

**Yeh, Hsin-Hung**  
Stanford University

## **L2 Spoken Chinese Corpus and the implications in CSL pedagogy and assessment**

Chinese as a Second Language (CSL) has been established as an academic discipline during the past decade, whereas the teaching methods at higher education level have not co-evolved concomitantly with the current context. Most efforts in current teaching practices have long been made to increase the variety of presenting and delivering language inputs to CSL learners, yet the effectiveness has yet to be carefully examined. As a result, a huge disconnect is taking place between teaching in the classroom and assessment outside the classroom, e.g. OPI and WPT developed by ACTFL. To abridge the widening gap, the present working project aims to construct a L2 Spoken Chinese Corpus with spoken data from CSL learners. To date, a well-constructed corpus of spoken data has yet to be found, and most corpuses were dedicated to written data, e.g. HSK (Hànyǔ Shuǐpíng Kǎoshì, *Mandarin Chinese Proficiency Test*) Dynamic Composition Corpus by Beijing Language and Culture University (BLCU). To make the data better reflect the spoken performance of CSL learners, e.g. what language tasks learners can do, the L2 Spoken Chinese Corpus collected data through OPI interviews with CSL learners conducted by certified testers. The benefits of adopting OPI to collect data are manifold. First, interviewees are not faced with random questions, but with coherent, adaptive questions based on learners' performance, while the interviewer, i.e. the certified rater, endeavors to elicit the best language sample from the learner according to the learner's performance in each language task. Second, each interview follows standard operating procedures and is about 20-30 minutes long, contingent on the learner's performance and level. Such procedures ensure that each interviewee has various chances to fully demonstrate their L2 Chinese skills to complete target language tasks. Third, as the number of certified OPI raters has been on the rise in recent years, the data collection process can potentially provide people who aspire to be certified raters with a platform for resources. We have collected data from learners at Middlebury Summer School and other institutes, and our results have shown that current CSL teaching practices, particularly at Advanced level do not fully prepare learners for OPI tasks at Advanced level, i.e. learners do not have enough chances to practice and improve their narration and description skills. What current CSL teaching practices deal with at Advanced level, however, is focused on OPI tasks at Superior level. Topics discussed in CSL classes at Advanced level are especially slated for political and environmental issues. As a result, the training does not help learners to lay a solid foundation at Advanced level, and even though they are trained to talk about certain topics at Superior level, their performance is much limited due to their weak foundation at Advanced level. It is proposed that current CSL class at Advanced level should seriously think about training learners to develop their narration and description skills, which are two fundamental skills which help learners to organize ideas and enrich vocabulary prior to more complicated and abstract tasks at Superior level.

## **References**

Adolphs, S., & Carter, R. (2013). *Spoken Corpus Linguistics: From Monomodal to Multimodal*. New York: Routledge.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

Renouf, A., & Kehoe, A. (2009). *Corpus linguistics: Refinements and reassessments*. Amsterdam: Rodopi.