



Developing English Speaking Skills of Grade 2 Learners through Captioned and  
Keyword-only Pictures

Nittaya Manuy

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of  
Master of Arts in Teaching English as an International Language

Prince of Songkla University

2016

Copyright of Prince of Songkla University

T

เลขที่	PE 1749	N58	2016
Bit. key	412943		
/ 03 ก.ค. 2560 /			

Thesis Title            Developing English Speaking Skills of Grade 2 Learners through  
Captioned and Keyword-only Pictures

Author                    Miss Nittaya Manuy

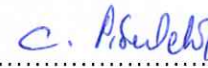
Major Program        Teaching English as an International Language


---

Major Advisor

  
.....  
(Asst. Prof. Dr. Chonlada Laohawiriyanon)


Examining Committee:

  
.....Chairperson  
(Assoc. Prof. Chaleosri Pibulchol)

  
.....Committee  
(Assoc. Prof. Dr. Adisa Teo)

  
.....Committee  
(Asst. Prof. Dr. Chonlada Laohawiriyanon)

The Graduate School, Prince of Songkla University, has approved this  
thesis as partial fulfillment of the requirements for the Master of Arts Degree in Teaching  
English as an International Language

  
.....  
(Assoc.Prof.Dr.Teerapol Srichana)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

.....*Chonlada L.*.....Signature  
(Asst. Prof. Dr. Chonlada Laohawiriyanon)  
Major Advisor

.....*Nittaya Manuy*.....Signature  
(Miss. Nittaya Manuy)  
Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

.....Nittaya Manuy.....Signature

(Miss. Nittaya Manuy)

Candidate

ชื่อวิทยานิพนธ์	การพัฒนาทักษะการพูดภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 2 โดยใช้รูปภาพที่มีคำบรรยายได้ภาพและรูปภาพที่มีเฉพาะคำสำคัญได้ภาพ
ผู้เขียน	นางสาวนิตยา มาน้อย
สาขาวิชา	การสอนภาษาอังกฤษเป็นภาษานานาชาติ
ปีการศึกษา	2558

### บทคัดย่อ

งานวิจัยนี้ใช้ทฤษฎีรหัสคู่และทฤษฎีภาระการทำงานของสมองเพื่ออธิบายผล การศึกษาของการใช้รูปภาพที่มีคำบรรยายได้ภาพและรูปภาพที่มีเฉพาะคำสำคัญได้ภาพต่อ ความสามารถด้านการพูดภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 2 รวมถึงศึกษาปัจจัยที่มี ผลต่อความสามารถทางการพูด โดยกลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 2 ของ โรงเรียนประถมศึกษาของรัฐแห่งหนึ่งในจังหวัดสงขลา ภาคใต้ของประเทศไทย จำนวน 49 คน หญิง 25 คน และชาย 24 คน ซึ่งได้มาจากการเลือกแบบเจาะจง และสุ่มให้กลุ่มหนึ่งเป็นกลุ่มที่ เรียนภาษาอังกฤษโดยใช้รูปภาพที่มีคำบรรยายได้ภาพและอีกกลุ่มหนึ่งเป็นกลุ่มที่เรียน ภาษาอังกฤษโดยใช้รูปภาพที่มีเฉพาะคำสำคัญได้ภาพ การเก็บรวบรวมข้อมูลได้มาจากการ ทดสอบก่อนเรียน หลังเรียน ความคงทน และการสัมภาษณ์ถึงโครงสร้าง เกณฑ์ที่ใช้ในการ ประเมินความสามารถในการพูดคือเกณฑ์การประเมินความสามารถทางการพูดของ CEFR ผลการวิจัยพบว่า กลุ่มที่เรียนภาษาอังกฤษโดยใช้รูปภาพที่มีคำบรรยายได้ภาพมีการพัฒนา ทางด้านการพูดเพิ่มขึ้นอย่างมีนัยสำคัญ จากระดับอ่อน เป็นระดับปานกลาง ( $Z = -2.236, p < .05$ ) โดยเฉพาะในด้านคำศัพท์และด้านความถูกต้องในการใช้ภาษา แต่กลุ่มที่เรียนภาษาอังกฤษโดยใช้ รูปภาพที่มีเฉพาะคำสำคัญได้ภาพมีความสามารถทางการพูดเพิ่มขึ้นจากระดับอ่อน เป็น ระดับปานกลาง และระดับเก่ง โดยมีทักษะย่อยที่เพิ่มขึ้นสูงที่สุดคือ ด้านการมีปฏิสัมพันธ์ ด้าน ความคล่องแคล่วในการใช้ภาษา และด้านคำศัพท์ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ( $Z = -3.145, p < .01$ )

Thesis Title	Developing English Speaking Skills of Grade 2 Learners through Captioned and Keyword-only Pictures
Author	Miss Nittaya Manuy
Major Program	Teaching English as an International Language
Academic Year	2015

### ABSTRACT

Drawing upon Dual Coding and Cognitive Load Theories, the purposes of this study were to examine the effects of using captioned and keyword-only pictures on Grade 2 learners' English speaking abilities and to examine factors affecting their English speaking abilities. The participants were recruited from 49 Grade 2 learners (25 female and 24 male) in a public primary school in Songkhla, southern Thailand. They were selected by purposive sampling and randomly assigned into two experimental groups: captioned picture group and keyword-only picture group. The data were collected from pre-test, post-test, retention test and semi-structured interview. Following CEFR oral assessment criteria, the results revealed that in the captioned pictures group, Grade 2 learners' English speaking abilities were significantly increased from poor to fair specifically in 'range' and 'accuracy' ( $Z = -2.236$ ,  $p < .05$ ). For keyword-only picture group, it was found that the learners' speaking skills increased from poor to fair and good at the significant difference 0.01 ( $Z = -3.145$ ,  $p < .01$ ). The three highly improved sub-skills were 'interaction', 'fluency' and 'range'.

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Asst Prof. Dr. Chonlada Laohawiriyanon who was extremely helpful and offered invaluable assistance, support, and guidance. Her willingness to motivate me contributed tremendously to my thesis. Deepest gratitude is also due to the members of the supervisory committee, Assoc. Prof. Chaleosri Pibulchol and Assoc. Prof. Dr. Adisa Teo for giving their precious time to carefully examine my thesis. Without their knowledge and assistance, this study would not be successful.

My grateful thanks are extended to Ms. Putthithada Chomchei for her help in statistical data analysis, Dr. Kathleen Nicoletti for kindly editing my lesson plans, Mrs. Sammireh Nagaratnam for editing my papers and giving good suggestions. I would like to extend my gratitude to all the lecturers of the Teaching English as an International Language program for their valuable lectures and support.

Special thanks also go to my close graduate friend, Ms. Titinan Ardhan, who is always there for me. I would like to express my sincere gratitude to my bosses, Mr. Simon Alaster Jack and Mrs. Mueanfan Jack for giving me time to collect research data and to forty-nine Grade 2 students for their participation.

Finally, and most important of all, I would like to express my love and gratitude to my beloved family for their understanding and endless love throughout the course of my study. This study would not have been successful without them.

Nittaya Manuy

## CONTENTS

ABSTRACTS.....	v
ACKNOWLEDGEMENTS.....	vii
CONTENTS.....	viii
LIST OF TABLES AND FIGURES.....	x
LIST OF PAPERS.....	xii
LIST OF ACCEPTANCE.....	xiii
<b>A SYNTHESIS REPORT</b>	
1. Introduction.....	1
2. Research Methodology.....	4
2.1 Target Population and Subjects.....	4
2.2 Instruments.....	5
2.3 Data Collection Procedures.....	10
2.4 Data Analysis Procedures.....	12
3. Findings.....	12
4. Discussion and Conclusion.....	21
5. Recommendations.....	23
<b>REFERENCES.....</b>	<b>24</b>
<b>APPENDICES</b>	
Appendix A Scope of Lesson Plans.....	27



## CONTENTS (CONTINUED)

## ACCEPTED MANUSCRIPTS

PAPER 1	Effects of Using Captioned Pictures on Grade 2 Learners' English Speaking Abilities.....	36
PAPER 2	Effects of Using Captioned and Keyword-only Pictures on Grade 2 Learners' English Speaking Abilities.....	45
VITAE.....		65

## LIST OF TABLES AND FIGURES

### FIGURE

#### A SYNTHESIS REPORT

1	Collecting Data Procedure.....	11
---	--------------------------------	----

### TABLE

#### A SYNTHESIS REPORT

1	Oral assessment criteria grid adapted from CEFR and Hirai & Koizumi (2013).....	7
2	Percentages of participants in the two groups who achieved different level of English Speaking Abilities in Pre, Post and Retention Tests.....	13
3	Wilcoxon Signed Ranks Test summary of English speaking abilities of the two groups.....	14
4	Mann-Whitney U Test summary of English speaking abilities between the two groups.....	15
5	The summary of interview in high and low achievers in CP and KP groups.....	17

#### PAPER 1

1	Wilcoxon Signed Ranks Test summary of pre-and post-treatment tests.....	40
---	---	----

#### PAPER 2

1	Percentages of participants in the two groups who achieved different level of English speaking abilities in pre, post and retention tests.....	55
---	---	----

## LIST OF TABLES AND FIGURES (CONTINUED)

2	Wilcoxon Signed Ranks Test summary of English speaking abilities of the two groups.....	56
3	Mann-Whitney U Test summary of English speaking abilities between the two groups.....	57

## LIST OF PAPERS

This thesis is based on the following papers, which will be referred to in the text:

Manuy, N., & Laohawiriyanon, C. (2016). Effects of using captioned picture on Grade 2 learners' English speaking abilities. *Proceedings of ICHiSS 2016: 8<sup>th</sup> International Conference on Humanities and Social Sciences*. National Defence University of Malaysia. (pp 217-225). Selangor, Malaysia.

Manuy, N., & Laohawiriyanon, C. (2016). Effects of using captioned and keyword-only pictures on Grade 2 learners' English speaking abilities, *Journal of International Studies*. (manuscript submitted for publication).

**8** **ICHISS 2016**  
INTERNATIONAL  
CONFERENCE ON  
HUMANITIES AND  
SOCIAL SCIENCES



**HOTEL ROYAL CHULAN  
DAMANSARA, SELANGOR,  
MALAYSIA  
MAY 27 - 29, 2016**

Our reference : UPNM (FPP)10.02/01 ( 144 )  
Date : 1 MAY 2016

Prof. / Assoc. Prof. / Dr. /Mr. /Mrs,

**Nittaya Manuy & Chonlada Laohawiryanon**

**Paper ID: A27**

**CONFIRMATION OF ABSTRACT ACCEPTANCE FOR 8<sup>th</sup> ICHISS 2016**

Thank you for your kind interest in participating in the 8<sup>th</sup> International Conference on Humanities and Social Sciences (ICHISS) 2016 "Innovation in Humanities and Social Sciences: Opportunities and Challenges" to be held on 27 – 29 May 2016 at The Hotel Royale Chulan Damansara, Selangor, Malaysia.

We are pleased to inform you that your abstract is accepted to be presented in the 8<sup>th</sup> ICHISS 2016.

The panel of evaluators is of the opinion that your paper does contain information linking your paper to one of the sub-themes of our conference. It will be better and more relevant if it shows linkage to the main theme of the conference which is "Innovation in Humanities and Social Sciences: Opportunities and Challenges". We hope, if possible, you will try to improve on this when you write your full paper.

Kindly submit your full paper via online at: [submission@upnm.edu.my](mailto:submission@upnm.edu.my) with the subject of e-mail: 'Your Paper ID: ICHISS 2016' before 7<sup>th</sup> May 2016. You are advised to edit and adhere to the format stated in our website when submitting your full paper.

For registration fees arrangements and payment method, please refer to our conference website at: <http://ichiss.upnm.edu.my/index.php/call-for-abstract/registration-fees>. Kindly take note on the advantage of the 'early bird' registration fee promotion.

Should you have any inquiries or require further clarifications, kindly contact us at: [amnah@upnm.edu.my](mailto:amnah@upnm.edu.my) and [norlaila@upnm.edu.my](mailto:norlaila@upnm.edu.my).

Thank you for your kind cooperation and attention on the above matter. We look forward to see you at the conference.

**"DUTY, HONOUR, INTEGRITY"**

With best regards,

**Organizing Committee of 8<sup>th</sup> ICHISS 2016**

จาก: ดร.ศุภชัย แจ่มใจ <supachai.je@phuket.psu.ac.th>

ชื่อเรื่อง: Re: ส่งบทความวิจัยเพื่อตีพิมพ์ในวารสารวิเทศศึกษา

วันที่: 16 พฤษภาคม 2559 15:51:10

เรียนคุณนิตยา มาน้อย

ตามที่ท่านได้ส่งบทความเพื่อรับการพิจารณาตีพิมพ์ในวารสารวิเทศศึกษา เรื่อง "Effects of Using Captioned and Keyword-only Pictures on Grade 2 Learners' English Speaking Abilities" ทางกองบรรณาธิการวารสารขอแจ้งให้ท่านทราบว่า กองบก.จะดำเนินการทบทวนผู้ทรงคุณวุฒิเพื่อประเมินบทความดังกล่าวจำนวน 2 ท่าน โดยระยะเวลาในขั้นตอนนี้ใช้เวลาประมาณ 3 เดือน และทางกองบก.จะแจ้งผลการประเมินบทความของท่านต่อไป

จึงเรียนมาเพื่อทราบ

ดร.ศุภชัย แจ่มใจ

บรรณาธิการบริหาร วารสารวิเทศศึกษา

คณะวิเทศศึกษา มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต

โทร 096-6469515

## 1. INTRODUCTION

Listening and speaking are two fundamental skills which children should master at a very young age before developing the abilities to read and write. In fact, children acquire their first language first through listening and then learn how to produce the language by imitating what they hear (Linse, 2005; Slattery & Willis, 2001). Similarly, it is believed that young L2 learners learn a language simply by absorbing what they hear and they are able to remember new language in "chunks". Children will have to rely on their language teachers to develop their listening and speaking skills (Pinter, 2006). To teach L2 to this group of learners, effective L2 input, mainly from teacher talk which incorporates language simplifications and repetitions has a crucial role to play. Teachers talk, therefore, is a good model of pronunciation and intonation. Teacher acting as a linguistic model coupled with effective gestures and other meaningful and interesting prompts are two important elements which would foster young L2 learners' listening and speaking abilities (Colon-Villa, 1997; Cook, 2000; Gibbons 2002; Linse, 2005). Well-established classroom routines conducted through simple tasks such as listening to stories or playing simple games can help young learners repeat what they hear in order to reinforce their language learning (Cameron, 2001).

Young learners are known to have short attention span, which lasts around 10-15 minutes, prefer physical activities, such as running, jumping, and dancing, learn by doing and playing, and benefit from repeating words, phrases, and sentences (Mackay, 2006; Slatterly & Willis, 2001). There are several teaching techniques designed to match the above characteristics and therefore are considered effective for this particular group of learners. First, storytelling is a popular technique that has been widely used because it can elevate their learning motivation (Rokhayani, 2012). Storytelling is useful in various ways. For example, it helps develop learning experiences of young children; it can stimulate learners' imagination from listening to stories; it

provides opportunities for them to exchange experiences with each other; and it helps develop learners' listening skills and concentration (Ellis & Brewster, 1991).

Among various teaching methods, it is evident that pictures have an important role to play in teaching young children, especially vocabulary, because they can motivate and capture learner's attention (Mansourzadeh, 2014; Wright, 1990). In the studies of Yoshii (2002) and Al-Ja Afari (2013), they have found that the use of pictures to teach vocabulary has positive effects on learners' attitude and vocabulary retention, while Rowe, Siverman and Mullan (2013) have found that the use of picture-word combinations enhanced four year-old learners' vocabulary knowledge. A study conducted with low proficiency learners reveals that pictorial aids assist learners to retain their knowledge (Yang & Chang, 2013).

Pictures can also be used to teach oral communication. Based on Bowen (1982), learners can describe pictures to their partners. Teachers can use pictures to encourage discussion in the classroom and motivate the learners to ask questions creatively based on the pictures. Alternatively, teachers can create interactive conversation using pictures as a prompt. If learners are interested in the details in the pictures, they can ask the teacher questions. Dobson (1992) also stated that the pictures can be good conversation starters and can create different discussions on various topics in the classroom, such as nature, food, classroom, and so on. There are many methods that can support the learners and the teachers to interact with each other by using pictures.

It can be said that the use of pictures in language classroom is supported by Dual Coding Theory (DCT) proposed by Paivio (cited in Clark & Paivio 1991). He maintained that effective learning takes place when learners receive bi-or multi-modal inputs through different sensory systems such as the visual (pictures, written texts) and the auditory (verbal). According to Clark and Paivio (1991) and Mayer (2009), different modalities of input would complement each other rather than interfere.



However, it is not always the case that bi/multi-modalities of messages would lead to better learning. Learners with different language proficiencies might have different reactions to such inputs. For example, young learners who begin to read might benefit more from listening only while adult learners would learn more effectively through reading and listening simultaneously (Sticht & James, 1984; Taylor, 2005).

However, there is a theory about information processing that seems to contradict to the DCT i.e. Cognitive Load Theory (CLT) proposed by Sweller (cited in Kalyuga, Chander & Sweller, 1999). It hypothesizes that the cognitive capacity in a learner's working memory is limited. If learners are required to perform heavy cognitive tasks, their ability to learn will be lowered (Kalyuga, Chander & Sweller, 1999). Three types of cognitive load are: 1) intrinsic, 2) extraneous, and 3) germane. Intrinsic cognitive load has to do with the complexity of content, being vocabulary or grammar. For learning to occur, learners should have sufficient prior knowledge. If instructional materials are unimportant, it is said to cause extraneous cognitive load in learners. Another source of this type of cognitive load is that the materials or inputs are presented in bi-or multi- modes and the messages are not highly relevant. This results in learners facing a cognitive overload situation which negatively affects learners' working memory, and therefore, comprehension ability. If teaching materials are at the right level of learners' ability and relevant to learning objectives, then germane cognitive load created by this combination would automatically enhance learners' schemas (Mayer, 2005).

Taken together that young learners have limited L2 ability due to limited linguistic exposure and that pictures only as against pictures with captions have been considered an effective teaching material to develop language ability especially vocabulary. For these reasons, this study aimed at using pictures to promote grade 2 learners' speaking ability. Two types of pictures (captioned and keyword-only pictures) were used to investigate if multi-modality input i.e. pictures, captions, and sounds, and pictures, keywords, and sounds had different effects on grade 2 learners' speaking ability. Two research questions were formulated as follows:

1. Are there any differences between the effects of using captioned-pictures and keyword-only pictures on Grade 2 learners' English speaking abilities? If so, how?
2. What are facilitating factors and barriers of using captioned-pictures and keyword-only pictures to teach young learners' speaking abilities?

## 2. RESEARCH METHODOLOGY

Quasi-experimental research design with two experimental groups was adopted to collect both quantitative and qualitative data. Quantitative data were obtained from three speaking tests: pre, post and retention while qualitative data were obtained from individual interviews with seven high and seven low achievers in each experimental group.

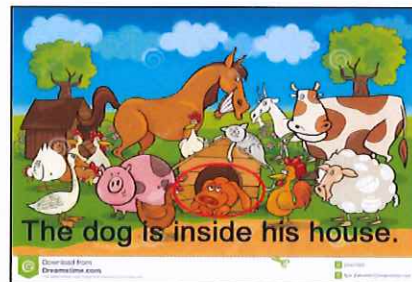
### 2.1 Target Population and Subjects

The study recruited 49 EFL Grade 2 students (25 female and 24 male), aged 7-9 on average. Three students were from Myanmar and 46 were Thais. They were in a public primary school in Songkhla, southern Thailand. The researcher collected the data in the second semester of the 2015 academic year. The participants represented a homogeneous group based on the pre-test scores i.e. all the participants in both groups were rated poor. They were chosen by purposive sampling and randomly assigned into two treatment groups: a group in which captioned pictures were used (n= 26) and a group in which keyword-only pictures were used (n = 23).

## 2.2 Instruments

The instruments used in this study were two types of pictures, speaking tests and a set of semi-structured interview questions.

### 2.2.1 Two types of pictures



Captioned Picture



Keyword-only Picture

Following the Basic Education Core Curriculum B.E. 2551 (A.D.2008), the researcher selected pictures which covered various topics such as numbers, animals, classroom items, sports, fruit and vegetable, weather, actions, food, body parts, days and seasons. The researcher made sure that the selected pictures were big enough for the whole class to see. Colors were attractive to children and details were clear (Bowen; 1982). All pictures were approved by three EFL experts prior to the experiment. They were piloted with students who had similar English proficiency in another primary school. The first type of picture included three modalities, i.e. pictures, captions, and sounds. The group that received this treatment was called captioned picture (CP). The second type of picture included 3 modalities, i.e. pictures, keywords, and sounds. This group was referred to as keywords-only picture (KP).

The techniques used by Lever and Sénéchal (2011) were adopted. The researcher teacher described the picture using basic English structure containing 5-7 words per sentence (Lutz & Huitt, 2003). For example, "The dog is inside his house". Each description was identical to the caption placed below the picture. The teacher

repeated each description several times so that the participants could repeat it after the teacher. After that, to initiate interactions, the teacher asked Wh-questions such as "Where is the dog?" or "What color is the dog?". Then, if the learners answered the questions correctly (The dog is inside his house.), the teacher would repeat that sentence again and asked the learners to repeat after her. If the learners provided wrong or incomplete answers (Dog inside house.), the teacher would model the correct answer and had learners repeat it immediately a few times. After that, both teacher and learners repeated the sentence several times together. Then, the teacher replaced a new caption and described the remaining part of the picture.

In short, the teacher used exactly the same teaching techniques to teach the two groups. The only difference was that the CP group saw the full captions while looking at the teacher and listening to the teacher's narration, whereas the KP group saw only the keyword such as "Dog" while looking at the teacher and listening to the teacher (The dog is inside his house). Finally, the participants were asked to retell the whole story by using connective words, such as "and" or "then".

### 2.2.2 Speaking tests



The same picture was used three times in the pre, post and retention tests. The pre and post tests were 15 weeks apart while the post and retention tests were only two weeks apart. Each participant was required to describe the picture in two

minutes. If he/she could not describe it, the teacher would elicit the answers by asking simple questions to assist him/her, for example, “ What is this? What are they doing?” All the tests were recorded individually for grading. Common European Framework of Reference (CEFR) was adapted to evaluate the participants’ speaking skills. To be specific, the five sub-skills of oral assessment criteria grid were employed: range, accuracy, fluency, interaction and coherence (University of Cambridge ESOL Examinations Research and Validation Group, 2009). Regarding the classification of the level of speaking skills, the researcher followed the idea of Hirai & Koizumi (2013) i.e. poor (0-20), fair (21-40), good (41-60), very good (61-80), and excellent (81-100), as seen in Table 1 below.

Table 1 Oral assessment criteria grid adapted from CEFR and Hirai & Koizumi (2013)

Levels	Range	Accuracy	Fluency	Interaction	Coherence
Excellent	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	Can initiate discourse, take his/her turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting other in, etc.	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some “jumpiness” in a long contribution.

Levels	Range	Accuracy	Fluency	Interaction	Coherence
Very good	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circumlocutions on topics such as family hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.	Can keep going comprehensibly, even through pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.	Can initiate, maintain and close simple face-to-face conversation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
Good	Uses basic sentence patterns with memorized phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.	Uses some simple structures correctly, but still systematically makes basic mistakes	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.	Can ask and answer questions and respond to simple statements. Can indicate when he/she following but is rarely able to understand enough to keep conversation going of his/her own accord.	Can link groups of words with simple connectors like "and", "but" and "because".
Fair	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Show only limited control of a few simple grammatical structures and sentence patterns in a memorized repertoire.	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.	Can link words with very basic linear connectors like "and" or "then".

Levels	Range	Accuracy	Fluency	Interaction	Coherence
Poor	Cannot use a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Cannot show limited control of simple grammatical structures and sentence patterns in a memorized repertoire.	Cannot manage very short, isolated, mainly pre-packaged utterances and cannot articulate familiar words, and to repair communication.	Cannot ask and answer questions about personal details. Cannot interact in a simple way.	Cannot link words with very basic linear connectors like "and" or "then".

Three EFL experts (one American and two Thais) were employed to obtain inter-rater reliability. To achieve this, the researcher teacher used a smart phone to video each participant while describing the picture. Each rater received a copy of the recording. They, then, rated each student performance, using the adapted assessment criteria grid. The results of the three raters were compared, and it was found that all three raters were unanimous in their judgment. Consequently, the three raters rated the remaining video clips. The results obtained from the three raters were averaged. Finally, the percentages were analyzed by using Wilcoxon Signed Ranks test and Mann- Withney U test.

### 2.2.3 Semi-structured interview questions

The purpose of the interview was to provide additional information to help interpret the quantitative data regarding factors contributing to the participants' speaking abilities as well as problems they encountered while studying. The participants in each experimental group were divided into two groups: high achiever and low achiever. Each group contained 7 participants, totaling 28 participants. Each participant was interviewed for 7-10 minutes. Also, the interview questions were in Thai to avoid language barriers. The interview questions were also approved by three EFL experts. The following is the set of questions used to elicit the participants' ideas about what helped and did not help in their learning how to speak English.

1. Did you look at captions/ keywords when you described pictures?
2. Did you get bored when the teacher asked you to repeat sentences many times?
3. Did you know some vocabulary before starting this class?
4. Did you learn to speak English before?
5. Were the lessons difficult?
6. Did you like the pictures used in each lesson? Why?
7. Did the captions/keywords below the pictures help you to speak English more? If yes, how?
8. Did the captions/keywords below the pictures help you learn more vocabulary? If yes, how?
9. When the teacher asked you many wh-questions, did the questions help you describe the pictures?
10. Do you have any suggestions about teaching methods?

### 2.3 Data collection procedure

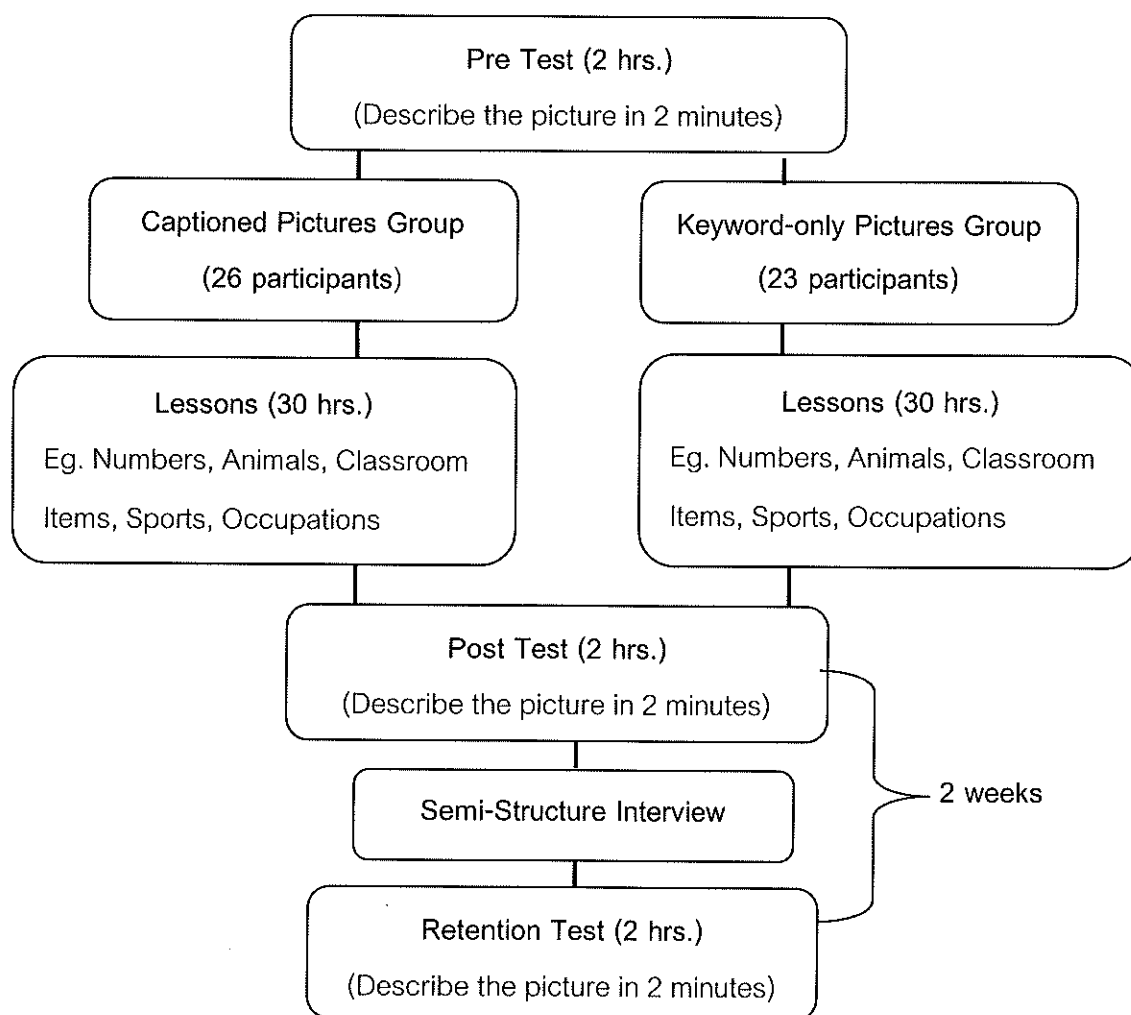
The data collection was carried out over the course of fifteen weeks. Prior to the experiment, each participant was required to take a pre-test by describing the given picture within two minutes. It took approximately two hours to have all of the research participants take part in the pre-test. This was done to assess their baseline speaking ability against which any change in oral production ability could be determined. Then, all the participants in the two intact groups were randomly assigned into a caption group (CP) and a keyword-only group (KP).

Regarding the treatment, the teaching technique employed in the study was based on Lever and Sénéchal's study (2011). The class met two periods a week and each period took fifty minutes. In total there were fifteen lessons (See appendix A); it took two periods to complete each lesson. After the completion of the teaching, the participants were required to take the post-test in which the procedure was identical to



that of the pre-test. The following week, seven high and seven low achievers from the CP and KP groups were selected to be interviewed, totaling twenty eight participants. The interview aimed at finding out the facilitating and hindering factors affecting their English speaking ability, such as contents, materials and learning processes. The interview was conducted in Thai and took 7-10 minutes for each interview. Two weeks after the post test, the participants were required to take the retention test, which was identical to the pre- and post-tests. Figure 1 below illustrates the data collection procedure.

Figure 1 Data collection procedure



## 2.4 Data analysis procedure

The data in this study were obtained from the participants' pre-, post, and retention test results, and their responses to the interviews. To answer the first research question, the test results were tallied to arrive at the percentages of participants whose performance was rated in different levels (poor, fair, and good). As the test results were presented in ordinal scale and there was also no normality, the researcher employed the Wilcoxon signed ranks test to examine the learners' English speaking ability in pre, post and retention test within group and the Mann-Whitney U test to analyze the data. The percentages were then analyzed to examine the differences of English speaking ability between the two groups.

To answer the second research question, qualitative data obtained from the interviews were analyzed for themes and then categorized to provide a further support in answering the research questions related to facilitating and hindering factors of learning how to speak English through the use of caption and keyword only pictures.

## 3. FINDINGS

### 3.1 Effects of using caption and keyword-only pictures on participants' speaking ability

To compare the results of the pre, post and retention tests between the two groups, descriptive statistics was used. Table 2 shows the percentages of participants in the two groups who achieved certain levels of English speaking ability in the pre, post and retention tests. The overall results reveal that 100% of the participants in both groups had poor speaking skills in the pre-test, suggesting the inability to use simple English to give details of people or concrete situations.

For the post test, the results reveal that the CP group had shown some apparent development in two sub-skills, namely, range and accuracy. To be specific, 35% of the participants reached fair level in vocabulary range and 27% of them

achieved fair level in accuracy. In comparison to the CP group, the KP group showed a more remarkable improvement than the CP group. That is, 52% of them obtained fair level and 9% reached good level in interaction. For 'range', 39% of the participants were in fair level and 26% obtained good level. Moreover, there were 39% of the participants who performed at the fair level and 17% at the good level in fluency. The results suggest that the participants who learned how to speak through looking at the keyword-only picture while listening to the teacher describing pictures could develop their vocabulary knowledge, answer questions, and were more fluent. However, they still could not use connective words, such as "and" or "then", to show the logical flow of their ideas. Regarding the retention test results, overall, there was a decline in the English speaking ability of the two groups of the participants, especially in the KP group. To be specific, those who reached good level in the post test were rated fair level in the retention test.

**Table 2** Percentages of participants who achieved different levels of English speaking ability in pre, post and retention tests

Skills	Levels	CP Group						KP Group					
		Pre-test		Post-test		Retention Test		Pre-test		Post-test		Retention Test	
		N	%	N	%	n	%	n	%	n	%	n	%
Overall	Poor	26	100	21	81	22	85	23	100	11	48	15	65
	Fair	0	0	5	19	4	15	0	0	6	26	8	35
	Good	0	0	0	0	0	0	0	0	6	26	0	0
Range	Poor	26	100	17	65	12	46	20	87	8	35	8	35
	Fair	0	0	9	<u>35</u>	14	54	3	13	9	<u>39</u>	15	65
	Good	0	0	0	0	0	0	0	0	6	<u>26</u>	0	0
Accuracy	Poor	26	100	19	73	20	77	23	100	12	52	15	65
	Fair	0	0	7	<u>27</u>	6	23	0	0	6	26	8	35
	Good	0	0	0	0	0	0	0	0	5	22	0	0
Fluency	Poor	26	100	21	81	23	88	23	100	10	43	17	74
	Fair	0	0	5	19	3	12	0	0	9	<u>39</u>	6	26
	Good	0	0	0	0	0	0	0	0	4	<u>17</u>	0	0

Interaction	Poor	26	100	21	81	21	81	23	100	9	39	12	52
	Fair	0	0	5	19	5	19	0	0	12	<u>52</u>	11	48
	Good	0	0	0	0	0	0	0	0	2	<u>9</u>	0	0
Coherence	Poor	26	100	21	81	24	92	23	100	13	57	15	65
	Fair	0	0	5	19	2	8	0	0	9	39	8	35
	Good	0	0	0	0	0	0	0	0	1	4	0	0

Table 3 below shows the results of Wilcoxon signed-rank test analysis to examine the English speaking of the participants within groups.

Table 3 Wilcoxon signed-rank test summary of English speaking ability of the each group

Skills	CP Group				KP Group			
	Test		Test		Test		Test	
	Pre and Post		Post and Retention		Pre and Post		Post and Retention	
	Z	Asymp. Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)
Overall	<u>-2.236*</u>	0.03	-.577	0.56	<u>-3.145**</u>	.00	<u>-2.352*</u>	.02
Range	-3.000**	0.00	-1.890	0.06	-3.286**	.00	-1.355	.18
Accuracy	-2.646**	0.01	-.447	0.65	-3.017**	.00	-1.890	.06
Fluency	-2.236*	0.03	-1.000	0.32	-3.314**	.00	-2.653**	.01
Interaction	-2.236*	0.03	.000	1.00	-3.557**	.00	-1.667	.10
Coherence	-2.236*	0.03	-1.342	0.18	-3.051**	.00	-1.342	.18

\* Significance at 0.05 and \*\* Significance at 0.01

As shown in Table 3, a similar pattern of improvements in English speaking ability within groups existed. That is, significant improvements in the post-test results of the CP and KP groups were found overall ( $Z = -2.236^*$  and  $-3.145^{**}$  respectively). The improvements in all sub-skills were also found in both experimental groups (CP, range =  $-3.000^{**}$ , accuracy =  $-2.646^{**}$ , fluency =  $-2.236^{**}$ , interaction =  $-2.236^*$ , coherence =  $-2.236^*$ ; KP, range =  $-3.286^{**}$ , accuracy =  $-3.017^{**}$ , fluency =  $-3.314^{**}$ , interaction =  $-3.557^{**}$ , coherence =  $-3.051^{**}$ ). Regarding the ability to retain their speaking ability, however, it was found that only the KP group retained the oral production ability overall. Moreover, fluency was the only speaking sub-skill that could be retained.

Table 4 demonstrates the results of the Mann-Whitney U test analysis to examine the between group differences in English speaking ability.

Table 4 Mann-Whitney U test summary of English speaking ability between the two groups

Tests	Skills	CP	KP	Z	Asymp. Sig. (2-tailed)	Effect Size
		Mean Rank	Mean Rank			
Pre-test	Overall	25	25	0	1.00	0
	Range	23.5	26.7	-1.881	0.06	-0.56
	Accuracy	25	25	0	1.00	0
	Fluency	25	25	0	1.00	0
	Interaction	25	25	0	1.00	0
	Coherence	25	25	0	1.00	0
Post-test	Overall	20.63	29.93	<u>-2.701**</u>	0.01	<u>-0.84</u>
	Range	20.44	30.15	-2.628**	0.01	-0.81
	Accuracy	21.92	28.48	-1.879	0.06	-0.56
	Fluency	20.33	30.28	-2.862**	0	-0.9
	Interaction	20.02	30.63	-3.039**	0	-0.96
	Coherence	22.12	28.26	-1.874	0.06	-0.56

Retention Test	Overall	22.77	27.52	-1.56	0.12	-0.46
	Range	23.69	26.48	-0.8	0.42	-0.23
	Accuracy	23.65	26.52	-0.896	0.37	-0.26
	Fluency	23.33	26.89	-1.299	0.19	-0.38
	Interaction	21.71	28.72	<u>-2.108*</u>	0.03	<u>-0.63</u>
	Coherence	21.88	28.52	<u>-2.324*</u>	0.02	<u>-0.7</u>

\* Significance at 0.05 and \*\* Significance at 0.01

As shown in Table 4, a different pattern of improvements in English speaking ability between the two experimental groups in the pre- and post-tests shows that the speaking ability of KP group significantly outperformed that of their counterparts. ( $z = -2.701^{**}$ ) with a 'large' effect size of -0.84. For sub-skills, 'range', 'fluency', and 'interaction' were the three areas that the KP group significantly surpassed the CP group with 'large' effect sizes ( $-2.628^{**}$ ,  $-2.862^{**}$ , and  $-3.039^{**}$ ,  $ES = -0.81$ ,  $-0.9$ ,  $-0.96$  respectively). However, the differences between the post and the retention test results of the two groups reveal that their English speaking ability dropped in most areas, except 'interaction' and 'coherence' where the KP group still significantly exceeded its counterparts with 'medium effect sizes ( $-2.108^*$  and  $-2.324^*$ ,  $ES = -0.63$  and  $-0.7$  respectively). The fact that the KP group outperformed the CP group in the post-test and still retained the ability to at least interact with the teacher with a logical flow of ideas could be attributed by the use of keyword-only pictures to teach Grade 2 learners to speak English.

### 3.2 Facilitating and hindering factors affecting Grade 2 Learners' English speaking ability

The semi-structured interview was conducted with 28 students from both experimental groups: 7 high achievers and 7 low achievers in each group. The results of the interviews were summarized below.

**Table 5** The summary of the semi-structured interview in high and low achievers in CP and KP groups

Topics	CP		KP	
	HA	LA	HA	LA
<b>Background Knowledge</b> - Vocabulary (some lessons) - Grammatical structure (some lessons) - Learning how to speak English through pictures in classroom	✓	✓	✓	✓
<b>Teaching Methods</b> - Picture + <i>caption</i> + sound (3 modalities) - Picture+ <i>keyword</i> + sound (3 modalities) - Pictures + sound (2 modalities)	✓	✗	NA	NA
- Teacher initiating interactions through the use of <i>wh</i> -questions based on pictures - Students repeat after teacher.	✓	✓	✓	✓
<b>Captions/keywords</b> - Captions/keywords were difficult in some lesson because they were too long.	✓	✓	✓	✓
<b>Benefits of using captions/keywords-only pictures on speaking abilities</b> - Pronounce words more clearly - Say short English words	✓	✓	✓	✓

As shown in Table 5, the participants who took part in the interview all shared similar views towards learning how to speak English through either treatment. When asked whether they had any background knowledge in English prior to the treatment, all of them agreed that they knew some vocabulary such as cats, apples, books as well as very basic grammatical structure, for example, there is.../ there are.../It is a/an..., thus facilitating them to speak English more easily during the early stage of instruction because they had some knowledge to rely on. They reported that they had learned some English from parents, books, extra class and English teachers at school, as illustrated in the excerpt below.

“I learned some vocabulary from Teacher Ganda (anonymous) at school, such as animals, fruits, vegetable.”

(Low achiever, KP group)

“In grade 2, Teacher Patty (anonymous) taught short grammatical structure to me in first semester. For example, there is.../there are.../it is a/an...”

(High achiever, CP group)

Regarding learning how to speak English through pictures, both high and low achievers in both groups reported that they had no such experience because they normally learned English through other methods, for example, they learned English from songs, games and their previous teachers focused on teaching all four language skills. Although there were some teachers who used flash cards, they used them to teach only vocabulary. This is illustrated in the excerpt below.

“Teacher Patty (anonymous) never used pictures to teach speaking English...in classroom.”

(High achiever, CP group)

“I never learned speaking English through pictures.”

(Low achiever, KP group)



As for using pictures to teach English speaking, all the participants taken part in the interview expressed their positive opinion towards such teaching method in that the teaching pace was not too fast nor too slow. While the high achievers in the CP group could develop their speaking abilities from looking at pictures, captions and sound simultaneously, the high achievers in KP group relied on three modalities as well i.e. pictures, keywords and sounds. The high achievers in both groups all remarked that they remembered English words and were able to pronounce the words more clearly, as reflected in an excerpt below.

"I looked at...pictures and caption together....looked at easy words....if there were some difficult words, I listened to the teacher. I could read, pronounce and speak out".

(High achiever, CP group)

"I looked at both pictures and keywords because I could read all words...repeat after the teacher...I understood and I could speak out."

(High achiever, KP group)

The same opinion holds true for low achievers in both groups. They said that looking pictures and listening sound simultaneously was essential for them to develop their speaking abilities because they had not yet mastered their encoding skills. As a result, they ignored reading either captions or keywords and only on pictures and listened to sound while learning how to speak English in the classroom. The excerpt below illustrated this point.

"I looked at the pictures and the captions but I could not read them. I listened to the teacher and I repeated the sentences."

(Low achiever, CP group)

"I looked at pictures and keywords but... I could not read. I listened to the teacher and repeated after her".

(Low achiever, KP group)

Moreover, following the technique used by Lever and Sénéchal (2011), during speaking activities the teacher asked Wh-questions such as “Where is the dog?” “What color is the dog?.” Such questions offered opportunities for more interactions between the teacher and the participants. This is illustrated in the interview data below.

“I could speak more clearly when I answered questions to teacher in classroom.”

(Low achiever, CP group)

“I could pronounce words more clearly when I answered questions...if I pronounced wrong, the teacher will show the correct one and...I repeat it again.”

(High achiever, KP group)

The other facilitating factor is the use of short and simple English structure consisting of 5-7 words per sentence, for example, “The dog is inside his house”, which was identical to teacher’s narration. This eased the participants to repeat after the teacher which in turn helped them remember and speak more easily. The excerpt below illustrated this point.

“It was very good when I repeated sentences many times after teacher's narration...I could remember and speak out more easily.”

(High achiever, KP group)

“I liked to repeat words after teacher...I could remember and speak English in words more.”

(Low achiever, CP group)

When asked about hindering factors, both high and low achievers in both experimental groups shared a common difficulty, namely, the length of either captions or keywords. They remarked that captions and keywords in some lessons were too long. For example, the CP group found that “The boy is taking photos.”/ “The girl is riding a bicycle.” / “There are four white chickens.” were too long to remember, and the KP group commented that “taking photos” or “riding a bicycle” were problematic.

As for the benefits of learning how to speak English through captions or keywords-only pictures, the interview data revealed that high and low achievers in both experimental groups could pronounce words more clearly because they could see pictures, captions or keywords and sound simultaneously. The other benefit was the ability to say short English words. The learners in this study had little exposure to English, so it was not unusual for them to have limited speaking ability and could handle only vocabulary at the beginning of their English education.

#### 4. DISCUSSION AND CONCLUSION

The aims of this study were to examine the effects of using captions and keywords-only pictures to teach speaking to Grade 2 learners and to study factors affecting their speaking ability. The results revealed that Grade 2 learners' English speaking ability in CP group significantly increased from poor to fair in the post-test but insignificant in the retention test. The two sub-skills that the learners improved were 'range' and 'accuracy'.

Interestingly, a significant increase in the speaking skills of the learners in KP group was found to improve from poor to good. The three improved sub-skills were interaction, range, and fluency.

The above finding indicated that 'range' was the common sub-skill that both groups of learners could improve. The result was congruent with previous studies which discovered the positive effect of using pictures on young learners' vocabulary learning ( Al-Jaafari, 2013; Mansourzadeh, 2010; Mohammed, 2007; Rasheed & Rowe, Silverman, & Mullan, 2013; Yoshii & Flaitz, 2002) One plausible explanation for this could be that young learners in this study had had little exposure to English (Chang & Read, 2006; Goh, 1999; Shang, 2008). Therefore, it was not unusual for them to have limited speaking abilities and could handle only vocabulary at the beginning of their English education (Hayati & Mohmedi, 2011).

Findings from the KP group revealed that the learners were better able to remember vocabulary from the three modalities (keywords, pictures, and sounds). When comparing the complexity of keywords and captions, it can be seen that the caption is far more complex than single words. Moreover, vocabulary presented in the experiment was concrete and related closely to their background knowledge. Such finding can be supported by the interview data in which the low English proficiency participants reported that they also looked at pictures and listened to the sound because they could not read. The finding is also in accordance with that in previous studies which posited that young learners who could not read books were good at listening (Sticht & James, 1984; Taylor, 2005). Viewed from CLT, the KP group had a lighter intrinsic cognitive load. According to Kalyuga, Chander & Sweller (1999), Mayer (2005), the lighter the intrinsic cognitive load, the easier for working memory which in turn automatically results in learning or creating learners' schemas (Pollock et al, 2002; Sweller, van Merriënboer and Paas, 1998). From the DCT point of view, it can be argued that different modalities of input would complement each other rather than interfere (Clark & Paivio, 1991; Danan, 2004; Mayer, 2009).

In conclusion, it could be said that the speaking ability of the CP group significantly developed from poor to fair. The two sub-skills that they developed the most were 'range' and 'accuracy', indicating that they had better vocabulary knowledge and were more able to use to construct simple English to describe pictures when compared to the pre-test results. In contrast, the KP group's oral production ability significantly increased from poor to good, with 'interaction', 'range', and 'fluency' that developed the most. This suggests that they could interact with the teacher more fluently and also had more knowledge in vocabulary. As far as Dual Coding Theory and Cognitive Load Theory is concerned, it could be concluded that for young learners with low English proficiency multi-modality input with picture, keyword and sound would yield better learning outcome than when picture, caption and sound are used because they have not yet master their reading ability.

## 5. RECOMMENDATIONS

### 5.2.1 Pedagogical Recommendations

Based on findings of the current study, it is recommended that the use of keywords below the pictures might be suitable for young learners aged 7-9 years old. They can learn well through pictures, keywords, and sounds because it is easy for them to remember and speak out. However, the use of captioned pictures could foster learners to speak accurately because they could see the whole sentences. Repetition is also meaningful for the beginners who started learning a language. The teacher should ask questions several times to motivate learners to interact in the classroom. The teacher should also repeat words, phrases, or sentences many times while he/she is teaching through the uses of pictures in the classroom because learners will remember and speak English more fluently. The teacher should also create speaking games by using pictures to promote English speaking ability in classroom. For example, the teacher might use pictures for whispering game because learners can see pictures and listen to sound from teacher. After that, they can practice English speaking when they whisper sentences to their classmates several times. Looking at pictures and listening to sound can support them to remember and speak out more easily. They also promote new vocabulary learning.

### 5.2.2 Recommendations for Further Study

For further studies, the researcher should use a larger sample size for better statistical analysis. Also, including a control group might yield a stronger finding. It may well be interesting to compare the effects of using multi-modalities (pictures, keywords and sounds) and bi-modalities (pictures and sounds) on learners' English speaking abilities. Moreover, it is recommended that future research might investigate the use of a series of 2 or 3 pictures instead of a single picture to find out whether and to what extent it can develop learners' speaking ability in connecting ideas or situations, and stimulate more teacher-learner interaction.

## REFERENCES


- Al-Ja Afari, I. S. (2013). Using pictures in vocabulary in Grades 5 and 6 classroom, Sharqiya South Region.
- Bowen, B. (1982). *LOOK Here! Visual aids in language teaching*. Macmillan, London.
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press.
- Chang, A. C. S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL quarterly*, 375-397.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, 3(3), 149-210.
- Cook, G. (2000). *Language play, language learning*. Oxford: Oxford University Press.
- Colon-Vila, L. (1997). Storytelling in an ESL Classroom. *Teaching Pre K-8*, 27(5), 58-59.
- Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators' Journal*, 49, 67-77.
- Dobson, M. J. (1992). *Effective techniques for English conversation group*. Washington: United States Information Agency.
- Ellis, G. & Brewster, J. (1991). *The storytelling handbook for primary English language teachers*. United Kingdom, Longman.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching ESL children in the mainstream classroom*. Portsmouth US: Heinemann.
- Goh, C. (1999). How much do learners know about the factors that influence their listening comprehension?. *Hong Kong. Journal of Applied Linguistics*, 4(1), 17-42.
- Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 42(1), 181-192.

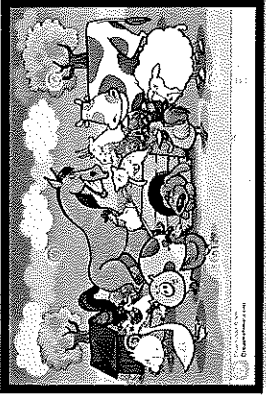
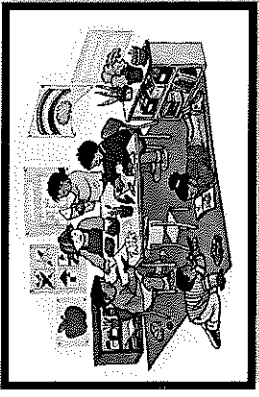
- Hirai, A., & Koizumi, R. (2013). Validation of Empirically Derived Rating Scales for a Story Retelling Speaking Test. *Language assessment quarterly*, 10(4), 398-422.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing splitattention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-371.
- Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, 108(1), 1-24.
- Linse, Caroline. T. (2005). *Practical English language teaching: Young learners*. McGraw Hill, New York.
- Lutz, S., & Huitt, W. (2003). *Information processing and memory: Theory and applications*. Educational Psychology Interactive. Valdosta, GA: Valdosta State University.
- Mackay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
- Mansourzadeh, N. (2014). A comparative study of teaching vocabulary through pictures and audio-visual aids to young Iranian EFL learners. *Journal of Elementary Education*, 24(1), 47-59.
- Mayer, R. E. (2005). *Cognitive theory of multimedia learning*. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.
- Mayer, R. E. (2009). *Multimedia learning*. New York: Cambridge University Press.
- Nolasco, R. & Arthur, L. (1991). *Large classes*. Hong Kong: Macmillan.
- Pinter, A. (2006). *Teaching young language learners*. Oxford: Oxford University Press.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and instruction*, 12(1), 61-86.
- Rasheed, W. & Mohammed, H. (2007). *The impact of using pictures in vocabulary achievement for fifth primary pupil*. College of Basic Education.
- Rokhayani, A. (2012). Motivating Students of Young Learners through Story Telling in English Class. *TEYLIN: from Policy to Classroom*.

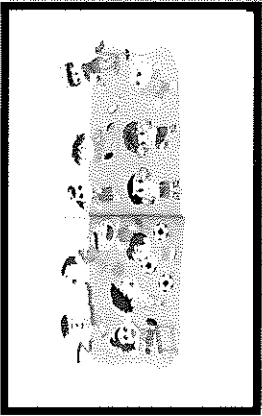
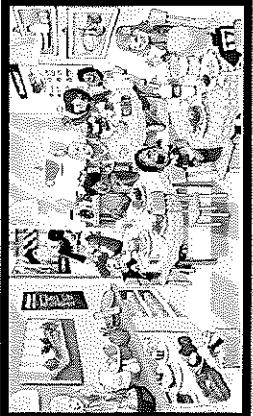
- Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology, 38*(2), 109-117.
- Shang, H.F. (2008). Listening strategy use and linguistic patterns in listening comprehension by EFL learners. *The Intl. Journal of Listening, 22*(1), 29-45.
- Slattery, M. & Willis, J. (2001). *English for primary teachers*. Oxford: Oxford University Press.
- Sticht, T. and Jamcs, J. (1984). *Listening and reading*. P.D.Perirson (Ed) Hancbook of Rcadine Rcsearch. New York: Longman. 293 -317.
- Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review, 10*(3), 251-296.
- Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals, 38* (3), 422-427.
- University of Cambridge ESOL Examinations Research and Validation Group. (2009). *Example of Speaking Performance at CEFR Levels A 2 to C 2*. Cambridge: Cambridge University Press.
- Wright, A. (1990). *Pictures for language learning*. Cambridge: Cambridge University Press.
- Yang, Y. T. C., & Chang, C. H. (2013). Empowering students through digital game authorship: Enhancing concentration, critical thinking, and academic achievement. *Computers & Education, 68*, 334-344.
- Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO journal, 33*-58.

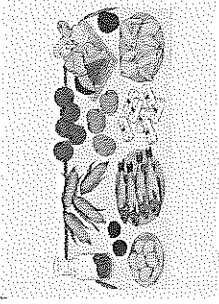
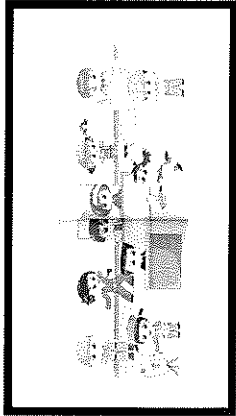


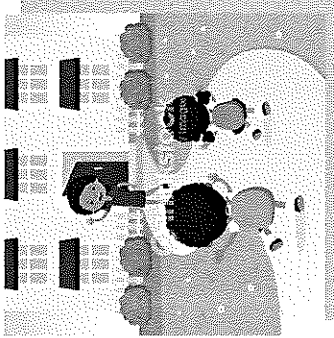
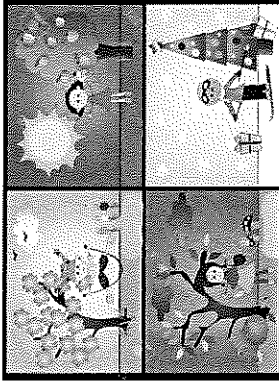
APPENDIX A  
(Scope of Lesson Plans)

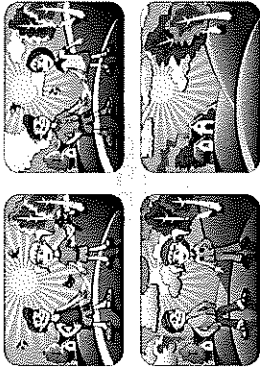
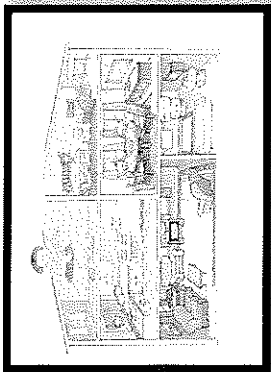
Lesson	Theme	Grammar	Vocabulary
Lesson 1	<p>Fruits &amp; Vegetable</p> 	<p>-What fruit do you like? / What fruit don't you like? I like (apples). / I don't like (watermelons). -How many (pineapples) are there? There are (two pineapples). -What color are (apples)? (Apples) are red. -What vegetable do you like? / What vegetable don't you like? I like (carrots). / I don't like (cucumbers). -What color are (carrots)? (Carrots) are orange.</p>	<p>Name of fruits and vegetables</p>

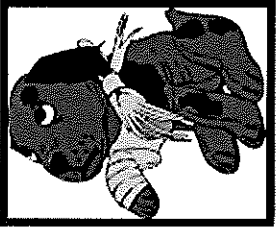

Lesson	Theme	Grammar	Vocabulary
Lesson 2	<p style="text-align: center;">Animals</p> 	<p>-What animal do you like? / What animal don't you like? I like a (pig). / I don't like a (cat).</p> <p>-How many (chickens) are there? There are (seven chickens).</p> <p>-What color is (dog)? / What color are (chickens)? A dog is (brown). / Chickens are (brown).</p>	<p>Different kinds of animals in a farm</p>
Lesson 3	<p style="text-align: center;">Classroom Items</p> 	<p>- What's this? / What's that? It's (color pencils).</p> <p>- Is this a (poster)? / Is this a (notebook)? Yes, it is. / No, it isn't. It's a (book).</p>	<p>Different types of classroom items</p>

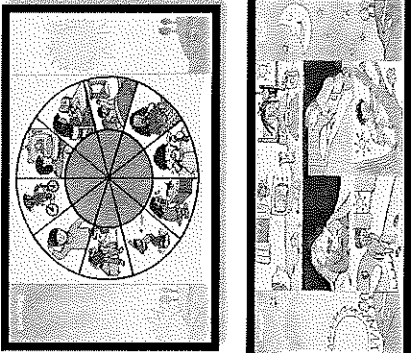
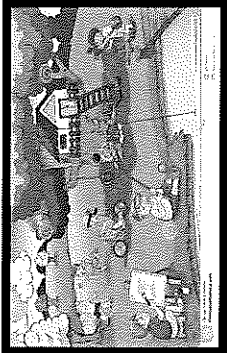
Lesson	Theme	Grammar	Vocabulary
<p>Lesson 4</p>	<p>Sports</p> 	<p>- What sport do you like? I like (football). - What sport can you play? I can play (golf). - Can you play (football)? Yes, I can. No, I can't. I can play (golf).</p>	<p>Different types of sports</p>
<p>Lesson 5</p>	<p>Foods</p> 	<p>- What food do you like? I like (hamburgers). - What food don't you like? I don't like (fried chickens). - Do you like (hot dogs)? Yes, I do. No, I don't. I like (hamburgers). - What do you want? I want (salad).</p>	<p>Different kinds of foods</p>

Lesson	Theme	Grammar	Vocabulary
Lesson 6	<p>Colors &amp; Shapes</p> 	<p>-What color is this? it's (yellow).</p> <p>-What shape is this? it's a (triangle).</p> <p>- What shape is orange? it's a circle.</p>	<p>Name of colors and types of shapes</p>
Lesson 7	<p>Occupations</p> 	<p>-Who is he/ she? He/She is a (lawyer).</p> <p>-Is he/she a (singer)? Yes, he/she is. No, he/she isn't. He/she is a (reporter).</p> <p>- What are you going to be in the future? I'm going to be a (cook).</p>	<p>Different kinds of occupations</p>

Lesson	Theme	Grammar	Vocabulary
<p>Lesson 8</p>	<p>Days</p> 	<p>-When do you go to school? I go to school on (Monday).</p> <p>- When don't you go to school? I don't go to school on Saturday and Sunday.</p> <p>- When is your birthday? My birthday is on Wednesday.</p>	<p>Days in a week</p>
<p>Lesson 9</p>	<p>Seasons</p> 	<p>- What season do you like? I like (winter).</p> <p>- What season don't you like? I don't like (autumn).</p> <p>- What season is hot? It is in summer.</p> <p>- Is winter hot? / Is winter cold? No, it isn't. It is cold. / Yes, it is.</p>	<p>Kinds of seasons</p>


Lesson	Theme	Grammar	Vocabulary
<p>Lesson 10</p>	<p>Weathers</p> 	<p>- How is the weather? It's (sunny). - What the weather do you like? I like (rainy). - What the weather don't you like? I don't like (windy).</p>	<p>Different kinds of weathers</p>
<p>Lesson 11</p>	<p>My House</p> 	<p>-Where is the (bed)? It is in the bedroom. -Is the television in the kitchen? Yes, it is. / No, it isn't. It is in the living room. - How many chairs are there in the dining room? There are four chairs.</p>	<p>Different rooms in a house</p>

Lesson	Theme	Grammar	Vocabulary
Lesson 12	<p>Body Parts</p> 	<p>- What is this? / What is that? This is a.../ That is a....</p> <p>- How many (eyes) do you have? I have two eyes.</p>	Parts of body
Lesson 13	<p>Clothing</p> 	<p>-What is this? / What is that? This is a.../ That is a...</p> <p>-What color is a dress? It's pink.</p> <p>- What does she wear? She wears a dress.</p>	Kinds of clothing

Lesson	Theme	Grammar	Vocabulary
<p>Lesson 14</p>	<p>Times</p> 	<p>-What do you do (in the morning)? I wash my face and comb my hair. - Do you read books in the evening? Yes, I do. / No, I don't. I watch TV.</p>	<p>Period of time, in the morning, in the afternoon, in the evening</p>
<p>Lesson 15</p>	<p>Actions</p> 	<p>What is he/she doing? -He/She is (drawing a picture). Is she walking? Is she drawing a picture? -Yes, she is./ No, she isn't. She is reading.</p>	<p>Different kinds of actions</p>



Example of Lesson Plan

Lesson	Theme	Grammar	Vocabulary	Expected Outcome
1	<p>Fruits &amp; Vegetables</p> 	<p>-What is this?            This is a/an _____. I like _____(s).            This is a/an _____. I (don't) like _____(s).            -What fruits do you like? / What fruits don't you like?            I like (apples). They are crunchy and sweet.            I don't like (bananas). They are soft.            I like (oranges). They are sour. I don't like (bananas). They are sweet.</p>	<p>Nouns: apples, oranges, bananas, grapes            Adj.: crunchy, soft, sweet, sour</p>	<p>This is an orange. I like oranges. They are sour. I don't like bananas. They are soft.</p>

## Effects of Using Captioned Pictures on Grade 2 Learners' English Speaking Abilities

Nittaya Manuy, M.A. Student, Teaching English as an International Language, Department of Languages and Linguistics, Prince of Songkla University, Songkhla, Thailand. Email: [LA\\_peung23@windowslive.com](mailto:LA_peung23@windowslive.com)

Asst. Prof. Dr. Chonlada Laohawiriyanon, Department of Languages and Linguistics, Prince of Songkla University, Songkhla, Thailand. Email: [chonlada.l@psu.ac.th](mailto:chonlada.l@psu.ac.th)

### ABSTRACT

Drawing upon Dual Coding Theory, learners better understand and remember information by receiving two different modes of inputs simultaneously. Some learners, however, learn well from single-mode inputs such as picture or sound. This article presented preliminary findings of a quasi-experiment study aiming at investigating 1) the effects of using captioned pictures on Grade 2 learners' speaking ability, and 2) the facilitating and hindering factors of using captioned pictures to developing speaking abilities of Grade 2 learners. The participants in this study were 26 Grade 2 learners in a public primary in Hat Yai District, Songkhla Province. Data were collected using a pre-post treatment test and semi-structured interview. It was found that the learners' speaking ability increased significantly, at a 0.01 significant level. The participants' speaking abilities improved from *Below A1 level* to *A1 Level* in two sub-skills: 35 % improved in their vocabulary range, and 27 % in language use accuracy. The subjects, nonetheless, could not interact in conversation. Nor could they combine sentences using simple connectors, such as 'and' or 'then'. It is recommended that future research investigate the use of a series of 2 or 3 pictures instead of a single picture to find out whether and to what extent it can develop learners' ability in connecting ideas or situations, and trigger more teacher-learner interaction.

**Keywords:** Captioned Pictures; Dual Coding Theory; Speaking Ability; Young Learners

### BACKGROUND AND RATIONALE

Teaching English to young learners first requires understanding about their characteristics. For example, they have short attention span (Bialystok, 2001). They are also active and love doing physical activities (Mackay, 2006). Learning activities which are suitable for them, therefore, should reflect their age and interest e.g. listening to stories repeatedly, learning from doing, and playing (Slattery & Willis, 2001). For language development, they can learn well from teachers, friends, and storytelling (Mackay, 2006). Picture has been used as an effective instructional tool to teach vocabulary as well as to develop all four language skills (Wright, 1990). For instance, pictures can be employed to enhance children to express ideas, to establish background for stories, to introduce new words, or as prompts when young learners cannot remember words. Subsequently, children can use all the experiences they have as a model to produce their own language. Finally, pictures can also enhance their imagination (Dobson, 1992).

Pictures used in language teaching come in at least three forms: pictures alone, annotated pictures (pictures with vocabulary placed underneath), and captioned pictures (pictures with a phrase or a sentence describing the pictures). According to Dual Coding Theory proposed by Paivio (1986), learners process information better if they are exposed to bi/multi-modal inputs e.g. picture and written text or picture and sound; one mode of information complement the other. The adoption of Dual Coding Theory in teaching

### Innovations in Humanities and Social Sciences: Opportunities and Challenges

language to children can be seen in the use of pictures to tell story which in turn can foster young learners' listening and speaking ability.

Young Thai learners are considered having limited language proficiency, as evident in Grade 6 learners' scores of the Ordinary National Education Test (O-NET) from 2013-2015 academic years (33.82%, 36.02% and 36.02% respectively) (The National Institution of Educational Testing Service, 2013, 2014, 2015). Their low performance can be attributable to different factors, namely, learners being unaware of the importance or taking the O-NET and lacking basic linguistic knowledge; teachers' reliance mainly on using rote learning technique to teach English speaking. Teachers responsible for teaching English to primary school learners should look for teaching aids and technique suitable to the age, interest, and language ability of the learners (Uamporn Lhinchareon, Sirisak Atwichai, and Pirapa Chaninn, 2009).

Storytelling has been widely used to develop children's speaking ability. Most empirical studies investigated the effectiveness of using pictures to develop learners' vocabulary knowledge. However, studies on using pictures to promote young learners' speaking ability are minimal (Ishell, 1994). The researchers, thus, were interested in studying developing Grade 2 learners' speaking ability through the use of captioned pictures.

#### LITERATURE REVIEW

Young learners or learners aged 7-9 years are unique in that their concentration lasts only 10-15 minutes (Bialystok, 2001). They love doing activities which involve physical movements, e.g. run, jump, dance, and sing. They are very active, get tired easily, but recover easily (Mackay, 2006). There are different learning activities appropriate for this group of learners: learning from doing and playing or learning from listening (Slattery & Willis, 2001). For foreign language development, young learners learn basically from frequent exposure and use the language as much as possible, which is the same way they acquire their mother language because learning from listening helps the learners take in more input (Slattery & Willis, 2001). They also learn a language better from teachers, peers, and storytelling.

Speaking and listening skills are interrelated; speaker sends voice messages while listener receives the messages. The listener not only receives the message, but also responds to the speaker; i.e. both the speaker and listener exchange the role of speaking and listening throughout the interaction (Clampitt, 2005). To learn a foreign language, at the beginning young learners gradually take in the language input from listening to the language repeatedly and then express verbally what they have learned (Slattery & Willis, 2001). Storytelling is a two-way communication activity because it provides learners opportunities to interact with the teacher about the contents of the story while listening to the story (Nation & Newton, 2009).

There are several teaching techniques which are considered effective for teaching children at a very young age. First, storytelling, a popular technique, has been widely used because it can elevate their learning motivation (Rokhayani, 2010). Storytelling is useful in various ways. For example, it helps develop learning experiences; it can trigger learners' imagination from listening to stories; it provides opportunities for young learners to exchange experiences with each other; and it helps develop the learners listening skill and concentration (Ellis & Brewster, 1991). Another teaching technique is using games to teach languages. It is suitable for children because young learners like challenges, competition, and enjoyment (Nolasco & Arthur, 1991). Through games, they can learn vocabulary and grammatical rules more effectively. Songs have been proven effective for developing learners' listening and speaking skills by listening to lyrics and pronunciation. They can also memorize vocabulary and grammatical structures. Moreover, songs can serve as a model for creative writing (Shen, 2009).

### Innovations in Humanities and Social Sciences: Opportunities and Challenges

Pictures are very useful for children to develop their linguistic abilities ranging from learning new words to describing pictures. Moreover, they are useful in teaching grammar, vocabulary, and the four language skills leading to the increase in teaching effectiveness (Wright, 1990). In addition, pictures provide pleasant classroom atmosphere and enhance learners' imagination (Dobson, 1992; Hill, 1990; Harner, 2001; Allen, 1983). Studies on using pictures in classroom language instruction demonstrated that pictures improved learners' speaking ability because teachers use pictures to prompt students to express their thoughts or feelings thereby they learn how to communicate in their real-world everyday situations. When teachers describe pictures, learners learn how to make meaning from what they see and hear simultaneously and especially when teachers repeat what they say over and over again (Al-Jafari, 2013; Mansourzadeh, 2010; Rasheed & Mohammed, 2007; Widyaningrum, 2014).

The fact that language learners can learn better by looking at the pictures while listening to the teacher can be explained by the Dual Coding Theory (Paivio; 1971,1986). Based on the theory, effective learning takes place when learners receive bi-or multi-modal inputs through different sensory systems such as the visual (pictures, written texts) and the auditory (verbal) in learner's working or short-term memory. According to Clark & Paivio (1991) and Mayer (2009), different modalities of input would complement each other rather than interfere. However, it is not always the case that bi/multi-modalities of messages would lead to better learning. Learners with different language proficiencies might have different reactions to such inputs. For instance, young learners who begin to read might benefit more from listening only while adult learners would learn more effectively through reading and listening simultaneously (Sticht & James, 1984; Taylor, 2005). It was also found that learners of different ages learn effectively through different kinds of inputs i.e. learners aged 4 learn better through listening. Those aged about 14, on the other hand, learn better through reading. Nonetheless, other factors should be taken into consideration, such as family background, gender, and language proficiency (Jerger, Damian, Spence, Tye-Murray & Abdi, 2009)

Storytelling through picture books is one of the interesting ways to teach speaking to young learners. Lever & Sénéchal (2011) suggested using picture books to develop the learners' speaking abilities. Teachers can tell a story slowly. During storytelling, the teacher interacts with learners while he/she is telling the story by asking relevant questions using Wh-questions e.g. "Where is/are X?" and "What does X do?". Most importantly, a teacher should constantly provide corrective feedback in the form of correct sentence to learners, whether learners answer correctly and incorrectly. This is to provide them with the correct model of language use. When the teacher finishes telling the story, the learners should retell the story again by using the connective words "and" or "then". Learners who are listening to a story while looking at pictures receive bi-modal inputs, which in turn strengthens their understanding. The results showed that storytelling and exchanging information helped young learners to develop grammar, vocabulary, as well as reading comprehension. Eventually, they were able to tell stories themselves.

#### RESEARCH METHODOLOGY

This quasi-experimental study investigated the effect of using captioned pictures on Grade 2 learners' speaking ability.

##### Participants

The target population of this study was 49 Grade 2 learners studying in a public primary in Hat Yai District, Songkhla Province. The sample group was 26 Grade 2 learners consisting of 12 girls and 14 boys, aged 7-9 years. Among those 26 learners, 24 were Thai and 2 Myanmar learners. The experiment was conducted in the second semester of the 2015 academic year. The sample group was selected purposively because they formed a heterogeneous class in terms language proficiency.

## Innovations in Humanities and Social Sciences: Opportunities and Challenges

### Research Instruments

Following the criteria set by Bowen (1982), the researchers selected pictures that were interesting, clear and colorful, large enough for every learner to see, and also culturally appropriate. Most important of all, the pictures had to be relevant to the objectives of the lessons. According to the Basic Core Curriculum of the Ministry of Education, 2551 B.E., the lessons should cover the following contents: 1) Numbers 2) Animals 3) Classroom Items 4) Sports 5) Occupations 6) Time 7) Clothing 8) Fruit and Vegetable 9) Weather 10) Actions 11) Food 12) Shapes & Colors 13) Body Parts 14) Days 15) Seasons. The experiment in this study took 30 periods; each period lasted 50 minutes. The class met twice a week over the course of 15 weeks. The number of words of captions of each picture ranged from 5-7 words which are considered appropriate for beginners (Lutz & Huitt, 2003). Below is a sample picture used in the present study.

### Captioned Pictures



Reference: [www.dreamstime.com](http://www.dreamstime.com)

The teaching procedures employed in the study followed the principles suggested by Lever & Sénéchal (2011). First, the teacher read the caption of each picture, e.g. "The dog is inside his house." while pointing at the dog in the picture. After reading the first caption, the teacher asked "Where is the dog?" If the students failed to produce the correct answer, e.g. "The dog in house!", the teacher would then supplied the correct sentence "The dog is inside his house". The students had to repeat the sentence several times. Even if the students produced the correct answer, the teacher still had to restate the answer and had the whole class repeat the sentence a few times. Then the teacher read the new caption in the same manner until all the captions describing the picture were introduced. Finally, the students were asked to collaboratively retell the whole story.

### Data collection

To collect data, the researchers carried out pre and post-tests. Each individual participant had 2 minutes to describe in English a picture about colors, number, characteristics, etc. The criteria used to examine speaking abilities were based on the Common European Framework of Reference (CEFR). There were 5 sub-skills of oral assessment criteria grid: range, accuracy, fluency, interaction and coherence. Speaking abilities were divided into 12 levels, namely, Below elementary (*Below A1*), elementary (*A1, A1+, A2, A2+*), intermediate (*B1, B1+, B2, B2+*) and advanced (*C1, C1+, C2*) (University of Cambridge ESOL Examinations Research and Validation Group, 2009). Three raters who were EFL experts (One American and two Thai) were employed to obtain inter-rater reliability. Each rater viewed the recorded videos and rated the participants' abilities following the grid. Percentage of participants placed in each level was calculated and Wilcoxon Signed Ranks Test was performed.

The first step of the data collection procedure was the pre-test in which each participant described in English a picture. Each had 2 minutes to give a detailed description. If he/she did not show any sign of comprehending what was the task in hand, the teacher would help by asking questions, such as "What is this? / What are they doing? / What color is his T-shirt? As for the treatment, the teacher researcher taught 15 classes according to the lesson plans. Each plan took 2 periods per week, totaling 30 periods. Finally, the post

### Innovations in Humanities and Social Sciences: Opportunities and Challenges

test was administered. The same picture was given to each participant to describe once again. Videos were recorded for subsequent evaluation.

#### FINDINGS

##### English Speaking ability

The pre- and post-test mean scores were analyzed based on the CEFR. Six aspects of speaking ability (overall, range of vocabulary used, accuracy, fluency, interaction, and coherence) are presented in Table 1.

**Table 1:** Wilcoxon Signed Ranks Test summary of pre-and post-treatment tests

Parts	Levels	Frequencies				Wilcoxon Signed Ranks Test		
		Pre-test		Post-test		Z	*	Asymp. Sig. [2-tailed]
		n	%	n	%			
Overall	Below A1	26	100	21	81	-2.236	*	0.03
	A1	0	0	5	19			
	A2	0	0	0	0			
Range	Below A1	26	100	17	65	-3.000	**	0.00
	A1	0	0	9	<u>35</u>			
	A2	0	0	0	0			
Accuracy	Below A1	26	100	19	73	-2.646	**	0.01
	A1	0	0	7	<u>27</u>			
	A2	0	0	0	0			
Fluency	Below A1	26	100	21	81	-2.236	*	0.03
	A1	0	0	5	19			
	A2	0	0	0	0			
Interaction	Below A1	26	100	21	81	-2.236	*	0.03
	A1	0	0	5	19			
	A2	0	0	0	0			
Coherence	Below A1	26	100	21	81	-2.236	*	0.03
	A1	0	0	5	19			
	A2	0	0	0	0			

**Note:** \* $p < .05$  and \*\* $p < .01$

### Innovations in Humanities and Social Sciences: Opportunities and Challenges

The figures in Table 1 show that 100% of the participants were at *Below A1* level in all skills, at a .05 significant level. In other words, none of them could give a detailed description of people or situations in simple words or phrase, using simple sentence structures. They could not exchange a conversation. Nor could they combine sentences using simple connectors, e.g. 'and' or 'then'.

The post experiment test shows that the top two sub-skills that the participants improved their speaking ability from the *Below A1* to *A1* at a significant level of .01. To be specific, 35% of the participants reached the *A1* level in the range of vocabulary and 27% in accuracy.

Results from this study show that the learners enlarged their vocabulary range and could use simple phrase and sentence structures more accurately. However, they could not take part in a conversation. Likewise, they could not combine two sentences using simple connectors, such as 'and' or 'then'.

#### Facilitating and hindering factors in learning speaking skill through captioned pictures

An interview after the post-treatment test was conducted to investigate the factors concerning using captioned pictures that helped or hindered the participants' English speaking development. Seven high achievers and seven low achievers were interviewed. It can be concluded that looking at captions while listening to the teacher narration helped them remember simple vocabulary, leading to the ability to produce correct pronunciation and speak more fluently. This is illustrated in the interview data below.

"I looked at both pictures and captions. I chose to look at simple words and listen to the teacher. They helped me able to read aloud and speak."

(female, high achiever)

"I looked at both pictures and captions. Because I looked at pictures and repeated after the teacher, I understood and could speak."

(male, high achiever)

However, if the teacher's narrations and the captions contained difficult words, the high achievers would be attentive to teachers' narration only. In contrast, the low achievers relied on pictures and teacher's narration. They reported that they ignored the captions. They also repeated after the teacher's narration. It can be concluded that low achievers needed language model from the teacher to learn how to speak. In addition, both high and low achievers further stated that captions containing 5-7 simple words were optimal and facilitated their memorizing and speaking abilities.

Regarding factors that hindered the speaking ability development, high achievers found lengthy captions too difficult to learn to speak. For example, captions describing animals or actions contained too many words e.g. "The girl is riding a bicycle." "There are four white chickens." as against "This is an apple." For low achievers, the only obstacle was their inability to read the captions, as illustrated in the interview data below.

"I looked at the pictures and the captions but I could not read them. I listened to the teacher and I repeated the sentences."

(male, low achiever)

It was also pointed out by the low achievers that their reading ability improved because they looked at captioned pictures, listened to the teacher and read the captions along.

### DISCUSSION

The present study aimed to find out the extent to which using captioned pictures to tell story could enhance Grade 2 learners' English speaking abilities. The post-test results revealed that one third of the participants reached the *At* level in comparison to the *Below At* level they possessed before the treatment. In other words, at the *At* level, learners could talk about themselves, and things around them using basic vocabulary, simple sentence structures, or phrases. However, learners were not able to converse or use basic connectors, such as *and* or *then*, to combine word groups or sentences.

Such findings were in agreement with the results of results previous studies which discovered that young learners could basically acquire vocabulary through the use pictures in English class (Al-Jaafari, 2013; Chai & Erlam, 2008; Mansourzadeh, 2010; Rasheed & Mohammed, 2007; Stewart & Pertusa, 2004; Widyaningrum, 2014; Yoshii & Flaitz, 2002). The fact that the subjects could improve their vocabulary from learning through captioned pictures together with the teacher's narration could be supported by the Dual Coding Theory. The participants in the present study received the multimodal inputs -- pictures, written texts in captions, and sound. Paivio (1986) and Danan (2004) both maintained that learners' working memory processes information better from bi/multimodal inputs than from single-modal ones. This is supported by the findings from the interview data from both high and low achievers who reported that they wanted to be exposed to the three modalities of inputs in order to develop their speaking abilities.

One plausible explanation as to why the learners were still unable to speak at sentence level could be attributed to the fact that they were only 7-9 years old and of course had very limited exposure to English (Hayati & Mohmedi, 2011). Thus, they were unfamiliar with linguistic system of the new language being learned (Chang & Read, 2006; Goh, 1999; Shang, 2008). Acquiring only English words is, therefore, considered realistic for learners learning a foreign language at a very young age.

### PEDAGOGICAL IMPLICATIONS

Based on the findings, captioned pictures can be used to foster young learners' speaking abilities if captions contain one simple sentence with approximately 5-7 simple words. Frequent exposure to verbal speech and written texts provides opportunities for them to repeat what they hear or see, which would result in increased speaking abilities. Other activities should be included to create more learning motivation.

### FURTHER RESEARCH

For future research, a series of 2-3 pictures instead of one single picture may be used to allow opportunities for learners to use connectors to combine ideas and support teacher-learner interaction. A control group with a larger number of participants is required to obtain more reliable results.

### REFERENCES

- Al-Jaafari, I. S. (2013). Using Pictures in Vocabulary in Grades 5 and 6 Classroom, Sharqiya South Region.
- Allen, V. F. (1983). Techniques in teaching vocabulary. New York: Oxford University Press.
- Bialystok, E. (2001). Bilingualism in Development. Cambridge: Cambridge University Press.
- Bowen, B. (1982). *LOOK Here! Visual Aids in Language Teaching*. Macmillan Publishers Ltd. London.



**Innovations in Humanities and Social Sciences: Opportunities and Challenges**

- Chai, J. & Erlam, R. (2008). The Effect and the Influence of the Use of Video and Captions on Second Language Learning. *New Zealand Studies in Applied Linguistics*.
- Chang, C.S. & Read, J. (2007). Support for foreign language listeners: Its effectiveness and limitations. *RELC*, 38(3), 375-395.
- Clampitt, P. G. (2005). *Communicating for Managerial Effectiveness*. California: Sage.
- Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149-170.
- Danan, M. (2004) Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators' Journal*, 49, 67-77.
- Dobson, M. J. (1992). *Effective techniques for English conversation group*. Washington: United States Information Agency.
- Ellis, G. & Brewster, J. (1991). *The Storytelling Handbook for Primary English Language Teachers*. United Kingdom. Longman.
- Goh, C. (1999). "How Much Do Learners Know about the Factors that Influence Their Listening Comprehension?". *Hong Kong Journal of Applied Linguistics*, 4, 17-40.
- Harmer, J. (2001). *The practice of English language teaching*. London: Longman.
- Hayati, A. & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 37, 301.
- Hill, D. A. (1990). *Visual Impact: Creative language learning through pictures*. Essex: Longman Group UK Limited.
- Isbell, R. (1994). The Effects of Storytelling and Story Reading on the Oral Language Complexity and Story Comprehension of Young Children. *Early Childhood Education Journal*. Vol. 32, No.3. East Tennessee State University.
- Jerger, S., Damian, M., Spence, M. J., Tye-Murray, N., Abdi, H. (2009). Developmental shifts in children's sensitivity to visual speech: A new multimodal picture word task. *Journal of Experimental Child Psychology*.
- Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, 108, 1-24.
- Lhinchareon, U., Atwichai, S., Chaninn, P. (2009). The causes of low O-NET scores in Grade 6 and Grade 12. National Institute of Educational Testing Service (Public Organization).
- Lutz, S., & Huit, W. (2003). *Information processing and memory: Theory and applications*. Educational Psychology Interactive. Valdosta, GA: Valdosta State University.
- Mackay, P. (2006). *Assessing Young Language Learners*. Cambridge: Cambridge University Press.
- Mansourzadeh, N. (2010). *A Comparative Study of Teaching Vocabulary through Pictures and Audio-visual Aids to Young Iranian EFL Learners*. University of Sistan and Baluchestan.
- Mayer, R. E. (2009). *Multimedia learning*. New York: Cambridge University Press.
- Nation, I. S. P. & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge/Taylor and Francis Group.

**Innovations in Humanities and Social Sciences: Opportunities and Challenges**

- National Institute of Educational Testing Service (Public Organization). (2013). Reporting O-NET scores results in 2013 academic year. Retrieved June 20, 2015 from <http://www.niets.or.th>. (2014). Reporting O-NET scores results in 2014 academic year. Retrieved June 20, 2015 from <http://www.niets.or.th>. (2015). Reporting O-NET scores results in 2015 academic year. Retrieved April 2, 2015 from <http://www.niets.or.th>.
- Nolasco, R. & Arthur, L. (1991). *Large Classes*. Hong Kong: Macmillan Publishers Ltd.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
- ..... (1986). *Mental Representations*. New York: Oxford University Press.
- Rasheed W. & Mohammed H. (2007). The Impact of Using Pictures in Vocabulary Achievement for Fifth Primary Pupil. *College of Basic Education*.
- Rokhayani, A. (2010). Motivating students of young learners through storytelling in English class. *TEYLIN 2: From Policy to Classroom*.
- Shang, H.F. (2008). Listening strategy use and linguistic patterns in listening comprehension by EFL learners. *The Intl. Journal of Listening*, 22(1): 29-45.
- Shen, C. (2009). Using English Songs: An enjoyable and effective approach to ELT. *English Language Teaching*, vol.2, No1. Zhejiang Gongshang University.
- Slattery, M. & Willis, J. (2001). *English for Primary Teachers*. Oxford: Oxford University Press.
- Stewart, W. A., & Pertusa, I. (2004). Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals*.
- Sticht, T. and James, J. (1984). Listening and reading. P.D.Perirson (Ed) *Handbook of Reading Research*. New York: Longman. 293-317.
- Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals*, 38 (3), 422-427.
- University of Cambridge ESOL Examinations Research and Validation Group. (2009). *Example of Speaking Performance at CEFR Levels A 2 to C 2*. Cambridge: *Cambridge University Press*.
- Widyaningrum, L. (2014). The Effect of Using Pictures on Students' Speaking Ability at Grade Eight of SMPN 2 Seputih Mataram. Bandar Lampung.
- Wright, A. (1990). *Pictures for Language Learning*. Cambridge: *Cambridge University Press*.
- Yoshii, M., Flaitz, J. (2002). *Second Language Incidental Vocabulary Retention: The effect of text and picture annotation types*. University of South Florida.

Effects of Using Captioned and Keyword-only Pictures on Grade 2 Learners'  
English Speaking Abilities

ผลของการใช้รูปภาพที่มีคำบรรยายได้ภาพและรูปภาพที่มีเฉพาะคำสำคัญได้  
ภาพต่อความสามารถด้านการพูดภาษาอังกฤษของนักเรียนชั้นประถมศึกษา  
ปีที่ 2

*Nittaya Manuy*<sup>1</sup>

*Chonlada Laohawiriyanon*<sup>2</sup>

Abstract

The purposes of this study were to examine the effects of using captioned and keyword-only pictures on Grade 2 learners' English speaking abilities and to examine factors affecting their English speaking abilities. The participants were recruited from 49 Grade 2 learners (25 female and 24 male) in a public primary school in Songkhla, southern Thailand. They were selected by purposive sampling and randomly assigned into two experimental groups: captioned picture group and keyword-only picture group. The data were collected from pre-, post-, retention tests and semi-structured interview. Following CEFR oral assessment criteria, the results revealed that in the captioned pictures group, Grade 2 learners' English speaking abilities significantly increased from

---

<sup>1</sup> M. A. student. Teaching English as an International Language Program, Faculty of Liberal Arts, Prince of Songkla University, Hat Yai Campus, Songkhla, Thailand  
E-mail: LA\_peung23@windowslive.com

<sup>2</sup> Asst. Prof. Dr. Department of Languages and Linguistics, Faculty of Liberal Arts, Prince of Songkla University, Hat Yai Campus, Songkhla, Thailand  
E-mail: chonlada.l@psu.ac.th

poor level to fair level, specifically in 'range' and 'accuracy' ( $Z = -2.236$ ,  $p < .05$ ). For keyword-only picture group, it was found that the learners' speaking skills increased from poor level to fair and good levels at the significant difference 0.01 ( $Z = -3.145$ ,  $p < .01$ ). The three highly improved sub-skills were 'interaction', 'fluency' and 'range'. For further studies, the researcher should use a larger sample size and add a control group. It may well be interesting to compare the effects of using multi-modalities (pictures, keywords and sound) and bi-modalities (pictures and sound) on learners' English speaking abilities.

**Keywords:** Captioned Pictures; English Speaking Ability; Dual Coding Theory; Young Learners

### บทคัดย่อ

วัตถุประสงค์ของงานวิจัยนี้คือเพื่อศึกษาผลของการใช้รูปภาพที่มีคำบรรยายได้ภาพและรูปภาพที่มีเฉพาะคำสำคัญได้ภาพต่อความสามารถด้านการพูดภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 2 และศึกษาปัจจัยเสริมและอุปสรรคของการใช้รูปภาพทั้งสองประเภทที่มีผลต่อความสามารถทางการพูด กลุ่มตัวอย่างคือนักเรียนชั้นประถมศึกษาปีที่ 2 ของโรงเรียนประถมศึกษาของรัฐแห่งหนึ่งในจังหวัดสงขลา ภาคใต้ของประเทศไทย จำนวน 49 คน (หญิง 25 คน และชาย 24 คน) ซึ่งได้มาจากการเลือกแบบเจาะจง และสุ่มให้กลุ่มหนึ่งเรียนภาษาอังกฤษโดยใช้รูปภาพที่มีคำบรรยายได้ภาพและอีกกลุ่มหนึ่งเรียนภาษาอังกฤษโดยใช้รูปภาพที่มีเฉพาะคำสำคัญได้ภาพ ข้อมูลได้จากผลการทดสอบก่อน-หลังเรียน ความคงทน และผลการสัมภาษณ์ถึงโครงสร้าง ใช้เกณฑ์การประเมินความสามารถทางการพูดของ CEFR ผลการวิจัยพบว่า กลุ่มที่เรียนภาษาอังกฤษโดยใช้รูปภาพที่มีคำบรรยายได้ภาพมีการพัฒนาทางการพูดเพิ่มขึ้นอย่างมีนัยสำคัญ จากระดับอ่อน เป็นระดับปานกลาง ( $Z = -2.236$ ,  $p < .05$ ) โดยเฉพาะในด้านคำศัพท์และด้านความถูกต้องในการใช้ภาษา ในขณะที่กลุ่มทดลองอีกกลุ่มหนึ่งมีความสามารถทางการพูดเพิ่มขึ้นจากระดับอ่อน เป็นระดับปานกลาง และระดับเก่ง โดยมีทักษะด้านการมีปฏิสัมพันธ์ด้านความคล่องแคล่วในการใช้ภาษา และด้านคำศัพท์เพิ่มขึ้นสูงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ( $Z = -3.145$ ,  $p < .01$ ) ในการทำวิจัยครั้งต่อไปผู้วิจัยควรเพิ่มจำนวนผู้เรียนในกลุ่มทดลองให้

มากขึ้น และเพิ่มกลุ่มควบคุม หรือเปรียบเทียบผลของการใช้สื่อที่ให้ข้อมูลภาพ คำศัพท์ และเสียง กับสื่อที่ให้เฉพาะข้อมูลภาพและเสียงต่อความสามารถทางด้านการพูดของผู้เรียนวัยเยาว์

**คำสำคัญ:** รูปภาพที่มีคำบรรยายได้ภาพ; ความสามารถทางด้านการพูดภาษาอังกฤษ; ทฤษฎีรหัสคู่; ผู้เรียนวัยเยาว์

## Introduction

Speaking is considered one of the four macro language skills necessary for effective communication in any language, especially when speaking to people from different language backgrounds (Madsen, Bowen & Hilferty, 1985). For young learners, aged 7-12 years, speaking is important for their language development. To do so, L1 learners experiment or play with words and sounds through meaning, making process such as interacting with parents, teachers, or peers or taking part in story telling activities (Colon-Villa, 1997; Cook, 2000; Linse, 2005). When teaching English or a second language, which co-occurs with the development of their mother tongue to young learners, EFL teachers should make sure that learning one language should complement the other by employing suitable teaching methods designed specifically for young learners.

Brown (2001) has suggested seven principles for teaching speaking as follows: 1) activities should cover a wide spectrum ranging from accuracy to fluency, 2) teachers should create learners' intrinsic motivation to learn how to speak, 3) authentic language should be used in a meaningful context, 4) feedback and correction are important elements to foster language learning, 5) teachers should bear in mind a natural link between speaking and listening, 6) learners require opportunities to initiate oral communication, and 7) teachers should encourage learners to use speaking strategies during communication.

To successfully teach children to learn how to speak, one should keep in mind the characteristics of children which play a key role in teaching. According to Slatterly and Willis (2001) and Mackay (2006), there are three main characteristics of young learners. First, their attention spans are around 10-15 minutes. Second, they prefer physical activities such as running, jumping, and dancing. Finally, while these learners like to be active, they are tired easily. Slatterly and Willis (2001) proposed that young learners can learn by doing and playing. They can learn languages from listening and being involved in activities or experiences in which they are using the languages. Finally, young learners benefit from repeating words, phrases, and sentences many times. With all of these factors in mind, it is suggested that young learners can learn languages from teachers, friends, and others through storytelling (Mackay, 2006).

Among various teaching methods, it is evident that pictures have an important role to play in teaching young children, especially vocabulary, because they can motivate and capture learner's attention (Mansourzadeh, 2010; Wright, 1990). In the studies of Yoshii (2002) and Al-Ja Afari (2013), they have found that the use of pictures to teach vocabulary has positive effects on learners' attitude and vocabulary retention, while Rowe, Siverman and Mullan (2013) have found that the use of picture-word combinations enhanced four year-old learners' vocabulary knowledge. A study conducted with low proficiency learners reveals that pictorial aids assist learners to retain their knowledge (Yang & Chang, 2013).

Pictures can also be used to teach oral communication. According to Bowen (1982) learners can describe pictures to their partners. Teachers can use pictures to encourage discussion in the classroom and motivate the learners to ask questions creatively based on the pictures. Alternatively, teachers can create interactive conversation using pictures as a prompt. If learners are interested in the details in the pictures, they can ask the teacher questions. Dobson (1992) also stated that the pictures can be good conversation starters and can create different discussions on various topics in the classroom, such as nature, food, classroom, and so on. There are

many methods that can support the learners and the teachers to interact with each other by using pictures.

Storytelling through picture books is one of the interesting ways to teach speaking to young learners. Lever and Sénéchal (2011) suggested using picture books to develop the learners' speaking abilities. Teachers can tell a story slowly. During storytelling, the teacher interacts with learners while he/she is telling the story by asking relevant questions using Wh-questions. For example, "Where is/are X?" and "What does X do?". Most importantly, a teacher should constantly provide feedback in the form of correct sentence to learners, whether learners answer correctly and incorrectly. This is to provide them with the correct model of language use. When the teacher finishes telling the story, the learners should retell the story again by using the connective words "and" or "then." Learners who are listening to a story while looking at pictures receive bi-modal inputs, which in turn strengthens their understanding.

Based on the Dual Coding Theory (DCT) proposed by Paivio (1971,1986), effective learning takes place when learners receive bi-or multi-modal inputs through different sensory systems such as the visual (pictures, written texts) and the auditory (verbal) in learner's working or short-term memory. According to Clark and Paivio (1991) and Mayer (2009), different modalities of input would complement each other rather than interfere. However, it is not always the case that bi/multi-modalities of messages would lead to better learning. Learners with different language proficiencies might have different reactions to such inputs. For instance, young learners who begin to read might benefit more from listening only while adult learners would learn more effectively through reading and listening simultaneously (Sticht & James, 1984; Taylor, 2005).

Another theory that accounts for learning is the Cognitive Load Theory (CLT). This theory influences instructional designs. The main principle of CLT is that the cognitive capacity in a learner's working memory is limited. If learners are required to perform heavy cognitive tasks, their ability to learn will be lowered (Kalyuga, Chander &

Sweller, 1999). There are three different types of cognitive load: 1) intrinsic, 2) extraneous, and 3) germane. Intrinsic cognitive load has to do with the complexity of content, being vocabulary or grammar. For learning to occur, learners should have sufficient prior knowledge. If instructional materials are unimportant, it is said to cause extraneous cognitive load in learners. Another source of this type of cognitive load is that the materials or inputs are presented in bi-or multi- modes and the messages are not highly relevant. This results in learners facing a cognitive overload situation which negatively affects learners' working memory, and therefore, comprehension ability. If teaching materials are at the right level of learners' ability and relevant to learning objectives, then germane cognitive load created by this combination would automatically enhance learners' schemas (Mayer, 2005).

Given that modalities of input or instructional materials contribute differently to linguistic performance, the present study applies both DCT and CLT in comparing how two different multimodal inputs (captions, pictures and sounds or CP and keywords, pictures, and sounds or KP) affect Grade 2 learners' English speaking abilities. This study is driven by two research questions.

1. Are there any differences between the effects of using captioned-pictures and keyword-only pictures on Grade 2 learners English speaking abilities? If so, how?
2. What are facilitating factors and barriers of using captioned-pictures and keyword-only pictures to teach young learners' speaking abilities?

### Research Methodology

Quasi-experimental research design with two experimental groups was adopted to collect both quantitative and qualitative data. Quantitative data were obtained from three speaking tests: pre, post and retention. Qualitative data were obtained from individual interviews with seven high and seven low achievers.



## Participants

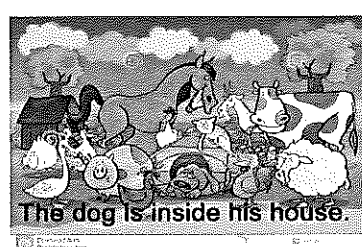
The study recruited 49 EFL Grade 2 students (25 female and 24 male), aged 7-9 on average. Three students were from Myanmar and 46 were Thais. They were in a public primary school in Songkhla, southern Thailand. The researcher collected the data in the second semester of the 2015 academic year. The participants represented a homogeneous group based on the pre-test scores. They were chosen by purposive sampling and randomly assigned into two treatment groups: a group in which captioned pictures were used (26 participants) and a group in which keyword-only pictures were used (23 participants).

## Instruments

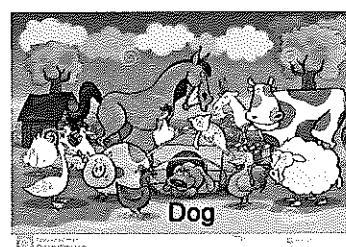
1. **Two types of pictures.** The researchers selected pictures, following the Basic Education Core Curriculum B.E. 2551 (A.D.2008). The contents of pictures covered numbers, animals, classroom items, sports, occupations, time, clothing, fruit and vegetable, weather, actions, food, shapes and colors, body parts, days and seasons. All pictures were approved by three EFL experts prior to the experiment. They were piloted with students who had similar English proficiency in another primary school. The first type of picture included three modalities, i.e. picture, caption, and sound. The group that received this treatment was called CP. The second type of picture included 3 modalities, i.e. picture, keyword only, and sound. This group was referred to as KP. Following the techniques used by Lever and Sénéchal (2011), during speaking activities, the researcher teacher described the picture using basic English structure containing 5-7 words per sentence (Lutz & Huitt, 2003). For example, "The dog is inside his house". The teacher's verbal description was exactly the same as that in the caption. Each description was repeated several times so that the participants could repeat it after the teacher. After that, to initiate interactions, the teacher asked Wh-questions such as "Where is the dog?" "What color is the dog?" or "How many dogs can you see in this picture?". In short, the teacher used exactly the same teaching techniques to teach the two groups. The only difference was that the CP group saw the full captions

while looking at the teacher and listening to the teacher uttering the same description as in the caption, whereas the KP group saw only the keyword such as “Dog” while looking at the teacher and listening to the teacher uttering the full description (The dog is inside his house). Finally, the participants were asked to retell the whole story.

### The Examples of Captioned and Keyword-only Pictures



Captioned Picture



Keyword-only Picture

2. **Speaking test.** The same picture was used three times in the pre, post and retention tests. The pre and post tests were 15 weeks apart while the post and retention tests were only two weeks apart. Each participant was required to describe a picture in two minutes. If he/she could not describe the picture, the teacher would elicit the answers by asking simple questions to help him/her. For example, What is this? What are they doing? All the tests were recorded for grading. The criteria used to examine speaking abilities were based on the Common European Framework of Reference (CEFR). There were five sub-skills of oral assessment criteria grid: range, accuracy, fluency, interaction and coherence (University of Cambridge ESOL Examinations Research and Validation Group, 2009). Three raters who were EFL experts (One American and two Thais) were employed to obtain inter-rater reliability. Each rater assessed the participants' abilities following the grid. The results obtained from the three raters were averaged.

3. **Semi-structured interview.** The purpose of the interview was to provide additional information to help interpret the quantitative data regarding factors contributing to the participants' speaking abilities as well as problems they encountered while studying. Examples of interview questions were as follows:

1. What do you think about listening to the teacher's description and looking at the caption?
2. Did you look at the caption when you wanted to describe the picture?
3. Can you remember what the caption says?

#### **Data collection procedure**

This study consisted of five stages. The entire procedure took 15 weeks. The procedure was as follows. Firstly, the participants took the speaking test individually to assess their baseline knowledge concerning speaking abilities. The participants were randomly assigned into two groups (CP and KP). Except for the different mode of captions, the sequence of teaching was all the same. Each class was 50 minutes long and met twice a week over 15 weeks. After that, the participants took the post test, followed by individual interviews. Finally, two weeks after the post test, the participants took the retention test.

#### **Data analysis**

To answer the first research question, the test results were tallied to arrive at the percentages of participants (poor, fair, and good levels). The percentages were then analyzed using Wilcoxon signed ranks test to examine the learners' English speaking ability in pre, post and retention tests for within group comparison and the differences of English speaking abilities between groups were analyzed by employing the Mann Whitney U test.

To answer the second research question, qualitative data were analyzed for themes and then classified.

## Findings

To compare the results of the pre, post and retention tests of the two groups, descriptive statistics were applied. Table 1 shows the percentages of participants in the two groups who achieved certain levels of English speaking abilities in the pre, post and retention tests. The results reveal that 100% of the participants in both groups were all at poor level in the pre-test, suggesting the inability to use English to give details of people or concrete situations by using simple words.

For the post test, the results revealed that the CP group had shown some apparent development in two sub-skills, namely, range and accuracy. To be specific, 35% of the participants achieved fair level of 'range' and 27% of learners reached fair level of 'accuracy'.

As for the KP group, they showed a more remarkable improvement than the CP group. That is, 52% of them reached fair level and 9% reached good level of 'interaction'. Thirty-nine percent achieved fair level and the 26% were at good level of 'range'. For 'fluency', 39% of the participants reached fair level and 17% reached good level. The results suggest that the participants who looked at the keyword only picture while listening to the teacher describing the picture developed their vocabulary knowledge, could answer questions when asked, and were more fluent. However, they still could not use connective words, such as "and" or "then", to connect the situations that they saw.

For the retention test, the overall results showed a decline in English speaking abilities, especially in the KP group. To be specific, those who demonstrated their abilities at good level in the post test only performed at fair level. It is suggested that the influence of the treatment is not long-term.

Table 1: Percentages of participants who achieved different level of English speaking abilities in pre, post and retention tests

Skills	Levels	CP Group						KP Group					
		Pre-test		Post-test		Retention Test		Pre-test		Post-test		Retention Test	
		n	%	N	%	n	%	n	%	n	%	n	%
Overall	Poor	26	100	21	81	22	85	23	100	11	48	15	65
	Fair	0	0	5	19	4	15	0	0	6	26	8	35
	Good	0	0	0	0	0	0	0	0	6	26	0	0
Range	Poor	26	100	17	65	12	46	20	87	8	35	8	35
	Fair	0	0	9	<u>35</u>	14	54	3	13	9	<u>39</u>	15	65
	Good	0	0	0	0	0	0	0	0	6	<u>26</u>	0	0
Accuracy	Poor	26	100	19	73	20	77	23	100	12	52	15	65
	Fair	0	0	7	<u>27</u>	6	23	0	0	6	26	8	35
	Good	0	0	0	0	0	0	0	0	5	22	0	0
Fluency	Poor	26	100	21	81	23	88	23	100	10	43	17	74
	Fair	0	0	5	19	3	12	0	0	9	<u>39</u>	6	26
	Good	0	0	0	0	0	0	0	0	4	<u>17</u>	0	0
Interaction	Poor	26	100	21	81	21	81	23	100	9	39	12	52
	Fair	0	0	5	19	5	19	0	0	12	<u>52</u>	11	48
	Good	0	0	0	0	0	0	0	0	2	<u>9</u>	0	0
Coherence	Poor	26	100	21	81	24	92	23	100	13	57	15	65
	Fair	0	0	5	19	2	8	0	0	9	39	8	35
	Good	0	0	0	0	0	0	0	0	1	4	0	0

Table 2 below shows the results of Wilcoxon signed-rank test analysis to examine the English speaking of participants within groups.

Table 2: Wilcoxon signed-rank test summary of English speaking abilities of the each groups

Skills	CP				KP			
	Test				Test			
	Pre and Post		Post and Retention		Pre and Post		Post and Retention	
	Z	Asymp. Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)
Overall	-2.236*	0.03	-.577	0.56	-3.145**	.00	-2.352*	.02
Range	-3.000**	0.00	-1.890	0.06	-3.286**	.00	-1.355	.18
Accuracy	-2.646**	0.01	-.447	0.65	-3.017**	.00	-1.890	.06
Fluency	-2.236*	0.03	-1.000	0.32	-3.314**	.00	-2.653**	.01
Interaction	-2.236*	0.03	.000	1.00	-3.557**	.00	-1.667	.10
Coherence	-2.236*	0.03	-1.342	0.18	-3.051**	.00	-1.342	.18

Note: \*p < .05 and \*\* p<.01

As shown in Table 2, a similar pattern of improvements in English speaking abilities within groups existed. That is significant improvements in the post test results of both groups were found overall (Z= -2.236, p <.05 and -3.145, p <.01 respectively), while there was an overall significant difference in the KP group, indicating that the

participants who looked at the keyword-only pictures while listening to the teacher could retain their English speaking abilities in the long run.

Table 3 demonstrates the results of the Mann-Whitney U test analysis to examine the differences in English speaking of participants between groups.

Table 3: Mann-Whitney U test summary of English speaking abilities between the two groups

Tests	Skills	CP	KP	Z	Asymp. Sig. (2-tailed)	Effect Size
		Mean Rank	Mean Rank			
Pre-test	Overall	25	25	0	1.00	0
	Range	23.5	26.7	-1.881	0.06	-0.56
	Accuracy	25	25	0	1.00	0
	Fluency	25	25	0	1.00	0
	Interaction	25	25	0	1.00	0
	Coherence	25	25	0	1.00	0
Post-test	Overall	20.63	29.93	-2.701**	0.01	-0.84
	Range	20.44	30.15	-2.628**	0.01	-0.81
	Accuracy	21.92	28.48	-1.879	0.06	-0.56
	Fluency	20.33	30.28	-2.862**	0	-0.9
	Interaction	20.02	30.63	-3.039**	0	-0.96
	Coherence	22.12	28.26	-1.874	0.06	-0.56
Retention Test	Overall	22.77	27.52	-1.56	0.12	-0.46
	Range	23.69	26.48	-0.8	0.42	-0.23
	Accuracy	23.65	26.52	-0.896	0.37	-0.26
	Fluency	23.33	26.89	-1.299	0.19	-0.38
	Interaction	21.71	28.72	-2.108*	0.03	-0.63
	Coherence	21.88	28.52	-2.324*	0.02	-0.7

Note: \*p < .05 and \*\* p<.01

As shown in Table 3, a different pattern of improvements in English speaking abilities between groups in pre-test and post-test, shows overall significant improvement of the KP group in post test results ( $z = -2.701, p < .01$ ). Further, the effect size value (-0.84) suggested high practical significance, meaning that the influence of using keyword-only pictures on learners' speaking abilities was 84%.

Based on the results of the post and retention tests, KP group outperformed CP group regarding interaction and coherence ( $z = -2.108, p < .05$  and  $z = -2.324, p < .05$ ). It showed that learners in KP group still retained their speaking abilities at A1 level.

#### Facilitating factors and barriers to the use of captioned-pictures and keyword-only pictures for young learners' speaking abilities

The results of the interviews reveal that HA differed from the LA in CP and KP groups in relation to the modality effects caused by the treatments employed in the present study. The HA in both groups reported that they relied on three modality language inputs i.e. the CP group relied on pictures, captions and sound while the KP group relied on pictures, keywords and sound. If some words in the captions or keywords were difficult, they relied on their auditory sense i.e. listening to the teacher, as reflected in an excerpt below.

"I looked at...pictures and caption together....looked at easy words....If there were some difficult words, I listened to the teacher. I could read, pronounce and speak out".

HA 1

"I looked at both pictures and captions because I could read all words...repeat after the teacher... I understood and I could speak out".

HA 2



In contrast, LA relied on only two modalities (pictures and teacher's description) because they could not read the captions, as shown in the next excerpt.

"I looked at pictures and keywords but I could not read. I listened to the teacher and repeated after her".

LA 1

They also reported that the fact that the teacher repeated the utterance several times helped their remembering, thus they could speak quite comfortably.

As for the KP group, HA preferred to be exposed to all three modalities simultaneously; they reported that they relied on listening to the teacher only when they found some vocabulary difficult. If they knew certain vocabulary, they relied only on pictures and described the picture immediately. By contrast, LA reported that they looked at the pictures and listened to the teacher and ignored the keywords, meaning that they preferred only two modalities. Repetitions were also found to be important to this group of participants to be able to provide an oral description of the pictures in English.

One problem that HA from both experimental groups had in common was the length of either captions or keywords. It might be difficult for them to remember and speak. For example, the CP group found that "The boy is taking photos." was too long to remember, and the KP group commented that "taking photos" was problematic.

## Discussion

The aims of this study were to examine the effects of using captions and keywords-only pictures to teach speaking to Grade 2 learners and to study factors affecting their speaking abilities. The results revealed that Grade 2 learners' English speaking abilities in both groups significantly increased from poor level to fair level in the post-test but insignificant in the retention test. The two sub-skills that the learners improved were 'range' and 'accuracy'.

Interestingly, a significant increase in the speaking abilities of the learners in keyword-only picture group was found to improve from poor level to fair and good levels. The three improved sub-skills were interaction, range, and fluency.

The above finding indicated that 'range' was the common sub-skill that both groups of learners could improve. The result was congruent with previous studies which discovered the positive effect of using pictures on young learners' vocabulary learning (Al-Jafari, 2013; Mansourzadeh, 2014; Rasheed and Mohammed, 2007; Rowe, Silverman, & Mullan, 2013; Yoshii and Flaitz, 2002). One plausible explanation for this could be that young learners in this study had had little exposure to English (Chang and Read, 2006; Goh, 1999; Shang, 2008). Therefore, it was not unusual for them to have limited speaking abilities and could handle only vocabulary at the beginning of their English education (Hayati and Mohmedi, 2011).

Findings from the KP group revealed that the learners were better able to remember vocabulary from the three modalities (keywords, pictures, and sounds). When comparing the complexity of keywords and captions, it can be seen that the caption is far more complex than single words. Moreover, vocabulary presented in the experiment was concrete and related closely to their background knowledge. Such finding can be supported by the interview data in which the low English proficiency participants reported that they also looked at pictures and listened to sounds because they could not read. The finding is also in accordance with that in previous studies which posited that young learners who could not read books were good at listening (Sticht and James, 1984; Taylor, 2005). Viewed from CLT, the KP group had a lighter intrinsic cognitive load. According to Kalyuga, Chander & Sweller (1999), Mayer (2005), the lighter the intrinsic cognitive load, the easier for working memory which in turn automatically results in learning or creating learners' schemas (Sweller, van Merriënboer and Paas, 1998; Pollock et al, 2002). From the DCT point of view, it can be argued that different modalities of input would complement each other rather than interfere (Paivio, 1986; Clark and Paivio, 1991; Danan, 2004; Mayer, 2009).

In conclusion, it could be said that the speaking ability of the CP group significantly developed from poor to fair. The two sub-skills that they developed the most were 'range' and 'accuracy', indicating that they had better vocabulary knowledge and were more able to use to construct simple English to describe pictures when compared to the pre-test results. In contrast, the KP group's oral production ability significantly increased from poor to good, with 'interaction', 'range', and 'fluency' that developed the most. This suggests that they could interact with the teacher more fluently and also had more knowledge in vocabulary. As far as Dual Coding Theory and Cognitive Load Theory is concerned, it could be concluded that for young learners with low English proficiency multi-modality input with picture, keyword and sound would yield better learning outcome than when picture, caption and sound are used because they have not yet master their reading ability.

#### **Pedagogical recommendations and implications for further studies**

Based on findings of the current study, it is recommended that the teacher might use only keywords below the pictures because young learners who are 7-9 years old can learn well through pictures, keywords, and sound. They can remember and speak out easily. Repetition is also meaningful for the beginners who started learning a language. The teacher should repeat words, phrases, or sentences many times while he/she is teaching through the uses of pictures in the classroom because learners will remember and speak English well. The teacher should also create speaking games by using pictures to promote English speaking ability in classroom. For example, the teacher might use pictures for whispering game because learners can see pictures and listen to sound from teacher. After that, they can practice English speaking when they whisper sentences to their classmates several times. Looking at pictures and listening to sound can support them to remember and speak out more easily. They also promote new vocabulary learning.

For further studies, the researcher should use a larger sample size for better statistical analysis. Also, including a control group might yield a stronger finding. It may well be interesting to compare the effects of using multi-modalities (pictures, keywords and sounds) and bi-modalities (pictures and sounds) on learners' English speaking abilities.

## References

- Al-Ja Afari, I. S. (2013). Using pictures in vocabulary in Grades 5 and 6 classroom, Sharqiya South Region.
- Bowen, B. (1982). *LOOK Here! Visual aids in language teaching*. Macmillan Publishers Ltd. London.
- Brown, H.D. (2001). *Teaching by principals*. Pearson Education, New York.
- Chang, C.S. & Read, J. (2007). Support for foreign language listeners: Its effectiveness and limitations. *RELC*. 38(3), 375-395.
- Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149-170.
- Cook, G. (2000). *Language play, language learning*. Oxford: Oxford University Press.
- Colon-Vila, L. (1997). Storytelling in an ESL Classroom. *Teaching Pre K-8*, 27(5), 58-59.
- Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators' Journal*, 49, 67-77.
- Dobson, M. J. (1992). *Effective techniques for English conversation group*. Washington: United States Information Agency.
- Goh, C. (1999). How much do learners know about the factors that Influence Their Listening comprehension?. *Hong Kong. Journal of Applied Linguistics*, 4(1), 17-42.
- Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 42(1), 181-192.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing splitattention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-371.
- Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, 108(1), 1-24.
- Linse, Caroline. T. (2005). *Practical English language teaching: Young Learners*. McGraw Hill, New York.
- Lutz, S., & Huitt, W. (2003). *Information processing and memory: Theory and applications*. Educational Psychology Interactive. Valdosta, GA: Valdosta State University.
- Mackay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
- Madsen, Harold, J. Donald Bowen and Ann Hilferty. (1985). *TESOL techniques and procedures*. Cambridge, MA: Newbury House Publishers.

- Mansourzadeh, N. (2014). A comparative study of teaching vocabulary through pictures and audio-visual aids to young Iranian EFL learners. *Journal of Elementary Education*, 24(1), 47-59.
- Mayer, R. E. (2005). *Cognitive theory of multimedia learning*. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia learning*. New York: Cambridge University Press.
- \_\_\_\_\_.(2009). *Multimedia learning*. New York: Cambridge University Press.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
- \_\_\_\_\_.(1986). *Mental Representations*. New York: Oxford University Press.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12(1), 61–86.
- Rasheed W. & Mohammed H. (2007). *The Impact of using pictures in vocabulary achievement for fifth primary pupil*. College of Basic Education.
- Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, 38(2), 109-117.
- Shang, H.F. (2008). Listening strategy use and linguistic patterns in listening comprehension by EFL learners. *The Intl. Journal of Listening*. 22(1): 29-45.
- Slattery, M. & Willis, J. (2001). *English for primary teachers*. Oxford: Oxford University Press.
- Sticht, T. and Jamcs, J. (1984). *Listening and reading*. P.D.Perirson (Ed) *Handbook of Reading Research*. New York: Longman. 293-317.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10 (3), 251-296.
- Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals*, 38 (3), 422-427.
- University of Cambridge ESOL Examinations Research and Validation Group. (2009). *Example of Speaking Performance at CEFR Levels A 2 to C 2*. Cambridge: Cambridge University Press.
- Wright, A. (1990). *Pictures for language learning*. Cambridge: Cambridge University Press.
- Yang, Y.T.C.& Chang, C.H. (2013). Empowering students through digital game authorship: *Enhancing concentration, critical thinking, and academic achievement*. *Computer & Education*, 68, 334-344.
- Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO journal*, 33-58.

## VITAE

Name: Nittaya Manuy

Student ID: 5711121055

Educational Attainment

Degree	Name of Institution	Year of Graduation
Bachelor of Arts (Language for Development) (Second Class Honors)	Prince of Songkla University (Hat Yai Campus)	2012

Scholarship Awards during Enrolment

Scholarship supported by Graduate School, PSU, Hat Yai Campus

Scholarship supported by Liberal Arts, PSU, Hat Yai Campus

List of Publication and Proceeding

Manuy, N., & Laohawiriyanon, C. (2016). Effects of using captioned picture on Grade 2 learners' English speaking abilities. *Proceedings of ICHISS 2016: 8<sup>th</sup> International Conference on Humanities and Social Sciences*. National Defence University of Malaysia. (pp 217-225). Selangor, Malaysia.

Manuy, N., & Laohawiriyanon, C. (2016). Effects of using captioned and keyword-only pictures on Grade 2 learners' English speaking abilities, *Journal of International Studies*. (manuscript submitted for publication).