evaluation of the test => Does your test measure the language ability it is designed to measure?*

Practical Aspects of Language Test Construction

1. The notion of testing, and practical consequences

Most people undertake scholastic tests, including language tests, with the utmost reluctance. In the case of language tests, they often have a strong expectation of failure, and they fear the consequences of failure. More often than not, their expectations become a self-fulfilling prophecy. If the ultimate success criterion is some level of useful functional ability in L2 as a foreign language (the story with immigrant acquisition of English is different), then the real failure rate in countries like Australia probably stands at over 90%; (e.g. see Asher's estimate for America).

Most people enjoy playing games. These may be physical sports, traditional indoor games of skill or luck, or more recently extended and extremely elaborate games through electronic media such as the internet. Many of these games involve extensive and extended learning. It is characteristic of nearly all games that success is measured in the form of some kind of test (it may be called a contest or competition). Players are rarely crushed or psychologically damaged in game failure. Rather, they tend to be spurred to greater effort and eventually succeed.

What is the key difference between academic language learning/testing and game playing/testing? It is not complexity, and it is not duration. It is not in the technical content of what is being learned or manipulated. The key difference is socio-political. That is, academic language learners and game players have utterly different perceptions of their roles in the process, and their beliefs are confirmed by an extensive cultural infrastructure of rewards and

punishments. These extend far beyond the learning process itself, and frequently have no intrinsic tie to it. It is this role perception which causes psychological damage, and it is the role perception which is at the heart of the overwhelming public failure of academic foreign language teaching. The language test in its various forms is a principal instrument of socio-political punishment.

2. Test Construction

Language teachers and language testers are generally well-intentioned groups of people. My thirty years of contact with these folk has led me to the general conclusion that most of them have an extremely simple idea of what a natural language is comprised of, what mental processes are involved in its generation, and how it is best taught to a learner. This does not mean that they are bad language teachers. If they can keep learners engaged in the target language they have a useful role, even if they don't know what is really going on. However, it is fair to say that nearly all language tests will be constructed as a compromise between teacher folk-wisdom and the political requirements of a workplace or culture.

Criteria like validity and reliability will be ultimately grounded in the conventional wisdom of language pedagogy as it exists within a local political culture; (thus conclusions could vary somewhat between, say, Australia and South Korea). That is the tester may maximise objective procedures and strive for consistency. However, the underlying decisions about what questions to ask and answer will always be subjective (this is true of all science).

The Practice Test of this assignment has been constructed in the conventional manner, based in its structure (but not its purpose) on

an international criterion-referenced public language test, the G-TELP.

3. Comparison of the Practice Test and G-TELP :

G-TELP	PRACTICE TEST
Proficiency test for employment	Diagnostic test of current proficiency, intended to encourage focused study
Mixture of direct and indirect measurement	Mixture of direct and indirect measurement
Discrete point except for speaking & writing tests; non-discrete elements integrative	Discrete point except for speaking & writing tests; non-discrete elements integrative

G-TELP	PRACTICE TEST
Criterion referenced (supposedly)	'Negotiable' : scoring is not compiled for the overall test. Individual elements are discussed and explored, post-test, with the testee
The test has been moderated by a body of native and non-native testing personnel, using test results from a large population of international students	The test has been designed and executed by a single tester on a minimum number of students
The test is conducted under controlled test conditions	For logistical reasons the test has been self-administered by testees at a location remote from the tester
Claims to be mainly objective; subjective judgements subject to moderation with other testers	Discrete point elements are mostly objectively marked, subject to some cultural interpretation. Subjective judgements are discussed post-test with the testee
Measurement scale : ordinal scales 1 to 5 for speaking, listening, reading (& vocabulary), writing and grammar. 5 on the scale represents beginner. Interval scale test scores are converted to an	A comprehensive measurement scale is irrelevant to the purpose of the Practice Test. The test content is modelled on subset taken from Level 3 G-TELP (intermediate level).

ordinal mastery matrix. "Mastery" is a score of 75% or greater in all skill areas; [this criterion is apparently arbitrary]	
Validity : G-TELP claims face validity based on the continued patronage of large Korean companies. Neither content validity or concurrent validity are publicly quantified however. There seem to be no published studies of the predictive validity of the test.	The sample of testees has been too small to establish any kind of formal validity. The tester is dubious that true construct validity can be established for any language test : there are too many interacting variables.
Reliability : G-TELP publishes no coefficients of reliability. However, as a large professional testing organization one would expect that their test design would be closely monitored for both formal validity and formal reliability	<ul style="list-style-type: none"> - The sample of testees has been too small to establish any measure of formal reliability. - With a single scorer there is no issue of inter-scorer reliability - Some of the elicitation tasks (e.g. writing) disclosed problems that are inherently intractable for test measurement, and hence resistant to measures of validity and reliability.

4. The Purpose of the Practice Test

The Practice Test is a diagnostic test of current proficiency, intended to encourage focused study. This purpose falls outside the normal range of formal test specifications recognized in text books.

However, in non-English speaking countries worldwide there is a major publishing and coaching industry making use of samples and imitations from international standard tests like TEOFL and IELTS. In net effect, these big tests are vehicles to justify the coaching industry. The Practice Test is an attempt to harness this phenomena.

5. The Application of the Practice Test applied to its purpose

Twelve individuals made some promise to take part in an initial trial of the Practice Test. In actuality only one participated, at a remote location under conditions of self-administration. The withdrawals were influenced by local conditions (exams mainly), but also by a perception that the Practice Test was long and time-consuming. Length is a dilemma in all testing, and becomes acute where voluntary participation is expected.

The Korean individual who did participate found the exercise, and a follow up analysis of his performance to be very useful. See the e-mail in Appendix 2. This was an indication that the Practice Test had succeeded in its explicit purpose as a “diagnostic test of current proficiency, intended to encourage focused study”.

6. The Content of the Test

The brief was to compose a “comprehensive ESL test”. As this tester understands the expression, a comprehensive ESL test is impossible in principle. No language test is able to test L2 comprehensively.

The Practice Test aimed to be a useful ESL test. Useful that is to both the testee(s) and the tester. In order to achieve this aim it had to be short enough to encourage participation, and focused enough to show up a pattern of weakness where one existed. One outcome was that it was perhaps too long to encourage participation under voluntary conditions.

The focus, which was effectively that of G-TELP level 3, was indeed fine enough to show up some patterns of weakness in the L2 usage of the participating testee. For example, by restricting itself to a fairly narrow band of syntactic constructions and probing these in some depth, it became clear that the testee had a problem with the logical content of English conjunctions. See Appendix 1 for a detailed examination of errors and infelicities, followed by tester comments.

The grammar test also included a small selection from Burt & Kiparsky's collection of 'goofs' (The Gooficon', 1972), with instructions to correct them. An open-ended exercise like this is anathema to the design of formal standardized tests (the outcomes are too variable), but are quite useful for diagnostic purposes.

The Writing and Speaking sections of the Practice Test could not in their nature be discrete point checks on L2 usage. The testee's performance in these sections did not show serious problems with the elements normally targeted in language tests. However there were serious discourse problems which could render the testee linguistically unfit in, for example, a business environment. The kind of discourse problems identified could probably not be reliably tested for in a formal, mass testing environment. In other words, the discourse problems showed up the limitations of the whole testing process. See Appendix 1 for an explicit analysis.

7. Test Specifications

The specifications of the Practice Test were broadly set by the G-TELP specifications. Since G-TELP is a very professional undertaking, the specifications were not likely to be badly astray. However the

adaptation of G-TELP specifications to a new diagnostic and incentive-generating purpose was experimental. Some small deviations from the G-TELP question patterning (e.g. the Gooficon entries) were clearly outside of standard specifications, but thought to be helpful. With only a single testee, post-evaluation is difficult, but there seem to have been no major problems.

8. Instructions to Candidates

Instructions to candidates took three forms. Firstly there was a Candidate Instruction Booklet. This summarized instructions in the actual test and provided some pre-test examples. Candidates were also referred to the G-TELP web site.

The second level of instruction was within the test itself, before each set of test items. The tester tried to make these instructions as direct and unambiguous as possible. The single testee who finally participated reported no problems with the instructions, although he did omit a whole section of the listening test. I understand that this was an oversight.

The third level of instruction was informal and pre-test. It was more general in nature than the test-internal instructions, and was necessary since most potential candidates were participating remotely and had to be reached by e-mail. Technical instructions on how to handle downloads, e-mail attachments, and (especially) recording were necessary. I kept it simple, but I believe that one techno-allergic individual, 500km away, was scared off by the mechanics of participating electronically.

9. The Quality of Individual Test Items

The tester would have liked both explicit feedback and statistical evidence on the quality of individual test items. The single participant did not report any problems. Post-test, the tester did discover that one multiple choice item had been misprinted, rendering it invalid. Since no overall test score was being compiled, this was not critical.

10. The Representativeness of the Sample

The sample of English language embedded in the Practice Test is a tiny fragment of possible English language content, and of the many kinds of patterns found in systems of natural language. It does not contain patterns or content which are deviant from general standard English, but cannot claim to represent the totality of those patterns in any comprehensive manner. The tester has already argued that a comprehensive language test is impossible in principle. However, a useful language test is possible, and the sample of language in the Practice Test has already served a useful function of checking, elicitation and motivation, if we are to believe feedback from the individual tested.

The tester feels from his own experience that the single candidate's English has a great deal in common with that of many other students in East Asia, especially South Korea. However, the present context offers no statistical base for establishing this.

11. Reliability and Validity

The comparison between the G-TELP test and the Practice Test earlier in this paper emphasized that the Practice Test, applied to one individual, offered no statistical base for calculating any kind of technical reliability or validity. To the extent that the Practice Test mirrors the G-TELP test it is likely to be defensible. However the Practice Test is not assessing any overall competence or achievement in L2. It is attempting to pinpoint and diagnose problem areas for attention by students and their tutors. In those terms, we would need the largest possible sample of participating students, testing over a period of time under closely observed conditions, to draw any credible conclusions about the real value (reliability and validity) of the Practice Test.

12. Scoring Procedures

The discrete items in the Practice Test could be scored more or less objectively, and we could expect the same kind of inter-marker consistency which is found in tests like G-TELP.

The evaluation of factors such as discourse etiquette, register and discourse content which this tester has attempted on the single test paper (Appendix 1), could probably not be scored in a fair numerical manner with any consistency. Further, such an evaluation would be beyond the competence of most non-native speaking teachers of English in countries such as South Korea without extensive re-education. For this reason alone, the Practice Test could not be scored on a numerical scale of language achievement.

The international sub-culture of formal language testing would probably have to dismiss the Practice Test as a valid exercise because of its inherent inability to score the kind of features mentioned in the previous paragraph. If testing is about numbers, one can understand this. Pedagogically, the rejection would be nonsense. The problem is with the testing concept and its criteria of value.

13. Aspects of Moderation

The discrete point elements in the Practice Test would not require moderation unless extensive use of the test showed up an ambiguity in certain questions. The qualitative discourse features mentioned in the previous section would need extensive moderation if more than one tester used the document, and even then numerical scoring could be dubious.

Given the purpose of the Practice Test, the real 'moderation' would come when the tester changed hats to become advisor, and the testee(s) became individual learners negotiating and coming to understand their problems in post-test consultation.

The initial application of the Practice Test with one testee in a remote location leaves little scope for post-test moderation of course.

14. Summary Evaluation of the Test

The Practice Test, as based on G-TELP, was tentatively pitched at G-TELP's concept of an intermediate standard of English language

acquisition. G-TELP provides descriptors for its scales (see earlier assignment papers), but does not justify them in any rigorous way.

The tester has 30 years of experience in assessing how ESL learners are likely to perform in a variety of environments – social, academic and occupational. The criteria applied in such subjective assessment are too various, subtle and largely sub-conscious to easily make explicit; (this is characteristic of all kinds of expert knowledge). However, life experience gives a fair amount of confidence in this kind of assessment (though surprises are always possible).

One kind of evaluation of the Practice Test would be to ask how it squared with the tester's subjective assessment of the candidate's L2 abilities. If we take the Practice Test in G-TELP numerical scoring mode, I think that in some sections, especially the grammar section, it would have badly underestimated the candidate. For example, the grammar score would have predicted the candidate to be incapable of handling tertiary study in English without great difficulty. I have taught the candidate at graduate level in English, and although he didn't find it easy, he proved perfectly able to handle the content. One of his graduate essays is forwarded for comparison.

Taken on its own terms as a diagnostic and motivational tool, the single application of the Practice Test seems to have been fairly successful. It showed up some issues that could be usefully addressed by the testee and his teachers. In alerting the testee to potential problems, it has apparently motivated him to look for more imaginative solutions than his prior, narrowly academic experience of English.

Appendix 1 : Error Summary of Kim Joon-woo's Practice Test

Thor May, 2005

Examples of errors and infelicitous usage for each section of the Practice Test are pasted directly into this document. The problems identified in each section are then followed by a tester comment.

Student responses are in bold type. Corrections are highlighted in yellow.

1. Listening Test

(part B not done)

=> No errors in Part A

2. Reading Test

1. The quality tag on clothing tells you

- a) that the item was made in South Korea
- b) that the item has been checked for faults
- c) that the item is a good buy

Tester comment

In general the testee had no problem interpreting both factual and inferential information in the reading test. The error recorded above

may have stemmed from inferring from the word 'quality', rather than studying the context carefully.

3. Grammar Test

2. If we take care, smallpox _____ not kill large numbers of people again.

- a. would
- b. will have
- c. will
- d. have

5. Smallpox has now ceased to be a major health problem.

[Test construction error]

- a. is
- b. _____ [nothing]
- c. was
- d. has

10. Logical inquiry together _____ advanced technology is needed to solve many modern problems.

- a. through
- b. by
- c. with
- d. and

17. Governments can always gamble that there will not be

epidemics _____ not prepare for emergencies.

- a. but
- b. or
- c. and
- d. because of

18. Doctors can only treat the most urgent cases first,
_____ try to give quick help to all patients.

- a. and
- b. but
- c. or
- d. before

19. Helping remote communities is emotionally rewarding
_____ often costs doctors in lost income.

- a. although
- b. however
- c. and
- d. but

20. Mass vaccinations cannot be conducted often _____ the
high costs involved.

- a. when
 - b. because
 - c. because of
 - d. whether
-

Grammar Section B

	A	B	C
21	In medicine, prevention is	Important xx	than trying to cure sick people
22	Careful planning eventually costs	Higher xx	in suffering than ignoring medical risks
23	Smallpox was one of the	III Xx?	successes of modern medicine

Tester comment :

The structures tested in the grammar section of the Practice Test are taken directly from Level 3, G-TELP. Level 3 is claimed to indicate intermediate mastery of English. The particular choice of structures is not explained by the G-TELP committee. If it is supposed to indicate some universal stage in language acquisition, then I think they have a problem. However the G-TELP tests overall, covering all levels, must clearly draw from a menu of language structures which seem significant, and divide them up for distribution across the various level tests. If this is the case, then the correlation between structural mastery at a particular level (as checked by multiple choice questions) and functional mastery in live language contexts is probably quite arbitrary. There is then a problem of deciding what 'intermediate' mastery actually refers to.

The testee in the Practice Test is living evidence (I think) that a problem with the grammar test at Level 3 G-TELP is a poor indicator

of functional mastery in the target language. If we took his performance on these grammar questions as final evidence he would be relegated to Level 2 G-TELP, which is basic social English. I know from extended contact that this testee is perfectly capable of handling a tertiary level course in an English speaking country. His is better equipped linguistically than a large number of individuals who undertake that step. To manage tertiary level courses, a student minimally needs to control a language at about 3 on a scale of 5 (ASLPR, or American Foreign Service scale).

The focused selection of items at each level of G-TELP does have diagnostic value, and that is true in the present case. The testee has some problems with conditionals, with the logic implicit in conjunctions, and with the choice of comparative forms. At his stage of learning, a good teacher should be able to provide useful explanation and coaching in all of these items.

4. Writing Test

Dear, Jackson

Hello, Mr. Jackson. I'm Joonoo, a chancellor of ChungJu Nat'l Univ. The reason I'm sending you this e-mail is to inform you about pick-up time, date, meeting place...things like that. You are going to meet one of our staff at main gate of Incheon Airport tomorrow. Your flight will arrive at 2p.m. and one of our staff will pick you up

as soon as you show yourself up in the arrival gate. Don't worry, most of staff speak fluent English. By the way, my man will guide you to a very fancy car which is very expensive. And it will take one hour to get you to Chungju. Chungju is very beautiful city which is located in middle part of South Korea. Chungju is the home of unique cultures and arts with abundant national cultural treasures, natural beauties and historic things. Central Tower, and Goguryeo Monument are the centers of Jungwon culture, boasting its unique culture and long history.

If you come to come to our college, I'll tell you details about our contraction and travel information.

Sincerely,

Chancellor.

Kim, Joonoo

Tester comments :

Grammatical errors apart, most of the issues raised below are rarely or never tested in any formal way. Indeed, many of them are never taught. The problems of mass testing for things like register and discourse etiquette may be insurmountable, especially where large norm-referenced tests and statistical 'reliability' are paramount. Yet the illocutionary effect of most spoken and written communication is far more likely to be corrupted by non-tested, non-taught elements than by formally recognized 'errors'.

a) Formal grammatical errors :

- Overall management of the formal grammar by the testee is quite good, and would match that of many native English users, except for a couple of characteristic L2 problems.

- the writer has an unstable control of the determiners 'a' and 'the'. They are omitted in a number of places in the e-mail, and once 'a' is inserted inappropriately. These elements are often amongst the last mastered by ESL learners (quite often never mastered) : they are unstressed in speech, and their application is frequently subtle. The subject & topic particles in Korean, and the object particle, have some functions which are analogous to English determiners, but the match is very approximate.

b) Register (the language of social marking)

The short e-mail by this testee contains a fair number of register violations which could be serious in a business environment.

i) The erratic use of register

The erratic use of register is typical of Korean L2 users of English (including Korean professors of English). Korean itself makes constant use of register by manipulating the formal set of honorific markings on verbs. This is a highly structured process. Korean learners of English, and their Korean teachers, tend to believe that because register is not rigidly marked in English, it doesn't exist. Thus the control of English modals (for example) is often poor. ESL

texts in Korea frequently fail to indicate whether words and expressions should be used in a formal or informal context.

ii) Terms of address

(the special register rules for addressing other people)

These are extremely important within the Korean cultural and linguistic context. However, they are hardly taught for English usage, and almost universally misused. There is no understanding that the honorifics 'Mr, Mrs, Miss' always go with a surname in second person, but never with a given name. Korean colleagues, greeting me informally in English, almost always say 'Hello May'. Korean students almost always address a foreign teacher as "Miss Susan" or "Mr Peter"; (I tell them that this is slave talk from old American cotton plantations and British colonies ! Foreign teachers find it quaint, and rarely correct the problem) .

The salutation opening the testee's e-mail shows considerable confusion over naming terms and honorifics, and so does the structure of the signature block (with job title preceding the signature). The Korean conventions are quite different here, and clearly no one has taught the testee the standard English patterns.

iii) Discourse etiquette

(what to appropriately raise in a text, and how to express it)

This is something which I have never seen formally tested anywhere. However, it can undermine the whole effect of a communication. For example, it may be appropriate in a Korean business e-mail (I don't know) to say : " my man will guide you to a very fancy car which is

very expensive” . In English the whole message is reduced to farce, or worse. The testee’s description of cultural attractions is similarly inappropriate, but so typical for East Asia that I have highlighted it in blue; (it might have been lifted from a tourist brochure).

c) Information Content

The general purpose of the kind the business e-mail requested in the Practice Test is to enable a recipient to confirm known facts, check for unknown facts, and provide the kind of information necessary for forward planning. At least, that is the convention in Western business practice.

By Western standards, the testee’s e-mail (in my view) is a very weak document. We know that Mr Jackson will be met by an English speaker. A flight arrival time is given, as well as a (very incorrect) travel time to Chungju. The rest is blah. Mr Jackson could justifiably feel exasperated, and his exasperation is shared on a daily basis by foreign businessmen and other visitors with engagements in East Asia. Mr Jackson has no way to plan his itinerary in detail, no way to check for alternatives, and no one to call if things go wrong (as they very often do).

The kind of failure I have just indicated is, again, never tested formally in Korea, and the backwash consequence is that it is no part of any curriculum, let alone a language curriculum. Yet this kind of failure is a constant and major cause of friction in relations between Koreans (as well as Chinese and Japanese) and the wider world.

The problem has a great deal to do with culture, and the organization of East Asian workplaces. The Korean “vice chancellor” of the testee’s e-mail would expect his subordinates to take care of all actual details, and it might never occur to him that Mr Jackson would want to make personal judgements about such minor matters. He would also believe instinctively (in accordance with his culture) that revealing the absolute minimum of information was the best way to maintain some control over the foreigner’s actions. (I have worked in East Asia for six years. The pattern is very, very clear).

5. Speaking Test

Testee transcript

“Thousands of years ago, so many people were dead by suffering from fatal infectious disease called smallpox. Many people around the world were suffering painful disaster because of that. But in 1797, almost at the end of the eighteenth century, a vaccine for the infectious disease was discovered by a doctor named Edward Jenner. From that moment doctors from all around the world gathered their power to destroy this bad kind of disease. The vaccine was delivered to the infected in patients through injection and soon the disease was conquered (conquered ?) by cooperation of doctors from all around the world. It was the victory for the medical science. Death rate of smallpox was dramatically decreased to small numbers and overall situations were improved by 1900. Time passed by. In 1965 WHO, the World Health Organization, decided to provide a cure to all around the world throughout whole part of the countries. So, many patients who lived in isolated areas like Africa or Asia could get vaccine injection. After active movement of the

truer (WHO ?) project, vaccine for 200 million vaccinations was stored in Geneva and New Delhi.”

Tester comment

Note : Problem segments have been highlighted in yellow.

Dialect

This speaker uses a slightly Korean-modified form of American English speech. It is consistent, and close enough to the standard to be acceptable in most social situations. Those mispronunciations which do occur (e.g. conquered, Geneva) are likely to be the consequence of sight-reading vocabulary which the speaker cannot recall having heard before.

Fluency

The speaker is quite fluent. Because the test was self-administered the tester has no way of checking if the testee worked from his own notes (a list of cues was provided). The pauses, breaks and anacoluthon which are characteristic of spontaneous speech by both native and L2 speakers are strikingly absent. The intonation is markedly flat (shows little pitch variation), and this is a common property of reading aloud.

On the other hand, the use of repeating phrases (underlined in the transcript) is one of the strongest differentiators between spoken and written texts. The repetition is typically a product of real-time mental processing, and evidence of its use in speech has underpinned much of my own doctoral research on formulaic

utterance (e.g. see my paper “Postsupposition and Pastiche Talk” at <http://thormay.net/lxesl/tech2.html>) .

Formal grammatical errors

Much informal spoken speech by native speakers contains a significant number of grammatical slips, but they tend to be qualitatively different and more random than errors by L2 speakers. The testee in the present context is speaking in a formal and self-conscious mode. Those errors which do occur are therefore likely to be a consistent problem in his repertoire, rather than simple slips. Predictably, the speaker’s most obvious grammatical problem (and which also showed up in other sections of the Practice Test) is in the use of determiners. Fortunately, this kind of error rarely has global consequences for the meaning of an utterance, although it may trigger negative social judgements.

Lexical choice

Inappropriate lexical choice is the factor which most strongly marks the testee as a non-native speaker. (The same problem was evident in the Writing Test segment of this Practice Test).

Examples : /1. so many people / 2. this bad kind of disease / 3. throughout whole part of the countries / 4. After active movement /

These phrases are quite acceptable in other contexts (except for #3). That is the nub of the problem. This kind of thing is difficult to teach (there are countless such phrases), and difficult to test. The L2 learner can only really master this issue through continued extensive exposure to the target language (especially through reading).

Appendix 2 Post-test e-mail from the Candidate

Hello, Thor

This is Joonoo. Teacher, how have you been up to lately? I hope you are well also... ^^;;

My summer vacation has just started ^^;; (from today to september)

I read your comments about my test result. It was so specific than I expected....

I didn't know that I missed some parts in the test.....

I checked many errors of mine.....what a shame..... ^^;;

I think I have to read more books(of course in English version)

And I realized that I'm lack of practice in speaking and writing English language

This G-TELP typed test made me study English harder than my past days ^^;;

I think it was a good and worthwhile test for me to check my English ability more specifically....

Teacher, I really want to use English language like native speakers.

When I speak English in front of some foreigners, I feel that I'm using different style of English...which is very strange.... ^^;;

Near future.....or soon....I'm going to take a national test for teacher candidates...and..if I pass the test, I will be a mid or high school English teacher. But I think I'm in very serious situation.....

For better writing skill, I think 'reading a lot 'and 'writing a lot' is the best way....and I'm ready to do that.....but I have no confidence in speaking part.....since I have no native partner that I could talk with....or talk to

Prof. Thor, I'm 25 years old this year, do you see any possibility that I could reach native speaker level?(in English).....I thought that I'm quite good at English but now....I don't think so...hmm...

I think I have to study harder.....

Happy always.....

from your student

Joonoo

References - Test Specification

[hyperlinks listed here were current at 19 April 2005]

AWEMAP - A Worldwide ELT EFL ESL EAL LEP ESOL Assessment Scales and Tests Mapping Project <http://www.geocities.com/esolscale/index.html?200510>

Birt, Marina & Carol Kiparsky The Gooficon, pub. Newbury House 1972

G-TELP website http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp04.asp

Hughes, Arthur Testing for Language Teachers (2nd edition) , pub. UK: CUP 2003

Kitao, Kenji and S. Kathleen Kitao Language Testing Resource Page <http://ilc2.doshisha.ac.jp/users/kkitao/online/www/test.htm>

Moskovsky, Christo Newcastle University, NSW; MATESOL course notes and testing sample, 2005

Parry, K., "Building vocabulary through academic reading". TESOL Quarterly, 25 (4), 629-653. 1991

Pimsleur Audio-lingual Courses

<http://www.pimsleurapproach.com/learn-chinese-cantonese.asp?> ;

<http://www.pimsleurdirect.com/languages/sample//>

Power, Ted "Language testing & methods of assessment - What is test reliability and validity? What aspects are most important for the teacher/tester?" <http://www.btinternet.com/~ted.power/esl0736.html>

Spolsky, Bernard *Measured Words*, pub. UK: OUP 1995

References - Part IV : Practical Aspects of Language Test Construction

Asher, James 2004 "TPR After Forty Years" <http://www.tpr-world.com/japan-article.html>

Birt, Marina & Carol Kiparsky *The Gooficon*, pub. Newbury House 1972

G-TELP website http://gtelp.co.kr/e_gtelp/gtelp/e_gtelp04.asp

Hughes, Arthur *Testing for Language Teachers* (2nd edition) , pub. UK: CUP 2003

Kitao, Kenji and S. Kathleen Kitao Language Testing Resource Page <http://ilc2.doshisha.ac.jp/users/kkitao/online/www/test.htm>

Power, Ted "Language testing & methods of assessment - What is test reliability and validity? What aspects are most important for the teacher/tester?" <http://www.btinternet.com/~ted.power/esl0736.html>

Spolsky, Bernard *Measured Words*, pub. UK: OUP 1995

Professional bio: Thor May's PhD dissertation, Language Tangle, dealt with language teaching productivity. Thor has been teaching English to non-native speakers, training teachers and lecturing linguistics, since 1976. This work has taken him to seven countries in Oceania and East Asia, mostly with tertiary students, but with a couple of detours to teach secondary students and young children. He has trained teachers in Australia, Fiji and South Korea. In an earlier life, prior to becoming a teacher, he had a decade of drifting through unskilled jobs in Australia, New Zealand and finally England (after backpacking across Asia in 1972).

contact: <http://thormay.net> thormay AT yahoo.com

All opinions expressed in this paper are entirely those of the author, who has no aim to influence, proselytize or persuade others to a point of view. He is pleased if his writing generates reflection in readers, either for or against the sentiment of the argument.

"International Language Testing Washback– standing the monster on its head " © copyrighted to Thor May; all rights reserved 2013