

The principal goal of the Mandarin Institute-Loyola Marymount University (LMU) STARTALK *Building the K-5 Chinese Readability Scale for Assessing Early Literacy* is to build the foundation for early Chinese reading instruction by creating lexiled vocabulary levels and assessments.

To design literacy instruction that addresses their students' reading levels, teachers must first be able to identify their students' language and literacy levels. In Chinese language immersion settings, literacy levels are aligned to grade-level content standards. This process is very difficult and time consuming since current assessments do not identify tools, measures or scales that can assist Chinese language teachers in selecting literature/informational texts appropriate for their students. Additionally, there are no standardized criteria to determine the readability of Chinese books/materials by reading level; as a consequence, teachers and schools use their own criteria and their own assessments that vary from among and between schools and programs.

### **The problem**

1. It is not easy to find materials in Chinese that are linguistically and culturally appropriate while providing engaging content.
2. There are no widely adopted reading texts or assessments in Chinese language classrooms. There are no defined reading levels for materials or students in alignment with content standards.
3. Without reading levels, it is difficult to apply a systematic approach for a balanced literacy program in Chinese

### **The challenge**

It is crucial for students to establish a strong reading foundation in Chinese and English to achieve success. Teachers (and parents) want to give their students accurate levels of books that advance and reinforce learning. However, curriculum-aligned books that are targeted to students' developmental reading levels and interests across different subjects are not easy to identify. Students want to be able to choose books that are of interest to them so that they will not be bored because they are too simple or frustrated because the books are too difficult.

### **The solution**

If we can provide a comprehensive literacy program, know how well a student can read, and how difficult a specific book is to comprehend, we can most likely predict how well that student will likely understand the book. This process will allow teachers to conduct assessments and use the results to guide their teaching in assisting their students. Thus, the CRS will serve as the following:

- ❑ *Reader Measure*: Industry standard assessment on student's reading ability in different subjects
- ❑ *Text Measure*: Industry standard measurement on difficulty of book
- ❑ *Balanced Literacy* program for reading in Chinese across curriculum

There are several phases and much research which needs to be completed before we can develop a fully functional Chinese Readability Scale that will meet the goals outlined above. The first critical step is the creation of a word frequency dictionary.

## K-5 WORD FREQUENCY DICTIONARY for L2 Chinese Language Learners

In 2015, STARTALK awarded an Infrastructure grant to the Mandarin Institute and Loyola Marymount University to *Build the K-5 Chinese Readability Scale*. Before this work can be done, the team needed to develop a baseline K-5 Word Frequency Dictionary for L2 Chinese language learners that will serve as the basis to build the K-5 Chinese Readability Scale (CRS).

### The Team

With the guidance of an interdisciplinary expert team consisting of Chinese linguists, language, literacy and assessment experts, Master Teachers from partial, dual and full immersion programs and 5 different states representing public, independent and charter schools spanning socio-economic and diverse student populations, and half a dozen research assistants, the team developed a high frequency word list for K-5 Chinese learners and defined word frequency scope for each grade level which can be used as a reference to determine text difficulty of a particular reading material.

### The Process

1. We focused our scope on K-5 Chinese immersion programs and on tackling the issue of analyzing text complexity. The most important factor in analyzing text complexity is determining the frequency of words. There is currently no Dictionary of High Frequency Words for Chinese L2 learners (learning Chinese as a second language). Texts that contain a large number of high frequency words will inherently be easier to comprehend than text which contains low frequency words. In order for this listing to be useful to classroom instructors, we needed to align the high frequency words with vocabulary requirements, Common Core and Content Standards for each grade level. These in combination, along with some other factors will serve as the foundation to creating a formula for measuring text complexity.
2. The Leadership Team determined that in order to create the K-5 Word Frequency Dictionary (WFD), we first needed to create and analyze a corpus of at least 2,500,000 Chinese characters in order to be statistically sound. In creating the corpus, we targeted a minimum of 100 books, narrative and informational, per grade levels by common core standards across multiple categories:
  - Published children's literature books
  - Online academic materials
  - Textbooks
  - Readers

Our Master Teachers categorized books by subjects aligned with Common Core State Standards and Subjects by grade and classified children's books by genre and grade levels. Teachers then selected and created sample sentences from books aligned with key vocabulary. \* We had to use caution so as not to skew the representative sample of materials targeted towards native speakers by comprehending emergent second language readers' literacy vocabulary especially at the kindergarten level. However, this selection of materials is representative of what is currently being used in K-5 Chinese immersion programs in the U.S. Based on these materials, we created a corpus of 2,595,956 characters.

3. We used a statistical software program that could be adapted to assist with frequency analysis of our corpus of nearly 2.6 million characters. The analysis was then manually calibrated and broken out into bands/levels. The CRS bands are equivalent to the level of text complexity that will be used to select appropriate texts for Chinese K-5 immersion classrooms, or high frequency dictionary bands aligned with K-5 grade levels.

The next phase involves validity and reliability testing, and the development of assessments using the K-5 Word Frequency Dictionary for L2 Chinese language learners. This work will begin in March 2016.

We consider the development of this word frequency list as a critical first step to establishing the Chinese Readability Scale. K-5 Chinese Word Frequency Dictionary will serve as a baseline for K-5 immersion curriculum in the U.S. and abroad.

When using the list to determine the text difficulty for a particular grade level, we recommend that 70% of the words in a text should fall into the range of the Word Frequency Dictionary band for a particular level.