A TAP Study of Thai Test-takers’ Processes of Taking a Chinese Vocabulary Test

Xiangdong Gu  PhD
Consultant, Cambridge English Language Assessment
Director/Professor, Chongqing University, China

Yuwen Shen
Assistant teacher, Chongqing University, China

11th April, 2014
Contents

- Think-Aloud Protocols
- Research Background
- Research Questions
- Research Methodology
- Findings and Discussions
- Conclusions
1. Think-Aloud Protocols (TAP)

1.1 Description of TAP

The participants in think-aloud studies must perform a particular task while continuously reporting whatever thoughts pass through their working memory. The verbalizations are recorded and then coded according to theory-driven and/or empirically driven categories.

(Ericsson 1988; Ericsson & Simon 1980, 1993)
1.2 Validity and Reliability of TAP

Validity

Information captured within verbal reports corresponds closely with Information actually heeded

Similar verbal protocols of the same task

Subject 1

Coder 1
Coder 2

Reliability

coding scheme accurately captures the behavior

Subject 1

Coder 1
Coder 2
2. Research Background

2.1 Our Teams’ Studies Using TAP


2.2 Brief Introduction to Hanyu Shuiping Kaoshi (HSK)

- A national standardized test launched in 1990 and revised in 2009

- To assess the Chinese language proficiency of non-native speakers (including foreigners, overseas Chinese and students from Chinese national minorities)
### 2.3 The Correspondence Between Levels of HSK and CEFR

<table>
<thead>
<tr>
<th>HSK (Revised Version)</th>
<th>Vocabulary</th>
<th>CEFR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Words &amp; Phrases</td>
<td>Characters</td>
</tr>
<tr>
<td>HSK-6</td>
<td>5000+</td>
<td>2663</td>
</tr>
<tr>
<td>HSK-5</td>
<td>2500</td>
<td>1685</td>
</tr>
<tr>
<td>HSK-4</td>
<td>1200</td>
<td>1064</td>
</tr>
<tr>
<td>HSK-3</td>
<td>600</td>
<td>617</td>
</tr>
<tr>
<td>HSK-2</td>
<td>300</td>
<td>347</td>
</tr>
<tr>
<td>HSK-1</td>
<td>150</td>
<td>174+</td>
</tr>
</tbody>
</table>

(see [http://www.hanban.edu.cn/tests/node_7486.html](http://www.hanban.edu.cn/tests/node_7486.html))
2.5 Brief Introduction to HSK-6 Vocabulary Test

- Second part of the reading test
- Altogether 10 short texts
- Each text with 3-5 blanks; each blank with four options
- Testees required to choose the most appropriate answer according to the context
3. Research Questions

- What types of information sources do test-takers use to answer vocabulary test of reading section in HSK-6? How do they use these information sources?

- What types of strategies do test-takers employ in their processes of taking vocabulary test? How do they use these strategies?

- Are there any differences among different proficient groups and in their use of information sources and strategies? If so, what are these differences?
4. Research Methodology

4.1 Research subjects

- **Subjects**
  
  12 **Thai** students taking the program of Master of Teaching Chinese to Speakers of Other Language in Chongqing University
## Three Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>Age</th>
<th>HSK Grade &amp; Score (total:300)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>M</td>
<td>27</td>
<td>HSK-6 (225)</td>
</tr>
<tr>
<td>H2</td>
<td>M</td>
<td>27</td>
<td>HSK-6 (218)</td>
</tr>
<tr>
<td>H3</td>
<td>F</td>
<td>28</td>
<td>HSK-6 (214)</td>
</tr>
<tr>
<td>H4</td>
<td>F</td>
<td>26</td>
<td>HSK-6 (206)</td>
</tr>
<tr>
<td>M1</td>
<td>M</td>
<td>25</td>
<td>HSK-6 (188)</td>
</tr>
<tr>
<td>M2</td>
<td>F</td>
<td>26</td>
<td>HSK-6 (192)</td>
</tr>
<tr>
<td>M3</td>
<td>F</td>
<td>24</td>
<td>HSK-5 (223)</td>
</tr>
<tr>
<td>M4</td>
<td>F</td>
<td>27</td>
<td>HSK-5 (221)</td>
</tr>
<tr>
<td>L1</td>
<td>M</td>
<td>25</td>
<td>HSK-5 (207)</td>
</tr>
<tr>
<td>L2</td>
<td>F</td>
<td>24</td>
<td>HSK-5 (183)</td>
</tr>
<tr>
<td>L3</td>
<td>F</td>
<td>26</td>
<td>HSK-4 (205)</td>
</tr>
<tr>
<td>L4</td>
<td>F</td>
<td>25</td>
<td>HSK-4 (211)</td>
</tr>
</tbody>
</table>
61. Compared with young people, the old tend to recall ____ . They will no longer ____ acquiring new knowledge but concentrate on their past experiences, and make a summary of their life-long ____.

A. legend, test, deeds  
B. the past, taste, time  
C. current situation, explore, career  
D. the past experience, try, years
4.3 Research Procedures

- **Instructional phase**
  - Objectives;
  - Written instructions; Training

- **Pilot phase**
  - Get familiar with the task;
  - Verbalization rehearsal

- **Simultaneous introspection phase**
  - Take the test;
  - Report simultaneously

- **Data transcription phase**
  - Literally transcribed;
  - Double check

- **Data analysis phase**
  - Units division;
  - Descriptive statistics;
  - Chi-square Tests

Intra-coder agreement 0.88; Inter-coder agreement 0.81
4.4 Analytical Framework

Information Sources

- Clause-level information
- Sentence-level information
- Text-level information
- Extra-textual information

Cognitive Strategies

- Inferencing
- Contextualization
- Reconstruction
- Repetition
- Planning
- Selective attention
- Comprehension monitoring
- Comprehension evaluation

Metacognitive Strategies

- Eliminating the options
- Matching the blank of an item with options
- Guessing blindly
- Missing

Test-wiseness Strategies

- Empirically driven (Gao & Gu 2008; Gu & Shi 2012)
5. Findings and Discussions

5.1 Information sources

Figure 1 Frequencies of information categories used by the three groups of test-takers
Clause-level information

- The high group used clause-level information in combination with sentence-level and text-level information in most cases.

Example 1 (H1)

Because when there are passengers in the car, the driver has bla bla. The driver has A. sign B. objective C. target D. focus. Not clear yet and just go ahead. Arriving at the crossroad, he hesitates to go left or right, hence awareness/attention/vitality/resolution is distracted. I think C is right. When passengers are in the car, the driver has a target and will focus on driving with caution against bla bla. However, when there are no passengers, he is impetuous. Arriving at the crossroad bla bla, his attention will be distracted.
• The low group tended to use clause-level information as the single source of information.

Example 2 (L3)

有乘客时，司机有…,目标，对象, 司机有对象，不合适，应该是司机有目标。

When there are passengers, the driver has a…target/objective. The driver has an objective. Not suitable. It should be that the driver has a target.
Sentence-level information

• The high group regarded the whole sentence as one unit.

Example 3 (H3)

团购就是团体购物，指的是认识的或者不认识的消费者，联合、结合、联络、合并起来，觉得是......联合起来，然后再看，来提升买家与商家的议价能力，只有消费者联合起来才能提升买家与商家的议价能力，所以选联合。

Tuangou is short for shopping in group. It refers to consumers known or unknown to combine/join/contact/merge together. I think the answer is “combine together”. Then, to enhance buyers’ bargaining ability with sellers. Only through combining customers together can enhance buyers’ bargaining ability with the sellers. So I choose “combine together”.

Example 4 (L1)

Tuangou is short for shopping in group. It refers to consumers known or unknown ...together. combine/join/contact/merge. Consumers known or unknown ...together. I think the answer is “combine together”, so I choose A.
Text-level information

• Every item of the vocabulary test consists of 2-3 sentences with 3-5 words deleted, i.e., a short text. This kind of text generally has centralized information. When test-takers can’t solve the problems with the obvious clause-level information, they tend to utilize the text-level information.

Example 5 (H1)

好，第68题，出租车在什么时候最危险？答案是没有乘客时。因为有乘客时，司机有什么什么，司机有A标志，B目标，C对象，D焦点。这个还不是很清楚，他会怎么怎么，他会，迫不及待，全神贯注，这是……我觉得B不对，与驾驶，什么什么，齐心协力，B想法设法，C小心什么什么，D千方百计，尽快到达目的地，而没有乘客时，他是什么什么的，他是茫然的，他是盲目的，他是抽动的，D他是急躁的，D不对，走到十字路口往左转右转犹豫不定，意识就被分散了，精力就被分散了，活力就被分散了，意志就被分散了。我觉得应该是C，因为有乘客时，司机有对象，他会专心致志于驾驶，小心什么什么的，而没有乘客时，他是冲动的，走到十字路口什么什么，活力就被分散了。
Example 5 (H1)

Well, item 68, When is it the most dangerous for a taxi? The answer is when no passengers are in the car. Because when there are passengers, the driver has bla bla. The driver has A. sign B. target c. objective D. focus. I am not sure yet. He will bla bla. He will, cannot wait for, concentrate on. This...I think B is incorrect, on driving, bla bla, make concerted efforts, B. try every means, C be cautious bla bla, D. make every endeavor to arrive at the destination as soon as possible. When there are no passengers, he is bla bla. He is confused/ blind/impetuous/irritable. D. He is irritable. D is wrong. Arriving at the crossroad, he hesitates to go left or right, awareness/attention/vitality/resolution is distracted. I think C is right. Because when passengers are present, the driver has a target and will focus on driving with caution against bla bla. However, when there are no passengers, he is impetuous. Arriving at the crossroad bla bla, his attention will be distracted.
Example 6 (M3)

Companies and people like this are too bla to make great progress. According to my experience and knowledge, stubbornness may not be good, and then they hardly make great progress. So D is correct.
5.2 Strategy use

Figure 2 Frequencies of strategy categories used by the three groups of test-takers
### Table 1 Cognitive Strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferencing</td>
<td>Infer or guess missing information or meaning of words or sentence</td>
<td>相对于年轻人来说，老年人更乐于什么什么，然后去看答案，A传说，B从前，C现状，D往事，这个呢，我觉得先排除A和C，因为老年人更乐于回忆，所以说，更可能说是以前的事情嘛，(Because the old are tend to recall, it is more likely to talk about the past events.) 所以啊，这个呢，应该选B或者选D，从前或者往事。</td>
</tr>
<tr>
<td>Contextualization</td>
<td>Relate new information to a wider familiar context</td>
<td>先有目标，他会，看B 的盲目，看不到路，也是看不到，也不知道该怎么做，后面也说了，(Which has been stated in the following part) 走到十字路口往往左转右转犹豫不定，不知道该往哪走，所以他的精力就被分散了。</td>
</tr>
<tr>
<td>Reconstruction</td>
<td>Recreate meaning using their own words according to their own understanding</td>
<td>然后我觉得焦点比目标好一些。这个意思就是说(That is to say) 把乘客当成焦点。然后会聚精会神。然后没有乘客时，精力会被分散。</td>
</tr>
<tr>
<td>Repetition</td>
<td>Repeat some words or sentences</td>
<td>是因为消费者 联络，联络(combine, combine) 起来，因为认识的或者不认识的得需要联络起来，以求得最优的一种购物手段，手段(means, means)？</td>
</tr>
</tbody>
</table>
Figure 3 Frequencies of cognitive strategies used by the three groups of test-takers
### Table 2 Metacognitive Strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Planning</strong></td>
<td>Preview contents in different forms; establish purpose for reading</td>
<td>相对于年轻人来说，老年人更乐于回忆什么，他们不再什么获得新的知识，而是专注于昔日的经历，为一生的什么做一个总结（一个读题的过程，明确目标）(A process of reading the text, in order to understand the requirments).</td>
</tr>
<tr>
<td><strong>Selective attention</strong></td>
<td>Notice specific aspects of reading materials; reading selectively according to purpose</td>
<td>第63题，我答B，是因为，这些选项里我只知道万一的意思(Of all the options I only know the meaning of in case)，把万一放进去，万一制定了基本方针，所以我就答B。</td>
</tr>
<tr>
<td><strong>Comprehension monitoring</strong></td>
<td>Monitor understanding while reading</td>
<td>老年人更乐于回忆，该是往事吧，这个意思我不太懂它的意思，先不做，往后猜(without choosing this blank and going on guessing the following ones)。</td>
</tr>
<tr>
<td><strong>Comprehension evaluation</strong></td>
<td>Check understanding for accuracy, completeness and acceptability</td>
<td>第63题，我选择也是D，因为我觉得，一旦制定了基本方针，就一步都不打算妥协，我觉得这样比较顺，别的选择不太顺，所以我选择D。再看其他的空(and then consider other blanks)，这样的公司及个人由于太固执，很难有大的发展。固执就是不太好的吧，就很难有大的发展。所以我确定D为正确答案(So I choose D as the right answer)。</td>
</tr>
</tbody>
</table>
Figure 4 Frequencies of **metacognitive strategies** used by the three groups of test-takers

- **Planning**: High Group (26), Medium Group (20), Low Group (40)
- **Selective Attention**: High Group (20), Medium Group (20), Low Group (18)
- **Comprehension monitoring**: High Group (40), Medium Group (30), Low Group (20)
- **Comprehension evaluation**: High Group (18), Medium Group (15), Low Group (12)
## Table 3 Test-wiseness Strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing</td>
<td>Do nothing about some blanks, relying the choices of other blanks</td>
<td>ABC最后的词我不知道什么意思，然后我再看，做生意时总会有一些企业或个人只知道坚持自己的立场，什么什么制定了基本方针，什么什么，后面有就，前面应该，可能万一，一旦，一旦……就，应该是一旦……就，后面就不管了(The following blanks will not be considered.)，选D。</td>
</tr>
<tr>
<td>Eliminating the options</td>
<td>Select an answer not because it is thought to be correct, but because the others do not seem reasonable or are not understandable</td>
<td>相对于年轻人来说，老年人更乐于什么什么，然后去看答案，A传说，B从前，C现状，D往事，这个呢，我觉得先排除A和C (In my opinion, A and C could be firstly get rid of.)，因为老年人更乐于回忆，</td>
</tr>
<tr>
<td>Matching the blank of an item with the options</td>
<td>Put the given options one by one in the blank in order to see whether it is suitable in the context</td>
<td>出租车在什么时候最危险？答案是没有乘客时。因为有乘客时，司机有什么什么，司机有A标志，B目标，C对象，D焦点(A. sign, B. destination, C. object, D. focus)。这个还不是很清楚，他会怎么怎么，他会，迫不及待，全神贯注，我觉得B不对，与驾驶，什么什么，齐心协力,B 想方设法，C小心什么什么，D 千方百计 (B. try every means to, C. be cautious about bla bla, D. make every effort to)。</td>
</tr>
<tr>
<td>Guessing blindly</td>
<td>Randomly choose from the given options without any particular consideration</td>
<td>这个63题，答案的词我不太明白，不知道它的意思，然后看起来，我觉得，额，它的答案就是C，这道题我就是猜的 (I worked out this question by mere guessing)</td>
</tr>
</tbody>
</table>
Figure 5 Frequencies of test-wiseness strategies used by the three groups of test-takers

- Missing: 33
- Matching the blank of an item with the options: 26
- Eliminating the options: 22
- Guessing blindly: 20

Chi-square = 13.358, Sig. = 0.038
6. Conclusions

6.1 Major Findings

Q1: *What types of information sources do test-takers use to answer vocabulary test of reading section in HSK-6? How do they use these information sources?*

- Test-takers employ all four categories of information sources, i.e. clause-level, sentence-level, text-level and extra-textual information, among which **clause-level information** is the most frequently used one, followed by text-level, sentence level and extra-textual information.
Q2: *What types of strategies do test-takers employ in their processes of taking vocabulary test? How do they use these strategies?*

- Test-takers employ **all three** strategy categories, i.e. cognitive, metacognitive and test-wiseness strategies.

- Among the four cognitive strategies, **inferencing and contextualization** are much more frequently used than reconstruction and repetition.

- In terms of the four metacognitive strategies, **comprehension monitoring** is most frequently used, followed by planning, selective attention and comprehension evaluation.

- As far as the four test-wiseness strategies are concerned, **missing** is used more frequently than matching the blank of an item with the options, eliminating the options, and guessing blindly.
Q3: Are there any **differences** among different proficiency groups in their use of information sources and strategies? If so, what are these differences?

- There exists some differences among different proficiency groups in their use of information sources and strategies, but **significant difference** only exists in the use of test-wisenedess strategies.

- The higher the proficiency of the group, the more information sources it uses, the more information from different levels it combines and the wider range of discourse it incorporates.

- The higher the proficiency of the group, the more cognitive and meta-cognitive strategies it uses. On the contrary, the lower the proficiency of the group, the more test-wisenedess strategies it uses.
6.2 Limitations

- The small item number of the instrument
- The small size of the participants
- The unrepresentativeness of the participants from a homogeneous cultural and linguistic background
- Thus the generalizability of the research findings remains to be further testified by triangulation methodology
6.3 Implications

- The text should be lengthened to provide more information and context for the test-takers.

- Care should be taken to avoid activating test-takers’ extra-textual information and to avoid their use of test-wise strategies to get the correct answers.

- More training and practice is needed for test-takers to improve their appropriate use of information sources and strategies.
6.4 Significance

➢ **Theoretically**, offered some insights into what the Chinese vocabulary test measures.

➢ **Practically**, put forward suggestions for the test designers on test development and test evaluation, and provided implications for language teachers and learners on how to improve Chinese language proficiency.

➢ **Methodologically**, proved think-aloud protocols as an effective approach to examine test-takers’ use of information sources and strategies.
References

Acknowledgements

- Organizers

- My MA team
Thank you

Contact information

0086-13983732737
0086-23-65102602
xiangdonggu@263.net