A quantitative approach to measuring CFL text difficulty

笪骏

Middle Tennessee State University

jda@mtsu.edu

ACTFL 2009

Nov. 21, 2009, San Diego

1. Chinese language teaching at MTSU

- Classroom instruction: 3 hours per week.
- Textbooks
  - Integrated Chinese (Yao, etc. 2005);
  - Chinese Link (Wu, etc. 2006).
- Supplementary materials
  - Chinese Reading World (CRW) (http://www.chinesereadingworld.org/, The University of Iowa Chinese Program)
- Problems
  - Students reported that CRW beginning level texts were still too difficult for them to comprehend when they finished IC (Level 1);
  - It is time consuming for instructors to find (graded) reading materials for beginning level students.
- Task: How do we select CFL reading materials?

2. Selecting texts for CFL instruction

- Past researches (e.g., Krashen 1985, Fielding 1990, Huckin and Coady 1993) have suggested that providing reading materials at a suitable level of difficulty allows learners to acquire vocabulary in the most efficient manner.
- Test of text difficulty
  - Judgment of instructors/learners;
  - The cloze procedure (Stevens, etc. 1992);
  - Readability formulas.

2.1 Readability

- Readability is how comfortably or easily a written passage or text can be read and understood. It is often measured using readability index, which is an objective estimate of text difficulty in terms of correlates between the reading ease of a piece of writing and an academic reading level.
- English:
  - Flesch-Kincaid Grade Level, Flesch Reading Ease (Flesch 1948, Kincaid, etc. 1975);
  - Smog (McLaughlin 1969);
  - Fry Readability Graph (Fry 1989);
  - …

2.2 Flesch-Kincaid Index

- Flesch–Kincaid Grade Level:

$$0.39\left(\frac{\text{total words}}{\text{total sentences}}\right)+11.8\left(\frac{\text{total syllables}}{\text{total words}}\right)-15.59$$

- Flesch-Kincaid Reading Ease:

$$206.835-1.015\left(\frac{total\ \ words}{total\ \ sentences}\right)-84.6\left(\frac{total\ \ syllables}{total\ \ words}\right)$$

2.3 Measuring Chinese text difficulty (I)

- Early studies on Chinese text readability were conducted by Yu (1960) and Yang (1971).
- According to Yang (1971), Yu's (1960) formula was an adaptation of Flesch's formula where sentence length and polysyllabic word counts (instead of single syllables) were used as predictors of text difficulty.
- Due to the date of publication, I have not been able to locate a copy of Yu's paper.

2.3 Measuring Chinese text difficulty (II)

- Yang (1971) was based on a text corpus containing 85 passages selected from books, magazines and newspapers and other publications for native speakers of Chinese;
- Yang found that number of strokes in a character, the presence of words in a basic word list and the proportion of full sentences in a passage have the most predictive power for text difficulty;
- He derived two formulas, where the first one contains 31 independent variables and the second 7 independent variables. Both formulas make use of the proportion of full sentences, the proportion of the words in his 5,600 basic word list, average number of character strokes, and the frequency of 5, 12, 22 and 23 stroke characters to predict text readability.

2.3 Measuring Chinese text difficulty (III)

- Jeng (2001) compared both linear and non-linear approaches to estimating readability of children's literature textbooks and other children's books in Traditional Chinese. His corpus contains 223 articles with a total of more than 82,000 words selected from children's books for native Chinese speakers in Grade 1 to 6;
- He demonstrated that both linear and non-linear algorithms produced similar results which are slightly less effective than human judgment. Further, variables such as ratio of familiar words, standard deviation of mean number of words per sentence, and mean total number of words per sentence were shown to be good predicators of text readability in a linear regression model.

2.3 Measuring Chinese text difficulty (IV)

- Shen (2005) finds that both sentence length and word frequency information are important factors that predict reading difficulty for American learners of Chinese as a Foreign Language.

2.3 Measuring Chinese text difficulty (V)

- Zhang (2000)
- Measurement 1:
  - Data: 29 textbooks, 5 paragraphs from each textbook.
  - Sentences/per 100 characters.
  - Findings:
    - 6-10 sentences/per 100 characters applies to beginning level textbooks;
    - No distinction between intermediate and advanced level textbooks.
- Measurement 2:
  - Data: 16 textbooks, 10 paragraphs from each textbook.
  - Average number of less frequent words per 100 characters (HSK's Levels 3 and 4 丙、丁级) for Intermediate level and advanced level.
  - Findings:
    - Index: Average sentence length + Average number of less frequent words per 100 characters
- Problems:
  - Sampling method;
  - The use of HSK word list;
  - Word segmentation in Chinese.

2.3 Measuring Chinese text difficulty (VI)

- What's known:
  - Sentence length (characters or words per sentence);
  - Character/word frequency.
- Problems
  - Difficulty in segmenting words in running texts;
  - Time consuming counting of character strokes;
  - Zhang (2000): Reliance on prescribed word list;
  - Samples:
    - Traditional characters (vs. Simplified);
    - Textbooks: Outdated or for native speakers (not adopted for American CFL learners).

3. This study

- Objectives
  - An alternative approach that uses less ambiguous stylistic features;
  - (Hopefully) an easier and less time consuming approach;
  - Samples that will reflect what is currently used in the U.S.;
  - A comparison between CFL texts and texts for Chinese native speakers.

3.1 Data

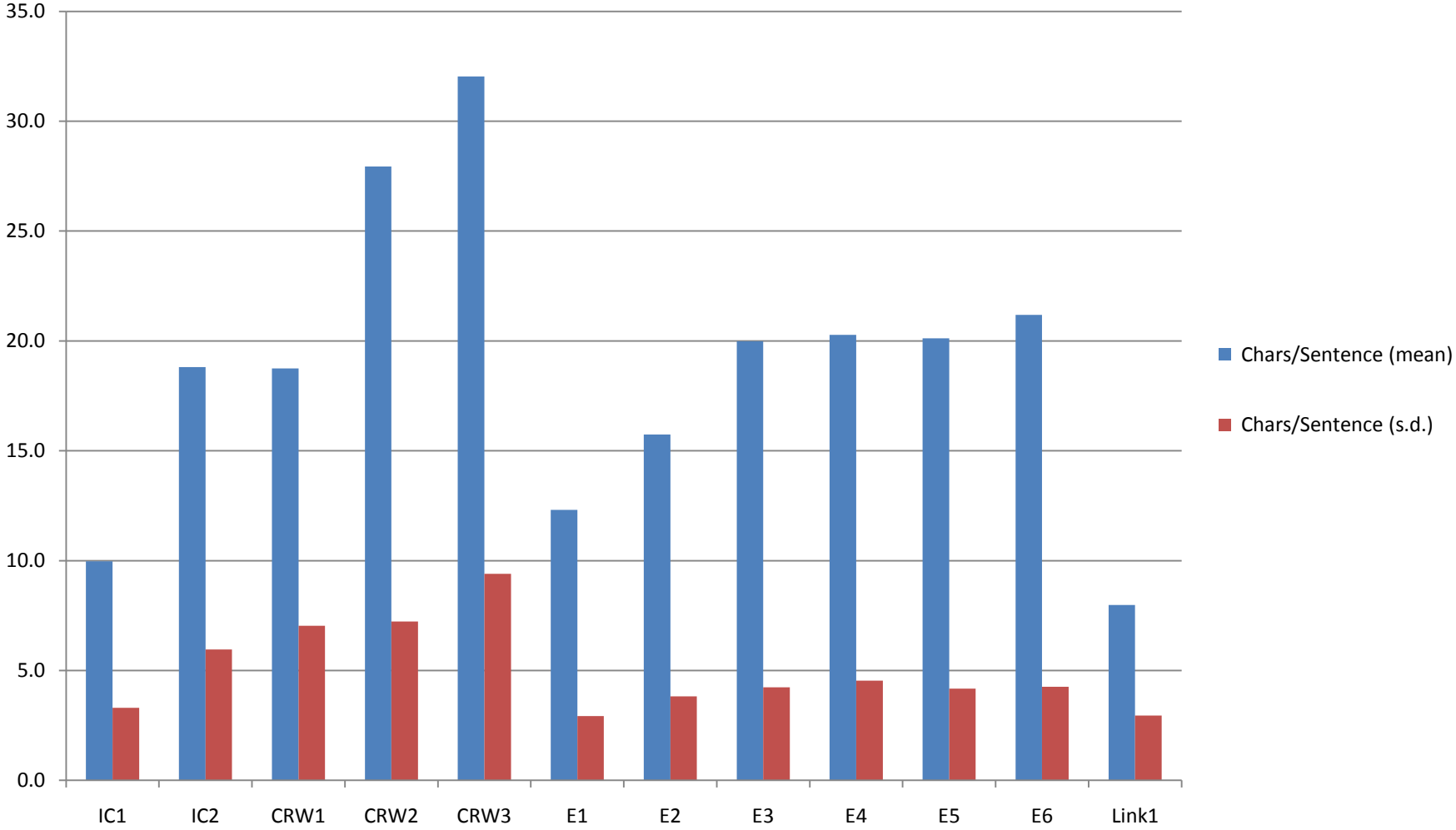
- Data
  - Integrated Chinese (Levels 1 and 2, Second edition): All texts;
  - Chinese Reading World, University of Iowa: 30 passages from beginning, intermediate and advanced levels, respectively;
  - 人教社《语文》一至六年级课本 (2001版): All texts;
  - Chinese Link: First 8 lessons (due to time constraint).

3.2 Measurement

- Intuitive: Average number of characters per sentence;
- Taking into consideration variations in character recurrence ratio: Average number of total characters per sentence + unique characters per sentence.

3.3 Results (I)

Figure 1 Average number of total characters per sentence and their standard deviations

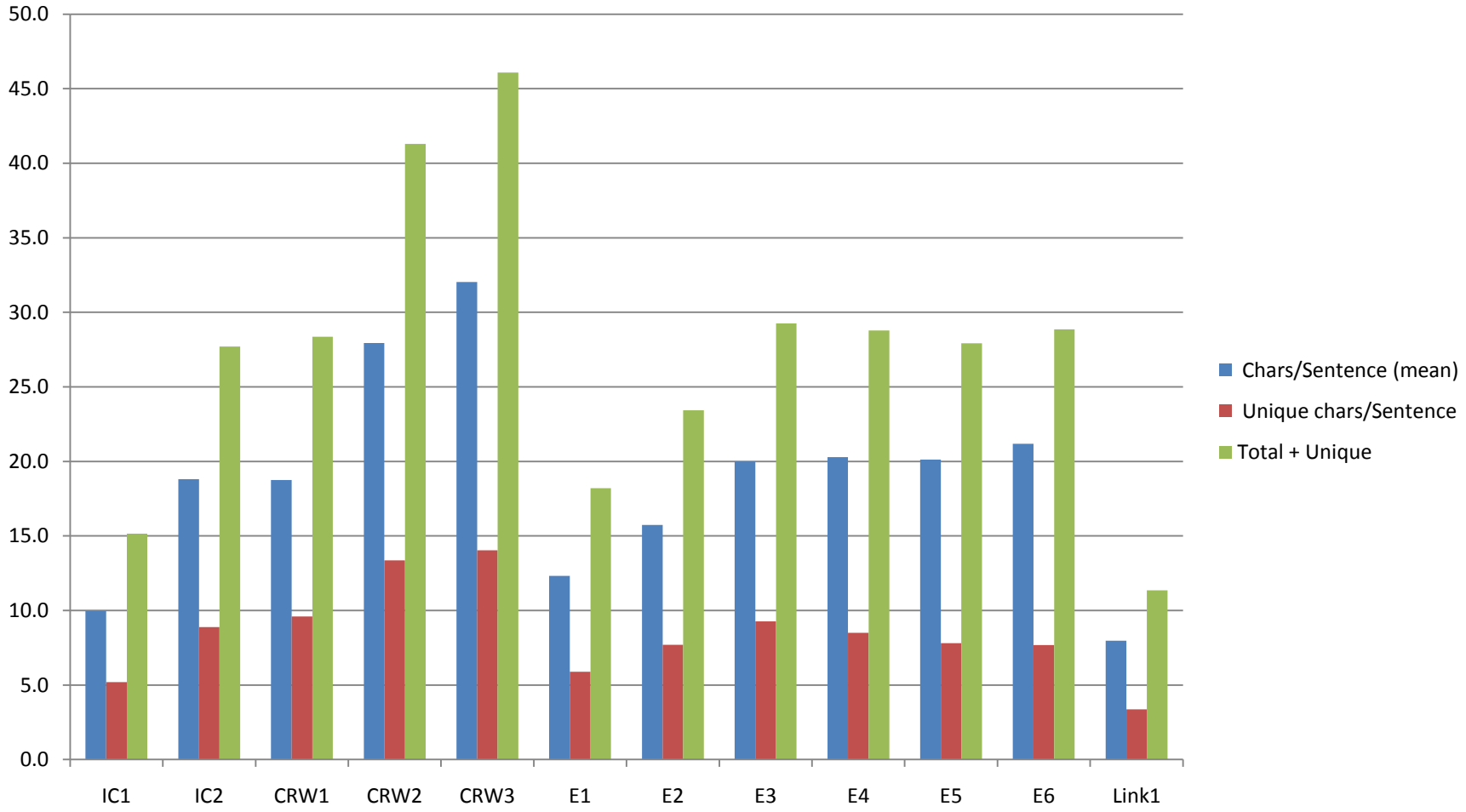


3.3 Results (II)

- Average number of total characters per sentence: t-tests among several pairs confirm previous findings.
  - t-test for IC 1 and 2: $p<0.01$
  - t-test for CRW
    - CRW1 and 2: $p<0.01$
    - CRW2 and 3: Not significant
  - t-tests for IC1 and CRW1: $p<0.01$
  - t-test for IC2 and CRW1: Not significant
  - t-tests for E1 and E2: $p<0.01$
  - t-test for IC1 and E1: $p<0.01$
  - t-test for IC1 and E2: $p<0.01$
  - t-test for E1 and CRW1: $p<0.01$

3.3 Results (III)

Figure 2: Characters per sentence + Unique characters per sentence



3.3 Results (IV)

- Average total characters+ unique characters per sentence t-tests:
  - t-test for IC 1 and 2: $p<0.01$
  - t-test for CRW
    - CRW1 and 2: $p<0.01$
    - CRW2 and 3: Not significant
  - t-tests for IC1 and CRW1: $p<0.01$
  - t-test for IC2 and CRW1: Not significant
  - t-tests for E1 and E2: $p<0.01$
  - t-test for IC1 and E1: $p<0.05$
  - t-test for IC1 and E2: $p<0.01$
  - t-test for E1 and CRW1: $p<0.01$

3.4 Discussions

- Beginning level CFL materials use shorter (and simpler?) sentences than those for native speakers at Grade 1. It is probably due to spoken vs. written language;
- There are more variations in sentence length among different levels of CFL materials than those for native speakers within the same textbook;
- For beginning level CFL learners, the targeted measurement can be
  - 10 characters/sentence, or
  - 15 total characters + unique characters/sentence;
- For intermediate level CFL learners, the targeted measurement can be
  - between 15 and 20 characters/sentence, or
  - between 25 and 30 token characters + unique characters/sentence;
- For advanced level CFL learners, the targeted measurement can be
  - above 20 characters/sentence, or
  - above 30 token characters + unique characters/sentence;
- CRW1 are good for intermediate level CFL learners.

4. Concluding remarks

- This is only a preliminary study. Due to time constraint, no character frequency information is included in the measurement;
- Statistics can be computed unambiguously;
- CFL samples are from current textbooks used in the U.S., but sample size is small and limited in coverage;
- We still need graded reading materials for beginning level CFL learners, whose difficulty can be measured using the characteristic statistics we found in this study.

References

- Flesch, R. A. 1948. A new readability yardstick. Journal of Applied Psychology. 32,3, 221–233.
- Fry, E. 1989. Reading formulas: Maligned but valid. Journal of Reading, 32, 292-297.
- Huckin, Thomas and James Coady. 1999. Incidental vocabulary acquisition in a second language: A review. Studies in Second Language Acquisition. 21, 181-193. http://journals.cambridge.org/article_S0272263199002028
- Jeng, Cheng-Chang. 2001. Chinese readability analysis using artificial neural networks. Ph.D. Dissertation. Northern Illinois University.
- Kincaid, J. P.; Fishburne, R. P., Jr.; Rogers, R. L.; and Chissom, B. S. 1975. Derivation of new readability formulas (Automated readability index, Fog Count and Flesch Reading Ease Formula) for navy enlisted personnel, Research Branch Report 8-75, Millington, TN: Naval Technical Training, U. S. Naval Air Station, Memphis, TN
- Krashen, S. 1985. The input hypothesis: Issues and implications. London: Longman
- McLaughlin, G. 1969. SMOG grading: A new readability formula. Journal of Reading. 12,8,639-646.
- Shen, Helen H. 2005. Linguistic complexity and beginning-level L2 Chinese reading. Journal of the Chinese Language Teachers Association. 40,1-28.

References (Contiuned)

- Stevens, Kevin T. , Kathleen C. Stevens and William P. Stevens. 1992. Measuring the Readability of Business Writing: The Cloze Procedure Versus Readability Formulas. Journal of Business Communication. 29,4, 367-382.
- The University of Iowa Chinese Program. Chinese Reading World. 2001. <http://www.chinesereadingworld.org>
- Wu, Sue-mei, Yueming Yu, Yanhui Zhang and Weizhong Tian. 2006. Chinese link. New Jersey: Pearson.
- Yang, Shou-jung. 1971. A readability formula for Chinese language. Ph.D. dissertation. The University of Wisconsin - Madison.
- Yao, Tao-chung, Yuehua Liu, Liangyan Ge, Yea-fen Chen, Nyan-ping Bi, Xiaojun Wang and Haohua Shi. 2005. Integrated Chinese (2nd Edition). Boston: Cheng and Tsui
- Yu, Chung-hsien. 1960. The readability of Taiwan newspapers. Pao Hsueh. 2,6,18-40.
- 张宁志. 2000. 汉语教材语料难度的定量分析. 世界汉语教学. 53,3,83-88.
- 课程教材研究所小学语文课程教材研究开发中心. 2001. 语文 (一 至六年级). 北京：人民教育出版社.