Assessment Tools for STARTALK Mandarin Language and Culture Program:

I. STAMP (Standard-Based Measurement of Proficiency)

http://casls.uoregon.edu/stamp.php
http://www.state.nj.us/education/aps/cccs/wl/g8assess/stamp.htm

The STAMP is a criterion-referenced test based on a set of standards consistent with ACTFL Proficiency Guidelines. Using computer-adaptive testing technology, STAMP identifies students’ proficiency from the Novice-Low through the Intermediate-Mid range. The key characteristics of STAMP are as follows:

**Content:** There are three parts in STAMP test: Reading, Speaking and Writing. When students log on to the test site, they first complete a short biographical questionnaire. They then proceed to the reading section instructions and sample items. Students may answer up to two sample items, and feedback is given for their responses to these samples. After this, they complete the reading assessment. Because STAMP is a computer-adaptive test, the delivery of STAMP test items is different for each student. Initially, all students get mid-level questions. As a student correctly or incorrectly answers the questions, the computer delivers higher- or lower-level test items, eventually establishing the appropriate benchmark level for the student. Throughout the test, the selection of test items is randomized, and the order of the four answer choices in the reading section is also randomized. In this way, students take individual, customized tests at their specific levels of competency.

After the reading benchmark level has been established, the students move on to the speaking section. Again, they begin with instructions and two sample items. The speaking section prompts are dependent on the benchmark level the student attained in the reading section. A speaking item at the student’s level of proficiency is randomly selected and presented to the student. As they speak, students may view the rubrics that will be used to grade their response so that they are aware of the criteria they must meet. This phase of the test does require that students have access to headset microphones to record their responses, which are streamed to an LLS Flashcom server. The STAMP test also utilizes voice recognition software that will be used to evaluate elicited response items that students will be asked to repeat and record into the system, thus allowing the system to machine score a certain amount of student speech samples and offer one more level of assurance and accuracy as connected to proficiency levels.

STAMP was developed to measure student proficiency or how well a student can actually use the language in communicative situations. Students interact with authentic and semi-authentic texts in a real world context. An understanding of second language and culture is implicit in the task and text. For instance, a reading item may present students with an actual bus schedule and ask them to select the appropriate route and departure time required to arrive at a particular location in time for a concert. In a speaking task, students may record a phone message to their friends about their plan for the day of the concert. All reading items provide students with a context, a piece of
authentic or semi-authentic text, and then a task. Students have a choice of four
responses. Speaking items are more open ended but share the quality of providing a
culturally authentic context. One of the advantages of Web-based testing is that tasks
such as obtaining information from a Web page are highly authentic and well
contextualized. The STAMP is currently being used by 6 states, 30 school districts and
15 universities, and the same technology is being used for universal ESL testing in
Minnesota and Oregon.

**Grading and Reports:** In order to facilitate curricular and instructional improvements,
teachers receive detailed feedback on their students’ performance. Reports based on
student and class performance are posted online for teachers to access. These reports
give teachers concrete feedback on students’ overall proficiency as well as breakdowns
of student performance according to topic and task type.

Reading scores are generated automatically by the computer and results are available
immediately after the test has been completed. The class or group report shows the
demonstrated proficiency of all members in the form of a 1-5 score. In addition, class or
group results may be viewed by topic. The speaking report, also posted online, shows
the speaking prompt and rubrics given. The student’s spoken response to the prompt is
accessible via a "play" button that allows the teacher to hear the student’s responses to
each prompt and view the scores assigned with the attached rubric. The teacher can also
print out individual reports that can be placed in a file or given to the student. These
reports detail the reading and speaking levels attained and include narrative
descriptions of what students are able to do at each of the levels.

**Facilities Required:** Computer and Internet connection, Headset Microphone

**Time Required:** The amount of time needed to complete STAMP test varies among
individuals. From our observation, the higher the student’s language proficiency is, the
longer STAMP test takes him/her to complete. Averagely, STAMP test takes about
50-80 minutes for our level II and level III students to complete.

**STAMP Features**

- Chinese, French, German, Italian, Japanese, and Spanish
- Available in reading, writing, and speaking
- Measures proficiency ranging from Novice-low to Intermediate-mid
- Age appropriate for grades 7 through 16
- Web-based
- Related to ACTFL Proficiency Guidelines
- Textbook independent
- Available through CASLS' partner company, Avant Assessment

**Benefits for Teachers**
• Provides validated, easy-to-use data
• Good for checking progress, placement, program and standards review, and instructional planning
• Access to longitudinal individual speaking and writing samples

Benefits for Students

• Engage in real-world language situations
• Emphasis on what students can do with the language
• Facilitates goal setting

Weakness

• Only categorical grades (1-6), not numerical grades, are given to each section of the test. Therefore, STAMP test is hard to measure students’ short-term progress
• Questions in the test might deviate from what students have learned in the textbook

II. SOPI (Simulated Oral Proficiency Interviews)
http://www.cal.org/topics/ta/sopi.html

The Simulated Oral Proficiency Interview (SOPI) is a type of tape-mediated test of speaking proficiency. All SOPI items are based on the speaking proficiency guidelines of the American Council on the Teaching of Foreign Languages (ACTFL). The test is presented to examinees via a test booklet and a master tape. It can be administered individually by anyone using two tape recorders. It can also be used in a language laboratory setting to test groups.

Procedure: During testing, the examinee listens to directions for speaking tasks from a master tape while following along in a test booklet. As the examinee responds to each task, his or her speaking performance is recorded on a separate response tape. Each examinee's response tape is later evaluated by a trained rater who scores the performance according to the proficiency guidelines developed by the American Council on the Teaching of Foreign Languages (ACTFL). Each examinee receives a proficiency rating on the ACTFL proficiency scale.

Content: The test begins with a warm-up that helps examinees relax and become accustomed to speaking in response to the master tape. The performance-based tasks that make up the remainder of the test are:

Picture-Based Tasks
To respond to these tasks, examinees use illustrations printed in the test booklet. Picture-based speaking tasks include giving directions, describing activities in a familiar setting, and telling a story.

**Topic-Based Tasks**

These tasks require examinees to explain or discuss a variety of topics. Topic-based speaking tasks include describing a procedure, presenting advantages and disadvantages, explaining and defending a point of view, or describing what would happen if a hypothetical situation were to come true.

**Situation-Based Tasks**

In these tasks, the examinee is placed in real-life situations and is asked to respond to them in a socially and linguistically appropriate manner. Situation-based tasks include giving advice to a friend, apologizing for having offended someone, and making a formal presentation to a group.

**SOPI & OPI:** The *SOPI* is not intended to replace the oral proficiency interview. It does, however, offer a standardized approach to oral proficiency testing in situations where it is not feasible to give an oral interview due to the unavailability of trained testing personnel or lack of time and budget for such testing.

**Time and Facilities:** The *SOPI* may be administered in a language laboratory setting and is thus ideally suited to large group testing. Alternatively, the test can be administered individually using two tape recorders. The administrator does not need to know the test language. There are several parallel forms of the *SOPI* in some languages. The test takes about 50 minutes to administer. Lower level proficiency students may stop after the first half of the test. In this case, the test administration time is about 25 minutes.

**SOPI Features**

- Available for Arabic, Chinese, French, German, Japanese, Hebrew, Hausa, Indonesian, Portuguese, Russian and Spanish
- Measuring speaking only
- Levels included in the guidelines are Novice (Low, Mid, High), Intermediate (Low, Mid, High), Advanced, Advanced High, and Superior. Novice Low means essentially no functional communicative ability whereas Superior means native-like skills.
- Age appropriate for grades 7 through 16
- Tape-mediated test
- Related to ACTFL Proficiency Guidelines
- Textbook independent
- Available through CAL (Center for Applied Linguistics)
Note: Because SOPI is only rated according to categorical grades (like STAMP), it is very hard to measure short-term progress. Therefore, we use the SOPI test materials but grade according to FLOSEM (Foreign Language Oral Skills Evaluation Matrix): http://www.asdk12.org/depts/sped/policy/part2/FLOSEM.pdf. FLOSEM does allow graders to distinguish grades of students at similar oral language proficiency.

III. Lingua-Folio (Self-assessment):

LinguaFolio is a portfolio assessment instrument designed to support individuals in setting and achieving their goals for learning languages. It includes these three components:

- Passport, where formal assessments and a student's self assessments (checklists) are documented,
- Biography, where information about a student's language background and intercultural activities are recorded, and
- Dossier, where samples of a student's work over time are archived.

This three-fold approach enables language learners at all ages and levels to document their language learning as they move along the continuum towards greater proficiency.

IV. PAUSD Pre- and Posttests

PAUSD pre- and posttests including two forms: (1) Pre- and posttest on listening, reading and writing; (2) pre- and posttests on cultural knowledge. They are both paper-based tests and are closely related to textbook and what students have learned during a certain period.