Outline

• Research Purpose
• Washback Models
• Research Design
• Results and Discussion
• Conclusion and Suggestions
I. Research Purpose (1)

• Impact & Washback (or backwash)  
  (Bachman & Palmer, 1996; Hamp-Lyons, 1997; Hughes, 1989)

Test taking and use of test scores

Macro: Society, education system
Micro: Individuals

Impact

• the effects of language tests on micro-levels of language teaching and learning, i.e. inside the classroom  
  (Bachman & Palmer, 1996; McNamara, 2000; Wall, 1997)
I. Research Purpose (2)

• Research gaps (Liu & Gu, 2013; Xu, 2012; Zhang, 2012)
  1. Limited empirical studies
  2. More concern on teaching than on learning
  3. More studies on CET than on TEM

⇒ an empirical study of the effects of TEM4-Oral on test takers’ learning
II. Washback Models

2. Hughes’ (1993) PPP Model
2.1 Washback Hypotheses

(Alderson & Wall, 1993; Alderson & Hamp-Lyons, 1996)

1. A test will influence teaching.
2. A test will influence learning.
3. A test will influence what teachers teach; and
4. A test will influence how teachers teach; and therefore by extension from (2) above
5. A test will influence what learners learn; and
6. A test will influence how learners learn.
7. A test will influence the rate and sequence of teaching; and
8. A test will influence the rate and sequence of learning.
9. A test will influence the degree and depth of teaching; and
10. A test will influence the degree and depth of learning.
11. A test will influence attitudes to the content, methods of teaching and learning.
12. Tests that have important consequences will have washback; and conversely
13. Tests that do not have important consequences will have no washback.
14. Tests will have washback on all learners and teachers.
15. Tests will have washback effects for some learners and some teachers, but not for others.
16. Tests will have different amounts and types of washback on some teachers and learners than on other teachers and learners.
2.2 PPP Model

- **Participants** include English teachers, students, administrators, materials developers and publishers, all of whose perceptions and attitudes towards their work may be affected by a test (Bailey, 1996: 262).

- **Process** refers to “any actions” taken by the participants which may contribute to the process of learning. Such processes cover material development, syllabus design, changes in teachings methodology, the use of learning and/or test-taking strategies, etc.

- **Products** involves ‘what is learned (facts, skills, etc.) and the quality of the learning (fluency, etc.)’. 
2.2 PPP Model (Hughes, 1993)

Participants  Processes  Products

- Test
- Students
- Teachers
  - Materials, writers and curriculum designers
  - Researchers
- Learning
  - Teaching
    - New materials and New curriculum
    - Research results
2.2 PPP Model in this study

• The nature of TEM4-Oral may first affect the perceptions and attitudes of students towards their learning tasks. These attitudes in turn may affect what students do in the learning process, including practicing authentic/simulated test items, increasing after-class exercises, etc. which will affect the learning outcomes, i.e. what is learned and the quality of learning.
III. Research Design

1. Research questions
2. TEM4-Oral
3. Participants
4. Instruments
5. Data collection and analysis
3.1 Research Questions

1. What are the effects of TEM4-Oral on students’ attitudes towards English learning?
2. What are the effects of TEM4-Oral on students’ learning processes?
3. What are the effects of TEM4-Oral on students’ learning products?
3.2 TEM4-Oral: Overview

- Graded Test for English Majors Band 4—Oral Test
- 21,000 test-takers from 167 schools in 2011
- An annual nationwide and tape-mediated EFL speaking test
- A syllabus-based performance test, assessing the speaking performance of sophomore English majors
3.2 TEM4-Oral: Tasks and Criteria

• 3 Tasks
  – Retelling a story
  – Talking on a given topic
  – Role-playing

• Rating Criteria

<table>
<thead>
<tr>
<th>Content-Related</th>
<th>Form-Related</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td>Fluency</td>
</tr>
<tr>
<td>Coherence</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Sufficiency</td>
<td>Naturalness/Intelligibility</td>
</tr>
<tr>
<td></td>
<td>Lexical variety</td>
</tr>
</tbody>
</table>
3.3 Participants

- 138 English major sophomores
- English Education: 61; English: 50; English Translation & Interpreting: 27
- Average age: 20
- 15 of 138 were interviewed voluntarily
  - High scorers (M=88.90)
  - Low scorers (M=77.58)
  - Average scorers (M=83.33)
3.4 Instruments

- **Questionnaire**
  - 2 pilot studies, several revisions; reliability check (Alpha=.946)
  - 31 items on 5-point Likert scale (5=Strongly agree, 1=Strongly disagree) + 2 MCs

- **Interview**: in-depth info; group difference

- **Class observation**: CE, Listening, Speaking
### Questionnaire framework

<table>
<thead>
<tr>
<th>Domains</th>
<th>Analytical categories</th>
<th>Subcategories</th>
<th>Number of Qs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>Students’ attitudes</td>
<td>-attitudes towards learning tasks</td>
<td>1 (MC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-attitudes towards TEM4-Oral</td>
<td>1*(negative/positive)</td>
</tr>
<tr>
<td><strong>Processes</strong></td>
<td>Students’ learning processes</td>
<td>-content of learning</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-method of learning</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-rate of learning</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-sequence of learning</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-depth of learning</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-degree of learning</td>
<td>5</td>
</tr>
<tr>
<td><strong>Products</strong></td>
<td>Students’ learning products</td>
<td>-quality of learning</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-learning skills and facts</td>
<td>1 (MC)</td>
</tr>
</tbody>
</table>
3.5 Data Collection and Analysis

• Data collection
  – Phase One: 15 Class observations
  – Phase Two: 3*3 pre-test group interviews (audio-recorded)
  – Phase Three: Questionnaire survey (138 sheets)
  – Phase Four: 15 post-test individual interviews (audio-recorded)

• Data analysis
  – Theme abstraction of class observation
  – Transcribing and coding of interviews
  – Descriptive statistics of questionnaires, SPSS 16.0
IV. Results and Discussion

4.1 Effects on learning attitudes

– Towards test: positive > negative (74%); totally positive (10.1%); no effect (13%)

⇒ “Tests will have washback effects for some learners, but not for others”.

– Towards score use: self-evaluation (74.6%); confidence in English study (73.9%); future development (66.7%); interest in English (62.3%)

⇒ “Tests that have important consequences will have washback”

⇒ Test functions: diagnostic; selective; motivational
4.2 Effects on Learning Processes (1)

- 70.4% students adapt their learning ($M=3.80$, $SD=1.11$)

<table>
<thead>
<tr>
<th>Learning Process</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content of learning</td>
<td>4.23</td>
<td>0.99</td>
</tr>
<tr>
<td>Methods of learning</td>
<td>3.67</td>
<td>1.02</td>
</tr>
<tr>
<td>Rate of learning</td>
<td>3.66</td>
<td>1.07</td>
</tr>
<tr>
<td>Sequence of learning</td>
<td>3.61</td>
<td>1.04</td>
</tr>
<tr>
<td>Depth of learning</td>
<td>3.74</td>
<td>1.08</td>
</tr>
<tr>
<td>Degree of learning</td>
<td>3.54</td>
<td>1.06</td>
</tr>
</tbody>
</table>
4.2 Effects on Learning Processes (2)

• Content of learning
  – Previous test papers/test-related exercises (87.7%)

• Methods of learning
  – Develop own test-taking/learning strategies (80.4%)
  – Seek suggestion from classmates (62.4%)/ teachers (35.5%)

• Rate of learning
  – Learning becomes more efficient (58.3%)
  – Spend more spare time on oral English practice (76.1%)
  – Oral practice/interaction in English becomes more frequent (74%; 66.7%)
4.2 Effects on Learning Processes (3)

• Sequence of learning
  – Sequence of class learning is not affected (60%)
  – Temporary priority over oral English skills/knowledge (51.1%)

• Depth of learning
  – Increase difficulty level of oral practice (50%)
  – Generalize narrative structure/interactive principles/expressions

• Degree of learning
  – A wider coverage of learning (vocabulary range, grammar, language input; pair work) (78.6)
4.3 Effects on Learning Products

• Quality of learning

<table>
<thead>
<tr>
<th>Options</th>
<th>Learning interest enhanced</th>
<th>More listening &amp; speaking</th>
<th>Learning goals clearer</th>
<th>Thinking in English</th>
<th>Self-confidence boosted</th>
<th>Oral proficiency improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/SD</td>
<td>4.45/0.90</td>
<td>3.96/0.98</td>
<td>3.75/0.97</td>
<td>3.50/1.09</td>
<td>3.43/1.03</td>
<td>3.30/1.02</td>
</tr>
</tbody>
</table>

• Knowledge & Skills

<table>
<thead>
<tr>
<th>Options</th>
<th>Coherence</th>
<th>Fluency</th>
<th>Interactiveness</th>
<th>Relevance</th>
<th>Appropriacy</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>82.6%</td>
<td>76.1%</td>
<td>70.3%</td>
<td>67.4%</td>
<td>61.6%</td>
<td>57.2%</td>
</tr>
<tr>
<td>Order</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Options</td>
<td>Accuracy</td>
<td>Vocabulary</td>
<td>Grammar</td>
<td>Pronunciation</td>
<td>Complexity</td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>47.1%</td>
<td>28.3%</td>
<td>25.4%</td>
<td>30.7%</td>
<td>26.8%</td>
<td></td>
</tr>
<tr>
<td>Order</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>
4.4 Effects on Learners of Different Levels

• Pre- & Post-test interviews with 15 learners (HS, AS, LS)

⇒ “Tests will have different amounts and types of washback on some learners than on other learners”.

– HSs are affected in all categories; LSs only affected in content and methods of learning
– HSs spent longer time preparing, more initiative, more hard-working, overall improvement; LSs less active, little initiative, little progress.
V. Conclusion (1)

1. TEM4-Oral influences all aspects of learning, confirming Alderson’s hypotheses:

(2) A test will influence learning.
(5) A test will influence what learners learn; and
(6) A test will influence how learners learn.
(8) A test will influence the rate and sequence of learning.
(10) A test will influence the degree and depth of learning.
(12) Tests that have important consequences will have washback.
(16) Tests will have different amounts and types of washback on some learners than on other learners.
V. Conclusion (2)

2. Effects of TEM4-Oral in three aspects
   - Overall: more positive than negative
   - On learning processes:
     content > depth > methods > rate > sequence > degree
   - On learning products:
     A. quality improved (enhanced interest, goal-targeted, confidence boosted, increased input & output)
     B. Improved knowledge & skills: content coherence, relevance, language fluency, appropriacy, strategy, interactiveness
V. Conclusion (3)

3. High-scorers are more affected than low-scorers, in both categories and the degree.
   – HSs are more active and initiative, knowing what to learn and searching for ways/strategies of learning.
   ➔ Individual differences should be considered in classroom teaching, encouraging peer interaction, increasing learner autonomy, etc.
V. Suggestions

• Larger sample size, covering learners from different types of colleges and universities;
• To analyze test scores/ tasks; compare language features;
• To collect comparative data, doing more longitudinal studies.
Thank you.