STRATEGIES USED IN TAKING A COMPUTER-BASED SPEAKING TEST: A RETROSPECTIVE VERBAL PROTOCOL STUDY[1]

Sutthirak Sapsirin[2]
Kanchana Prapphal[3]
Lyle F. Bachman[4]

Abstract

This study investigated the strategies used by examinees when performing a computer-based speaking test. The data were collected from nine university students through a retrospective interview. The results showed that the examinees used several strategies, ranging from goal setting, assessment, planning to communication strategies. These strategies seemed to be associated with the constructs the CBST aimed to measure. It can be concluded that the findings provide evidence that supports the validity of the CBST score interpretations.

[1] This study is part of the first author's doctoral dissertation.

[2] Instructor, Language Institute, Chulalongkorn University, Bangkok, Thailand
[3] Professor, Language Institute, Chulalongkorn University, Bangkok, Thailand
[4] Professor, Department of Applied Linguistics & TESL, University of California, Los Angeles, USA

Introduction

A speaking test can currently be delivered via computers with various sources of input such as text, visual and audio (Flewelling and Snider 2001, Kenyon and Malabonga 2001). The characteristics of these computer-based speaking tests (CBSTs) differ to a great extent from other types of oral proficiency tests such as face-to-face oral interviews and group oral discussions. The major differences lie in the absence of interlocutors with whom the examinee interacts. In a computer delivered format, examinees usually respond to preset questions presented with visual and/or audio input and record their answers which will be rated afterwards.

As test method facets have long been acknowledged as one of the factors that influence language test performance (Bachman 1990, Chalhoub-Deville 1996, Skehan 1998), the implementation of such computer-based speaking tests has raised concerns about the validity of score interpretations (Chapelle 2001, Norris 2001). This is because little is known about the extent to which the test methods of such tests may alter candidates' performance, which in turn may affect the assessment of their oral language ability. Therefore, there is a need to investigate the effects of these test methods in order to better understand what the test is actually measuring. This will also help determine the degree to which we can justify the score interpretations made from CBSTs.

Literature review

Strategies used in a speaking task

Strategies hypothesized to occur when a person is performing a speaking task may include metacognitive strategies (Bachman and Palmer 1996), and communication strategies (Færch and Kasper 1983). Metacognitive strategies or strategic competence are defined as "higher order executive processes that provide a cognitive management function in language use, as well as in other cognitive activities" (Bachman and Palmer 1996: 70). In carrying out a language as well as non-language activity, these metacognitive strategies act as a mediator among topical knowledge, language knowledge, personal characteristics and affective schemata as well as between these components and the features of language use and setting. The areas in which metacognitive strategies operate include goal setting, assessment and planning.

Goal setting involves establishing a communicative goal, that is, making a decision about what one intends to do by using language. This typically involves identifying one or more language use tasks, choosing a task if choices are provided and deciding whether to carry out the task(s) or not. Assessment involves assessing the characteristics of the language use setting, determining the topical knowledge, affective schemata, and areas of language knowledge that will be needed to complete the language use task. Planning involves selecting the components of language knowledge, topical knowledge, and affective schemata to be used in a plan, formulating and selecting a plan to achieve the task successfully and evaluating the effect of one's utterance in achieving the intended communicative goal.

Communication strategies are "potentially conscious plans for solving what to an individual presents itself as a problem in reaching a particular communicative goal" (Færch and Kasper 1983: 36). Communication strategies are used when communication is problematic and they occur internally, without the interlocutor's help.

According to Cohen and Dornyei (2002), communication strategies can be classified into four major groups: avoidance or reduction strategies; achievement or compensatory strategies; stalling or time-gaining strategies; and interactional strategies. First, avoidance or reduction strategies include message abandonment, topic avoidance, and message replacement. Message abandonment occurs when the speaker does not finish a message due to some language difficulty. Topic avoidance may be used when the speaker avoids some topic areas which s/he has difficulty in talking about. Message replacement is a strategy of substituting the original message with a new one when the speaker feels incapable of delivering it.

The second group of communication strategies is achievement or compensatory strategies which consist of several subcategories. Circumlocution is describing words the speaker cannot remember. Approximation refers to using an alternative form which is as close in meaning as the target word. The use of all-purpose word is another strategy in which the speaker uses a general 'empty' word when s/he cannot think of the specific word. Word-coinage occurs when the speaker creates a new L2 word based on a supposed rule. The speaker may also use non-linguistic means such as mime, gestures and facial expression. The next subcategory is literal translation in which the speaker translates literally from L1 to L2. Foreignizing is the

case when the speaker uses an L1 word but with L2 phonology and/or morphology. Finally, code switching is including L1 words in L1 pronunciation when producing L2 speech.

The next group of communication strategies is stalling or time-gaining strategies. This consists of the use of fillers or other hesitation devices and repetition. In the first subcategory, the speaker uses fillers to gain time to think. As for the second subcategory, the speaker repeats a word or expression right after it was said.

Finally, interactional strategies comprise various subcategories. First, the speaker may appeal for help from his/her conversation partner directly or indirectly. S/he may also ask for repetition from the conversation partner when s/he does not understand. The speaker may ask for clarification, that is, requesting explanation of something s/he does not understand. Asking for confirmation is another subcategory in which the speaker requests confirmation that s/he understood something correctly. Expressing non-understanding refers to expressing that the speaker does not understand something. Finally, interpretive summary may be used when the speaker paraphrases what the conversation partner said to check that s/he understood it correctly.

Few studies have investigated speaking strategies which involve both metacognitive and communication strategies. For example, Cohen and Olshtain (1993) studied the strategies in speech act formulation. The participants, 15 advanced English language learners, role-played with a native speaker of English in six speech act situations, that is, two of apologies, complaints, and requests. It was found that the learners half of the time made a general assessment of the utterances they would use, but did not plan specific vocabulary or grammatical structures. They used English or their first language in planning and executing the utterances. They also employed a variety of strategies in searching for, retrieving, or selecting language forms. Finally, they did not pay much attention to grammar or to pronunciation. Another study by Cohen, Weaver and Li, (1998) (cited in Cohen 2000) examined strategies used in monologic tasks: self-description, story retelling and city description in two groups of learners: a control group (N = 23) and a group receiving strategies-based instructional treatment (N = 32). One of the purposes of the study was to explore the rationale for the responses of the sub sample of participants to the post-test Strategy Checklist. The results showed that in the experimental group at least one participant practiced a particular response several times before recording it. A student also reported avoiding using words he was uncomfortable with; one paraphrased when not remembering a word; and another spoke without pauses to sound more fluent. On the other hand, in the control group, one student reported translating into his native language before responding. One thought it would be ‘cheating’ to write down the response before speaking, and another showed his frustration at his limited language abilities.

These previous studies have illustrated a variety of speaking strategies that were employed in a speaking task. Thus, in the present study, the strategies investigated were those hypothesized to occur when a participant was performing an oral language task. The working definition of strategies, then, includes metacognitive strategies (Bachman and Palmer 1996), and communication strategies (e.g. Færch and Kasper 1983, Cohen and Olshtain 1993).

Verbal protocol analysis

Although quantitative analyses of test and item scores can provide important evidence for the validity of score-based interpretations, they may not yield information about the processes underlying test performance (Bachman 1990). Thus, in order to obtain such information, which can complement the quantitative analysis, the use of verbal protocol analysis has been recommended (Bachman 1990). Verbal protocol analysis is 'a methodology which is based on the assertion that an individual's verbalizations may be seen to be an accurate record of information that is (or has been) attended to as a particular task is (or has been) carried out' (Green 1998: 1–2). The informant usually either talks or thinks aloud what s/he does to complete a task. Verbal protocols can be categorized into three types: self-report, self-observation and self-revelation (Cohen 1984). Self-reports are learners' description of generalized language behavior. Self-observation is the inspection of specific and contextualized language behavior. This can be reported by the learner either introspectively or retrospectively. Finally, self-revelation is think-aloud, the description of thought processes while the information is being heeded.

Verbal protocol analysis has been used increasingly in cognitive psychology, educational psychology, psychology of assessment, cognitive science, and social psychology (Green 1998). As noted earlier, in the area of language testing, verbal protocol analysis has been used to supplement data obtained from quantitative techniques to investigate the validity of assessment methods and reliability of judgments (Bachman 1990, Banerjee and Luoma, 1997, Cohen 2000, Green 1998). For example, verbal report data have been used to explore the processes and strategies examinees use to complete a reading comprehension task (Anderson, Bachman, Perkins and Cohen 1991, Cohen 1984) to investigate what a cloze test measures (Storey 1997) and to compare the strategies used in first and second language reading comprehension tests (Nevo 1989). This technique has also been employed to analyze the sequence of rating, the interpretations the raters made while using an analytic rating scale, and the difficulties raters had in rating written language performance (Lumley 2002). Finally, insights into strategies in oral production tasks have been provided through verbal reports in studies by Cohen and Olshtain (1993), Cohen (1998) and Cohen, Weaver and Li, (1998) (cited in Cohen 2000).

Verbal protocol analysis can yield data on the cognitive processes of informants; however, to ensure that the technique and the data it provides are valid and reliable, researchers should be aware of the following aspects (Green 1998). First, the informant should be trained to give a protocol report beforehand. Appropriate instructions should be used to guide the production of the verbal report. Moreover, the informant should be encouraged to express her or his thoughts, but not to rationalize them. During the report, the researcher should not intervene in the process but should only encourage the informant to keep speaking. Finally, the report should be conducted while the informant is taking the test because a delay may introduce errors in the data.

In the assessment of speaking, little research has been conducted using verbal protocols. However, Green (1998) suggests that a retrospective reporting technique should be applied in this situation. This is

because an examinee has to converse with or without another interlocutor when performing a speaking test. To avoid disrupting this process, a retrospective report is recommended. Also, it should be conducted very soon after the test.

Verbal protocols can yield data on what the examinees actually do when performing a test, which can indicate what the test measures. This led to the research question of the present study: What strategies do the examinees employ in taking the CBST?

Methodology

Participants

Participants in this study were nine first-year university students in a large university in Thailand. The criteria for selection were based on students' university Test of English Proficiency scores and the faculties they were studying in so that they would represent students with various English language proficiency levels and academic fields. Their scores from the Test of English Proficiency, which included tests of Listening, Structure and Reading Comprehension, showed that their general English proficiency levels differed. The mean score of 4,969 first-year university students in the same academic year as these nine students was 473.34 and the standard deviation was 49.79 (the maximum value was 672, and the minimum value was 367). Three of the study's participants received scores between +1S and +2S from the mean; three between -.5S and .5S from the mean; and three between -1S to -2S from the mean. In terms of academic fields, an attempt was made to select participants from the full range of academic disciplines as classified by the university: biological science, social science, physical science and technology, and humanities. However, since the students' class schedules varied and were in conflict, only those from the following faculties were eventually contacted: the Faculty of Commerce and Accountancy, Economics, Education, Law, Science and Allied Science. These faculties were generally in the social sciences and sciences. In each faculty, students whose scores fell in the three proficiency groups were identified. They were approached and asked for their participation in the research study. Nine participants agreed to participate. Within each proficiency level, one was from the Faculty of Science or Allied Science, one from the Faculty of Education or Law, and one from the Faculty of Commerce and Accountancy or Economics.

Instrumentation

The Computer-Based Speaking Test (CBST) was developed by the researcher based on the framework of test development described by Bachman and Palmer (1996) and that for computer-based test design described by Fulcher (2003). The CBST is a multi-media English oral proficiency test delivered by a computer that is intended to elicit and measure knowledge of pronunciation, syntax, vocabulary, cohesion and functions. The computer presented test instructions and test tasks, controlled the preparation and response time, and stored participants' responses. There was both text and audio input. The participants wore headsets and spoke into microphones. The test was administered in a language lab by the researcher and lasted approximately 25 minutes. There were four test tasks, representing four different task types: narrating, expressing opinions, describing an imaginary ideal world, and persuading. In each task, the participants had two minutes to prepare their answers and one and a half minutes to respond. During the

preparation time, pieces of paper were provided so the participants could note down their ideas.

After being developed, the test was evaluated by three content experts and tried out with participants similar to those for whom it was intended. Based on this trial, some changes were made; the latest version of the test is presented in Appendix A. The reliability of the test, as estimated by Cronbach's coefficient alpha from the scores of 158 first-year university students in the same academic year with the nine participants, was 94.

Procedure

The procedure for collecting the verbal protocol data was tried out and revised in a pilot study. Initially, the procedure was adapted from the method used by Cohen & Olshtain (1993), and O'Loughlin (2001) in which the participants were asked some preset questions. Another method that was tried out was a retrospective self-observation in which the participants were trained to give a verbal protocol with no follow-up questions from the researcher. On the basis of trying out these two methods, the final procedure combined both methods and a retrospective interview was used.

Before the retrospective interview was conducted, the participants were told the purpose of the session, given a description of the CBST, and informed that their participation was voluntary, and that their responses would be treated confidentially. They were next trained how to give the verbal reports and did some practice tasks which were adapted from Ericsson and Simon (1993) and Green (1998). (See Appendix B for the procedures used for collecting the verbal protocols.). After doing the practice tasks, the participants could ask questions if they did not understand the procedures. After they said they were ready, the session began with the administration of the CBST. The participants completed the test individually in a computer laboratory. In addition to being recorded by the computer, their responses to the test tasks were audio recorded so that they could be played back during the follow-up interview.

Upon completing each task, test takers gave their verbal reports on the strategies they used in Thai and these were audio recorded. They were asked to begin by describing their thoughts during the preparation time and then to describe those during the response time. In order to help them recall their thoughts, the participants could review their notes, listen to the audio recording of their own responses to the test task, and stop their response tapes at any time. At the beginning of the report, the participants were asked to start talking about the first thing they thought of. They were encouraged to continue talking when they were silent. When the researcher speculated that there might be some missing information, she would help them recall by asking them to review these points. The researcher also asked them to clarify their thoughts if these were unclear. At the end of each report, the researcher asked the fixed questions regarding the participants' thoughts about English grammar, vocabulary and pronunciation. After they finished reporting their thoughts for the first task, they then went on to do the second task, reporting their thoughts and so on. The verbal reports lasted from 45 minutes to two hours, depending on how long the participants provided the reports.

Data Analysis

The audio-recorded verbal reports were transcribed by the researcher and then coded according to taxonomies of speaking strategies to deal with communication problems as suggested by previous research in strategic competence (Bachman & Palmer 1996) and communication strategies (e.g., Cohen & Olshtain 1993, Dörnyei 1995, Færch & Kasper 1983) and by the data from the pilot study. To try out the coding scheme, a part of the verbal reports from three participants, each from a different proficiency group, was selected. These reports were analyzed by the researcher and another coder, an experienced applied linguistics researcher and lecturer. The two coders worked independently in segmenting and coding the data. First, the data were segmented into words, phrases or sentences, each of which represented a distinct process or strategy. Then, they were assigned a code or category related to strategies used in taking the test. After that, the coders discussed the results. In general, they agreed on the coding but there were also some limitations and discrepancies. The coding scheme was then revised to include more taxonomies which were found in this tryout. (See Appendix C for the coding scheme that was used.) As for the discrepancies, some were not resolved between the coders. These data could be interpreted in more than one way because the context may not be specific or may be too little to pinpoint the exact code. Thus, they had to be left with two interpretations.

To investigate the consistency of the coding, the rest of the reports from the three participants were coded independently by the coders using the revised scheme. There were a total of 273 segments. The coders' agreement on identifying the segments was 254 segments (93.04%). Of the 254 segments, the coders' agreement on coding was 178 segments (70.08%) and the disagreement was 76 segments (29.92%). The discrepancies may be due to the lack or inadequacy of the context in the verbal reports, leading to more than one interpretation of the strategies the participant was employing. Thus, they were disregarded.

Results

The results indicate the strategies the participants reported using while preparing for and responding to the CBST tasks. They are organized by strategies. The transcripts which represented exact verbatim produced by the participants were translated from Thai to English by the researcher. In the transcripts below, the participants are identified as "P1" to "P9", while the letter "R" refers to the researcher. The abbreviations for the test tasks are "Nar", "Op", "Im", and "Per" for the narrative, opinion, imaginary and persuasive tasks, respectively. When there were more than one strategy in the excerpt, the italics correspond to that being discussed. The data in parentheses are information provided by the researcher to make the transcripts more understandable.

The main strategies that were found included goal setting, assessment, and planning. Other strategies used to cope with communication problems were also reported.

Metacognitive strategies

Goal Setting

Goal setting involves identifying the test tasks, choosing one or more tasks to do when given a chance, and making a decision whether to attempt to complete the task one has chosen. The verbal protocol analysis showed that none of the

participants reported using the first two processes. For the last one, only one examinee (P2) said that when she was doing the imaginary task, she did not understand the prompt "what this imaginary ideal world would be like in general", so she decided not to answer this part. Her report is presented below.

Im R What were you thinking at this point?
P2 I was rereading the prompt.
R The first prompt?
P2 (I was thinking) what it wanted me to do. *I didn't understand so I stopped.*
R Did you say anything else after that?
P2 No.
R So your answer ended here?
P2 Yes.

Assessment

Assessment includes assessing the characteristics of the task, assessing one's own topical and language knowledge as well as affective schemata, and assessing the correctness or appropriateness of their responses.

Assessing the characteristics of the task

Most participants reported that they assessed the characteristics of the test tasks to understand what the task requirements were. Some used the translation provided on the screen to understand the prompt. Example:

Op P4: During the preparation time which lasted 2 minutes, I read the questions to see what they were. Then, I looked at the translations. When I understood the question, I started thinking about the answer.

Another example showed that P3 understood that the situation in the narrative task involved talking to a close friend, so she decided to use the phrase "long time no see" in her response to make it correspond to the task situation.

Nar P3: It said that "(imagine that) I am your close friend" so I put "long time no see" to make it like she was a friend I haven't seen for a long time.

Other examples from the imaginary and persuasive tasks revealed that some participants paid attention to the task requirements before performing them.

Im P9: (During the preparation time) I was thinking what the main issue of this task was. When I looked at the word 'world', I guessed it may be about the world in my imagination, something like that.

Per P5: I started to think of greeting words. I wanted to make it match the situation that I am an invited guest (to a TV show).

Assessing one's own topical and language knowledge as well as affective schemata

In this stage, speakers assessed themselves to see if topical and language knowledge related to the tasks was available for them. They also evaluated their affective schemata or feelings toward the tasks they were performing.

Assessing topical knowledge and affective schemata

The data showed that this process was employed when the participants evaluated their knowledge to see if they had related world knowledge or similar experience to the task requirements. A few participants mentioned that they were aware that they had previous experience or thoughts about the narrative and opinion tasks. In this case, the knowledge about these topics was already available for them. For example, P1 said that she had the most enjoyable day on campus so it was easy to think about the story.

Nar P1: Luckily, there was a real event like this that happened to me. So it was easy to think about the story.

P3 mentioned that she had thought about studying for a Bachelor's degree abroad before, so when she was doing the opinion task, she recalled this thought.

Op P3: In this task I didn't have to think much because *it happened to me last year. It was about a scholarship. I was deciding whether to get a scholarship to study abroad, or to study in Thailand* because I just got accepted into this university. So I could think of the content quickly. *The part that they asked whether it was a good idea or not. I used to think about it. I used to think about it a long time ago.*

Assessing language knowledge and affective schemata

A few participants evaluated their ability in speaking English when they did the narrative and imaginary tasks, and expressed their feelings about the experience of taking the CBST. One participant (P1) felt comfortable and confident that she would be able to do a part of the test even though she didn't have enough time to prepare.

Nar P1: The (preparation) time was running out but this part "how you felt about it" *I felt that I would probably be able to say it* that I felt a bit sad, disappointed. But I didn't have time to note it down so I would just say it (without preparation).

Another participant reported feeling worried when doing the imaginary task. She felt that she might not be able to do it because the vocabulary appropriate to the task may be formal and difficult to use.

> Im P6: The first thing I thought of was the vocabulary must be difficult. I would not be able to think of the words.

She further explained that "I felt that the vocabulary must be at another level, not like those used in telling stories".

Another examinee said that she felt depressed about her English language ability after doing the narrative task.

> Nar P9: *I felt that my English was very bad.* (I thought) Why do I feel I don't know much about English? If I don't study more, I will keep feeling like this. *I felt depressed.* I couldn't think of anything. It's like I had known these vocabulary before. *I should have been able to tell the story.*

Assessing the correctness or appropriateness of the responses

This process included monitoring and evaluating the correctness or appropriateness of one's answers to the test questions. It may occur before, during or after responding. All nine participants reported that they evaluated their own responses in terms of content, grammar and pronunciation.

For example, P7 was assessing what the next content should be to make the response appropriate.

> Per P7: When I finished speaking this part, I was deciding whether to continue with the reasons (to visit Thailand) or start talking about the example (of places to visit or activities to do).

Another example of assessing the appropriateness of the content was from P6. She stated that she used *Phuket*, *Panga*, and *Koh Chang* as examples of tourist attractions in the South in the persuasive task. After she gave these three examples, a thought came to her mind, "At this point I thought 'is Koh Chang in the South?'"

In addition to content, the participants monitored and evaluated their grammar and pronunciation. For example, P5 was concerned about the verb tense and form.

> Im P5: I was worried about the verb tense, whether to use 'is' or 'was' or what.

Or in another instance:

> Im P5: I didn't know whether I should use 'has' or 'have' with the word 'population'.

Another participant was concerned about the structure of an utterance:

> Op P2: In this part I thought whether 'because' and 'when' can be put next

to each other. Later I felt it didn't matter because it was speaking.

Some participants monitored their pronunciation:

Nar P8: I was thinking about how to say 'Saturday 14th'. I must have said it wrong earlier.

Although all the participants reported monitoring their speech, they did not do so in all tasks. For example, P4 and P7 mentioned that they did not pay attention to grammar or pronunciation in some tasks.

Nar P4: I didn't care about my grammar. I just wanted to give the answer, just to answer before the time was running out.

Per P7: In this task I didn't think much about grammar. I thought about the vocabulary instead. For the pronunciation it was automatic. I didn't pay specific attention to it.

Planning

The third stage, planning, included selecting elements from topical and language knowledge to be used in a plan, formulating a plan, and selecting a plan. Planning also included the strategies the participants employed when faced with difficulties in communication, for example, topic avoidance, formal avoidance, and paraphrase.

Selecting elements from topical and language knowledge to be used in a plan

The verbal reports revealed that all participants used this process when performing the tasks.

Selecting topical knowledge

Participants decided to choose some specific elements from their previous experience or world knowledge to be used in their answers. The elements they selected were, for example, a specific story that happened to them (the narrative task), knowledge about the world environment (the imaginary task), previous experience in giving the answer to the opinion task, and knowledge about Thai culture (the persuasive task). The excerpts below illustrate this point.

P3 selected a specific experience she had earlier as her story in the narrative task:

Nar P3: I spent a lot of time thinking which day to talk about because there were a lot of (enjoyable) days but they were not quite remarkable. *So I chose the thank-you party for the seniors* because it was the most recent and remarkable.

P5 referred to her knowledge of the world and used it in her response to the imaginary task.

Im P5: After that I was thinking about the current environment: "Now there is a lot of pollution. There should

be some trees."

P2 explained how she came up with the answer to the opinion task. She thought about her past experience and selected to answer this task the same way she did before.

Op P2: I was thinking that during my high school, a teacher once asked me to answer a question in English. The question was whether I would like to study abroad, something like that. My answer was "no."

Finally, P1 used her world knowledge to create the answer to the persuasive task:

Per P1: I was thinking about the Thai culture that was beautiful and well-known.

Selecting language knowledge

All participants mentioned that they selected some elements from their language knowledge to be used in a plan. These elements included knowledge of syntax, vocabulary, and cohesion. For example, some stated that they were selecting specific verb tenses and modals to be used.

Examples:

Op P7: I was thinking about the tense I should use.
R: In this task, which tense did you think of?
P7: The future and the present tense.

Im P8: I was thinking about the grammar, like "my ideal world would be".
R: Which part that you were thinking about the grammar?
P8: The word 'would be'. I remembered that for an imaginary situation 'would be' should be used.

Some participants reported selecting specific words to be included in their responses. The examples are:

Nar P1: I used this phrase 'and you know' to make it like the way I talk to friends. So I added it (in my talk).

Per P3: I was thinking of words used in the advertisement about visiting Thailand like 'world with smile' so *I copied these words and used them in my talk.*

Some participants stated that they deliberately selected only simple or common words in their responses.

Example:

Im P4: I wrote down simple words 'green' and 'clean'. I was trying not to use difficult words because I couldn't

think of one.

In addition to content words, some participants were selecting which function words to link their utterances:

> Per P5: I was thinking of the word used after that (previous utterance). I thought if I spoke without this word, it would sound strange. I should find a word, a conjunction. The word which meant 'in addition to' or something like that.
>
> Op P7: I was thinking about the conjunctions, for example, 'owing to', 'due to', 'by the way', something like that.

Two participants noted that they took some language forms found in the test prompt and used these in their talk. For example, P2 said she copied 'good idea' from the prompt. P4 was thinking about the English word for 'abroad'. He felt that there may be many words which could express this meaning. He decided to use 'abroad' because it was already present in the prompt.

Formulating a plan

After selecting which elements of the topical and language knowledge would be used, the participants formulated a plan in which they made a decision about how to most successfully combine these elements to form a response. The plans they reported making involved content and language plans. For content planning, the participants planned the message or topic that would be talked about. For language planning, they reported formulating English utterances from the message. For example, some planned the organization of the whole talk and some the word order in an utterance. Examples of making a content and language plan are presented below.

Making a content plan

An excerpt from P5 showed how she planned the reasons why foreigners should visit Thailand.

> Per P5: I was thinking about the things in Thailand that foreigners liked. I started writing the note because I was afraid I would forget. At first, I didn't think about Thai people (as one of the reasons). I was thinking of nature or Thai food. After thinking a while, I thought the characteristics of Thai people would be more prominent so *I combined these two (nature and food) and made the reason about Thai people a separate one.*

Another participant explained how she planned the organization of the story she wanted to talk about:

Nar P7: I was thinking of how to tell the story in English and about the story. (I was thinking of) the two things at the same time. *I was thinking which part of the story I should talk about first. What we did first and what was next.* At the same time I was thinking of how to put it in English.

Similarly, P1 reported her thoughts about how she planned the organization of the story.

Nar P1: I would start with telling when the story happened.

Making a language plan

The language plans found in the verbal reports involved planning effective pronunciation, vocabulary, word order, and English utterances that could express the idea the participant had in mind. For instance, P1 planned to make her talk convincing by delivering it with appropriate rhythm (the persuasive task).

Per P1: (I was thinking about) the presentation style. How I should say it to make it interesting. I don't want to talk like (speaking slowly, word by word) "Thailand-is-a-". I want to make it more exciting, interesting than that.

She also made a plan about the choice of word:

Per P1: I was thinking that if I put 'welcome to', I mean, if I could complete my talk like what I intended and then say "so welcome to Thailand", it would sound good. It would be like a persuasion, something like that.

In another example, P4 planned how to put words into an appropriate sentence. Two excerpts illustrated this point:

Nar P4: I was thinking that ok I would say what happened was a football competition. The Thai language for this part was 'football pra-pei-nee' (traditional football) but what is it in English? So I started thinking of each word. I know the word football but what is 'pra-pei-nee'. Ok, 'tradition'. Then I put them in an order and made a sentence.

Im P4: I was writing the answers for the sub questions 1 and 2. I was thinking of the content in Thai. Then *I thought of how to put it in*

English and wrote it down.

Selecting one plan

The third stage in planning is selecting one plan. The data analyses showed that after formulating the plan as shown earlier, the participants usually selected that plan as their response to the task. Thus, it may be seen that these two were not separable.

Communication Strategies

In addition to metacognitive strategies, the participants reported using several communication strategies when faced with a communication problem. There were two options they chose: formulating a plan to solve it or avoiding the problem. The plans to solve the problems included impromptu, paraphrase and direct translation. However, some gave up and avoided the difficult content or language forms; in other words, they employed the strategies of topic avoidance and formal avoidance, respectively.

Impromptu

There were times when some participants did not have enough time to prepare for their talk. Therefore, they had to plan out the content and language while speaking. This strategy, which we called impromptu, was used in all tasks. The examples are shown below.

> Op P7: The first thing I did was writing my idea in Thai. After that I would put it in English if there was some time left. If I couldn't finish this before the (preparation) time ended, *I would think of what to say and speak at the same time.*

Paraphrase

Paraphrase includes using description, circumlocutions or exemplifications to express an idea when the speaker could not find the right words. Examples are:

> Nar P1: I couldn't recall the word 'numb' so I said something like "I felt I couldn't control my body and had a stomach ache", something like that. I couldn't think of the word. I just described other symptoms that I had to make it as close as possible.

In this case, P1 described other symptoms that really happened to her to express the idea that she felt numb. In another example, P6 described the word 'beggar' as 'someone who sits under the bridge'.

> Im P6: I wanted to say 'beggar' but I couldn't think of the word so I was thinking that beggars usually sit under the pedestrian bridge so I said 'someone who sits under the bridge' because I couldn't recall the word.

In addition to describing, the speaker may paraphrase by giving examples. An example

was found in P4's excerpt. He used examples to describe the word 'scenery'.

> Per P4: I was thinking of words to use like the word 'scenery'. But after a while, I suddenly got stuck. So I gave examples instead. I used 'sea, mountain' instead.

Direct translation from Thai to English

There was one instance in which a participant used direct translation to help her continue the talk. She wanted to say 'senior' but she got stuck so she translated the Thai word 'ruen-pee' ('senior') as 'brother', which has a sense of 'someone older'.

> Nar P2: This word 'brother' I actually I wanted to say 'senior' but I didn't know what it was in English so I used the word 'brother' instead.

Topic avoidance

Some participants stated that they avoided talking about a particular topic or message when they had difficulties expressing those thoughts. For example, P1 did not know how to describe a department store as she planned so she did not include the description in her talk.

> Nar P1: I wanted to say that the (Siam) Paragon was a grand department store but I didn't know how to say it so *I decided not to talk about it.*

In another example, P5 thought about the environmental problems in the world but she was not able to express the idea. So she skipped using it in her talk.

> Im P5: I was thinking about the environment, the air. There were CFC, greenhouse effect but I didn't know how to say it *so I didn't mention it.*

Other examples from P7 and P8 showed they decided to avoid elaborating their idea because they got stuck.

> Op P7: I was thinking about what I had noted down earlier. It was about (studying abroad would give you) new experience. *I wanted to elaborate this idea but I didn't know what to say* so I started the second reason. I just cut it short.

> Per P8: After saying the phrase 'beautiful place' I was trying to elaborate it but I got stuck so I changed to talk about the first sub question. At first, I was planning to give some examples (of beautiful places) or elaborate how beautiful they were but I couldn't think of anything to say.

Formal avoidance

Some participants mentioned that they avoided pronouncing some words and replaced them with others they thought easier to use. For example, P2 avoided telling the time '5.30 pm' because it was difficult to pronounce.

> Nar P2: I was thinking of the time the event happened. It wasn't five o'clock. I felt lazy. Actually, I wanted to say 'five thirty' but it was difficult to say so I chose to say 'five o'clock' instead.

A similar example is from P1:

> Nar P1: I remembered seeing the word 'the Faculty of Pharmaceutical Science'. It was difficult to pronounce and I didn't know the correct pronunciation. So I said 'the pharmacist' instead.

In summary, the participants were engaged in several strategies when preparing for and doing the CBST. The strategies they reported included making a decision whether to attempt to complete the task; assessing the characteristics of the test tasks; assessing one's own topical and language knowledge as well as one's feelings; assessing the correctness or appropriateness of the responses; selecting specific elements from topical and language knowledge; and formulating and selecting a plan as a response to the tasks. Finally, they employed several communication strategies to deal with communication problems through the use of impromptu, paraphrase, direct translation, and topic and formal avoidance.

Discussion

The analyses of the participants' verbal reports show that they employed various strategies when doing the CBST. They went through goal setting, assessment and planning in which their topical and language knowledge were involved. Several strategies were used to solve problems in communication such as paraphrase and topic avoidance. These findings correspond to the metacognitive strategies proposed by Bachman and Palmer (1996) and previous research on communication strategies (e.g. Færch & Kasper 1983). Strategic competence, as defined by Bachman and Palmer (1996), consists of executive processes, or metacognitive strategies, which operate in language use in both non-test and test settings, as well as other cognitive activities. Communication strategies, on the other hand, have come into play when the speaker has difficulties in reaching a particular communicative goal (Færch and Kasper 1983). Similarly, they can be utilized in any language use situation. Thus, the presence of these strategies in the CBST may suggest that the tasks could create a situation in which the participants are allowed to demonstrate their oral language proficiency. As the verbal reports revealed, the participants had a particular communicative goal in mind when performing the tasks. They employed both their language, topical and strategic abilities to accomplish their goals.

It can be seen that not all components of strategic competence were found in the verbal protocols. Only one component in goal setting was reported, and the strategies of formulating and selecting a plan were

inseparable. The reason why all areas in the goal setting process did not appear in the verbal report may be due to the characteristics of the task prompts (Bachman & Palmer 1996). The test tasks did not allow the participants to choose the task to perform, and the tasks were clearly specified as four different items. Thus, the participants did not need to identify the tasks or choose which one to attempt. The reason why formulating and selecting a plan cannot be separated may be because of the time limitation. Since the participants had limited preparation time, they may not have been able to create several plans from which to choose. They may have had to decide to make only one plan to complete the tasks in time.

In addition, the strategies employed seem to be associated with the constructs the CBST aimed to measure. That is, the CBST was designed to elicit language performance from which we can draw inferences about the speaker's five areas of language knowledge: knowledge of pronunciation, syntax, vocabulary, cohesion and function. The verbal report analyses showed that the participants did make use of these types of knowledge as expected. Thus, their test scores may allow us to make inferences about their speaking ability in the five areas. In other words, the results seem to yield evidence that supports the construct validity of the CBST score interpretation.

Conclusion

The present study investigated the strategies that participants used when performing the CBST. To fulfill the task requirements, the participants revealed in their verbal reports that they employed not only their language knowledge but also their topical knowledge through the operation of their strategic competence. This seems to provide evidence to support the construct validity of the test-based interpretations, since the test tasks seem to have engaged the participants' language ability that the test was intended to measure.

The results of this study suggest some theoretical and practical implications for language testing. One theoretical implication is that examinees do apparently employ a variety of metacognitive and communication strategies while taking a test of speaking, thus supporting the notion that these are part of language ability. The practical implication for developers of speaking assessments is that they should consider the extent to which these are included in their definition of the construct to be assessed. More specifically, developers may need to include, as evidence for the validity of score-based interpretations, empirical research on the extent to which such strategies are or are not part of the construct to be assessed. In addition, the study illustrates a way that verbal protocol analysis may be used to collect evidence for the validity of score-based interpretations of language ability, and to help test developers and test users better understand what the test actually measures.

Limitations

One of the limitations of this study is that the number of the participants in the study was rather small. Therefore, the results of the study may not generalize beyond this group. The construct being investigated includes only five areas of language ability. Thus, it is recommended that further research should examine other types of oral language knowledge such as sociolinguistics knowledge and knowledge of language functions apart from those in this study. In addition, the tasks in this

study generated only monologic responses. Thus, there should be other studies which investigate the effects of computer-based speaking tests on the strategies that examinees use in other types of speech such as dialogues and group discussions. In terms of research methodology, the verbal protocol approach used could be further refined to better fit a variety of speaking tasks. For example, it would be interesting to see a different approach to elicit speaking strategies. Finally, the replication of this study with other groups of examinees with different proficiency levels, first languages, and so forth, would provide more general insights into the kinds of strategies examinees employ in speaking assessments, nature of oral language ability, and the extent to which one can claim the construct validity of an oral test.

Acknowledgements

The researchers would like to acknowledge the Royal Golden Jubilee PhD Program of the Thailand Research Fund for providing the grant to make this project possible.

References

Anderson, N. J., L. Bachman, K. Perkins, And A Cohen. 1991. An exploratory study into the construct validity of a reading comprehension test: Triangulation of data sources. *Language Testing* 8: 41–66.

Bachman, L. F. 1990. *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bachman, L. F. and A. S. Palmer. 1996. *Language testing in practice.* Oxford: Oxford University Press.

Banerjee, J. and S. Luoma. 1997. Qualitative approaches to test validation. In *Encyclopedia of language and education volume 7: Language testing and assessment*, edited by C. Clapham and D. Corson, pp. 275–287. Dordrecht: Kluwer Academic Publishers.

Chalhoub-Deville, M. 1996. Performance assessment and the components of oral construct across different tests and tater groups. In, *Studies in language testing 3: Performance testing, cognition, and assessment: Selected papers from the 15th Language Testing Research Colloquium, Cambridge and Arnhem*, edited by M. Milanovic and N. Saville, pp. 55–73. Cambridge: Cambridge University Press.

Chapelle, C. 2001. *Computer applications in second language acquisition: Foundations for teaching, testing, and research.* Cambridge: Cambridge University Press.

Cohen, A. D. 1984. On taking language tests: what the students report *Language Testing* 1: 70–81.

Cohen, A. D. 1998. *Strategies in learning and using a second language.* London: Longman.

Cohen, A. D. 2000. Exploring strategies in test taking: Fine-tuning verbal reports from respondents. In *Learner-directed assessment in ESL*, edited by G. Ekbatani and H. Pierson. pp. 127–150. London: Lawrence Erlbaum Associates.

Cohen, A. D. and Z. Dörnyei. 2002. *An introduction to applied linguistics.* London: Arnold.

Cohen, A. D. and E. Olshtain. 1993. The production of speech acts by EFL learners. *TESOL Quarterly* 27: 33-56.

Cohen, A. D., S.F. Weaver and T. Li. 1998. The impact of strategic-based instruction on speaking a foreign language. In *Strategies in learning and using a second language*, edited by A.D. Cohen, pp. 107-56. London: Longman.

Dörnyei, Zoltán. 1995. On the teachability of communication strategies. *TESOL Quarterly* 29: 55-85.

Ericsson, K. A. and H. Simon. 1993. *Protocol Analysis.* Cambridge, Mass: MIT Press.

Færch, C., and G. Kasper. 1983. Plans and strategies in foreign language communication. In *Strategies in interlanguage communication*, edited by C. Færch and G. Kasper, pp. 20-60. London: Longman.

Flewelling, J. and D. Snider. 2001. *CARLA,* Computer-assisted recorded language assessment: software for evaluating student oral responses. 25 January, 2001 *<http://www.icte.org/T01_Library/T01_166.PDF>*.

Green, A. 1998. *Verbal protocol analysis in language testing research: A handbook.* UK: Cambridge University Press.

Kenyon, D.M. and V. Malabonga. 2001. Comparing examinee attitudes towards computer-assisted and other oral proficiency assessments. *Language Learning & Technology* 5: 60-83.

Lumley, T. 2002. Assessment criteria in a large scale writing test: What do they really mean to the raters? *Language Testing* 19: 246–276.

Nevo, N. 1989. Test-taking strategies on a multiple-choice test of reading comprehension. *Language Testing* 6: 199–215.

Norris, J. M. 2001. Concerns with computerized adaptive oral proficiency assessment. *Language Learning & Technology* 5: 99–105.

O'Loughlin, K. 2001. *The equivalence of direct and semi-direct speaking tests.* Cambridge: Cambridge University Press.

Skehan, P. 1998. *A cognitive approach to language learning.* London: Oxford University Press.

Storey, P. 1997. Examining the test-taking process: A cognitive perspective on the discourse cloze test. *Language Testing* 14: 214–231.

Appendix

Appendix A: Computer-based speaking test

Instructions: This test consists of four parts and it will last approximately 25 minutes. In each part, you will speak for 1:30 minutes about a general topic. Before

you speak, you will have *two minutes* to prepare. During the preparation time, you can make some notes on the paper given if you wish.

Part 1

Imagine that I am **your close friend**. We haven't seen each other for some time. Today, we met and you want to tell me what has happened to you recently.

Situation: Talking to a friend

Topic: The most enjoyable day on campus

Please be sure to talk about:

- *when the story happened*
- *where it happened*
- *what happened first*
- *what happened next*
- *how it ended*
- *how you felt about it.*

Part 2

Imagine that I am **one of your friends**. I asked you what you think about studying for a Bachelor's degree abroad.

Situation: Talking to a friend

Topic: What do you think about studying for a Bachelor's degree abroad?

Please be sure to talk about:

- *whether you think studying for a Bachelor's degree abroad is a good idea or not*
- *reasons to support your opinion.*

Part 3

Imagine that you are invited to **a television show**. The host asks you to persuade the viewers to visit Thailand. The viewers are **foreigners**.

Situation: Persuading foreigners to visit Thailand

Topic: How would you persuade foreigners to spend their vacation in Thailand?

Please be sure to talk about:

- *at least two reasons why they should come to Thailand*
- *at least two examples of what they can do or see in Thailand.*

Part 4

Suppose that your teacher asked you to present your idea in class about an imaginary situation.

Situation: Talking in front of the class

Topic: What would your ideal world be like?

Please be sure to talk about:

- *what this imaginary ideal world would be like in general*
- *at least two characteristics of this ideal world*
- *the ways in which this world would be a better place for everyone.*

End of the test.

Thank you very much for your participation.

Appendix B: Procedures for verbal protocol collection in the main study

The instructions for the tutorial of the stimulated verbal report procedures were adapted from Ericsson and Simon (1993) and Green (1998). They were read to and given to the participants in print as follows:

In this study I am interested in your thoughts when you work on the tasks I am going to give you. To do this, I am going to ask you to do the first task. After you finish, I would like you to tell me all that you can remember about your thinking from the time you started the task until you completed it.

- Please tell me about your memories in the sequence in which they occurred while preparing and giving the answers. Please start your report saying "I first thought of…"
- You can talk in English, Thai or a mixture of the two.
- To help you remember your thoughts, you can look at the question and your notes that you took when you prepared for the answer. Also, you can listen to your response that will be recorded. You can play and pause it as you like.
- If you are uncertain about any of your memories, please let me know.
- I don't want you to work on the task again, just report all that you can remember thinking about from the time you read the question until you gave the answer. Also, don't plan out or try to explain to me why you thought in a certain way.
- If you are silent for any long period, I will ask you to continue your talk.
- Your talk will be audio recorded so please speak loudly.

Do you have any questions?
Now let's do some practice tasks.

Practice Task 1
Add these two numbers. After finishing the task, tell me all that you can remember about your thinking.

15 and 27

Practice Task 2
Judge whether two letters are in alphabetical order or not. After finishing the task, tell me all that you can remember about your thinking.

B S

Judge whether two letters are in alphabetical order or not. After finishing the task, tell me all that you can remember about your thinking.

M L

Practice Task 3
Answer the question below. You have 15 seconds to prepare your response and 30 seconds to answer. You can use the paper

provided. Your answer will be recorded. After finishing the task, tell me all that you can remember about your thinking when you prepared and gave the answer. You can use your notes and play back the tape to help refresh your memories.

What do you plan to do after you graduate from the university?

Now are you ready to start the first test task?

Appendix C: Coding scheme for the verbal reports

The coding scheme for the verbal reports consisted of several taxonomies which are based on the literature in the model of language ability by Bachman and Palmer (1996), and the research on communication strategies by, for example, Færch and

Kasper (1983), Cohen and Olshtain (1993)
and Dornyei (1995).

Taxonomies of speaking test taking strategies and processes	**Definition/description**
Metacognitive strategies	
1. Goal setting	Deciding what one is going to do
1.1 Identifying the test tasks	
1.2 Choosing one or more tasks to do when given a choice	
1.3 Deciding whether or not to attempt to complete the task(s)	
2. Assessment	Taking stock of what is needed, what one has to work with, and how well one has done
2.1 Assessing the characteristics of the test task	Identifying the characteristics of the test task to determine - the desirability and feasibility of successfully completing the task, and - what elements of topical knowledge and language knowledge likely to require the test taker to do
2.2 Assessing one's own topical knowledge, language knowledge and affective schemata	
2.2.1 Assessing one's own topical knowledge and affective schemata	Determining the degree of relevant topical knowledge that is available, and if available, which might be used. Also, determining one's affective schemata or feelings for coping with the task requirement.
2.2.2 Assessing one's own language knowledge and affective schemata	Determining the degree of relevant language knowledge that is available, and if available, which might be used. Also, determining one's affective schemata or feelings for coping with the task requirement.
2.3 Assessing the correctness or appropriateness of the response to the test task	The areas to be assessed include the grammatical, textual, functional and sociolinguistic features of the response and its topical content.
3. Planning	Deciding how to use what one has
3.1 Selecting specific elements from topical knowledge and language knowledge that will be used in a plan	

3.1.1 Selecting specific elements from topical knowledge that will be used in a plan	These elements include real world knowledge or personal experience.
3.1.2 Selecting specific elements from language knowledge that will be used in a plan	These elements include syntax, pronunciation, vocabulary, cohesion, organization, function and socio linguistic features.
3.2 Formulating a plan	
3.2.1 Making a content plan	Planning the topic or message in the talk
3.2.2 Making a language plan	Planning how to formulate English utterances from the message: syntax, pronunciation, vocabulary, cohesion, organization, function and sociolinguistics features
3.3 Selecting one plan	
Communication Strategies	
- Message abandonment	Leaving a message unfinished
- Topic avoidance	Avoiding topics or concepts
- Message replacement	Substituting the original message with a new one
- Code switching	Switching from L2 to L1 or another foreign language
- Paraphrase	Using descriptions, circumlocutions or exemplification
- Approximation	Using a more general word for an unknown word e.g. ship for sailboat
- Word coinage	Construction of a new interlanguage word
- Directly translating Thai words into English	Using an English word that is direct translation of the Thai word. Although the English word exists, its meaning is different from the actual intended meaning.
- Using prefabricated patterns/ formulaic speech	Using memorized stock phrases
- Foreignizing	Using an L1 word by adjusting it to L2 phonology and/or morphology
- Formal avoidance	Avoiding using rules/items in phonology, morphology, syntax and vocabulary.
- Impromptu	No planning of the content or language in advance
- Use of fillers	Using fillers to gain time to think
- Repetition	Repeating a word or expression right after it was said