Ogata, H. et al. (Eds.) (2015). Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education

A Pilot Study of Explicit Reading – Representation of Implicit Chunking in the Readers' Mind

Pin-Fang CHEN[a*], Chi-Te KAO[a], Cheng-Ting CHEN[b], Kuo-Chen LI[c], Shien-Yung CHENG[a] & Jia-Sheng HEH[a]

[a]*Department of Information and Computer Engineering, Chung Yuan Christian University, Taiwan*
[b]*Department of Applied Linguistics and Language Studies, Chung Yuan Christian University, Taiwan*
[c]*Department of Information Management, Chung Yuan Christian University, Taiwan*
*peter5922749@gmail.com

Abstract: In this paper, we describe a pilot study of Explicit Reading (ER), which stands for explicit representation of the implicit chunking process in the readers' mind. A chunk is a meaningful group of words. According to psychological linguistics, reading is a cognitive process, which includes chunking words of a sentence into meaningful groups of words. In this study, a web-based reading system was built for students to read assigned short articles and manually chunk each sentence into a hierarchical chunk, which is the parsing tree of the sentence. The chunking actions are recorded and timed for further analysis. Students are clustered according to the similarity of their chunking results. Two classes of non-native speaking college students of an English course participated in this study and 22 feasible cases were collected. For a particular syntax structure, the chunking style and unit quiz score are correlated. Compared with the reference answer, even the students have high unit quiz score may conduct chunking mistakes.

Keywords: explicit reading, chunking, prefabricated chunk, EFL/ESL reading comprehension

1. Introduction

How do we help non-native poor readers to improve their reading skill? Why a good reader may not be a good writer? What happens in readers' mind when they read or "misread" a sentence? Is there any method to represent the implicit process of reading or "misreading"? How do we diagnose misreading? Kao et al. (2011) designed a system to help Chinese speaking EFL students to "explicitly represent a sentence as a parsing tree (chunking)" by answering to a series of predefined parsing questions, which are designed according to a referenced parsing tree created by teacher. However, during our previous study, we found that different student may have different chunking style, which is probably not the same as the teacher's. Furthermore, an abnormal chunking style may indicate a misread of the sentence. How do we discover and represent these misreads which happens implicitly in the reader's mind. What will we see if we just let the readers to present their own chunking style? In this study, we designed and implemented a web-based reading system, called "Explicit Reading for English", which let students to read assigned short articles and manually chunk each sentence into a hierarchical chunk, which is the parsing tree of the sentence. The method will be described in Section 3. The parsing trees are scored by calculating precision, recall, accuracy, and f-measure of each chunking action with respect to a reference answer, created by teacher. Students are clustered according to their chunking styles to see the correlation with respect to their quiz score of the assigned articles.

2. Literature Review

The concept of chunk was proposed by Miller (1956) and is highly related to working memory (Baddeley & Hitch, 1974). Li (2010) indicated that, since the capacity of working memory is limited, if the reader possesses better knowledge of chunking, their capacity of working memory would be used more efficiently. Chunk has been named by different researchers, such as phrasal lexicon (Becker, 1975), lexical phrase (Nattinger & DeCarrico, 1992), prefabricated chunk (Skehan,1998), formulaic language (Wray, 2002). Nattinger & DeCarrico (1992) said that lexical phrase is the multiword lexical phenomena that exist somewhat between the traditional poles of lexicon and syntax, and the ability to use lexical phrase is important. Zhang & Xu (2007) suggested that teachers could encourage students to find prefabricated chunks in the reading material to increase their awareness to chunk. Later, Li (2010) conducted research on prediction reading comprehension performance by the chunk uses of the students. She concluded that reading comprehension is better predicted by chunk uses in students' writing than by their grammatical accuracy. However, her research measured chunk uses in language production, while reading is a process of language reception, and she indicated that the measurement of chunk uses in language reception is still a challenge. Our purpose is to response to that challenge.

3. Method of Chunking

An example is shown in Table 1. Words are broken into several lines, and each line is a chunk. The indentation before each line represents the hierarchical structure. Chunk with more indentation belongs to the chunk with less indentation to form a chunk/sub-chunk structure. With these chunking strategies, readers can manage their working memory when they read long sentences. We provided some rules to build referenced answer and to demonstrate to students. If it is not chunked properly, some lines will be not integrated with their meanings. Building improper chunks may indicate that the reader may not comprehend that sentence, therefore, a "misread" may occur in the reader's mind. Teacher of the English class chooses some articles from a college English textbook as the reading materials, and parses each article to be the reference answer. Then, the articles are transformed into web pages. Students can login the system by their student ID, then choose article to read. Click between words of the article will break the line, click at the end of a line will un-break it. Click on any word of the line will add indentation, click on one of the leading indents will decrease indentation. At the end of this process, students can save their result. Once students submit the result of an article, they can see the reference answer shown side by side with their own result.

Table 1: An example of parsing tree.

```
Parsing Tree:
It
is essentially
    a set–up
        where
            we
            have
                a warehouse
                    for our products
                that
                serves
                    a certain number of retail outlets
                        close to that warehouse
.
```

4. Experiment and Results

Two classes of college first-year students (N=90) from management school of a university at northern Taiwan are participated in this study. We explained and demonstrated the system and let them practice freely for two hours, and they can practice online anytime after the class. After two weeks,

we let them voluntarily to join the reading of the last article. Finally, 22 feasible cases were collected. Data is analyzed by RStudio version 0.98. During each reading session, the action and elapsed time of break/un-break lines and indent/un-indent lines are recorded. We cluster students into groups according to their chunking results. An interesting result is found when we investigate the different chunking styles of the phrase "close to that warehouse". The clustering result and the quiz score is correlated to the correctness of the chunking results, as shown in Figure 1. The cluster of correct chunking result has higher average quiz score.

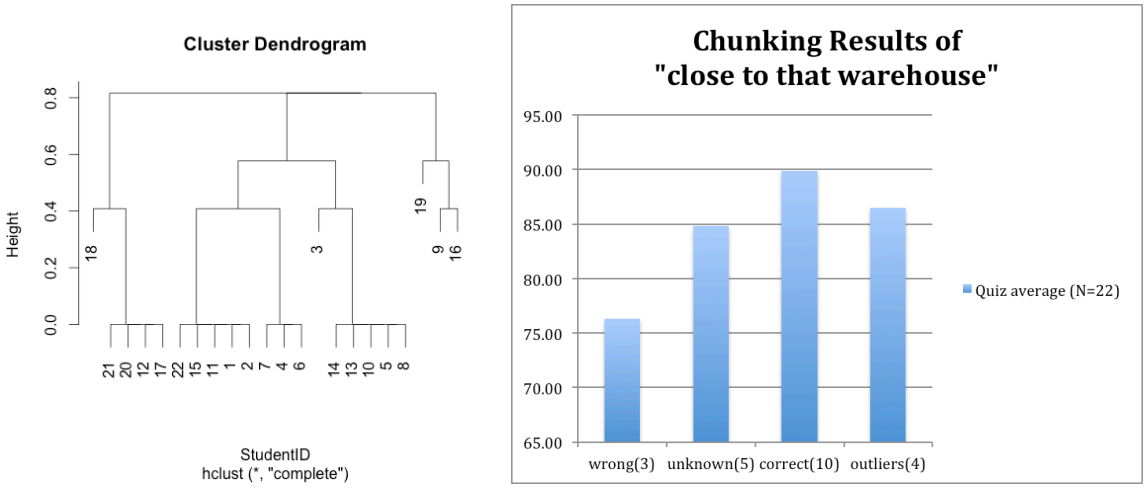

Figure 1. (Left) Clustering Analysis of Different Chunking Style of the phrase "close to that warehouse". (Right) Correctness of Chunking Results vs. Average Quiz Score..

5. Conclusion and Future Works

We designed and implemented the reading system, "Explicit Reading for English", and conducted a pilot experiment to collect data for the improvement of the system. With the explicit representation of chunking, the incorrect chunking results of students can be shown and be diagnosed. In the future, there are many other aspects of research can be investigated through this system, such as the individual consistency of their chunking style, more clustering analysis can be done, more data can be collected to span a broad range of different proficiency levels. And the relation between the ability of chunking in reading and the ability of writing is also very interesting. With a way to represent the implicit reading in the readers' mind, we think it opened a door to connect the aspects of psychological linguistics of reading to the aspect of language acquisition and teaching.

References

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, *8*, 47-89.

Becker, J. D. (1975, June). The phrasal lexicon. In *Proceedings of the 1975 workshop on Theoretical issues in natural language processing* (pp. 60-63). Association for Computational Linguistics.

Kao, C. T., Chang, J. C., Li, K. C., Chen, C. T., & Heh, J. S. (2011). An interactive tools for English reading comprehension by providing questions and hints to student during reading and parsing. *Proceeding of the 19th International Conference on Computers in Education*. Chiang Mai, Thailand.

Li, F. (2010). Chunk use as a predictor of reading comprehension performance. *Journal of Yunmeng*, 31(3), 135-138.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Nattinger, J. R., & DeCarrico, J. S. (1992). Lexical phrases and language teaching. Oxford University Press.

Skehan, P. A. (1998). Cognitive approach to language learning. Oxford University Press.

Wray, A. (2002). Formulaic language and the lexicon (Vol. 5). Cambridge: Cambridge University Press.

Zhang, L. I., & Xu, Y. (2007). Chunk approach for English teaching. *Sino-US English Teaching*, 4(11), 14-18.