# Syntactic Processing in Read-Aloud Performance of Second Language Learners

**Soisithorn Isarankura**[1]
*Faculty of Arts, Dhurakij Pundit University, Bangkok, Thailand*

## Abstract

This study investigated the syntactic processing of Thai EFL learners with high and low English proficiency by examining pause patterns in a read-aloud task. It further examined the relationship between syntactic processing and English proficiency of L2 learners. Participants included 30 Thai undergraduate students who were classified into high and low proficiency groups based on their English proficiency test scores. Participants were asked to digitally record their voices while reading an Aesop fable out loud. Auditory and acoustic analysis was then applied to these individual recordings to identify their pause locations. The results revealed that the high proficiency group processed language in longer and more syntactically unified units than the low proficiency group. With insufficient syntactic knowledge, pause positions of lower proficiency learners were more varied. Namely, a large number of pauses occurred at locations which did not correlate with English syntactic units. The findings suggest that there is a relationship between L2 proficiency and the ability

[1] Email: soisithorn.isa@dpu.ac.th

to process linguistic representations at the syntactic level.

**Key words**: Information processing; syntactic processing; read-aloud; pause-defined unit (PDU)

**บทคัดย่อ**

งานวิจัยชิ้นนี้ศึกษาการประมวลทางไวยากรณ์ของผู้เรียนชาวไทยที่มีความสามารถทางภาษาอังกฤษในระดับสูงและต่ำโดยวิเคราะห์การหยุดเว้นระยะจากการอ่านออกเสียงภาษาอังกฤษ นอกจากนี้ยังมุ่งที่จะตรวจสอบความสัมพันธ์ระหว่างการประมวลทางไวยากรณ์กับความสามารถทางภาษาอังกฤษของผู้เรียน ผู้เข้าร่วมการวิจัยประกอบด้วยผู้เรียนไทยจำนวน 30 คน ซึ่งถูกจำแนกออกเป็นกลุ่มที่มีความสามารถทางภาษาอังกฤษสูงและต่ำโดยใช้เกณฑ์ในการจำแนกจากคะแนนที่ได้จากแบบทดสอบความสามารถทางภาษาอังกฤษ ผู้เข้าร่วมการวิจัยบันทึกเสียงการอ่านออกเสียงนิทานอิสป จากนั้นจึงวิเคราะห์เสียงที่บันทึกไว้โดยใช้การฟังร่วมกับการวิเคราะห์ด้วยซอฟท์แวร์โปรแกรมเพื่อหาตำแหน่งการหยุด ผลจากการวิจัยพบว่าผู้เรียนที่มีความสามารถทางภาษาอังกฤษในระดับสูงจะประมวลข้อมูลทางภาษาเป็นหน่วยทางไวยากรณ์ที่ยาวกว่าและมีความสัมพันธ์ทางวากยสัมพันธ์ได้ดีกว่ากลุ่มผู้เรียนที่มีความสามารถต่ำกว่า ซึ่งจากการที่มีความรู้ทางวากยสัมพันธ์ที่จำกัดทำให้ตำแหน่งของการหยุดเว้นระยะในกลุ่มผู้เรียนที่มีความสามารถต่ำกว่ามีตำแหน่งที่หลากหลาย อีกทั้งการหยุดเป็นจำนวนมากไม่สอดคล้องกับหน่วยทางไวยากรณ์ จึงอาจสรุปได้ว่ามีความสัมพันธ์กันระหว่างความสามารถทางภาษาที่สองกับความสามารถในการประมวลข้อมูลทางไวยากรณ์ในระดับประโยค

**คำสำคัญ**: การประมวลข้อมูล การประมวลทางไวยากรณ์ การอ่านออกเสียง หน่วยระหว่างช่วงหยุด

# Syntactic Processing in Read-Aloud Performance of Second Language Learners

**Soisithorn Isarankura**
*Faculty of Arts, Dhurakij Pundit University, Bangkok, Thailand*

## 1. Research background

### *1.1 Information processing*

When information is perceived, the human brain normally activates to receive, decode, and produce a piece of information in a process referred to as "information processing." The idea of information processing has been applied to human thought process by many cognitive psychologists who view the individual as a processor of information who, much like a computer, takes in information and follows a program to produce output (McLeod, 2008). Like a computer, the brain codes, stores, and uses information to produce output. Cognitive psychologists assume that information is made available by the environment, which is processed by a series of processing systems (e.g., attention, perception, short-term memory). The received information can be stored, retrieved and transformed using "mental programs", resulting in behavioral responses. For example, in processing information through reading, the eye receives visual information, decodes information and sends it to the brain where it is stored and coded. The information is then transformed systematically, and the output (i.e., behavioral response) might merely be reading silently or orally, what appears on the printed page.

Cognitive psychologists claim that people can process only a limited amount of information at a time without becoming overloaded. Information processing is governed in part by the amount of working memory capacity available in the individual for language comprehension processes. Working memory, an alternative model of short-term memory developed by Baddeley and Hitch (1974, cited in McLeod, 2008) (Figure 1), works with new incoming information during the decoding and encoding processes, while long-term memory (LTM) stores information in a more permanent fashion. Although the relationship between working memory and long-term memory is not understood entirely, it is clear that working memory and long-term memory interact with one another to exchange and match the received and the stored information before an external response is transmitted.

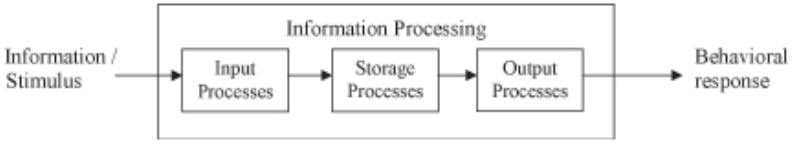



Figure 1. The information processing system
(*Adapted from: McLeod, 2008*.).

### *1.2 Syntactic processing*

Acquiring a language presupposes the ability to analyze and process linguistic input in appropriate ways. Linguistic representations are divided into levels of linguistic knowledge based on the size of linguistic units. These may include phonetics, phonology, morphology, syntax, semantics and pragmatics. These levels of linguistic knowledge are stored in the LTM (Field, 2003). In the perception process, it is believed that information is processed from the smallest linguistic into

the largest units, or bottom-up processing. Here, processing begins with sounds, then proceeds to sentences and finally meanings. In the production process, on the other hand, information tends to be processed from larger to smaller units or from the top-down processing. Here processing begins from meanings to sentences, words, and sounds. as shown in Figure 2 (Figure 2).

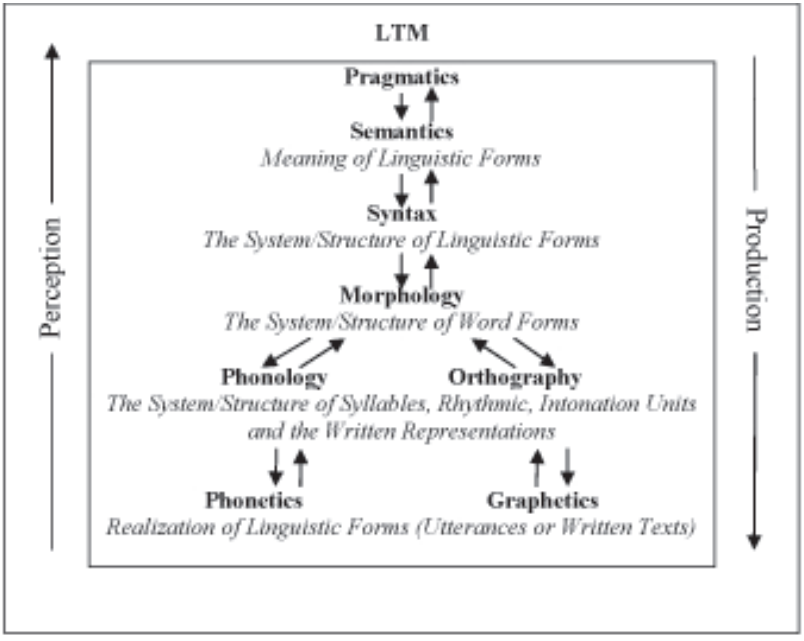



Figure 2. Perception and production linguistic processing *(From Luksaneeyanawin, 2007).*

Clahsen and Felser (2006) asserted that linguistic processing of an L2 is less automatic than L1 processing. Yet, L2 processing may eventually become automatic and can be native-like in highly proficient learners. According to Clahsen

and Felser, L2 processing at the word and morpho-syntactic levels, which involves adjacent or locally-related words, may be more easily mastered as learners' proficiency increases. Syntactic processing is more complex, as it involves the construction of syntactic representations constrained by principles specifying the hierarchical constituent structures at the phrase, clause and sentence levels. Thus, sufficient syntactic knowledge is a prerequisite for successful processing. Research suggests that the degree of automaticity may increase with L2 proficiency in the processing of locally related constituents, but less so in the processing of units involving structurally complex phenomena. Clahsen and Felser (2006) maintain that adult L2 learners' processing of complex syntax continues to be nonnative-like even after many years of L2 usage and exposure. Learners' difficulties are assumed to be attributable to a lack of relevant syntactic knowledge. It may then be claimed that learners with different levels of L2 proficiency have different amounts of syntactic knowledge stored in their long-term memory, resulting in different amounts of working memory capacity available for linguistic processes. This brings to mind a crucial question: Do learners with higher L2 proficiency process information at the syntactic level better than lower proficiency learners given that the more proficient learners have greater amounts of syntactic knowledge? To answer this question, an experiment was conducted to investigate how learners at relatively high and low English proficiency levels divided sentences into syntactically-related units by utilizing pause patterns in a read-aloud task as explicit evidence.

### *1.3 Language processing during a read-aloud task*

To understand how language is processed when a reader is performing a read-aloud task, refer to Figure 3 below (Figure 3).

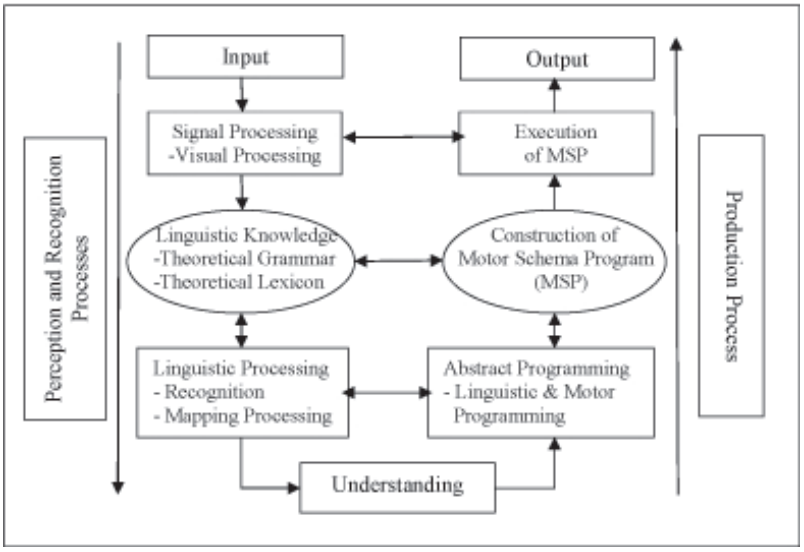



Figure 3. Information processing in the read-aloud task
*(Adapted from Luksaneeyanawin, 2007)*.

As displayed in t Figure 3, language processing, when performing a read-aloud task, involves both perception and production processes. Once new linguistic information (e.g., what is seen in the text) is received, the reader begins processing the first chunk of information in the perceptive domain. The working memory then matches the received information with the linguistic knowledge (e.g., theoretical grammar and theoretical lexicon) stored in the long-term memory in order to understand the message. After that, the reader starts the linguistic and motor programming in the productive domain to create the output by reading the first chunk of processed information out loud. To get new information from the text, the reader starts reading again and processing the next chunk of information. The reader repeats the same process until the whole text has been read.

### *1.4 Pauses and syntactic processing*

One of the basic themes in studying pauses in spoken language has been the relationship between pause occurrence and syntax. Multiple studies (Wu, 2003; Zvonik, 2004; Strangert, 2004) have demonstrated a correlation between pauses and syntactic structure, suggesting that the use of pauses entails syntactic processing of the speaker. Grosjean et al. (1979) asserte that pause occurrence is strongly correlated with the degree of cohesion between words in the utterance. The results of those studies show that pause location tends to occur at major syntactic boundaries like clauses and sentences. In addition, it was found that the complexity and the length of a syntactic constituent affected pause placement. It can then be said that in second language learning that knowledge of L2 constituent grammar, or the ability to recognize relationships between structural parts of a sentence, plays a crucial part in the ability to read in appropriate *pause-defined units* (i.e., chunks of information between pauses).

### *1.5 Factors affecting pauses*

Pause occurrence is constrained by a number of factors. Examples of these factors may include; physiological constraints such as articulation rate and respiratory capacity, linguistic factors like syntactic structure, boundary strength, and phrase length; different speaking tasks; and genres of spoken language—whether the utterance is spoken spontaneously or read. According to Goldman-Eisler (1972), spontaneous speech is much more conducive to pauses of cognitive origin than read-aloud speech. Pauses in spontaneous speech are used primarily to allow additional time needed for the cognitive planning process such as deciding on a topic, selecting ways to present his topic, sorting out appropriate syntactic structures, managing turn-taking, selecting lexical

items, and so on. One can imagine that variables related to cognitive constraints vary depending on each individual speaker. As a result, many speakers may produce non-fluent speech with more frequent pauses than others.

Pauses in read-aloud speech are less constrained by cognitive factors. Findings from many studies suggest that they are affected to a large extent by syntactic structures. Read-speech pauses mainly coincide with certain syntactic boundaries like sentence, clause and phrase boundaries (Goldman-Eisler, 1972; Krivokapi, 2008), whereas pauses in spontaneous speech can be present elsewhere in the utterance. In light of this, the present study utilized a read-aloud task as an instrument to examine pauses as evidence of syntactic processing in learners of two different English proficiency levels.

### *1.6 Relevant research studies*

Several research studies have been conducted to investigate pause patterns of both L1 speakers and L2 learners. Many studies on pauses in native speakers' speech mainly examined the size (or duration) of a pause between words or phrases while others investigated the likelihood of producing a pause (or pause placement) within sentences. One of the basic themes for L2 pause studies has been to study the influence of syntactic structure and semantic coherence on pause location and duration at hierarchical levels of syntactic boundaries.

Strangert (1997) investigated how pause behavior depends on syntactic structure. She found that pause duration tended to be influenced by the complexity of noun phrases (NPs) and verb phrase (VPs) in a sentence. As NP complexity increased, pause duration increased accordingly. In addition, the length of the words immediately preceding the boundary had significant influence on pause duration, i.e., the length of a

pause tended to increase when longer words preceded the boundary.

Wu (2003) examined the relations between syntactic structure and pause placement in read speech of 120 isolated declarative sentences by a male native speaker. The findings suggest that there is a close relationship between the placement of pause and the types of syntactic structure. For example, pauses at the subject-predicate (SP) structure varied depending on whether the subject is a pronoun, a noun, or a noun phrase. The results indicate that, while there was hardly any pause after the subject pronoun, the frequency of pause distribution substantially increased when the subject was a noun phrase. In the verb-object (VO) structure, it was found that pauses hardly occurred. However, the rate of pauses occurrencing at the VO structure tended to increase when there were no pauses in the preceding complex SP juncture or adverbial-head (AH) structure. The results of Wu's study supported Strangert's findings that the pauses are largely influenced by the complexity of the structure.

In the study examining pause patterns preceding and following '*that*' in '*that*-noun clauses', Wongchompoo (2013) found that all five native English speakers produced significantly longer pauses at the preceding position in all 10 target sentences. The results of 50 Thai graduate students, in contrast, showed that pauses were significantly longer at the following '*that'* in nine sentences. However, in the remaining sentence, the duration of pauses preceding and following '*that'* did not differ statistically. Wongchompoo asserted that the pattern performed by Thai learners, which deviated from the native English speaker norm, could be due to L1 transfer. Thai speakers normally pause after '*waa*', which is a Thai translation of '*that*' in English. The pause after '*that*' among Thai learners in his study could be hypothesized to be the

evidence of the carry-over of '*waa*' in L1 to the '*that*-noun clause' in L2 English.

Rungrojsuwan (2007) investigated Thai learners' information processing at different levels of linguistic representations by examining '*pauses in writing*' in two text copying tasks. Participants included 18 Thai first-year students who were classified into groups of high and low based on their scores on the university's placement test. In the first experiment, participants were asked to copy the text by silently reading the original text then writing the first chunk of information that they could remember on a piece of paper. It is claimed that during this process the short-term memory (STM) activates and interacts with LTM in order to retrieve and match information in the text with linguistic knowledge stored in the LTM. To copy a new chunk of information from the text, the participants then repeated the same process until all information was written down. The results show that students with low English proficiency divided the original text into information chunks (PDU) which were greater in number but smaller in length than those of learners with high English proficiency. When examining the positions of '*writing pauses*', it was found that students in the low group had difficulty processing language phonologically, morphologically, and syntactically. These learners paused more frequently within words, between base words and their derivational or inflectional morphemes; and between word boundaries or small groups of words that did not correlate with syntactic units. Comparatively, students in the high group had problems with inflectional morphemes at the morphological level. Learners' problems in morphological complexity in English could hypothetically result from morphological differences between Thai and English. At the syntactic level, however, the '*writing pauses*' in the high group tended to correlate with

syntactic units. The results imply that the level of linguistic representations (i.e., the size of linguistic units) plays a role in L2 acquisition to a large extent.

In the second experiment, Rungrojsuwan repeated the same text copying task process, except this time participants were given a time limit of five seconds per glance. This method of time control was employed in order to compare the size of PDU each participant could process within the time limit. It was found that participants with high English proficiency could recall more words than those in the low group, suggesting that there was a relationship between learners' English proficiency levels and their abilities to process linguistic information. Rungrojsuwan therefore concluded that, even during a simple language related task like text copying, linguistic knowledge is used in language learning. Thus, language teachers could use a simple task like this to aid learners in developing their English language skills.

The discussion hereby brings us to the assumption that learners' ability to process information in an L2 setting depends largely on varying amounts of knowledge in the second language that have been stored in their long-term memory. In the read-aloud task, working memory was also necessary for learners to recognize what was written in the text at the time of scanning to perceive information. After which, working memory interacted with long-term memory to retrieve linguistic knowledge and match information from the two sources. However, the question of whether or not learners with higher L2 proficiency can better process language at the syntactic level in their working memory than lower proficiency learners, given that more proficient learners have greater amounts of syntactic knowledge, still stands. It is assumed, in this study, that the explicit evidence of the processed information could be measured in terms of the number of

pauses as well as the size of pause-defined units (PDU's) that learners produce in reading the text aloud. The size of a PDU (or the number of words per pause) is assumed to indicate the number of words a learner can process at a time. It is hypothesized that lower-level proficiency learners will produce smaller chunks of PDU's than those with higher English proficiency and thus make a higher number of pauses. Moreover, learners with lower proficiency will experience limitations in syntactic processing due a lack of knowledge. As a result, many of their information chunks may not correlate with English syntactic units.

As there have been few studies investigating pause patterns in read-aloud speech in relation to syntactic processing among Thai learners, it is hoped that the findings of this study will provide some evidence for language teachers and that a read-aloud activity can be used as a valuable tool to identify learners' language problems at various levels of linguistic representations for reference in instruction and development of instructional material. It is well noted that children learning to read in their first language (L1) normally start with reading aloud. L1 children need to be trained to associate the sound and word meaning and then to recognize prosodic features in language to become aware of the predictable language patterns of syntax. It is quite surprising, though, that reading aloud tends to receive little attention in L2 classes. In order to correct the errors learners make, it is extremely important for teachers to understand what kind of problems they encounter. A read-aloud task can be used regularly to explore learners' problems as they interrelate with the target language in an EFL classroom.

## 2. The experiment

### *2.1 Objectives of the study*

(1) To investigate L2 processing, at the syntactic level, from the read-aloud performance of Thai EFL learners with relatively high and low English proficiency levels

(2) To examine the relationship between L2 proficiency and the ability to process linguistic representations at the syntactic level

### *2.2 Participants*

Participants involved 30 Thai students drawn from a pool of 195 first-year and fourth-year English major students at a university in Bangkok, Thailand. Based on their scores on an in-house test of English proficiency, 15 students with the highest scores and 15 students with the lowest scores were selected and classified as respondents of the *high* and *low* groups, respectively. The test of English proficiency was a computer-based test consisting of three sections: reading, writing, and listening, with multiple-choice, gap-filling, and sentence completion test-type items. Scores of the high group ranged from 74 to 88 out of 100, and scores of the low group ranged from 30 to 38. The high group consisted of 5 males and 10 females, ages 21 to 26, with 10 to 18 years of formal English instruction. The low group comprised of 3 males and 12 females, ages 17 to 22 with 6 to 13 years of formal English instruction.

### *2.3 Research instrument*

Two instruments were used for data collection:

(1) An Aesop fable, *The North Wind and The Sun*, a U.S. version, transcribed by Dr. Lucinda Hart-Gonzalez and adapted for the purpose of the experimentation

(2) The PRAAT sound analyzing software, Version 5.1.15

The Aesop fable contained 119 words, including one title phrase, and five sentences divided into three paragraphs. The selection of the reading text was based on the following criteria. First, the vocabulary and content were at an appropriate level for learners of lower English proficiency to read without much difficulty. Second, the sentences contained various types of structures (e.g., compound and complex sentences with independent and dependent clauses). Finally, the fable was recommended by the International Phonetic Association (IPA) for use in phonetic descriptions as an illustration of spoken English language.

Prior to the experiment, the text was analyzed for potential pauses at the phrase, clause, and sentence boundaries. Following syntactic analysis, the text was marked with double slashes (//) between words at strong syntactic boundaries and a single slash (/) at weak syntactic boundaries.

### *2.4 Data collection*

The participants were given a copy of the printed text and instructed to read the fable for themselves silently with no time constraint so as to keep information load to the minimum. After the familiarization period, the participants were instructed to start reading the text aloud once at their normal speaking rate. During this time, their speech was digitally recorded on a computer. The average time participants took to complete the task was clocked at 12 minutes.

### *2.5 Data analysis*

To identify pause locations, researcher and two native Thai listeners, who had earned their Master's Degrees in the field of English, listened to each recording and marked the

symbol (/) on a copy of the printed transcript where a pause was heard. To support the preliminary auditory analysis, the PRAAT sound analyzing software was used for acoustic analysis to automatically detect pause locations. Following the criterion set by Campione and Veronis (2002), this study applied the threshold of pause duration at 200 milliseconds (ms) or longer as the acoustic correlate for pausing. At locations where inconsistencies existed in the auditory analysis among the three listeners, the PRAAT program was conducted manually to verify the results and obtain consistency.

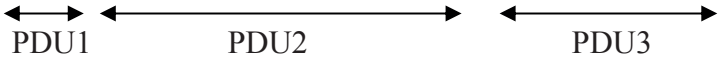
A series of quantitative analyses was conducted to measure the number and size of PDU's. Pause locations were presented and discussed qualitatively to illustrate the extent to which PDU's correlated syntactic units.

## 3. Results and discussion

### *3.1 Number of PDU's*

To give an overall picture of the language-processing abilities of the high and low groups, all chunks between pauses (referred to as '*pause-defined units*' or PDU's) in Luksaneeyanawin [1988]) were counted, per participant, as represented by the following example:

Then, / the North Wind blew as hard as he could, / but the harder he blew, / ….




The PDU numbers for all participants in each group were calculated for mean values and standard deviation as shown in Table 1 (Table 1).

Table 1. Number of PDU's of the high group and the low group.

| Participants | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | $\overline{x}$ | Std. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| High | 24 | 25 | 24 | 19 | 24 | 21 | 22 | 27 | 22 | 16 | 19 | 32 | 23 | 22 | 27 | **23.13** | 3.852 |
| Low | 31 | 34 | 37 | 26 | 19 | 47 | 32 | 37 | 19 | 26 | 38 | 19 | 33 | 42 | 20 | **30.67** | 8.918 |

As can be observed from the data in Table 1, learners in the high group and low group had different numbers of PDU's. The high-learner group read the 119-word text in an average of 23.13 PDU's, whereas the low group read the same text in 30.67 PDU's. This meant that learners in the low group paused more frequently than those with higher proficiency levels, suggesting that the low-learner group had to pause a greater number of times to process the 119-word text than learners in the high group. When testing differences across the groups, statistical results showed significant difference in the number of PDU between the two groups as shown (Table 2).

Table 2. Comparisons of the mean values of PDUs across the groups.

| Group | N | $\overline{x}$ | Std. | t | Sig. |
|---|---|---|---|---|---|
| High | 15 | 23.13 | 3.852 | -3.004 | .007* |
| Low | 15 | 30.67 | 8.918 | | |

*p < .05

### *3.2 Size of PDU*

The size of PDU calculated per participant involved obtaining the average mean value of each group (Table 3) by taking the total proportion of words divided by the number of PDU's.

Table 3. Size of PDU of the high group and the low group.

| Group | N | $\bar{x}$ | Std. | t | Sig. |
|---|---|---|---|---|---|
| High | 15 | 5.28 | .899 | 2.520 | .019* |
| Low | 15 | 4.23 | 1.336 | | |

*p < .05

Table 3 shows that the high learner group read the test at the average of 5.28 words per PDU, whereas the low learner group read the text at the average of 4.23 words per PDU. This suggests that the high group had the ability to process chunks of bigger size better than the low group. The results imply that learners with lower English proficiency tend to have smaller amounts of working memory capacity available for linguistic processes, resulting in their limited ability to process fewer numbers of words at a time. When testing differences across the groups, statistical results revealed that the performance between the high group and the low group differed significantly, as shown in Table 4 (Table 4).

Table 4. Comparisons of the mean values of PDU size across the groups.

| Group | N | $\bar{x}$ | Std. | t | Sig. |
|---|---|---|---|---|---|
| High | 15 | 5.28 | .899 | 2.520 | .019* |
| Low | 15 | 4.23 | 1.336 | | |

*p < .05

### *3.3 Pause positions*

The data set below shows pause positions in the read-aloud performance of the high group and the low group. The numbers in parentheses indicate the percentage of participants producing pauses, and the single slash (/) and double slash (//)

indicate pauses predicted on the basis of syntactic structures (refer to Section 2.3). In the data set, 14 locations (i.e., six at sentence finals and eight at major clause boundaries) are predetermined to be strong boundaries and are marked with (//). Seventeen other locations, marked with (/) for weak boundaries between phrases, are predicted as potential but optional pause sites. The asterisk (*) indicates pauses made by the participants at positions regarded as syntactically inappropriate.

| **The High Group** (n=15) | **The Low Group** (n=15) |
|---|---|
| [**Title**] The North Wind (13.3)/ and The Sun (**100**)// | [**Title**] The North Wind (6.7)/ and The Sun (**100**)// |
| [**S1**] One day, (**86.7**)/ the North Wind (6.7) and the Sun (33.3)/ were arguing/ about (**53.3**)* which of them (13.3) was stronger, (**100**)// when a traveler came along (**53.3**)/ wrapped up (33.3) in an overcoat. (**100**)// | [**S1**] One day, (**66.7**)/ the North Wind (6.7) and the Sun (40)/ were (13.3)* arguing (13.3)/ about (13.3)* which (33.3)* of (13.3)* them (40) was (20)* stronger, (**100**)// when (20)* a traveler (33.3) came along / wrapped (46.7) up (**60**)* in (13.3)* an (6.7)* overcoat. (**100**)// |
| [**S2**] They agreed // that (**93.3**)* the one (6.7)/ who could make the traveler (13.3) take his coat off (**100**)// would be considered (40)/ stronger (13.3)* than the other one. (**100**)// | [**S2**] They (13.3)* agreed (33.3)// that (33.3)* the one (46.7)/ who (6.7)* could (33.3)* make (6.7) the (6.7)* traveler (40) take (13.3)* his (13.3)* coat (20)* off (**53.3**)// would (6.7)* be (6.7)* considered (**60**)/ stronger (13.3)* than (20)* the other one. (**100**)// |
| [**S3**] Then (**80**)/ the North Wind / blew (6.7)* as hard as he could, (**100**)// but (33.3)* the harder he blew, (**100**)// the tighter (20)/ the traveler (6.7)* wrapped | [**S3**] Then (13.3)/ the North (6.7)* Wind / blew (40)* as (20)* hard (**60**)* as (26.7)* he (13.3)* could, (**100**)// but |

| | |
|---|---|
| (6.7)* his coat (13.3)* around him; (**100**)// and at last (**100**)/ the North Wind (6.7)/ gave up trying. (**100**)// | (46.7)* the harder (26.7)* he blew, (**100**)// the (13.3)* tighter (46.7)/ the (13.3)* traveler (33.3)* wrapped (**66.7**)* his (20)* coat (6.7)* around him; (**100**)// and (13.3)* at last (40)/ the North (13.3)* Wind (26.7)/ gave (6.7)* up trying. (**100**)// |
| [**S4**] Then (**66.7**)/ the Sun began (6.7)* to shine hot, (**100**)// and (6.7)* right away (**66.7**)/ the traveler (20)* took his coat off. (**100**)// | [**S4**] Then (13.3)/ the Sun (26.7)* began (20)* to shine hot, (**100**)// and (6.7)* right (6.7)* away (33.3)/ the (6.7)* traveler (40)* took (26.7)* his (33.3)* coat (13.3)* off. (**100**)// |
| [**S5**] And so (**66.7**)/ the North Wind (6.7)/ had to admit (13.3)// that (**86.7**)* the Sun (13.3)* was stronger (26.7)/ than he was. (**100**)// | [**S5**] And (13.3)* so (20)/ the North (13.3)* Wind (33.3)/ had (40)* to admit (**60**)// that (46.7)* the Sun (13.3)* was (6.7)* stronger (46.7)/ than (13.3)* he (6.7)* was. (**100**)// |

(*) *inappropriate pause positions*

From the data set, it can be observed that learners in both groups paused at sentence finals (marked with a full stop) without exception. Variability existed at various locations within sentences. In the high group, pauses of 100% agreement occurred at 13 locations, whereas uniformity among learners in the low group existed at 11 locations. Pauses at syntactically inappropriate positions made by varying percentages of participants occurred at 13 locations in the high group and 56 locations in the low group. As a whole, pause positions of the low group were substantially more varied. For ease of examining how learners processed language at the syntactic

level, the performance of learners in each group will be discussed qualitatively.

*3.3.1 The high group*

The data of the high group show that pauses produced by varying percentages of participants occurred at 45 locations. Among these, pauses with more than 50% of agreement from participants (as highlighted in bold font) accounted for 22 locations, 13 of which were unanimous. These 13 unanimous pauses occurred at the title end and all sentence finals (marked with a full stop), and at major clause boundaries within sentences (five of which were marked with a semi-colon or a comma), which coincided with the predetermined pauses as described in section 2.3. Six other pauses had more than 50% of agreement from the high-group learners corresponding to predicted locations, while the remaining three did not correlate with syntactic units. As indicated in the data set above, pauses occurring at 13 locations, marked with asterisks, were syntactically inappropriate at1 pause in Sentence 1, 2 pauses in Sentence 2, 5 in Sentence 3, 3 in Sentence 4, and 2 in Sentence 5. Of these inappropriate pauses, there were only 3 locations worth discussing wherein more than 50% of the participants paused.

In Sentence 1, a syntactically inappropriate pause made by 53.3% of the participants occurred within a prepositional phrase (between the preposition '*about*' and its object '*which of them*…'). This location was an inappropriate pause site because the preposition, which functions as the head of the prepositional phrase, should not be separated from its complement by a pause. This faulty division of syntactic units among learners in the high group may have been caused partly by the transfer of the learners' first language. The Thai equivalent of '*about*' is normally attached to the verb in

spoken Thai language. Consequently, many Thai speakers tend to pause after the word '*about*' in English sentences.

Two other inappropriate pauses were found at the '*that-noun clause*' structure in Sentences 2 and 5. In Sentence 2, the position where 93.3% of learners paused was found after the word '*that*' in the '*that-noun clause*'. Syntactically, the word '*that*', which functions as a complementizer, is considered an integral part of a noun clause in English, and thus it should not be separated from the clause by a pause. A syntactically appropriate pause site for a sentence with a noun clause structure should be before the complementizer '*that*'. When learners pause before '*that*', it may be hypothesized to result from the transfer of the Thai language in a similar manner as the word '*about*' discussed above. The word '*waa*' is a Thai translation of '*that*', but it performs a different function in a Thai noun clause. While '*that*' is part of a noun clause in English, '*waa*' is a particle in a Thai serial verb construction belonging to the verb it is attached to. These verbs include, for example, verbs of utterance such as *say*, *state*, *argue*, *announce*, and mental verbs such as *think*, *know*, *wonder*, and *understand*. When Thai speakers say a sentence with the '*waa*-noun clause', a pause will appear after the particle '*waa*', and not within the serial verb construction.

The same phenomenon also existed in the pause pattern at the '*that-noun clause*' juncture in Sentence 5. The data shows that 86.7% of learners in the high group placed a pause after the word '*that*'. The consistent pattern of placing a pause after the word '*that*' tends to confirm the assumption that L1 transfer may account for the syntactically inappropriate pause at the '*that*-noun clause' structure among the majority of Thai learners with higher English proficiency. These results are consistent with Wongchompoo's (2013) findings that pauses of Thai graduate students were significantly longer at the position

following the complementizer '*that*' of the '*that*-noun clause' structure. Wongchompoo asserted that it was likely the result of the L1 transfer.

Considering inappropriate pauses in the 5 sentences, it can be seen from the data above that the high learner group made the highest number of inappropriate pauses when reading Sentence 3 at 5 locations: *between verb and adverb*, *after a conjunction*, *after a simple subject*, *between verb and its direct object*, and *before a prepositional phrase*. This could be due to the fact that Sentence 3 is the longest and most complex sentence as it contains 33 words made up of five independent and dependent clauses. One may observe that pauses also occurred most frequently, at 13 words per PDU, in this sentence. This is in line with Wu's (2003) study that the frequency of pause occurrence is largely influenced by the complexity of the structure of the sentence.

*3.3.2 The low group*

The data for the low group shows that pauses were substantially more varied with participants pausing at 90 locations in total. However, pauses with more than 50% of agreement among learners (as highlighted in bold font) occurred only at 18 locations. Among these, pauses of 100% agreement were found at 11 locations—6 at the title end and all sentence finals (marked with a full stop), and 5 at major clause boundaries (marked with a semi-colon or a comma). At 7 other locations, where more than 50% of learners paused, it was found that four of these pauses coincided with the predetermined pause sites, whereas the remaining three did not. From the data set of the low group, syntactically inappropriate pauses, marked with asterisks, occurred at 56 locations (i.e., 10 in Sentence 1, 12 in Sentence 2, 17 in Sentence 3, 9 in Sentence 4, and 8 in Sentence 5). Almost all

of these inappropriate pauses were made by less than 50% of the participants. The following are inappropriate pause locations and frequency of occurrence in the low group as opposed to those in the high group.

| *Inappropriate Pause Locations* | *Frequency of Occurrence* | |
|---|---|---|
| | *Low Group* | *High group* |
| 1. Within a verb phrase | 10 sites | 1 site |
| 2. Within a prepositional phrase | 4 sites | 1 site |
| 3. Within an adverb phrase | 1 site | - |
| 4. Between an adjective and the noun it modified | 3 sites | - |
| 5. Between a noun and its determiner | 8 sites | - |
| 6. Between a verb and its direct object | 3 sites | 1 site |
| 7. Between a verb and an adverb | 2 sites | 1 site |
| 8. Between a copula verb and a complement | 2 sites | - |
| 9. Between a simple subject and a verb | 8 sites | 3 sites |
| 10. Between conjunctions | 2 sites | - |
| 11. Before a prepositional phrase | 3 sites | 2 sites |
| 12. After a conjunction | 10 sites | 4 sites |
| | --- | --- |
| Total | 56 sites | 13 tokens |

From the above data, it can be seen that learners in the low group made inappropriate pauses more frequently than those in the high group at the following locations: *after a conjunction* (10 positions), *within a verb phrase* (10 positions), *between a noun and its determiner* (eight positions), and *between a simple subject and a verb* (eight positions). The high group, on the other hand, paused at fewer inappropriate locations. In addition, the data shows that some high-group learners tended to have problems pausing inappropriately at

the following locations: *after a conjunction* (four positions), *between a simple subject and a verb* (three positions), and *before a prepositional phrase* (2 positions).

It should be acknowledged that although pause occurrence in read speech is constrained by fewer factors than are pauses in spontaneous speech, this qualitative analysis will by no means provide a thorough account of what processes were actually involved when the learners paused. An obvious limitation of this study is that the participants were not asked to state why they had paused at particular locations. Thus, the ability to draw any conclusive assumption with regard to learners' actual sources of inappropriate pausing is limited. Since understanding why such pauses occurred might reveal valuable insights into any linguistic elements that may be associated with a pause, further investigation of the participants' metalinguistic knowledge governing their pause patterns could be pursued.

Another notable point was observed in the pause pattern of the low group at the '*that-noun clause*' levels. The data shows a different pattern from that of the high group. While the majority of learners in the high group paused after '*that*', those from the low group paused before and after in Sentences 2 and 5. Such difference could be attributed to limited syntactic knowledge of learners with lower English proficiency. As supported by the data, learners in the low group made a larger number of pauses when reading the entire text. However, several pauses did not correlate with syntactic units.

From the data of both high and low groups, one can see that of the 12 boundaries marked by punctuation, 6 at the title end and sentence finals, 5 marked with a comma, and 1 marked with a semi-colon, both the high and low groups

paused unanimously at 11 locations. The results suggest that learners in the high and low groups paused in a similar manner at locations where punctuation was indicated. Variation existed at only 1 location in Sentence 1. A pause after '*One day*', which was marked with a comma, had an agreement of 86.7% from the learners in the high group, and 66.7% from those in the low group.

A greater degree of variations between the two groups can be observed at 3 other strong boundaries where no punctuation existed. Two locations occurred at the '*that-noun clause*' juncture in Sentences 2 and 5 as discussed earlier. The third location was a predicted pause at the complex subject-predicate juncture in Sentence 2 (…*the one who could make the traveler take his coat off* // *would be considered*…). At this location, a pause of unanimous agreement occurred in the high learner group, while it had only 53.3% agreement from the low group. Additionally, different pause patterns between the high group and the low group can also be observed at potential pause sites after sentence adverbs which encode a change of topic, as in: *Then* / (in Sentences 3 and 4), *and at last* / (in Sentence 3), *And so* / (in Sentence 5). Although there was no punctuation at these locations, the majority of learners in the high group made a pause at each location, whereas only a small percentage of learners in the low group paused at these sites. This suggests that learners in the low group relied more heavily on punctuation as a pause marker than the high-group learners who appeared to be more aware of the syntactic and semantic ties between phrases and clauses.

## 4. Conclusion

Using a read-aloud task, this experiment was conducted to investigate the L2 syntactical levels of processing of Thai EFL learners with relatively high and low English proficiency

levels. It further examines the relationship between L2 proficiency and the ability to process linguistic representations at the syntactic level.

In relation to the number of pauses and the PDU size, the results reveal that learners with lower English proficiency paused substantially more frequently than those with high proficiency. In other words, low level learners read the 119-word text in a larger number of chunks than the higher proficiency learners. Consequently, their chunks of information, referred to in this study as PDU's, contained a smaller number of words. The findings of the present study are in line with Rungrojsuwan's (2007) results of '*writing pauses*', which demonstrated that lower English proficiency learners divided the original text into a higher number of chunks, and thus their chunks were shorter than those of learners with higher English proficiency. This result likely suggests that learners with varying levels of language proficiency possessed different levels of language processing ability.

Findings of the study investigating pauses and syntactic processing, are supported by Rungrojsuwan's conclusion that pause positions of learners in the high group learners generally occurred between major syntactic boundaries—such as sentences and clauses,and were in agreement with predicted pause sites. Additional pauses were found among the high-group learners between phrases at predictably weaker boundaries. Pauses made at syntactically inappropriate pause locations were substantially fewer in number than those made by low group learners. As can be seen, learners in the low group made syntactically appropriate pauses mainly at sentence boundaries and major clause boundaries, which were signaled by specific punctuation, namely, a full stop, a semi-colon, or a comma. Several other pauses found in the data of the low group occurred between words or within the

syntactically related groups of words that should not be interrupted by a pause. Thus, it can be concluded that learners with higher English proficiency seem to have more syntactically related pauses between sentences, clauses and phrases, while low group learners' pauses were more varied. Furthermore, lower proficiency learners tend to make numerous pauses at word boundaries which did not correlate with English syntactic units probably because they had limited knowledge with regard to the relationships between structural parts of sentences.

Based on the first objective of the study, the number and size of linguistic units indicated the learner's level of linguistic processing and L2 acquisition. It was found in the experiment that syntactic processing obviously occurred in the high group. The results also implied limited syntactic knowledge of learners in the low group. In relation to the second objective, it was found that learners in the low group produced fewer syntactically related pauses than those in the high group, suggesting that acquisition of syntactic knowledge may be achieved at a later stage of L2 development. With increasing proficiency, learners tended to become more aware of syntactic and semantic ties between words and phrases. As a result, it may be claimed that there is a relationship between L2 proficiency and the ability to process linguistic representations at the syntactic level.

## 5. Pedagogical implications

The read-aloud task can be used in the classroom to create awareness of the learner's ability to process information at various levels of linguistic representation through the use of evidence from pause placement.. Not only is the read-aloud activity useful for helping students develop their L2 skills at various linguistic levels, it is also essential for enhancing the

students' pronunciation ability. Teachers can show their students how they can apply their knowledge of parts of speech and syntactic structure to developing their pronunciation skills. As reading aloud has received little attention in EFL classes, it is hoped that the findings will shed some light on how a read-aloud task can be a valuable tool to assess the learner's linguistic processing ability and enhance the learner's second language skills.

## 6. Recommendations for future research

Based on the findings of the study, the following recommendations are made for further research.

Firstly, the participants were not asked to state the reasons for pausing at particular locations. Therefore, the ability to explicate their performance or provided a thorough account of what processes are involved when the participants pause is limited. Further research is thus recommended to investigate the reasons behind each pause by using either think-aloud protocols or retrospective interviews.

Secondly, this study has provided preliminary analyses of pause patterns as explicit evidence of syntactic processing by Thai learners with two different levels of English proficiency. Continued research can be conducted to examine, in greater depth, the characteristics of information chunks produced by learners at varying developmental stages in order to shed some light on common errors made by learners during their inter-linguistic development.

## Acknowledgements

This research could not have been completed without the help of many people. I am particularly grateful to Dhurakij Pundit University and the DPU Research Committee for

providing me with funding for this project. My profound appreciation goes to Assistant Professor Dr. Sudaporn Luksaneeyanawin, Assistant Professor Dr. Pasinee Sornhiran, and Assistant Professor Dr. Namtip Pingkarawat for their insightful suggestions and continuing support. I also wish to thank Thomas Edward Smith, who gave valuable suggestions and edited the first draft, and all participants for their cooperation and willingness to participate in this study.

## Biodata

Soisithorn Isarankura received a B.A. in English (first-class honors) from Thammasat University, Thailand. She holds an M.A. in Linguistics and English Language Teaching from the University of Leeds, U.K., and a Ph.D. in English as an International Language from Chulalongkorn University, Thailand. She is now an Assistant Professor of English at Dhurakij Pundit University. Her research interests have been in the areas of Interlanguage and Second Language Acquisition. Her papers include L2 acquisition of English articles, tenses and aspect, and pronunciation of English prosody. She can be reached at Faculty of Arts, Dhurakij Pundit University, telephone number 0-2954-7300-29, email address: soisithorn.isa@dpu.ac.th.

## References

Campione. E., & Veronis, J. (2002). *A large-scale multilingual study of silent pause duration*. In *Speech Prosody 2002*. Aix-en-Provence, France, April 11-13, 2002. Available from http://www.isca-speech.org/archive

Clahsen, H., & Felser, C., (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*, 564-570.

Field, J. (2003). *Psycholinguistics: A resource book for students*. London: Routledge.

Goldman-Eisler, F. (1972). Pauses, Clauses, Sentences. *Language and Speech, 15*, 103-147.

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology*, *11*, 58-81.

Krivokapi, J. (2008). *Prosodic planning: Effects of phrasal length and complexity on pause duration*. Retrieved from http://www.ncbi.nlm.nih.gov/ pmc/articles/PMC 2131696

Luksaneeyanawin, S. (1988). *The characteristics and functions of pause in Thai*. (A research report). (In Thai). The Faculty of Arts, Chulalongkorn University.

Luksaneeyanawin, S. (2007). Unfolding Linguistics. In W. Aroonmanakun (Ed.), *Unfolding Linguistics*. Bangkok, Chulalongkorn University Printing House.

McLeod, S. A. (2008). *Information Processing - Simply Psychology*. Retrieved from http://www.simplypsychology.org/informationprocessing.html

Rungrojsuwan, S. (2007). Linguistic information processing of second language learners with different levels of language proficiency. In W. Aroonmanakun (Ed.), *Unfolding Linguistics*. Bangkok, Chulalongkorn University Printing House.

Strangert, E. (2004). Speech chunks in conversation: Syntactic and prosodic aspects. *Proceedings of Speech Prosody 2004*, 305-308.

Wongchompoo, N. (2013). *Pausing patterns: Preceding and following 'that' in 'that-noun clauses' between native English speakers and Thai learners of English*. (Unpublished master's paper). Language Institute, Thammasat University, Thailand.

Wu, H. (2003). Distribution of pause in read speech. *The Humanities Study Forum*. Bureau of International Cooperation, Chinese Academy of Social Sciences.

Zvonik, E. (2004). *Pausing and temporal organization of phrases: An experimental study of read speech*. (Unpublished doctoral dissertation). Department of Computer Science, University College Dublin, National University of Ireland.