

# ทฤษฎีต่าง ๆ ในการทดสอบทางภาษา และปัญหาที่ยืดเยื้อในการทดสอบสัมฤทธิ์ผล

(Theories in Language Testing and Persistent Problems in Testing the Achievement)

ผ.ศ. สุวัฒน์ สุขมลสันต์

การทดสอบทางภาษานั้นมักต้องอิงอยู่กับแนวความคิดหรือความเชื่อเกี่ยวกับทฤษฎีทางภาษาเสมอ ดังนั้นก่อนที่จะกล่าวถึงทฤษฎีต่าง ๆ ในการทดสอบทางภาษา จึงมีความจำเป็นที่ต้องกล่าวถึงทฤษฎีทางภาษาควก่อนไปด้วย ก่อนที่จะกล่าวถึงปัญหาในการทดสอบในภายหลัง

ก. ทฤษฎีต่าง ๆ ทางภาษาศาสตร์

1. ทฤษฎีทางภาษานั้นมีอยู่มากมาย แต่พอจะแบ่งออกได้เป็นกลุ่มใหญ่ ๆ ได้อย่างไรบ้าง?

ทฤษฎีทางภาษามีอยู่มากมายก็จริงแต่พอที่จะแบ่งออกเป็น 3 กลุ่มใหญ่ได้ดังนี้ (Morrow, 1977:1):

1.1 กลุ่มที่นิยมการแยกภาษาออกจากสังคมอย่างเด็ดขาด (Strict-Separatists)

นักภาษาศาสตร์กลุ่มนี้มีทัศนะต่อภาษาว่า ภาษามีลักษณะเหมือนกับวัตถุชิ้นหนึ่งซึ่งแยกตัวออกจากสังคมเป็นอิสระได้อย่างเด็ดขาด ดังนั้นการเรียนภาษาคือการเรียนความจริงต่าง ๆ เกี่ยวกับภาษา เช่น การเรียนวรรณคดีสมัยเก่า หรือการศึกษาที่มาของคำศัพท์ต่าง ๆ เป็นต้น โดยคิดว่าส่วนย่อย ๆ เหล่านี้ประกอบกันเป็นภาษา โดยไม่คำนึงถึงความแตกต่างในการใช้ภาษาในสังคมที่แตกต่างกัน พวกนี้มุ่ง “ศึกษาภาษา” (Language Study)

### 1.2 กลุ่มที่นิยมการแยกภาษาออกเป็นส่วนย่อย ๆ ที่มีฐานะเท่าเทียมกัน (Separate-But-Equalists)

นักภาษาศาสตร์กลุ่มนี้คิดว่าภาษานั้นสามารถแยกออกเป็นส่วนย่อย ๆ ที่มีฐานะเท่าเทียมได้ คือ ด้านการฟัง พูด อ่าน และเขียน แต่ละส่วนมีลักษณะอิสระ แต่มีความสัมพันธ์ซึ่งกันและกัน และมีความสัมพันธ์กับสิ่งแวดล้อมด้วย โดยคิดว่าภาษาเป็นพฤติกรรมอย่างหนึ่งของมนุษย์ ที่อาจเรียนรู้ได้ด้วยการตอบสนองต่อสิ่งเร้า นักภาษาศาสตร์พวกนี้มีแนวความเชื่อเกี่ยวกับการเรียนภาษา เช่นเดียวกับพวกนักพฤติกรรมนิยม (Behaviorists) นักภาษาศาสตร์พวกนี้อาจเรียกได้ว่าเป็นพวกโครงสร้างนิยม (Structurists) พวกนี้มุ่งศึกษา “การใช้ภาษา” (Language Usage)

### 1.3 กลุ่มที่คิดว่าภาษานั้นมีสภาพรวมกันอยู่กับสิ่งอื่น ๆ (United-Linguists)

นักภาษาพวกนี้คิดว่า ภาษานั้นมีความสัมพันธ์กับสิ่งอื่น ๆ อีกมาก เช่น วัฒนธรรม ฐานะทางสังคม สภาพของจิตใจ และความพร้อม เป็นต้น และคิดว่าภาษานั้นไม่สามารถแยกออกเป็นส่วนย่อย ๆ ได้อย่างอิสระรวมทั้งคิดว่าภาษามีการเปลี่ยนแปลงอยู่เสมอ นักภาษาศาสตร์พวกนี้มุ่งศึกษา “ภาษาที่ใช้” (Language Use) กลุ่มนี้รวมเรียกว่าพวกปฏิวัติ (Transformationists) จากแนวความคิดของนักภาษาศาสตร์กลุ่มนี้ทำให้เกิดมีศาสตร์ทางภาษาอีกหลายอย่างเช่น จิตภาษาศาสตร์ (Psycholinguistic) สังคมภาษาศาสตร์ (Sociolinguistic) และ Paralinguistic เป็นต้น ซึ่งศาสตร์ใหม่ๆ เหล่านี้อาจเรียกรวม ๆ กันว่า “มหภาษาศาสตร์” (Macro-Linguistics) ส่วนศาสตร์ทางภาษาดั้งเดิมนั้นรวมเรียกว่า “จุลภาษาศาสตร์” (Micro-Linguistics)

## 2. แนวความเชื่อทางภาษาเหล่านี้เกี่ยวข้องกับการเรียนการสอน และการสอบทางภาษาอย่างไร?

แนวความเชื่อของนักภาษาศาสตร์ทั้ง 3 กลุ่มดังกล่าวแล้วเกี่ยวข้องกับการเรียนการสอน และการทดสอบมาก แต่ผู้เขียนจะขอกล่าวแต่พอเข้าใจที่แสดงไว้ในตารางเปรียบเทียบของทั้ง 3 กลุ่มดังต่อไปนี้

ความสัมพันธ์ระหว่างทฤษฎีทางภาษากับการเรียนรู้การสอนและการสอบ

กลุ่ม	ทฤษฎีทางภาษา	ลักษณะของหลักสูตร	ลักษณะของการเรียน
Strict-Separatists	<p>ภาษา มีลักษณะที่ชัดเจนตั้งแต่ต้น ซึ่งมีความเป็นเอกเทศของตนเอง และไม่เกี่ยวข้องกับสิ่งแวดล้อม หรือพฤติกรรมอื่นใด</p>	<p>หลักสูตรโดยมากจะเป็นการศึกษาวรรณคดีชิ้นใดชิ้นหนึ่ง การศึกษาความเป็นมาของคำศัพท์ และการศึกษาไวยากรณ์ของภาษา โดยมากหลักสูตรไม่มีความสัมพันธ์กัน และไม่มีการวางแผนต่อเนื่องกัน</p>	<p>การเรียนภาษาเป็นปัจจัยที่ชี้พัฒนาสมอง ดังนั้น การเรียนภาษาจึงเป็นการเรียนเรื่องราวของภาษา</p>
Separate-But-Equalists	<p>ภาษา มีลักษณะที่ชัดเจนตั้งแต่ต้นหนึ่ง แต่มีความสัมพันธ์กับสิ่งแวดล้อม จึงถือว่าภาษาเป็นพฤติกรรมอย่างหนึ่งของมนุษย์ และสามารถแยกภาษาออกเป็นส่วนย่อยๆ ได้ โดยแต่ละส่วนมีความสัมพันธ์กัน</p>	<p>หลักสูตรเป็นแบบ form-based grammatical syllabus (หลักสูตรที่มุ่งศึกษาโครงสร้างทางไวยากรณ์) เพราะเป็นพวกโครงสร้างนิยม มุ่งสอนลักษณะย่อยๆ ของภาษาทีละจุดๆ แล้วนำแต่ละจุดมาสัมพันธ์กับสิ่งแวดล้อม</p>	<p>การเรียนภาษาเป็นการเรียนรู้ของนิสัย โดยการเรียนรู้แบบตอบสนองต่อสิ่งเร้าอยู่เสมอๆ ด้วยการฝึกฝนการใช้กระบวน (Pattern) ทางภาษาเพื่อให้สามารถใช้ชุดของนิสัยที่ต้องการได้อย่างอัตโนมัติ</p>
United-Linguists	<p>ภาษา มีลักษณะรวมผลกับพฤติกรรมอื่น ๆ แต่ละสิ่งแวดล้อมที่แยกออกจากกันไม่ได้ และภาษามีการเจริญเติบโตหรือเปลี่ยนแปลงอยู่เสมอ</p>	<p>หลักสูตรเป็นแบบ rule-based grammatical syllabus (หลักสูตรที่มุ่งศึกษากฎเกณฑ์ทางไวยากรณ์ที่แฝงอยู่ในโครงสร้างเล็ก) โดยการศึกษาร่วมกับสิ่งแวดล้อมไปพร้อมๆ กัน</p>	<p>การเรียนภาษามุ่งให้เกิดมี "ความสามารถในการสื่อสาร" (Communicative Competence) ในตัวผู้เรียน เพื่อให้ผู้เรียนสามารถสื่อสารและเข้าใจประโยคต่างๆ ได้อย่างไม่จำกัดจำนวน รวมทั้งสามารถใช้ได้ อย่างเหมาะสมกับสถานการณ์ต่างๆ ที่เป็นจริง</p>

<p>กลุ่ม</p> <p>Strict-Separatists</p>	<p>ลักษณะของการสอน</p> <ul style="list-style-type: none"> <li>- สอนวรรณคดี การแปล การตีความข้อความในวรรณกรรมที่เห็นว่าเป็น</li> <li>- สอนให้ศึกษาที่มาของคำศัพท์</li> <li>- อธิบายกฎเกณฑ์ต่างๆ ทางภาษา เพื่อให้ผู้เรียนจำกฎเกณฑ์เหล่านั้นไปประยุกต์ใช้</li> <li>- สอนแบบนิรนัย (Deduction) คือ สอนส่วนย่อยๆ ไปหาส่วนใหญ่ ในลักษณะสอนไวยากรณ์ (Grammatical Approach) โดยเน้นหนักเรื่อง “การศึกษารายวิชา” (Language study)</li> </ul>	<p>ตัวอย่างแนวความเชื่อ</p> <p>– ภาษาคือสิ่งที่บันทึกไว้ (A language is what is written.)</p>
<p>Separate-But-Equalists</p>	<ul style="list-style-type: none"> <li>- สอนตามแนวความเชื่อของพวกเขาพฤติกรรมนิยม โดยการใช้ในการฝึกเป็นกระสวน (Pattern drills) เป็นต้น บางครั้งก็ใช้บทสนทนาด้วย แต่ก็ควบคุมโครงสร้างทางภาษา ซึ่งมักจะทำให้ง่ายขึ้น (Simplified text)</li> <li>- หลีกเลี่ยงการท่องจำกฎเกณฑ์ทางภาษา แต่ให้จำประโยคไปใช้ได้อย่างอัตโนมัติ</li> <li>- สอนแบบอุปนัย (Induction) โดยการเรียนจากตัวอย่าง เพื่อให้สรุปกฎเกณฑ์ทางภาษาได้เอง ในลักษณะสอนโครงสร้างของภาษา (Structural approach) โดยเน้นหนักเรื่อง “การใช้ภาษา” (Language Usage)</li> </ul>	<ul style="list-style-type: none"> <li>- ภาษาคือชุดของนิสัย (A language is a set of habits.)</li> <li>- จงสอนภาษา ไม่ใช่สอนเรื่องราวของภาษา (Teach the language, not about the language.)</li> </ul>

กลุ่ม	ลักษณะของการสอน	ตัวอย่างแนวความเชื่อ
United-Linguists	<ul style="list-style-type: none"> <li>- คัดค้านการฝึกเหมือนเครื่องจักรยนต์ (Mechanical Drills) แต่ใช้การแก้ปัญหาแทน</li> <li>- นิยมใช้ การเล่นสมมติ (Role-Play) เพื่อการสนทนา</li> <li>- อธิบายกฎเกณฑ์ทางภาษาเมื่อจำเป็นเท่านั้น</li> <li>- เน้นเรื่องความหมาย (Meaning) มากกว่ารูปแบบ (Form) และนิยมใช้ภาษาต้นแบบ (Authentic text)</li> <li>- การสอนใช้แบบอาศัยหัวข้อเรื่องหรือหน้าที่ของภาษา (Notional/Functional approach) หรือแบบอาศัยสถานการณ์ (Situational approach) โดยมุ่งให้ผู้เรียนเกิดมี “ความสามารถกึ่งในสื่อสาร” (Communicative competence)</li> <li>- มุ่งศึกษา “ภาษาที่ใช้” (Language use)</li> </ul>	<ul style="list-style-type: none"> <li>- ภาษาคือคำพูด ไม่ใช่สิ่งที่บันทึกไว้ (Language is speech, not writing.)</li> <li>- ภาษาย่อมแตกต่างกัน (Languages are different.)</li> <li>- ภาษาคือสิ่งที่เจ้าของภาษาพูด ไม่ใช่สิ่งที่คนอื่นคิดว่าเจ้าของภาษาเขาจะพูด (A language is what its native-speakers say, not what someone thinks they ought to say.)</li> </ul>

กลุ่ม	ทฤษฎีการทดสอบ	สิ่งที่ทดสอบ	วิธีการทดสอบ
Strict-Separatists	<p>- ทฤษฎีการทดสอบของกลุ่มไม่ได้บ่งบอกไว้อย่างชัดเจน แต่เป็นที่เข้าใจกันว่า "ทดสอบเรื่องราวของภาษา" (Test about the language)</p>	<p>- ทดสอบกฎเกณฑ์ต่าง ๆ ทางภาษา โดยมุ่งถึงเรื่องความถูกต้องตามหลักไวยากรณ์ เช่น การใช้ Shall และ Will เป็นต้น</p> <p>- ทดสอบการแปล และการวิจารณ์วรรณกรรมที่ศึกษารวบรวมไว้</p> <p>- ทดสอบข้อความที่กำหนดให้ท่องจำ เช่น บทอาขยาน หรือคำศัพท์</p>	<p>- ทดสอบด้วยแบบทดสอบย่อย เช่น การแต่งประโยคจากคำที่กำหนดให้ การแปลข้อความ การวิจารณ์วรรณกรรม หรือการเขียนข้อความที่กำหนดให้</p>
Separate-But-Equalists	<p>- ผู้นำคนสำคัญทางการทดสอบของกลุ่มนี้ คือ Robert Lado (โดยเฉพาะการทดสอบภาษาที่สอง) ผู้ซึ่งเชื่อว่า : (Lado, 1960)</p> <p>- ผู้เรียนมีปัญหในการเรียน ทำานจนพบปัญหานั้น แล้วทำการทดสอบ</p> <p>- การวิเคราะห์ความแตกต่างมีความจำเป็นสำหรับการค้นหาปัญหา</p> <p>- จึงทดสอบสิ่งที่สอน แต่จึงทดสอบเฉพาะปัญหาในการเรียนครั้งละปัญหา ๆ</p>	<p>- ทดสอบส่วนประกอบทางภาษา 4 อย่าง คือ การออกเสียง ไวยากรณ์ โครงสร้างและศัพท์ ทั้งในความหมายทางวัฒนธรรม ด้วย ทงนี้ โดยการทดสอบ 4 ทักษะ คือ การฟัง การพูด การอ่าน และการเขียน</p> <p>- ทำการทดสอบส่วนประกอบย่อย ๆ ของส่วนประกอบทางภาษาทาง 4 อย่าง จากการจับประโยค จนถึงระดับหน่วยเสียง</p>	<p>- ทดสอบด้วยแบบทดสอบปรนัย เช่น การเลือกตอบ การเติมคำในช่องว่าง การจับคู่ การเรียงส่วนของประโยค และเลือกข้อถูกผิด เป็นต้น</p> <p>- ทดสอบส่วนประกอบของภาษาด้วยข้อสอบวัตถุจตุรคัย (Discrete-Point items) โดยเน้นหนักที่การทดสอบปัญหาในการเรียนภาษาที่รู้ได้จากการเปรียบเทียบระหว่างภาษาที่หนึ่งและภาษาที่สอง</p>

กลุ่ม	ทฤษฎีการทดสอบ	สิ่งที่ทดสอบ	วิธีการทดสอบ
	<ul style="list-style-type: none"> <li>- ความตรงเชิงสัมพันธ์ (Concurrent Validity) ของแบบทดสอบ เป็นสิ่งจำเป็น</li> <li>- ผู้เรียน ไม่สามารถที่จะเรียก ได้ว่ารู้ภาษา จนกว่าเขาจะใช้ภาษานั้นได้ อย่างอัตโนมัติและจับใจ ในสถานการณ์การปฏิบัติ</li> </ul>		<ul style="list-style-type: none"> <li>- ทดสอบด้วยข้อทดสอบจำนวนมาก เพราะต้องการ ให้ความเที่ยงตรง (Validity) และต้องการทดสอบความรวดเร็วในการใช้ภาษาด้วย</li> <li>- นิยมทำการทดสอบด้วยข้อสอบแบบเลือกตอบ (Multiple-choice test) ที่มี ความเที่ยงตรง (Validity) และความเชื่อมั่น (Reliability) สูง</li> </ul>
United-Linguists	<ul style="list-style-type: none"> <li>- ผู้นำของกลุ่มนี้เกี่ยวกับ การทดสอบทางภาษา คือ Carroll ผู้ซึ่งเชื่อว่า: “เราต้องตระหนักว่าการแสดงออกทางภาษาของบุคคลมีความเกี่ยวข้องกับความสามารถ ทัศนคติ และความ สามารถจริง ๆ ทางภาษาของเขาในลักษณะที่ผสมผสานกัน เช่น การเข้าใจความ” (Carroll, 1968: 56) นักทดสอบกลุ่มนี้มีความเชื่อว่าภาษานั้นแยกออกเป็นส่วนย่อยๆ ไม่ได้เว้นแต่ จะเป็นภาษาในระดับ “จุลภาษา” (Micro-Language)</li> </ul>	<ul style="list-style-type: none"> <li>- ทดสอบความสามารถทางภาษา (Linguistic Competence) ในระดับ “จุลภาษา”</li> <li>- ทดสอบความสามารถทาง การสื่อสาร (Communicative Competence) ในระดับ “มหภาษา”</li> <li>- ในทางปฏิบัติเน้น การทดสอบ คำหนึ่งถึงเรื่อง ปฏิสัมพันธ์ (interaction) เนื้อเรื่อง (context) ฉาก (setting)</li> </ul>	<ul style="list-style-type: none"> <li>- ทำการทดสอบในระดับ “จุลภาษา” ด้วยข้อสอบวัดจุดย่อย (Discrete-point items) ที่มีลักษณะเป็นปรนัย</li> <li>- ทำการทดสอบในระดับ “มหภาษา” ด้วยข้อสอบวัดรวมสรุป (Global items) ที่มีลักษณะเป็นปรนัย หรือ อัตนัย</li> <li>- การทดสอบทั้งสองระดับ ทำการ ทดสอบทั้งความสามารถศึกษา และ ความสามารถฉันทน์</li> </ul>

กลุ่ม	ทฤษฎีการทดสอบ	สิ่งที่ทดสอบ	วิธีการทดสอบ
	<p>แต่ในการสื่อความหมายนั้น เราต้องใช้ภาษาในระดับ “มหภาษา” (Macro-Language) ซึ่งต้องใช้ในลักษณะรวมกันหลาย ๆ อย่าง (integrated skills) ดังนั้น การทดสอบก็ต้องมุ่งทำการทดสอบทักษะรวมเหล่านี้พร้อม ๆ กัน เพราะจะทำให้สามารถทดสอบตัวอย่างของส่วนประกอบทางภาษาได้กว้างขวางยิ่งขึ้น</p> <ol style="list-style-type: none"> <li>1. การทดสอบรวมทักษะนี้ ทำให้สามารถทดสอบตัวอย่างของส่วนประกอบทางภาษาได้กว้างขวางยิ่งขึ้น</li> <li>2. การทดสอบแบบนี้ทำให้การตีความหมายของคะแนนการสอบง่ายขึ้น</li> <li>3. การทดสอบแบบนี้ไม่จำเป็นต้องทำ “การวิเคราะห์ความแตกต่าง (Contrastive analysis) ระหว่างภาษาให้เสียเวลา”</li> </ol> <p>ทฤษฎีการทดสอบนี้ได้แนวความคิดมาจากความหมายของคำว่า “ความสามารถ” (Competence) และ “ความสามารถ” (Performance) ของ Chomsky (1965)</p>	<ul style="list-style-type: none"> <li>- เรื่อง (topic)</li> <li>- หน้าที่ (function)</li> <li>- ความน่าจะเป็นไปได้ (modality)</li> <li>- บทบาท (role)</li> <li>- การรู้ล่วงหน้า (pre-supposition)</li> <li>- พิธีการ (formality)</li> <li>- สถานะ (status)</li> <li>- อารมณ์ (mood)</li> <li>- การยอมรับได้ (acceptability)</li> <li>- ทางด้านภาษา (linguistically)</li> <li>- ทางด้านสังคม (socially)</li> <li>- ทางด้านเนื้อเรื่อง (contextually) และ</li> <li>- ทางด้านประเพณี (culturally)</li> </ul>	<p>วิธีทดสอบ</p> <ul style="list-style-type: none"> <li>- นิยมทำการทดสอบด้วยการเขียนตามคำบอก และข้อสอบแบบปิด (close tests)</li> <li>- โปรดดูตัวอย่างแบบทดสอบตามแนวความคิดของนักภาษาศาสตร์กลุ่มทฤษฎีภาษามัน</li> <li>- นิยมใช้ภาษาต้นแบบ (authentic text) ในการทดสอบ</li> </ul>

กลุ่ม	ข้อดีของการทดสอบตามทฤษฎี	ข้อเสียของการทดสอบตามทฤษฎี
Strict-Separatists	<ul style="list-style-type: none"> <li>- ออกข้อสอบง่าย</li> <li>- ทำการทดสอบความรู้และความสามารถได้อย่างกว้างขวาง</li> <li>- มีประโยชน์ในการทดสอบ ความรู้ ความจำ การวิเคราะห์ และการสังเคราะห์ทางภาษา</li> </ul>	<ul style="list-style-type: none"> <li>- ออกไม่ครอบคลุมความคิดรวมยอดของ “ การรู้ภาษา ”</li> <li>- ให้คะแนนยุติธรรมได้ยาก</li> <li>- เสียเวลาในการตรวจแบบทดสอบมาก</li> <li>- ทาค่าทางด้านสถิติ เช่น ความเที่ยงตรง (Validity) และความเชื่อมั่น (Reliability) ได้ยาก</li> </ul>
Separate-But-Equalists	<ul style="list-style-type: none"> <li>- ตรวจสะดวก</li> <li>- มีความยุติธรรมในการให้คะแนน</li> <li>- ออกคลุมเนื้อหาตามโครงสร้างลึก (Deep Structures) ได้ไม่ยาก</li> <li>- ทาค่าความเที่ยงตรงและความเชื่อมั่นของข้อสอบได้ง่าย และสามารถปรับค่าสถิติเหล่านั้นได้ด้วย</li> <li>- มีประโยชน์มากในการทดสอบระดับ “ จุดรู้ ”</li> <li>- มีประโยชน์มากในการทดสอบ ความรู้ ความจำ และการวิเคราะห์</li> </ul>	<ul style="list-style-type: none"> <li>- ค่าความเที่ยงตรง และค่าความเชื่อมั่น เป็นค่าทางสถิติที่แน่นอน จะยิ่งแม่นยำไม่ได้ ต้องคำนึงถึงที่มาของค่าเหล่านั้นด้วย (เพราะอาจบอกความหมายได้ผิด ๆ และอาจปรับปรุงได้ค่าย)</li> <li>- ออกคลุมเนื้อหาตามโครงสร้างผิว (Surface Structures) ได้ยาก</li> <li>- ภาษาไม่ใช้ผลผลิตของส่วนย่อย ๆ ทางภาษา</li> <li>- เสียเวลาในการสร้างแบบทดสอบนั้นมาก และมีปัญหาเรื่อง Security ของข้อทดสอบ</li> <li>- แบบสอบแต่ละชุดใช้ได้ ในวงจำกัดเท่านั้น เพราะต้องสร้างจากปัญหาที่ได้จากการเปรียบเทียบระหว่าง 2 ภาษา ผู้สอบอาจมีพื้นฐานทางภาษาที่ 1 แตกต่างกันมาก</li> <li>- ใช้ทดสอบระดับ “ มหภาษา ” ได้ยาก</li> <li>- ทำการทดสอบความสามารถด้านการสังเคราะห์ได้ยาก</li> <li>- แนวความคิดทางภาษาขัดแย้งกับแนวความคิดเชิงทฤษฎีทางภาษาศาสตร์และทางจิตวิทยา ในปัจจุบันที่ว่า “ ภาษามีการเปลี่ยนแปลงอยู่เสมอ และการเรียนรู้ภาษาเกิดจากการสังเคราะห์ทางภาษา ไม่ใช่เกิดจากการวิเคราะห์ (Morrow, 1977 : 17-18)</li> </ul>

กลุ่ม	ข้อดีของการทดสอบตามทฤษฎี	ข้อเสียของการทดสอบตามทฤษฎี
United-Linguists	<ul style="list-style-type: none"> <li>- สามารถใช้วัดความรู้และความสามารถทางภาษาได้ทั้งระดับ “จุดภาษา” และ “มหภาษา”</li> <li>- การตีความหมายของคะแนนทำได้สะดวก</li> <li>- ข้อสอบบางชนิด เช่น การเขียนตามคำบอก และข้อสอบแบบปิด (Close tests) ออกได้ง่าย และทดสอบความสามารถทางภาษาได้อย่างกว้างขวาง</li> <li>- การทดสอบมีแนวคิดคล้ายตามแนวความคิดทางภาษาศาสตร์ และจิตวิทยาสมัยใหม่</li> <li>- การใช้ข้อสอบแบบปรนัยและอัตนัย ในการทดสอบพร้อมกัน ทำให้สามารถวัดความสามารถด้านวิเคราะห์และสังเคราะห์ทางภาษาได้ด้วย (นอกจากความสามารถด้านอื่น ๆ)</li> <li>- การใช้ภาษาต้นแบบ (Authentic text) ทำให้สามารถวัดความรู้ความสามารถทางภาษาได้ตรงตามความเป็นจริงยิ่งขึ้น</li> </ul>	<ul style="list-style-type: none"> <li>- การทดสอบแบบปรนัยทำให้ทราบแต่ความแตกต่างของเกณฑ์ปกติ (Norm) ของภาษา ของผู้สอบ และของผู้ออกข้อสอบเท่านั้น ไม่ใช้เกณฑ์ปกติ (Norm) ที่แท้จริงของผู้สอบ</li> <li>- แบบทดสอบบางอย่างตรวจยาก เช่น การเขียนตามคำบอก หรือข้อสอบแบบปิด (Close tests) อาจทำให้ขาดความยุติธรรม และเสียเวลามาก (ปัจจุบันนิยมใช้ modified close test แทน)</li> <li>- การสร้างข้อทดสอบที่ติดเนน ผู้สร้างต้องมีความรู้เกี่ยวกับภาษา (ท 2) เป็นอย่างดี และต้องรู้ศาสตร์หลายด้าน เช่น ต้องมีความรู้ด้านไวยากรณ์ประวัติ ถึงคัมภีร์ และจิตวิทยาภาษาศาสตร์ด้วย จึงจะเป็นการดี เพราะการเรียนรู้และการทดสอบทางภาษาเกี่ยวข้องกับศาสตร์เหล่านั้น ซึ่งอาจเป็นการยากที่ครูทั่ว ๆ ไปจะพึงปฏิบัติได้ เช่น ความรู้ ความเข้าใจเรื่องโครงสร้างลึก (Deep Structures) โครงสร้างผิว (Surface Structures) การยอมรับได้ (Acceptability) การพูด (Speech acts) หน้าที่ของการพูด (Speech functions) ความสามารถถักถัก (Competence), ความสามารถถักถัก (Performance) และระบบใหม่ไรท์สัน (Idealized system) ต่างมีส่วนเกี่ยวพันในการสร้างข้อสอบที่ใช้วัดความสามารถถักถักในการสื่อสาร (Communicative competence)</li> <li>- แนวความคิดบางอย่างของทฤษฎีนี้ยังไม่แพร่หลาย ทำให้ผู้ประสงค์จะนำไปใช้มีความรู้ยังไม่เพียงพอ</li> </ul>

### 3. ข้อทดสอบแบบ Discrete-Point Test และ Global Test แตกต่างกันอย่างไรรบ้าง ?

ข้อทดสอบแบบสอบวัดจุดๆ (Discrete-point Test) และแบบสอบวัดรวมสรุป (Global Test) นั้น แตกต่างกันอย่างมาทั้งในด้านรูปแบบและจุดประสงค์ในการใช้ ทั้งนี้เพราะแบบสอบชนิดแรกมุ่งทำการทดสอบทางภาษาตามแนวความเชื่อของนักภาษากลุ่มที่ 2 ส่วนแบบที่ 2 นั้นเป็นแนวการทดสอบของนักภาษากลุ่มที่ 3 ซึ่งบางครั้งยังคงอาศัยแบบทดสอบตามแนวแบบสอบวัดจุดๆ (Discrete-Point Test) มาใช้ด้วยสำหรับการทดสอบความสามารถระยะแรก เพื่อให้เข้าใจความแตกต่างของแบบทดสอบทั้ง 2 ชนิดดังกล่าวแล้ว พร้อมทั้งแนวความเชื่อเบื้องหลังไปรคศึกษาจากตารางที่ 2 ต่อไปนี้

ตารางที่ 2

ข้อเปรียบเทียบระหว่าง Discrete-Point Test และ Global Test

Discrete-Point Test (แบบทดสอบจุดย่อย)	Global Test* (“แบบทดสอบรวมสรุป”)
<ol style="list-style-type: none"> <li>1. เป็นการทดสอบเกี่ยวกับกรวิเคราะห์ภาษา (Analytical)</li> <li>2. เป็นการทดสอบปัญหาทีละจุดๆ (Discrete point)</li> <li>3. นิยมใช้ประเมินผลด้วยการอิงกลุ่ม (Norm Referenced)</li> <li>4. มักใช้ทำการทดสอบด้านการรับรู้ (Reception) เช่น การฟัง และไวยากรณ์</li> <li>5. ใช้ทดสอบ “ความสามารถด้านภาษา” (Linguistic Competence) เช่น ไวยากรณ์ ระบบการออกเสียง และระบบหน่วยคำ เป็นต้น</li> <li>6. เป็นการทดสอบระดับ “จุลภาษา” (Micro-Language)</li> <li>7. นิยมใช้เพื่อ “ประเมินผลรวมยอด” (Summative Evaluation) เช่น การทดสอบใช้แบบสอบสัมฤทธิ์ผล (Achievement Tests) เป็นต้น</li> <li>8. มุ่งทำการทดสอบด้านโครงสร้าง (Form) ของภาษา</li> <li>9. เน้นหนักในการทดสอบ “การใช้ภาษา” (Language Usage)</li> <li>10. การประเมินผลอิงอยู่กับทฤษฎีทางสถิติ (Idealization)</li> <li>11. มักใช้เพื่อทดสอบโครงสร้างลึก (Deep Structures)</li> <li>12. แบบทดสอบคำนึงถึงเรื่องความเชื่อมั่น (Reliability) มากกว่าอย่างอื่น</li> </ol>	<ol style="list-style-type: none"> <li>1. เป็นการทดสอบเกี่ยวกับกรสังเคราะห์ภาษา (Synthetical)</li> <li>2. เป็นการทดสอบภาษาหลายๆ ด้านรวมกัน (Wholistic)</li> <li>3. นิยมใช้ทำประเมินผลด้วยการอิงเกณฑ์ (Criterion Referenced)</li> <li>4. มักใช้ทำการทดสอบด้านการแสดงออก (Production) เช่น การพูด และการเขียน</li> <li>5. ใช้ทดสอบ “ความสามารถด้านสื่อสาร” (Communicative Competence) เช่น การพูด และการเขียน เป็นต้น</li> <li>6. เป็นการทดสอบระดับ “มหภาษา” (Macro-Language)</li> <li>7. นิยมใช้เพื่อ “ประเมินผลความก้าวหน้า” (Formative Evaluation) เพื่อรู้ความบกพร่องของการเรียน เช่น การทดสอบใช้แบบสอบวินิจฉัย (Diagnostic Tests)</li> <li>8. มุ่งทำการทดสอบด้านหน้าที่ (Function) ของภาษา</li> <li>9. เน้นหนักในการทดสอบ “ภาษาที่ใช้” (Language Use)</li> <li>10. การประเมินผลอิงอยู่กับคะแนนดิบ (Raw Data)</li> <li>11. มักใช้เพื่อทดสอบโครงสร้างผิว (Surface Structures)</li> <li>12. แบบทดสอบคำนึงถึงเรื่องความเที่ยงตรง (Validity) มากกว่าอย่างอื่น</li> </ol> <p>* หมายถึง แบบทดสอบนี้อาจเรียกว่า Wholistic Test หรือ Integrated Test ก็ได้</p>

ผู้เขียนเชื่อมั่นว่าผู้อ่านคงจะเคยชินกับแบบทดสอบตามแนวนักทฤษฎีทางการทดสอบ กลุ่มที่ 1 (Strict-Separatists) และกลุ่มที่ 2 (Separate-Equalists) แล้ว จึงขอเสนอแต่เพียง ตัวอย่างแบบทดสอบของพวกกลุ่มที่ 3 (United-Linguists) เท่านั้น แม้ว่ารูปแบบของแบบทดสอบจะคล้ายกับของพวกกลุ่มที่ 2 แต่ลักษณะคำถามแตกต่างกัน ผู้อ่านที่สนใจโปรดศึกษาได้จาก ภาคผนวก ท้ายบทความนี้

โดยมากในการทดสอบทางภาษาโดยเฉพาะภาษาต่างประเทศ ครูผู้สอนมักจะสนใจที่จะทำการทดสอบผลสัมฤทธิ์ในการเรียนมากกว่าเพื่อจุดประสงค์อื่น ปัญหาที่ผู้เขียนต้องการจะกล่าวต่อไปก็คือปัญหาที่ยึดถือในการทดสอบสัมฤทธิ์ผลในการเรียนภาษาเพื่อที่ว่าท่านผู้อ่านจะได้ตระหนักไว้บ้างในการทดสอบ

## ข. ปัญหาที่ยึดถือในการทดสอบสัมฤทธิ์ผลทางภาษา

เพื่อให้เข้าใจปัญหาในการทดสอบสัมฤทธิ์ผล ผู้อ่านควรเข้าใจเกี่ยวกับแบบทดสอบสัมฤทธิ์ผลเสียก่อน

### 1. แบบทดสอบสัมฤทธิ์ผลคืออะไร ?

แบบสอบสัมฤทธิ์ผล (Achievement Test) คือ “เครื่องมือที่สร้างขึ้นเพื่อวัดผลจากการสอนหรือการอบรมรายวิชาใดรายวิชาหนึ่ง” (Anastasi, 1963 : 33) ในการใช้แบบทดสอบสัมฤทธิ์ผลนั้น “เราสนใจแต่เพียงผลที่เกิดขึ้นภายในกรอบของหลักสูตรที่เรารู้แล้วเท่านั้น” (Davies, 1972 : 22) ทั้งนี้ เพื่อทดสอบว่านักเรียนเรียนรู้มากน้อยเท่าใด ดังนั้นในการทดสอบทางภาษาเราถือว่า แบบทดสอบสัมฤทธิ์ผลนั้นคือ “เครื่องมือที่จะได้มาซึ่งข้อมูลที่บ่งบอกว่า นักเรียนเรียนรู้มากน้อยเท่าใดจากการเรียนรายวิชาใดรายวิชาหนึ่ง” (Clark, 1972 : 3)

ในทางทฤษฎีแล้ว แบบทดสอบสัมฤทธิ์ผล (Achievement test) และแบบทดสอบวินิจฉัย (Diagnostic test) นั้น แตกต่างกันใน 3 ลักษณะคือ (Diagnostic test คือแบบทดสอบที่มุ่งในการทดสอบการเรียนรู้ในอดีตและในอนาคตว่ามีอะไรบ้างที่ผิดพลาดในการเรียนรู้ และเพื่อจะได้ใช้เป็นเครื่องมือในการวางแผนการเรียนซ่อมเสริม-Davies, 1972 : 23) :

1. หน้าที่
2. เนื้อหา
3. วิธีการวิเคราะห์ข้อสอบ

ลักษณะเฉพาะ	แบบทดสอบสัมฤทธิ์ผล	แบบทดสอบวินิจฉัย
1. หน้าที่	- เพื่อใช้ทดสอบว่าผู้เรียนได้เรียนรู้อะไรบ้าง	- เพื่อใช้ทดสอบว่าผู้เรียนยังไม่ได้เรียนรู้อะไรอีกบ้าง
2. เนื้อหา	- มีเนื้อหาเพื่อการทดสอบอย่างกว้างขวาง	- มีเนื้อหาเพื่อการทดสอบในวงแคบแต่ใช้จำนวนข้อทดสอบมาก
3. วิธีการวิเคราะห์ข้อทดสอบ	- เน้นหนักในการวิเคราะห์หาความเป็นวิวิธพันธ์ (heterogeneity) ของข้อสอบ (ข้อสอบที่แต่ละข้อมุ่งทดสอบสิ่งที่ต้องการสอบต่างกัน)	- เน้นหนักในการวิเคราะห์หาความเอกพันธ์ (homogeneity) ของข้อสอบ (ข้อสอบที่มุ่งทดสอบสิ่งเดียวกัน แต่ถามในหลายๆ ด้านต่างกัน)

อย่างไรก็ตามในการใช้แบบสอบถามเพื่อวัดว่า “นักเรียนได้เรียนรู้มากเท่าใด” ก็หมายความว่าในอีกนัยหนึ่งว่า “นักเรียนยังไม่ได้เรียนรู้อีกมากน้อยเท่าใด” ดังนั้น นักทดสอบบางท่านเรียก แบบทดสอบทุกชนิดที่วัดผลการเรียนรู้ภายในกรอบของหลักสูตรที่วางไว้ว่า แบบสอบวินิจฉัย (Davies, 1972 : 23) แต่นักทดสอบบางคนเช่น Clark (Clark, 1972 : 3) กลับเรียกแบบทดสอบที่มีลักษณะดังกล่าวแล้วว่า แบบทดสอบสัมฤทธิ์ผล ซึ่งเขาแบ่งออกเป็น 2 ชนิดคือ แบบทดสอบสัมฤทธิ์ผลเชิงวินิจฉัย (diagnostic achievement test) และแบบทดสอบสัมฤทธิ์ผลทั่วไป (general achievement test) แบบทดสอบทั้ง 2 ชนิดนี้แตกต่างกันดังนี้

ลักษณะเฉพาะ	แบบทดสอบสัมฤทธิ์ผลเชิงวินิจัย	แบบทดสอบสัมฤทธิ์ผลทั่วไป
1. หน้าที่	- เพื่อใช้วัดสัมฤทธิ์ผลในการเรียน “จุลภาษา” micro-language ได้แก่ ความรู้พื้นฐานเกี่ยวกับระบบหน่วยเสียง หน่วยคำ วากยสัมพันธ์ และความสัมพันธ์ระหว่างระบบเหล่านี้ กับระบบความหมายของคำ หรือ ความสามารถศักยภาพทางภาษา (linguistic competence ; Rivers, 1972 : 33)	- เพื่อใช้วัดสัมฤทธิ์ผลในการเรียน “มหภาษา” (macro-language) ได้แก่ การใช้ภาษาเพื่อประโยชน์ในการสื่อสารรวมทั้งการให้และการรับ โดยการพูดและการเรียนหรือความสามารถจลน์ทางภาษารวม (integrated language performance ; Rivers, 1972 : 33)
2. โครงสร้าง	- ประกอบด้วย “จุลภาษา”	- ประกอบด้วย “มหภาษา”
3. การทดสอบ	- ทดสอบด้วยข้อทดสอบแบบวัดจุดย่อยเพื่อทดสอบว่าผู้เรียน “สัมฤทธิ์ผล” หรือว่า “ไม่สัมฤทธิ์ผล” ในการเรียนแต่ละจุดย่อยของเนื้อหาที่ต้องการทดสอบ เช่น ทดสอบด้วยแบบทดสอบแบบเลือกตอบ เป็นต้น	- ทดสอบด้วยข้อทดสอบแบบวัดรวมสรุป โดยมีได้มุ่งหวังว่าจะทดสอบแต่ละจุดย่อยของเนื้อหา เช่น การทดสอบด้วยการเขียนเรียงความ และแบบสอบปิด (close tests) เป็นต้น

ดังนั้น เราอาจสรุปได้ว่าแบบทดสอบสัมฤทธิ์ผลคือ เครื่องมือที่มุ่งทำการทดสอบว่านักเรียนได้เรียนรู้อะไรบ้างในระดับ “จุลภาษา” โดยการสอบด้วยแบบทดสอบสัมฤทธิ์ผลเชิงวินิจัย และได้เรียนรู้อะไรบ้างในระดับ “มหภาษา” โดยการทดสอบด้วยแบบทดสอบสัมฤทธิ์ผลทั่วไป ทั้งนี้การทดสอบทำในกรอบของหลักสูตรที่วางไว้เท่านั้น

ปัญหาที่ตามมาในการทดสอบสัมฤทธิ์ผลในการเรียนก็คือ “สัมฤทธิ์ผลในการเรียนภาษาคืออะไร?”

## 2. สัมฤทธิผลในการเรียนภาษาคืออะไร ?

ตามลักษณะของหน้าที่ของแบบทดสอบสัมฤทธิผลนั้น เราสามารถที่จะกล่าวได้ว่า สัมฤทธิผลในการเรียนภาษาในรายวิชาใดรายวิชาหนึ่งนั้น ก็คือ สิ่งที่นักเรียนได้เรียนรู้จากการเรียนวิชานั้นๆ ในระดับ “จุลภาษา” หรือ “มหภาษา” หรือในทั้งสองระดับ แล้วแต่จุดมุ่งหมายในการสอนแต่ละรายวิชา

## 3. เราสามารถจะวัดสัมฤทธิผลในการเรียนภาษาหรือไม่ ?

คำถามนี้ตอบยาก เพราะมีปัจจัยหลายอย่างที่เกี่ยวข้อจนกลายเป็นปัญหาที่ยึดเยื้อในการทดสอบสัมฤทธิผลในการเรียนภาษา ก่อนที่จะตอบคำถามนี้เราควรทำความเข้าใจกับปัญหาที่ยึดเยื้อดังต่อไปนี้ คือ

### 3.1 ปัญหาเกี่ยวกับความเที่ยงตรงในการทดสอบ

ในการทดสอบทางภาษานั้นเป็นการยากมากที่จะรู้ว่า สิ่งที่เราทดสอบนั้นคือสิ่งที่เราต้องการทดสอบอย่างแท้จริงหรือไม่ หรือพูดอีกประการหนึ่งก็คือ เป็นการยากที่จะรู้ว่าข้อทดสอบของเรานั้นมีความเที่ยง (Validity) หรือไม่ แม้ว่าแบบทดสอบจะมีความเชื่อมั่น (Reliability) สูงมากเพียงใดก็ตาม (เพราะว่าความเชื่อมั่นสามารถเพิ่มหรือลดได้โดยวิธีสถิติ) การหาความเที่ยงตรงของข้อสอบนั้น เราไม่มีวิธีการที่จะหาได้อย่างแท้จริง โดยวิธีการทางวิทยาศาสตร์ แต่เราหาได้แต่เพียงการอ้างอิงเกณฑ์อื่น หรืออาศัยตรรกวิทยาเท่านั้น (Carroll, 1972 : 8) ที่เรารู้แน่ไม่ได้ก็คือสิ่งที่เราทดสอบนั้น คือ สิ่งที่เราต้องการทดสอบหรือไม่ เพราะว่าเป็นการยากที่จะรู้ว่า สัมฤทธิผลในการเรียนภาษาแต่ละด้านนั้นประกอบด้วยอะไรบ้าง โดยเฉพาะอย่างยิ่งการทดสอบในระดับ “มหภาษา” เช่น การอ่านเข้าใจความ และการฟังเข้าใจความ (reading comprehension and listening comprehension) เป็นต้น เพราะว่าเป็นการยากที่จะรู้ว่า “การเข้าใจความ” นั้นประกอบด้วยส่วนประกอบทางภาษาอะไรบ้าง? นักทดสอบทางภาษาก็มีความเห็นแตกต่างกันในเรื่องส่วนประกอบของภาษาในแต่ละด้าน ยกตัวอย่างเช่น Davies และ Widdowson (Davies and Widdowson, 1974 : 168-176) เชื่อว่า การอ่านเข้าใจความสามารถในเรื่องเกี่ยวกับ

ก. การอ้างอิงโดยตรง (direct reference)

ข. การสรุปนัยของข้อความ (inference)

ค. การสร้างสมมติฐาน (supposition)

ง. การประเมินผล (evaluation)

แต่ว่าการทดสอบการอ่านเข้าใจความของ TOEFL (Test of English as a Foreign Language) นั้นจะทำการทดสอบเกี่ยวกับความรู้และความสามารถในการเรื่องดังต่อไปนี้ด้วย คือ

ก. การตีความหมายของเนื้อหา (paraphrasing)

ข. การทดสอบศัพท์ในและนอกเนื้อเรื่อง (Vocabulary testing in and out of the context) และ

ค. การสรุปผลของเนื้อเรื่อง (conclusion)

ดังนั้น เป็นการยากที่จะกล่าวได้ว่า สิ่งที่เราทดสอบนั้นคือสิ่งที่เราต้องการทดสอบอย่างแท้จริงหรือไม่ เพราะเป็นการยากยิ่งที่จะรู้ส่วนประกอบที่แท้จริงของเนื้อหาบางอย่างทางภาษา โดยเฉพาะอย่างยิ่งการทดสอบระดับ “มหัศจรรย์” แม้ว่าเราจะหาความเที่ยงตรงของแบบทดสอบได้หลายวิธีก็ตาม (Cronbach, 1960 : 106) เช่น

1. ความเที่ยงตรงเชิงเนื้อหา (Content validity) โดยการเปรียบเทียบข้อทดสอบกับเนื้อหาของหลักสูตร
2. ความเที่ยงตรงเชิงสัมพันธ์ (Concurrent validity) โดยการหาความสัมพันธ์ระหว่างผลของการทดสอบกับผลของการทดสอบของแบบทดสอบที่มีอยู่แล้ว
3. ความเที่ยงตรงเชิงพยากรณ์ (Predictive validity) โดยการหาความสัมพันธ์ระหว่างผลของการทดสอบนั้น กับผลผลิตสุดท้ายที่ใช้เป็นเกณฑ์ในอนาคต
4. ความเที่ยงตรงเชิงความเชื่อ (Construct validity) โดยการเปรียบเทียบของการทดสอบกับแนวความเชื่อหรือทฤษฎีที่ตั้งขึ้น

แต่ว่าการหาความเที่ยงตรงโดยวิธีทางสถิตินี้ก็ไม่ใช่หลักประกันว่า สิ่งที่เราทดสอบนั้นคือสิ่งที่เราต้องการทดสอบอย่างแท้จริง แม้ว่าจะเป็นแนวทางที่บ่งให้รู้ถึงแนวโน้มว่าสิ่งที่เราต้องการทดสอบนั้นคืออะไร (Carroll, 1972 : 9)

ดังนั้น เราจึงสามารถสรุปได้ว่า การหาความเที่ยงตรงของสิ่งที่เราต้องการวัดนั้นอาจสามารถทำได้ในการทดสอบภาษาระดับ “จุลภาษา” ส่วนในระดับ “มหัศจรรย์” นั้นทำได้ยาก โดย

เฉพาะอย่างยิ่งการทดสอบ “ความสามารถศักยภาพในการสื่อสาร” (Communicative Competence) คือความสามารถในการใช้ภาษาต่างประเทศ โดยการเลือกใช้รูปแบบของภาษา (registers) และ สำนวนที่ใกล้เคียงกับเจ้าของภาษาได้อย่างเหมาะสม และถูกต้องตามความหมายของเนื้อหาที่สื่อสารในสถานการณ์ที่เป็นจริง (Jakobovits, 1970)

### 3.2 ปัญหาเกี่ยวกับขอบข่ายของเนื้อหา

นอกจากนี้ปัญหายืดเยื้อในการทดสอบสัมฤทธิ์ผลในการเรียนภาษาก็คือ เราไม่ทราบแน่ชัดว่า “เราจะต้องทำการทดสอบเนื้อหาในการสอนนั้นมากน้อยสักเท่าใด” จึงจะครอบคลุมกับความคิดรวบยอดเกี่ยวกับความสามารถอย่างแท้จริงในการใช้ภาษาต่างประเทศ “และเราก็ไม่รู้แน่ชัดว่าภาษานั้นสามารถที่จะแยกออกเป็นส่วนประกอบย่อย ๆ เพื่อการทดสอบได้หรือไม่ (Carroll, 1972 : 11) ที่เป็นเช่นนี้ก็เพราะว่า ค่าสหสัมพันธ์ของผลการทดสอบทักษะต่าง ๆ อาจจะแตกต่างกันได้แล้วแต่ว่าการสอนและการสอบจะเน้นหนักไปในแนวทางใด ยกตัวอย่างเช่น ค่าสหสัมพันธ์ระหว่างการอ่านเข้าใจความ และการฟังเข้าใจความสูง ไม่ได้หมายความว่า ผู้เรียนมีสัมฤทธิ์ผลในการเรียนทักษะทั้งสองเท่า ๆ กัน (เพราะค่าเฉลี่ยของการทดสอบนั้นอาจแตกต่างกันอย่างมีนัยสำคัญก็ได้)

นอกจากนี้ “ความสามารถในการใช้ภาษาต่างประเทศ ไม่ใช่ผลิตผลของส่วนย่อยต่าง ๆ ทางภาษา” (Rivers, 1972 : 32) ทั้งนี้ก็เพราะว่า ผู้ที่มีความสามารถในระดับ “มหภาษา” สามารถที่จะมีสัมฤทธิ์ผลในระดับ “จุลภาษา” ได้ด้วย แต่ผู้ที่มีความสามารถในระดับ “จุลภาษา” อาจจะไม่มีสัมฤทธิ์ผลในระดับ “มหภาษา” ในระดับที่ยอมรับได้ก็เป็นได้

### 3.3 ปัญหาเกี่ยวกับประสิทธิภาพในการทดสอบ

ปัญหายืดเยื้ออีกประการหนึ่งก็คือ ปัญหาเกี่ยวกับการทำการทดสอบ เช่น การเก็บรักษาข้อทดสอบไว้เป็นความลับ ค่าของเวลาและพลังงานในการสร้างข้อทดสอบ การทำการทดสอบ การให้คะแนน และการให้ระดับคะแนน เป็นต้น สิ่งต่าง ๆ เหล่านี้ล้วนแต่มีผลโดยตรงกับความเที่ยงตรง ความเชื่อมั่น และประสิทธิภาพของแบบทดสอบด้วย ยกตัวอย่างเช่น “การที่จะทำให้แบบทดสอบมีความเที่ยงตรง และความเชื่อมั่นมากขึ้น เวลาในการทดสอบต้องมีเพียงพอ และควรต้องใช้วัสดุอุปกรณ์ช่วยด้วย” (Carroll, 1972 : 14)

### 3.4 ปัญหาเกี่ยวกับการทดสอบเนื้อหาที่ทำการสอน

มีบ่อยครั้งที่ผู้สอน “สอนตามข้อสอบ” ทำให้ผลที่ได้คือความไม่สมบูรณ์ของการเรียน เพราะข้อสอบนั้นอาจไม่ครอบคลุมเนื้อหาที่สอน หรืออาจมีเป้าประสงค์ที่แตกต่างกับเป้าประสงค์ของรายวิชาที่สอนก็ได้ ปัญหายืดเยื้อเช่นนี้เกิดขึ้นได้ง่าย โดยเฉพาะอย่างยิ่งการสอนเพื่อสอบข้อสอบที่ผู้อื่นสร้างขึ้น

### 3.5 ปัญหาเกี่ยวกับการประยุกต์ใช้ทฤษฎี

มีบ่อยครั้งเช่นกันที่ทฤษฎีเกี่ยวกับการทดสอบยากแก่การนำมาประยุกต์ใช้ได้ อาจจะเป็นเพราะผู้สอนขาดความรู้ในทฤษฎี หรือขาดความชำนาญในการประยุกต์ใช้ทฤษฎี นอกจากนี้ปัญหานี้อาจเกี่ยวกับเวลาและค่าใช้จ่ายในการใช้แบบทดสอบด้วย เลยทำให้ยากแก่การนำมาประยุกต์ใช้ เช่น การสัมภาษณ์เพื่อทดสอบการฟังและการพูดนักเรียนจำนวนมาก ทำให้เสียเวลามาก ถ้าหากว่าจะใช้เครื่องบันทึกเสียงห้องปฏิบัติการภาษาก็ราคาแพง และผู้สอนบางคนไม่มีความรู้ในการควบคุมเครื่อง เป็นต้น

## 4. บทสรุป

ดังนั้น เราสามารถที่จะพูดได้ว่า การทดสอบสัมฤทธิ์ผลในการเรียนภาษานั้นยากมาก แต่ก็ไม่ใช่ว่าจะทำไม่ได้ วิธีหนึ่งที่เราจะทำได้ในขณะนี้คือการทดสอบสัมฤทธิ์ผลทางภาษาผ่านการทดสอบ ส่วนประกอบของความสามารถจริง (performance) และความสามารถศักยภาพ (competence) ทางภาษา ยกตัวอย่างเช่น การทดสอบ การอ่านเข้าใจความด้วยการทดสอบผ่านส่วนประกอบ 5-6 อย่างดังกล่าวแล้ว หรือในการทดสอบผลผลิตทางภาษา (language production เช่น การเขียน และการพูด เป็นต้น) อาจทำได้ด้วยการวัดความสามารถในการใช้ ส่วนประกอบทางภาษาได้อย่างยืดหยุ่น หรือความสามารถในการให้ความหมายแก่ข้อความหรือเนื้อความได้อย่างมีประสิทธิภาพ เป็นต้น

นอกจากนี้ เนื้อหาทางภาษาที่ทำการทดสอบจะต้องมีความเหมาะสมกับสภาพความเป็นจริงด้วย กล่าวคือ ผู้สอนจะต้องคำนึงว่าภาษาที่ใช้เป็นภาษาที่ง่ายขึ้น (simplified text) หรือว่าเป็นภาษาที่ทำให้ยากขึ้น (complexified text) หรือว่าเป็นภาษาต้นแบบ (authentic text) ภาษาที่ใช้ประกอบด้วยส่วนประกอบทางภาษา (linguistic components) ที่ใช้กันอยู่เสมอ หรือว่า

เป็นภาษาที่ไม่ค่อยมีคนนิยมใช้ ในการสอนนั้นควรจะเน้นหนักไปในทางส่วนประกอบใด และจะต้องตระหนักเสมอว่าการเรียนรู้ภาษานั้นจะต้องคำนึงถึงความสามารถในเชิงของความเร็วและความถูกต้องด้วย ไม่ใช่แต่เพียงความรู้ความสามารถในเชิงเนื้อหาเท่านั้น เป็นที่น่าสังเกตว่าแนวโน้มในการทดสอบในปัจจุบันนี้ มุ่งทำการสอนและสอบ “ภาษาต้นแบบ” ที่ประกอบด้วยส่วนประกอบทางภาษา (linguistic components) ที่นิยมใช้กันอยู่อย่างแพร่หลาย แทนที่จะเป็นเนื้อหาจาก “ภาษาที่ทำความง่ายขึ้น” หรือ “ภาษาที่ทำให้ยากขึ้น”

ในการทดสอบสัมฤทธิ์ผลในการเรียนนั้น ควรทำการทดสอบทั้งในระดับ “จุลภาษา” ด้วยแบบทดสอบสัมฤทธิ์ผลเชิงวินิจฉัย (diagnostic achievement tests) และในระดับ “มหภาษา” ด้วยแบบทดสอบสัมฤทธิ์ผลทั่วไป (general achievement tests) แบบของการทดสอบควรจะเป็นทั้งแบบทดสอบจุดย่อย (discrete-point tests) และแบบทดสอบรวมสรุป (global tests) ทั้งแบบอัตนัยและปรนัยด้วย ทั้งนี้เพื่อให้ครอบคลุมเนื้อหาที่ต้องการทดสอบหลายๆ ด้าน และถ้าเป็นไปได้ควรเป็นการทดสอบแบบอิงเกณฑ์ (Criterion-reference tests) มากกว่าจะเป็นการทดสอบอิงกลุ่ม (Norm-reference tests) ซึ่งรายละเอียดเกี่ยวกับเรื่องเหล่านี้ผู้สอนควรศึกษาและทำความเข้าใจด้วย ทั้งนี้เพราะว่า อย่างน้อยจะทำให้แบบทดสอบที่สร้างขึ้นใช้นั้นมีความเที่ยงตรงเชิงเนื้อหา (Content validity) และอาจมีความเชื่อมั่น (Reliability) ในระดับสูงได้ด้วย

ดังนั้น เราจะเห็นว่า เราต้องเผชิญกับปัญหาที่ยืดเยื้อในการทดสอบสัมฤทธิ์ผลในการเรียนภาษาต่อไปอีกนานเท่านั้น “จนกว่าเราจะรู้แน่ชัดว่าสิ่งที่เราตั้งใจสอนนั้นคืออะไร ไม่เช่นนั้นแล้วเราจะไม่สามารถวัดผลความสำเร็จในการเรียนได้” (Rivers, 1972 : 33 by Valette)

## ภาคผนวก

### ตัวอย่างข้อทดสอบของ United-Linguists

#### I. Discrete-Point Tests

- \* 1. Multiple-Choice Item (แบบเลือกตอบ)
- \* 2. Matching-Item (แบบจับคู่)

#### II. Global Tests

1. Dictation (การเขียนตามคำบอก)
  2. Close Test (แบบทดสอบ “ปิด”) และ Modified-Close Test (แบบทดสอบ “ปิดชนิดปรับปรุง”)
  - \* 3. Answering Question Test (แบบตอบคำถาม)
    - อาจมีการตอบคำถามโดยตรง หรือ
    - ตอบคำถามจากเครื่องบันทึกเสียงหรือข้อมูลที่ให้
  4. Asking Question Test (แบบถามคำถาม)
    - ให้ดูภาพหรืออ่านเรื่องแล้วให้ถามคำถามที่เกี่ยวข้อง
  5. Interaction Participation Test (แบบทดสอบที่ผู้สอบและผู้ถูกทดสอบมีกิจกรรมร่วมกัน “แบบทดสอบกิจสัมพันธ์”)
    - ครูอ่านบทสนทนาแล้วให้นักเรียนพูดตอบ (เรียกว่า แบบโครงสร้างหรือ Structured Test)
    - \* - ให้บทสนทนาที่ไม่สมบูรณ์แล้วให้ผู้สอบเติมให้สมบูรณ์ เช่น อาจละบทของใครคนหนึ่งในบทสนทนา (เรียกว่าแบบกึ่งโครงสร้าง Semi-Structured Test)
    - แบบสัมภาษณ์
  6. Narration Test (แบบบรรยายความ)
    - แบบเรียงความอิสระ (Free-Composition Test)
    - แบบเรียงความชนิดควบคุม (Controlled-Composition Test) นิยมใช้สถานการณ์ หรือภาพเป็นสิ่งควบคุม แทนที่จะใช้โครงสร้างหรือศัพท์ (ซึ่งเป็นแนวความคิดของพวกเขาโครงสร้างนิยมในกลุ่มที่ 2)
- ต่อไปนี้เป็นตัวอย่างแบบทดสอบที่มีเครื่องหมาย \* นอกนั้นคิดว่าท่านคงจะคุ้นเคยแล้ว

## ตัวอย่างข้อทดสอบ

### 1. Multiple-Choice Test

Presentation : Oral (tape-recorded) or Written

Type : Identification

Present :

Excuse me, do you know where the nearest post-office is, please ?

1. (Setting)           Where might somebody ask you that question?
  - a. In your house.
  - b. In your office.
  - c. In the street.
  - d. In a restaurant.
  
2. (Topic)            What is the person asking you about?
  - a. The price of stamps.
  - b. The age of the post-office.
  - c. The position of the post-office.
  - d. The size of the post-office.
  
3. (Function)        Why is the person speaking to you?
  - a. To ask your advice.
  - b. To thank you.
  - c. To ask for information.
  - d. To warn you.
  
4. (Modality)        Does the speaker think that you know the answer to the question?
  - a. He is sure that you do.
  - b. He is sure that you don't.
  - c. He doesn't know if you do or not.
  - d. He thinks you really ought to know the answer.
  
5. (Pre-supposition)    Would you understand from the question
  - a. that you are the first person the speaker has asked?
  - b. the speaker already knows where there is a post-office?
  - c. the speaker has already asked somebody else but has not found out the information?
  - d. the speaker just wants to check what he already knows? (N.B. stress : Excuse me, do you know....)
  
6. (Role)            Who is the speaker?
  - a. A friend of yours.
  - b. A person you met recently.
  - c. A person you have never met before.
  - d. A very close friend.

7. (Formality) How would you describe the attitude of the speaker?  
 a. Very, very polite.  
 b. Very friendly.  
 c. Polite.  
 d. Unfriendly.
8. (Status) Is the speaker treating you  
 a. as a friend?  
 b. as somebody very important?  
 c. as an equal?  
 d. as somebody who is less important than he is?
9. (Mood) How would you describe the mood of the speaker?  
 a. Bored.  
 b. Anxious.  
 c. Depressed.  
 d. Excited.
- (Note : With appropriate intonation, etc. This could only be used with a taped presentation.)

## 2. Matching Test

Presentation : Oral (tape-recorded) or Written

Type : Matching

Present :

The manager will see you now, sir.

1. (Setting) Where might you hear this? Put a tick by any of the following expressions you might hear in the same place.  
 a. Stop talking.  
 b. Would you care to take a seat?  
 c. I'm afraid he's not in at the moment.  
 d. Look, I've told you before. Don't do that.
2. (Topic) Listen to/Look at these comments. In some of them the speaker is talking about a train journey he has just made. Put a tick next to the ones which refer to this  
 a. ...and we took off on time, despite the fog in New York.  
 b. So by the time we got there we were running over half an hour late.  
 c. The guard was very helpful. He found us a seat in a non-smoker.  
 d. The trouble was we got a puncture, so that held us up.

3. (Function) Look at these two columns. Join appropriate sentences from each together to make one utterance:

I'm sorry but it just isn't  
good enough  
I really do appreciate what  
you've done  
I'm not sure if I like it  
or not  
It really was fantastic  
Excuse me please

I can't thank you enough  
Is this the way to the town  
centre?  
It should have been ready yester-  
day  
What do you think?  
I enjoyed every minute of it

### 3. Close Tests

A close test consists simply of a passage of language from which certain words are deleted and must be supplied by the learner. The words to be deleted \_\_\_\_\_ chosen at random by deleting \_\_\_\_\_ the word. Originally this procedure \_\_\_\_\_ developed to assess the "readability" \_\_\_\_\_ a given passage, i.e. how \_\_\_\_\_ or difficult it is for \_\_\_\_\_ native speaker to read and \_\_\_\_\_. But its potential as a \_\_\_\_\_ of overall language proficiency was \_\_\_\_\_ noted. Oller (1973) claims that \_\_\_\_\_ tests of L<sub>2</sub> proficiency have \_\_\_\_\_ following characteristics:

1. Internal consistency;
2. Reliability;
3. Validity;
4. Ease of construction and administration;
5. Ease of standardisation.

### 4. Answering Question Test

Prior to the test the candidate is given a card which outlines a situation. He must invent a number of questions he would want to ask in that situation. (The situations being derived from the specified exam, syllabus.)

e.g. you arrive at Victoria Station in London to catch the train to Paris. You want to find out something from each of these people. What would you say?

1. A passer-by in the station entrance.
2. The booking clerk.
3. Your friend who is travelling with you.
4. The ticket collector.

### 5. Semi-Structured Test

A sample of the type of dialogue which might be used is as follows:

A: Hello, I haven't seen you for ages. How are you?

B: \_\_\_\_\_

A : That is a shame. What happened?

B : \_\_\_\_\_

A : Was anyone else in the car?

B : \_\_\_\_\_

A : Was she hurt?

B : \_\_\_\_\_

A : Well I'm glad to hear that. Look, I have to go now but why don't we have lunch together on Friday?

B : \_\_\_\_\_

A : Right, I'll see you at 12.30 on Friday then. Bye.

### หนังสืออ้างอิง

1. Allen, J.P.B. and Davies, A., *Testing and Experimental Methods*, (The Edinburgh Course in Applied Linguistics, Volume 4.) London : OUP 1977
2. Anastasi, A., *Psychological Testing*. N.Y. : The Macmillan Co., 1965
3. Carroll, J.B., "Foreign Language Testing : Will the persistent problems Persist?" in *Testing in Second Language Teaching New Dimensions* by M.C. O'Brien (ed.), Ireland, ATESOL, 1972.
4. Clark, J.L.D., *Foreign Language Testing: Theory and Practice*, Pa: The Center for Curriculum Development, 1972
5. Cronbach, L.J., *Essentials of Psychological Testing*, N.Y.: Harper and Brothers, 1960
6. Davies, A., "Language Proficiency Testing and the Syllabus" in *Testing in Second Language Teaching: New Dimensions* by M.C. O'Brien (ed.), Ireland, ATESOL, 1972
7. Davies, A. and Widdowson, H.G. "Reading and Writing" in *Techniques in Applied Linguistics*, London: OUP 1974
8. Davies, A., *Language Testing Symposium A Psycholinguistic Approach*, London : OUP 1976
9. Davies, A., "Survey article" in *Language and Linguistics*, 1979.
10. Jakobovits, L.A., *Foreign Language Learning: A Psycholinguistic Analysis of the Issues*, Rowley, Mass: Newbury House, 1970
11. Lado, R., *Language Testing*, London: Longman, 1977. (Ninth Edition)
12. Morrow, K., *Techniques of Evaluation for a Notional Syllabus*, University of Reading, 1977.
13. Rivers, W.M., "Testing and Student Learning" in *Testing in Second Language Teaching: New Dimensions* by M.C. O'Brien (ed.), Ireland, ATESOL, 1972