The Relationship between Verbal Working Memory, Self-assessment, and EFL Oral Ability

Yi-Chun Christine Yang1, Ph.D. (楊逸君)
Shou-Der David Tseng2, Ph.D (曾守得)

ABSTRACT
Verbal working memory has been central in SLA research because of its role in language learning. Studies show that verbal working memory is related to students’ perceptions of teacher’s corrective feedback such as recasts, for example. However, little research has been conducted as to the relationship between EFL students’ verbal working memory, self-assessment, and EFL oral ability. This study examined whether there was any relationship between the three factors. 28 senior English majors participated in this study. The methods included a listening span test to measure students’ current level of verbal working memory, an oral presentation, and students’ self-assessment of their performance afterwards. A rubric from Chen (2008) was utilized for students’ self-assessment and teacher’s evaluation. The results showed that there was no correlation between students’ verbal working memory and their self-assessment. Students with high scores of the listening span test scored variously in their self-assessment. However, students’ self-assessment was highly correlated with teacher’s evaluation. Although students’ verbal working memory cannot function as a predictor of students’ self-assessment, the results of the listening span test still provide language teachers with perspectives on students’ ability to process or storage verbal or written language for production later. The majority of the students obtained more than 10 points out of the 20 full score which also entails that senior students’ verbal working memory has reached a certain level. Furthermore, senior students can self-assess their current language ability more accurately according to the match between their self-assessment and teacher’s evaluation.

Key words: self-assessment, verbal working memory, teacher evaluation, oral English

Introduction
Theories in second language acquisition have been proposed in the past few decades and research in SLA has been attempting to explore and even identify probable learning paths so as to provide language learners with principles in making foreign or second language learning more effective (e.g., Krashen, 1985; Schmidt, 1990; Swain, 1995). However, due to the complexity of language learning which may have been the result of both implicit and explicit processing of language input along with other factors such as sociolcultural influence and learners’ individual differences, many mysteries are still unsolved. Verbal working memory or working memory, one of individual differences related to language learners’ aptitudes, has been prominent in understanding how well learners can learn a second/foreign language. Research in the field shows that verbal working memory is connected to learners’ attainment of language input and the amount of the input that they receive that can be internalized into their long-term memory. (e.g., Ardila, 2003; Baddeley, 2003; Mackey and Sachs, 2012; Papagno, Valentine, and Baddeley, 2004; Wright, 2012).

Working memory, as Baddeley (2003) defined, “involves the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (p. 189) and verbal working memory refers to temporary storage and manipulation of verbal information (Baddeley, 1986)(cited in Acheson and MacDonald, 2009,
As such, learners’ ability to utilize the system was also connected to that of their self-assessment because when self-assessing one’s performance, learners have to perceive what they have spoken and storage the utterances in their verbal working memory system for later reflection and evaluation. Harris and Brown (2013) claim that “self-assessment …have been shown to engage and empower students, develop…self-regulation and metacognition.” (p. 101). Self-assessment, hence, can foster EFL students’ metacognition, which also entails their higher level of awareness and devotion in monitoring and evaluating their language learning. Additionally, students are able to self-evaluate their learning progress though self-assessment has not been regarded as a formal assessment instrument in language classrooms (Oscarson, 2013). As can be seen, self-assessment is associated with EFL students’ levels of verbal working memory and it is likely that students with better working memory may be more accurate in self-assessing their oral ability. Additionally, recent studies on verbal working memory suggest that working memory plays a critical role in L2 acquisition (Revesz, 2012). Thus, this study was to examine the match or mismatch between tertiary EFL learners’ verbal working memory and their self-assessment of the oral ability. Additionally, the accuracy of students’ self-assessment was further validated by teacher’s evaluation when comparing with learners’ levels of verbal working memory.

**Verbal working memory**

In 1960s Atkinson and Shiffrin (1968) proposed that the short-term memory (STM) system which has limited capacity is critical in the internalization of the input in the long-term memory. This two-system model suggested that this short-term storage memory also functioned as a working memory for other complex activities to proceed, such as reasoning. However, this proposition has drawn a lot of criticism. Later in the 1970 and the 80s, researchers brought forth findings to speak against this two-component model, stating that patients with impaired STM were still able to do many cognitive tasks, such as comprehending people’s speech or even run a shop. Baddeley and Hitch (1974) postulated the multi-component working memory, arguing that short-term memory system was not equal to working memory, in which it contained three components: visuospatial sketchpad, central executive, and phonological loop. Furthermore, Baddeley, in his review of working memory and language (2003), indicated that when the performance of the phonological loop was impaired, it disrupted foreign but not native language learning.

Research on working memory or verbal working memory has been focusing on whether it contributes to L2 acquisition of different language skills or aspects (O’Brien, Segalowitz, Freed, and Collentine, 2007; O’Brien, Segalowitz, Collentine, and Freed, 2006; Osaka and Osaka, 1992), particularly grammatical structures and vocabulary learning (Masoura and Gathercole, 1999, 2005; Papagno and Vallar, 1992; Revesz, 2012), and older learners (Mackey and Sachs, 2011). The results suggest that working memory and language aptitude are closely related and its noticing feature played a crucial role L2 acquisition (Abrahamsson and Hyltenstam, 2008; DeKeyser, 2000, 2010; Robinson, 2005; Sawyer and Ranta, 2001). One of the main findings from these studies is that verbal working memory plays a predictive role in second or foreign language learning. Baddeley (2003) proposed a three-component model explicating that working memory is not the same as the short-term memory system. Table 1 shows the model:
In this model, the phonological loop plays a permanent role in language learning in accordance with previous research findings (Baddeley, Papagno, and Vallar, 1988) and it seems that the loop “can be a useful aid in learning new words” (Baddeley, 2003, p. 194). The phonological loop is especially closely associated with L2 acquisition and it includes two subcomponents: a temporary storage system and a subvocal rehearsal system. The information received in the temporary storage cannot be retained unless it is refreshed by the subvocal rehearsal system when the rehearsal system also concurrently registers the visual information within the store of the name of the item. Hence, it is concluded that students’ working memory, particularly verbal working memory is critical to their success in language learning.

As have been abovementioned, research in working memory has been concerned with language learning of different language areas, learners of different ages such as old learners (Mackey and Sachs, 2012) and those who have difficulty learning native or foreign languages have also be examined (e.g., Ardila, 2003; Holmes, Gathercole, and Dunning, 2010; Masoura, 2006). Weissheimer and Mota (2009) conducted a study about EFL students’ individual differences in working memory and how the differences may have been associated with their oral production in L2. The results show that the speaking span test could predict students’ fluency and complexity in spoken English. The speaking span test used in this study, though not exactly the same as the listening span test invented in the current study, is also based on a similar theoretical framework postulated by Daneman and Carpenter (1980) and Daneman (1991). It is expected that through the administration of the listening span test, the researcher in the current study is able to understand EFL college students’ current levels of verbal working memory and how the cognitive activity contributes to our knowledge regarding students’ individual differential ability to store and retrieve what they produce orally.

Self-assessment

Self-assessment has been widely employed in language learning of different areas as an alternative to grading (Borwn and Abeywickrama, 2010). Though there are criticism with respect to the subjectivity essentially involved in self-assessment, Oscarson (2013) argues that self-assessment, with a careful designation of specific measuring criteria, can still be reliable and valid for the purpose of formative assessment and “should be viewed...as much a part of the process of learning as a mechanism for assessment.” (p. 4). Studies in self-assessment have been about the comparison between students’ self-assessment and teacher’s assessment to see whether students’ perceptions of their learning correspond to teachers’ evaluation of their learning progress (e.g., Brantmeier, 2005, 2006, 2012; Brantmeier and Vanderplank, 2008; De Saint-legar, 2009; Hargan, 1994; Krausert, 1991; LeBlanc and Painchaud, 1985; Oscarson, 1978; Ross, 1998; Wan-a-rom, 2010). Chen (2008) conducted a study to investigate EFL learners’ learning of self-assessing their oral ability and how their self-assessment became matched with teacher’s assessment after a period of
training in executing self-assessment in a class. Chen (2008) argues that studies on self-assessment place too much emphasis on students’ ability to evaluate or predict their language performance without considering self-assessment as a learning tool. Then she indicated factors that influence the effectiveness of self-assessment included clear criteria, training, intervention and feedback, and sufficient practice (Chen, 2006; Orsmond et al., 2000; Patri, 2002; Stefani, 1998; Taras, 2001) (p. 239). Among these factors, training, intervention and feedback, and practice are beneficial to the enhancement of students’ cognitive processing in evaluating their language performance so as to score more accurately and to be more compatible with the authority (teacher) assessment. However, learners’ verbal working memory may as well affect their perception of teachers’ training, intervention and feedback, and the further practice of self-assessment implicitly but substantially. Likewise, EFL learners’ ability to accurately self-assess their oral ability may have been connected to their verbal working memory albeit that the emphasis of studies on the notion has been mainly about how self-assessment assists in language learning. The current study assumed that learners with better verbal working memory may be able to more precisely self-assess their oral ability and in turn perform better in the speaking tasks. Thus, to understand tertiary EFL learners’ verbal working memory and how it has affected their self-assessment of their own oral ability, this study investigates the following research questions:

1. What are tertiary EFL learners’ current verbal working memory levels?
2. What is the match or mismatch between tertiary EFL learners’ verbal working memory and their self-assessment of the oral ability?
3. What is the match or mismatch between students’ self-assessment and teacher’s assessment in relation to learners’ verbal working memory?

Method

Participants

Twenty-eight tertiary EFL students who were senior English majors taking Language Testing and Assessment class participated in this study. Their current proficiency level approximately ranged from high intermediate to advanced level. Before taking this class, they would have taken most of the requirements and electives. This class was an elective so that this cohort of students were composed of students from three different classes, which entails that students’ English proficiency may have varied to some degree but not in a great deal.

Instruments

A demographic questionnaire to obtain students’ information and learning experience was invented to collect the data to understand other factors that brought forth students’ current learning situations. Then a listening span test based on the concept of Daneman and Carpenter (1980) was utilized to measure students’ working memory capacity. Students had to listen and repeat directions and the test was known to be cognitively demanding yet was designed with simple English words so as not to confound the WM test with students’ current language proficiency (Wright, 2012). There were five sets of sentences, including direction words in each sentence. Students were asked to repeat the direction words and then to recall as many sentences as they could remember. There were 2 seconds intervals between each sentence. The sentences were 9 to 13 words long and will be in English although working memory is language independent, which implies that the language used in the test would not affect the reliability of the test (Osaka and Osaka, 1992). As Wright (2012) addressed in her study investigating the effects of working memory on the accuracy and the fluency of
forming questions in spoken English, “Greater capacity on this span task has been found to correlated with higher scores in general measures of L2 proficiency”\( (p.\ 5) \). Put differently, students’ accuracy on self-assessment may as well interrelated with their test results of the span test. Furthermore, research findings also suggest that greater working memory entails the facilitation of word retrieval and utterance formation in L2 speech (Fortkamp, 1999; Kormos, 2006). The span task was conducted at the beginning of the semester. The full score of the listening span test was 20.

To elicit learners’ performance in spoken English, the researchers requested the learners to do one oral presentation during the semester as the main source of data to collect in terms of students’ language performance and whether their self-assessment became equivalent with teacher’s assessment. Chen’s (2008) grading rubrics was adopted in this study for both students’ self-assessment and for the instructor to evaluate learners’ oral presentations. Table two presents the procedure of the experiment in this study. Table 1 illustrates the procedure for data collection.

**Table 1**

*The procedure of data collection and analysis*

<table>
<thead>
<tr>
<th>Instruments/approached</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic questionnaire</td>
<td>● Students were asked to complete the demographic questionnaire in the second session of the class.</td>
</tr>
<tr>
<td>Verbal working memory test</td>
<td>● Students were asked to do the span test after the third or the fourth session in a computer lab during the semester.</td>
</tr>
<tr>
<td>Oral presentation</td>
<td>● Students did the oral presentation.</td>
</tr>
<tr>
<td>Teacher’s assessment</td>
<td>● The instructor evaluated students’ performance and compared students’ self-assessed grades with her own assessment.</td>
</tr>
</tbody>
</table>
| Data analysis          | ● The instructor collected all of the data obtained through these instruments and approaches to analyze the relationship between students’ verbal working memory, self-assessment, and teacher assessment.  
                        | ● Students’ scores of self-assessment and teachers’ scores were analyzed using Pearson Production Coefficient to see whether there was any correlation between these variables. |

**Results**

Among the 28 students in this class, 13 students obtained 15 points out of the 20 points on the listening span task. None of the 28 students got the full score on the test. One very interesting phenomenon after analyzing the test scores is that some students did not gain many points on these short sentences but did obtain several points on the longer sentences. Table 2 presents the means and the standard deviations of the scores from the three sources.

**Table 2**
Means and standard deviations of verbal working memory (VWM) scores, students’ self-assessment (SA) and teacher assessment (TA)

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWM score</td>
<td>13.5357</td>
<td>1.68835</td>
</tr>
<tr>
<td>SA</td>
<td>79.4643</td>
<td>6.11302</td>
</tr>
<tr>
<td>TA</td>
<td>82.1429</td>
<td>11.37481</td>
</tr>
</tbody>
</table>

Though the means from students’ self-assessment and teacher assessment are close, the standard deviations between the two illustrate that teacher assessment is more able to distinguish students of different oral abilities. However, the mean score of students’ self-assessment is lower than that of teacher’s assessment. It is likely that EFL students have had the tendency of grading themselves in self-assessment whereas teachers may have been more lenient when scoring students’ oral performances.

According to the results of the working memory test, it is suggested that almost half of the students in this class have quite high levels of verbal working memory. Only one student got 9 points out of the 20 full score in the test whereas the rest of the class scored more than 10 points (M=13.53). Table 3 shows the correlation between verbal working memory scores, students’ self-assessment, and teacher assessment.

Table 3
Correlations-Verbal working memory scores (VWM scores), self-assessment (SA), and teacher assessment (TA)

<table>
<thead>
<tr>
<th>VWM scores</th>
<th>SA</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWM scores</td>
<td>1</td>
<td>-.154</td>
</tr>
<tr>
<td>SA</td>
<td>-.154</td>
<td>1</td>
</tr>
<tr>
<td>TA</td>
<td>-.035</td>
<td>.542**</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed); n=28.

As have been presented in table 3, there was no correlation between students’ verbal working memory and their self-assessment, which did not reject one of the null hypotheses in this study. However, there is a quite high correlation between students’ self-assessment and teacher’s assessment of students’ presentations (r=.542, p<.01), and the result also corresponds to Chen’s (2008) finding. In other words, students’ self-assessment of their oral presentations is closely associated with teacher’s evaluation and it can be seen that students are able to objectively measure their performance as their language ability has reached a certain level and they may have been more aware of their current language proficiency.

According to the results, students who tended to self-assess their own performance lower had higher scores on the listening span test. Thus, it is suggested that students’ verbal working memory may not predict students’ self-assessment. One conclusion which can be reached in the current study is that students’ verbal working memory is independent of their ability of self-assessment. Furthermore, there is no correlation between students’ working memory test scores and teacher’s evaluation.

Discussion

With regard to the findings of the study, EFL students’ working memory was not
related to their self-assessment. However, students’ self-assessment was highly correlated with teacher’s evaluation which implies that self-assessment could be used as a tool to help teachers understand students’ current language proficiency in addition to teacher’s own assessment. In addition, the application of self-assessment enables students to have a grasp of their current language ability and know what aspects of language skills they still need to improve. Put differently, self-assessment functions as a reminder to raise students’ awareness of their oral ability at present and provides them with more perspectives regarding what they should work harder in oral English. Although the result of the listening span test did not show significant correlation with students’ self-assessment scores, the test revealed important information. For example, students who had high scores in the listening span task did not perform well as some students who scored lower in the span test. Additionally, students who did not get points in recalling short sentences were able to recall longer sentences. Hence, it is likely that the sentences used in the span task needed to be revised to make these long sentences harder so as to explore whether students’ ability to recall long sentences instead of short ones was derived from their cognitive processing ability or their verbal working memory.

**Pedagogical implications**

First, teachers can use self-assessment as a tool to obtain more insights concerning students’ own perceptions of their current language ability as well as to understand the match or the mismatch between students’ self-assessment and teacher’s evaluation.

Second, using self-assessment in a language classroom can also enhance students’ motivation because when students are more conscious of their strengths and weaknesses in language learning, they will be more able to seek for the solutions to solve the problems. For further studies, a questionnaire after the experiment is conducted can be invented to understand how self-assessment helps student to learn the target language and what they can do with the results of the assessment. Furthermore, in the process of self-assessing their own performance, they could learn how to judge themselves more objectively.

Third, using working memory test, though not helping teachers predict students’ current language ability, offers teachers more insights into student’s current cognitive processing capacity or their verbal working memory. Furthermore, it can assist teachers in finding how this factor may have affected other areas of students’ learning, such as pronunciation or lexis.

**Conclusion**

The findings of the study indicate that EFL college students’ verbal working memory is independent from their self-assessment of L2 speech whereas students’ self-rating is highly correlated with teacher’s assessment. Nevertheless, due to the small sample size of the study (N=28), the generalizability of the results is limited. It is also suggested that students’ utterances be further analyzed in terms of speech production in various types so as to have a better understanding of the relationship between students’ oral performance and sub-scores of the listening span test. Because of this very skewed spread of the scores of the listening span test, the sentences in the listening span test should be revised to retest their validity and reliability in future research.

**References**


Harris, L. R., & Brown, G. T. (2013). Opportunities and obstacles to consider when using peer-and self-assessment to improve student learning: Case studies into teachers’ implementation. Teaching and Teacher Education, 36, 101-111.


**Appendix**

**Oral presentation grading rubrics** (Adapted from Chen, 2008, p. 259)

<table>
<thead>
<tr>
<th>Scoring criterion/level</th>
<th>Excellent (90% &amp; above): Performed at a very high level, i.e. at a level that makes it exceptional for the class</th>
<th>Good (80%–89%): Performed at a high level, i.e. at a level that clearly exceeds competency</th>
<th>Fair (70%–79%): Performed without being exceptional in any way but thought of as competent.</th>
<th>Ok (60%–69%): Performed at a minimally acceptable level with flaws that are not serious enough to merit a failing mark.</th>
<th>Poor (below 60%): Performed at a very low level, without showing any positive quality. Having serious flaws.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content (35%) Language (35%)</td>
<td>35-30</td>
<td>29-24</td>
<td>23-18</td>
<td>17-13</td>
<td>13-</td>
</tr>
<tr>
<td>Delivery (20%)</td>
<td>20-18</td>
<td>17-15</td>
<td>14-12</td>
<td>11-9</td>
<td>8-</td>
</tr>
<tr>
<td>Manner (10%)</td>
<td>9-10</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Content</td>
<td>Wonderful, surprising, interesting, fascinating, flat, boring, not organized, not having a point</td>
<td></td>
<td>Delivery</td>
<td>Fluent, good-guidance, with expressive voice, with various gestures, with visual aids, with eye contact with the audience, average, not lively at all, dull</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Precise, error-free, clear pronunciation, appropriate, good use of transitions, poor grammar, unacceptable use of words, lack of necessary vocabulary</td>
<td></td>
<td>Manner</td>
<td>Calm, polite, graceful, dramatic, appropriate, no indication of nervousness or irrelevant body language</td>
<td></td>
</tr>
</tbody>
</table>