In Conversation

Asian Association for Language Assessment
Student Committee
Issue 2 (Fall/Winter 2016)

Contents

Editors' Note

In continuation with the aims of The Asian Association for Language Assessment (AALA) Student Committee to engage our student members and graduate students of the field, this issue of the *In Conversation* series presents the interviews with our two plenary speakers and workshop presenters at the 3rd AALA Conference in Bali, Indonesia. The interviews provide a snapshot of their sharings at the conference, as well as their invaluable advice to students. The two academics interviewed are Dr Luke Harding and Professor Nathan Carr.

The interview with Dr Luke Harding, from the University of Lancaster, is a follow-up to the workshop on *Assessing Listening* that he had conducted, as well as his plenary speech on *Diagnostic language assessment: From theory to practice*. In a similar vein, the interview with Professor Nathan Carr, from the California State University, Fullerton, focuses on his experiences in research on automated scoring and his thoughts about the current status as well as future trends of such research. In this interview, Professor Carr also shared his experiences in academic publication and dispensed precious advice for professional development.

In addition, we are honored and grateful to have the Executive Board member, Assistant Professor Jirada Wudthayagorn, pens her message to our student members. We conclude this issue with a highlight on the student award winners at the 3rd AALA Conference which we hope would enthuse fellow students in their academic journey.

Warm Regards,

Carol Dabarera & Sha Liu (Editors)

From left to right at the 3rd AALA Conference, Bali Indonesia: Sha Liu (Editor), Professor Antony Kunnan (AALA Founding President), Dr Luke Harding (Workshop and plenary presenter), Carol Dabarera (Editor), Professor Nathan Carr (Workshop and plenary presenter)

Insights on Assessing Listening & Diagnostic Language Assessment: An Interview with Dr Luke Harding

Interviewed by Carol Dabarera

Much of your research revolves around assessing listening and you have conducted a workshop in this area at the 3rd AALA Conference held in Bali in May 2016. In addition, you have also published papers and gave presentations on communicative language testing and validation. How do you see these themes coming together to respond to the current needs in language testing?

This is an interesting question. Generally, my thinking at the moment is that there is a need in language testing/assessment – particularly communicative language testing – to broaden our constructs by developing a better understanding of the communicative demands of real-world language use. One of the key themes of my research is that communicative language assessment might benefit from a focus on adaptability: assessing what the learner/test-taker is able to do in novel situations. This connects with my research on listening assessment; the use of a wider variety of accents in listening assessment would go some way to tapping into these important competences that language learners need to develop. For example, at the moment I'm working on a study investigating listeners' sources of difficulty and strategy use when completing listening tasks featuring a totally unfamiliar accent. I hypothesise that the accent will pose some challenge to the listeners, but that there will be evidence of strategy use in dealing with the novel accent. This would help to demonstrate that the nature of the listening input might lead to the assessment of a very different type of listening construct.

With diagnostic assessment, the view of broadening the construct will still be relevant but we are also looking at breaking the constructs down to specific elements that would enable us to do this type of fine-grained assessment. One of the things I am interested in, in communicative testing, is the role of variability in speech. When diagnostic testing is brought in, a diagnostic task could be determining how good someone is in adjusting to a new accent. So you can still bring those new understandings of the construct into the diagnostic tests but it would be at a more fine-grained level. So I think there is definitely crossover between the two things – broadening and breaking down of the construct.

> "My thinking at the moment is that there is a need in language testing/assessment – particularly communicative language testing – to broaden our constructs by developing a better understanding of the communicative demands of real-world language use."

I have noticed your recent focus on diagnostic assessment. How did you get interested in and involved with this? And how have you been influenced by the earlier efforts in the language testing community, broader education assessment context, and psychometric community?

Before I moved to Lancaster University I had some experience developing diagnostic listening and reading tests at the Language Testing Research Centre in Melbourne. But it was not until I moved to Lancaster University (in 2010) and started talking about these issues with my colleagues Charles Alderson and Tineke Brunfaut that I became more interested in diagnostic assessment from a research/theory perspective. It was at Lancaster that I learned more about the development of DIALANG (which is still hosted by Lancaster University) and the history of that particular project. I started to read more in this area on topics such as Cognitive Diagnostic Assessment and dynamic assessment, but my influence has primarily been to continue the tradition of DIALANG's mission: to develop diagnostic instruments specifically for diagnostic purposes.

How do you see the field of language testing progressing in this area of diagnostic assessment in the near future? Which are the aspects that are in urgent need for more research?

In the future, as I've covered in my presentation at the AALA conference, I think there are two important areas for progress/research. One is the development of more purpose-built diagnostic tools for teachers (and others) to use in the diagnostic assessment process. The other is to develop a deeper understanding of diagnostic competence among teachers, and to understand how this might best be promoted.

You have together with some colleagues proposed a theory of diagnosis in second and foreign language (SFL) assessment in 2014, which you later contextualized in the domains of reading and listening in 2015. Could you share briefly the 5 principles for diagnosis in this framework, and the associated 4-step diagnostic process?

The five principles of diagnostic language assessment are (briefly):

1. It is not the test which diagnoses, it is the user of the test.

2. Instruments themselves should be designed to be user-friendly, targeted, discrete and efficient in order to assist the teacher in making a diagnosis. They should provide rich and detailed feedback. Most importantly, useful testing instruments need to be designed with a specific diagnostic purpose in mind.

3. The diagnostic assessment process should include diverse stakeholder views, including learners' self-assessments.

4. Diagnostic assessment should ideally be embedded within a system that allows for all four diagnostic stages: (1) listening/observation, (2) initial assessment, (3) use of tools, tests, expert help, and (4) decision-making

5. Diagnostic assessment should relate, if at all possible, to some future treatment.

The four-step process, as explained in principle #4, is:

1. The listening/observation stage
2. The initial assessment stage
3. The hypothesis checking stage (which may often involve the use of tools, tests, etc.)
4. The decision-making stage (where feedback is provided, and treatment discussed)

How would you evaluate the perspectives of schools on diagnostic testing in the classroom and their current state of readiness to implement it?

I recognise that the diagnostic procedure – as described – is an "ideal", and could be difficult to implement in some educational contexts. But a key to making this approach more attractive and manageable is the development of a repository of tasks/tools that teachers can use at Stage 3, and also the increasing development of teachers' diagnostic competence (assuming that teachers will most commonly take on the role of "diagnosticians" in the classroom).

And how would you envision the applicability of this theory in contexts other than the classroom?

Beyond the classroom, I think the principles and process could be adapted to fit, for example, computer-based diagnostic assessment systems. The important aspects of the theory: that diagnosticians and learners are centrally involved in problem identification, that tools and tasks are purpose-built for diagnostic purposes, and that there is detailed and targeted feedback, could certainly be applied beyond the classroom.

> "The important aspects of the theory (of diagnosis in second and foreign language (SFL) assessment) that diagnosticians and learners are centrally involved in problem identification, that tools and tasks are purpose-built for diagnostic purposes, and that there is detailed and targeted feedback, could certainly be applied beyond the classroom."

You have mentioned in Harding et al. (2015) that insights into the underlying constructs is crucial to the diagnostic process, as well as shared challenges such as the extent of specificity (grain size) in diagnostic testing and the number of test items required for a reliable diagnosis. How would you suggest for these needs to be prioritized and navigated?

With respect to diagnostic constructs, I think that deep knowledge of construct and diagnostic assessment must go hand-in-hand. My own view of the listening construct has been influenced a lot by the work of John Field who has proposed a processing model for listening comprehension. It seems to me that the processing model fits very neatly into a diagnostic assessment framework as it helps to understand where listening has broken down in the comprehension process. I think the process-based view is certainly applicable to the receptive skills – listening and reading. These skills cannot be observed and hence the process-based view of reading and listening can be useful because it can show us where the comprehension process breaks down.

But I think the challenge, as I've mentioned above, is to ensure that those educators who are doing diagnostic work are working with a deep and clear understanding of the construct. Regarding grain size, number of items to form reliable judgements, etc. – these are ongoing questions for diagnostic assessment. To some extent, I think grain size could be determined by what sort of treatment is available. In terms of reliability, I would say that the process we have developed means this is less of a problem as tasks will be discrete, targeting a particular element of the ability, selected based on Stages 1 and 2.

Following up with your recommendation for a deep and clear understanding of the construct, how do you feel this could be achieved in consideration of the very limited literature on second language (L2) and foreign language (FL) learning as noted by Alderson (2015)? It is understood that much L2 and FL research taps largely on the literature on first language (L1) learning.

We are to some extent, constrained. However, more people are doing second language research now and a lot of that research revolves around the acquisition of grammatical structures, etc. Searching for information on the skills is a bit harder though. I would say that there is still much more information out there on reading than on listening. The situation has improved over the last ten years or so, but we still need to know a lot more, particularly about how listening develops.

I personally think that we need more longitudinal studies. One of the things I was talking about this morning is to look at the longitudinal development of discrete measures. If a recently submitted funding application is successful, I would be working on a longitudinal study of speaking development among international university students with Judit Kormos at Lancaster. It is one thing to know about the construct, and another to know about what you might expect to develop. Central to diagnostic assessment is the knowledge of the problem which has to be situated in the given context (e.g. a developmental path). In the case of the development trajectory, knowing more about the skills and how they develop will be really useful for understanding the assessment of language learning.

What is your research agenda for the next 5 years?

Good question! There are lots of things I'd like to do. With respect to language testing and assessment, I would like to continue looking at the assessment of English as a Lingua Franca competences, particularly in purpose-built tasks which require communicative problem solving. I would also like to look at how listeners (first and second-language) might be trained to listen to speakers with different accents, and what variables might influence their success in adapting. I have an ongoing interest in the assessment of oral language skills generally, particularly pronunciation/intelligibility assessment, so there is likely to be some further research on that front. I also have some other non-assessment-related projects that I would like to carry-out.

"I would have three pieces of advice for students who wish to work in these areas. One is to engage with fields beyond language testing and assessment. The second piece of advice I would have is to consider diverse methodologies, both quantitative and qualitative. The final piece of advice is to maintain a critical position"

What advice would you have for students who are interested in pursuing research in the areas that you have explored? (E.g. Listening assessment, diagnostic assessment, communicative language testing, validation, etc.)?

I think there is still a lot of work to be done in all of these areas, and it would be great to see more students addressing these issues. I would have three pieces of advice for students who wish to work in these areas. One is to engage with fields beyond language testing and assessment. For example, in researching listening there is much to gain from psychology, phonetics, sociolinguistics, etc. Similarly, for diagnostic assessment, we learned a lot by talking with diagnosticians in a number of different professional fields. New insights and innovations rely on researchers stepping outside of their field and looking at issues in new ways. The second piece of advice I would have is to consider diverse methodologies, both quantitative and qualitative. The final piece of advice is to maintain a critical position. Increasingly, I'm coming to realise that the key question in assessment is not "how can we make this test better/more reliable/more effective/etc.?" but rather "why does this test exist at all? What purpose does it serve? And whose interests are promoted by its existence?" One of the reasons I find diagnostic language assessment a particularly valuable area of study is that I see it as assessment designed purely in the interests of the learner: to promote learning, and to provide the learner with valuable, useful and accurate feedback. And that makes it a worthwhile topic to explore.

Improving Authenticity and Validity in Teaching and Assessment in the Age of Automated Scoring: An Interview with Professor Nathan Carr

Interviewed by Sha Liu

Professional Background:

You majored in Russian Studies and Russian as an undergraduate student and focused on TESOL when pursuing your Master's degree in University of Oklahoma. For a start, could you share with us how you entered the field of language assessment?

I got my master's degree from Oklahoma City University (OCU). The University of Oklahoma did not offer a master's in TESOL. I had wanted to be in Oklahoma, since I'd just spent three years teaching in Taiwan. I thought about applying to Oklahoma State University, which was the other school in the state with a TESOL program, but decided Oklahoma City would present more options for finding a job teaching ESL while I was in school. OCU also did not require the GRE, and I had problems finding out how to register for the test while in Taiwan. As it turned out, I think I made the right choice.

My first involvement with language assessment came when I was asked to create a placement test for my English program in Taiwan. I did the best I could, but I didn't really know what I was trying to do. A couple of years later, I started helping out with administering and scoring the placement test at the ELS center in Oklahoma City during my master's program. But I had no plans at that time to make that my career focus.

What made you decide to switch to language assessment?

I was applying to three doctoral programs at three schools, including UCLA. My master's advisor, Dilin Liu (now at the University of Alabama), told me that since UCLA had a language testing track, I should do that if I was admitted there, not SLA. He said that since I was not afraid of math (and I'd done pretty well on the GRE math section), it would be a good choice, and I could always use that training to produce better instrumentation for SLA research later on if I wanted to. It was great advice, and I've never regretted moving into assessment. Once I started at UCLA, and started taking my first class with Lyle Bachman, I very quickly decided that I'd found my proper place in the larger field of applied linguistics.

In your opinion, what is the defining moment in your career?

That's a hard question. Was it finding out about the job opportunity in Taiwan, which got me into TESOL? Was it deciding to do a Ph.D., something I'd never considered before being encouraged to during my master's? Deciding to focus on assessment? It's hard to say.

About the research on Automated Essay Scoring:

Much of your recent research resolves around Automated Essay Scoring (AES). Could you summarize for us how research in this area has matured over the years?

Well, I've written about AES, and I'm certainly interested in it, but it has not really a research focus of mine thus far. I'd say things have come pretty far, but they still have some distance to go, especially in the area of assessing second language writing and diagnosing learner errors. I believe that they do a fairly good job (depending on the scoring engine) of predicting human ratings, but I think they are generally looking at something different from what humans would look at. And the limited research available (including your study!) tends to show that they are not good at accurately diagnosing language errors. My research, though, focuses on automated scoring for limited production tasks, rather than on AES.

The use of AES, especially in high-stakes assessment settings, has met with strong criticisms among composition researchers and teachers (e.g., CCCC position statements). What do you think are the main reasons for such criticism? What are your views on the operationalization of AES as a second rater in large scale high-stakes tests such as TOEFL iBT and GRE?

I think some of the objections are philosophical as in the case of the CCCC. Some people really do feel that writing should always be done for an audience that is actually reading, understanding, and responding to it. I'm not quite in that camp, myself—frankly, I'm not terribly interested in what 500 or 5,000 different strangers have to say about a writing prompt on a test. I don't intend any offense against these hypothetical test takers, and I'm sure most of them are not terribly interested in the topic themselves, either. So for testing and pedagogy, especially testing, I don't find the philosophical argument that compelling.

I think there is also an element of fear of being put out of work by AES systems. As I said earlier and will explain further in a moment, I don't believe they can replace human rating, at least not yet. In a few years, though, who knows. And certainly we can all imagine the temptation to cut costs by using an AES system that is not as accurate as human rating, but is much cheaper to operate and produces results. Those results might suffer from validity and accuracy issues, but any reasonably cynical person can imagine that not stopping a lot of decisionmakers.

"I think there is also an element of fear of being put out of work by AES systems. As I said earlier and will explain further in a moment, I don't believe they can replace human rating, at least not yet."

More compelling, though, I believe, is the issue that the main AES engines on the market presently were all created for native speaker writing, rather than non-native language, and they are still not able to diagnose learner errors very accurately. As Hoang and Kunnan (2016) and Liu and Kunnan (2016) have found, there are a lot of both false negatives (where the system misses a language error) and false positives (where it mistakenly counts something as an error) in their output. In my own anecdotal experience in an ESL writing course I recently taught, my campus learning management system has an option to get AWE (automated writing evaluation) feedback on student papers, but after looking at a couple of papers and their diagnostic feedback, I decided not to even bother encouraging students to look at it. This feedback is provided by a big-name commercially available scoring engine, too, I should add.

Another issue is that these systems are not, at present, good at identifying illogical statements or errors of fact. However, this is a less serious problem in my view. To some extent, I think that common problems of content could get caught via the training sample, to the extent that they occur in the training sample and lower scores. If facts and logic are not part of the construct definition, too, then we do not want an AES system to penalize for them, any more than we want human raters to. And frankly, how many test takers are going to deliberately get facts wrong or spout illogical nonsense on a test?

All that being said, though, I believe that in time, whether it is in 1 year or 20, the language issues will be solved, and we will be able to get reasonably accurate diagnostic feedback from AES and AWE systems. In the meantime, however, I think that a good system is probably good enough to use in conjunction with a human rater, as ETS does on the TOEFL iBT Writing tasks, but I would not be comfortable with using one at present to score a high-stakes writing test without human intervention. I think it might be good for the profession to give consideration now to what an acceptable rate of language misdiagnosis would be for AES and AWE in high-stakes situations. One place we might start is by looking at how accurate human judgements are of language errors, and use that as a target.

In the last few decades, AES systems have been gradually integrated into second language writing classrooms. How do you understand such instructional application of AES systems? What are your suggestions for the line of research that evaluates the pedagogical effectiveness of such AES systems in EFL/ESL writing instruction?

I think that AES and AWE are probably most useful for second drafts. They could be useful with first drafts if there has been some feedback on content and organization based on prewriting, such as on an outline. In terms of feedback on content and organization, these systems can probably give some useful estimate of quality, and possibly some useful recommendations on revision. In terms of language feedback, though, they can flag things to look at, but the students will then have to give critical consideration to each point, one at a time. It seems about as much use to have them practice editing for specific language points one at a time, such as a pass through the paper looking for fragments and run-ons, then a second run-through for singular/plural, another run-through for verb tenses, etc. And for late drafts, when the focus is shifting to editing for language, rather than content and organization, I think AES/AWE is of much less use, given the current state of the art.

What do you think of the recent calls for the use of argument-based validation frameworks to conceptualize AES research?

I think that is probably just as appropriate as it is for any other area of language assessment research. Argument-based frameworks are the dominant validation paradigm in our field, and I am not aware of any serious alternatives, really. Cronbach proposed something similar to argument-based validation in 1988—that is the first example I am aware of, although there might have been something else before of which I am unaware—and I am not sure how quickly it caught on, but Kane published his article "An Argument-Based Approach to Validity" only four years later, in 1992, making a strong and comprehensive argument (no pun intended) for the approach. In my view, the only question, then, is whether to follow the full assessment use argument model (Bachman, 2005; Bachman & Palmer, 2010) with claims and warrants and rebuttals etc., or to do something similar but with less emphasis on the formal logic and argument structure. I think the answer to that question really depends on the context—if you are talking about a low-stakes small-scale decision, obviously you do not have or need to deploy the same sort of validation resources that you would for a large-scale high-stakes context.

"In my view, the only question, then, is whether to follow the full assessment use argument model (Bachman, 2005; Bachman & Palmer, 2010) with claims and warrants and rebuttals etc., or to do something similar but with less emphasis on the formal logic and argument structure. I think the answer to that question really depends on the context—if you are talking about a low-stakes small-scale decision, obviously you do not have or need to deploy the same sort of validation resources that you would for a large-scale high-stakes context."

What are you currently researching on?

I am working on one or two projects involving automated scoring, and on a large validation project. On the automated scoring, I had been developing something about the generalizability of scoring keys for short-answer CAS, something I mentioned earlier. However, I presented the initial findings at LTRC in 2011, just before my wife got pregnant with our first child. When I had a chance to catch my breath, we'd already had a second, and I'm not sure the data are recent enough to be interesting to a competitive journal. If I can come up with a few more ideas for chapters, I might make that the topic of a book on CAS for short-answer questions. Alternatively, I might adopt the methodology to a different dataset and move on from there. The other topic I'm pursuing in CAS is the question of how well it can be made to work with short-answer questions that were designed for human scoring. The particular test I'm working with has answers that are usually one or two sentences long, which makes the keyword/regular expression matching method even more challenging. I co-presented a paper on this with Antony Kunnan at the 2016 AAAL conference, and we are hoping to continue developing the study, for now mostly by trying to tweak the scoring key. Even if the answer proves to be that the keyword matching approach does not work very well with such items, or that it only works with an unreasonable amount of effort, that will be a useful finding.

The other project involves the Vietnam Standardised Test of English Proficiency (VSTEP). The main test assesses test takers in reading, writing, speaking, and listening, and certifies them as being at the B1, B2, or C1 level of proficiency. All test takers are assessed in writing and speaking—they are not optional components—and one of the desired consequences of the test is to promote positive washback in English teaching and learning, getting teachers and learners to focus on learning to communicate in English. I was brought in to do a two-week series of training workshops for English faculty and test developers under the auspices of a U.S. State Department English Language Specialist project in August. This was followed by a series of online consultations for planning and such, and as of this writing, I am in Hanoi again for two weeks to advise on validation research. There is an assessment use argument framework in place now, with various elements prioritized in terms of studies to provide backing for the claims. That framework will guide their validation research agenda for the next few years. At the moment, I am advising in particular on procedures for item and dependability analyses, standard setting, and scaling, as well as on a couple of validation studies that are still in their early phases. They have a very promising team working on all this, and I am excited to be able to help out on such an ambitious project. I think it has the potential to really help push English instruction in a positive direction there, and that should in turn lead to some useful benefits for Vietnam internationally, particularly economically.

About scholarly publication:

You started to write for scholarly publications even before you embarked on your PhD study and you have published widely. Could you share with us your publishing experiences? What are the major challenges that you encountered? How did you overcome these challenges?

Well, before starting my doctorate, I had one accidental publication—I had taken an independent study class as an undergraduate, and my professor, Dragan Milivojevic, liked my paper on Tolstoy as a Daoist sage so well that he decided to include it in a collection he was editing. This happened while I was in Taiwan, though, and back in the days before e-mail, it was extremely easy to lose touch with someone. So, I did not know it was even happening until the book was published! I was dumbfounded to learn my paper was in an actual book—I had thought it was pretty decent, but not that good. Other than that, though, I only had some presentations, not actual publications, until my doctoral studies at UCLA.

I think that in terms of challenges to publication, the greatest one I have faced has been having time to work on research. In the California State University system, the focus is very heavily on teaching, and I normally teach four courses per semester. Also, once you have tenure and even more so once you are promoted to full professor, the pressure to publish decreases and the pressure to serve on or chair committees, including at the college or campus level, increases. That is all work that faculty need to contribute to—really, the work of running the university has to be done, and much of it truly requires faculty involvement—but at the same time, I did not go to graduate school so I could become chair of the departmental personnel committee! So, all the competing demands for time, along with my meager attempts to have a personal life, make it hard to squeeze in time for research. But I keep trying, and my advice would be that students do so too. If you can, schedule one specific day, even one specific morning or afternoon, for working on research projects. I also find that having a looming deadline, like an upcoming conference, helps to concentrate the mind—it is amazing how much you can do when you have to.

> "If you can, schedule one specific day, even one specific morning or afternoon, for working on research projects. I also find that having a looming deadline, like an upcoming conference, helps to concentrate the mind—it is amazing how much you can do when you have to."

Beyond all that, though, I think that I should probably mention that it is important to think through the study before you start it. Have clear research questions written, ones that grow out of what has been done before on that topic or on related topics. Make sure you are not doing the 10th version of the same study: That is, make sure you are contributing something new to the field. (That was a criticism of a conference proposal based on my dissertation—one reviewer said he or she did not see why we needed yet another paper on that topic! So, before submitting that same paper for publication, I revised it to explain more clearly how the results mattered to anyone besides me. It worked.) Choose the research approaches that will answer those questions, and use them appropriately. When you report your results, explain how you evaluated any statistical assumptions, and if you are reporting a quantitative study in a testing journal, remember to include complete descriptive statistics for all the variables included in the analysis. Also, try presenting your results at a major conference, and getting feedback. If experts have tough questions for you, or tell you that you have done something wrong in your analyses or interpretation, don't be offended. Either take their advice and try to incorporate it, or find convincing evidence or arguments for why they are mistaken. Experts can make mistakes, after all, but you should not assume they are wrong just because being right would mean having to redo your analyses!

Advice for AALA Student Members:

What advice would you give to our AALA student members with regards to their professional development?

I would say there are five pieces of advice I would like to give:

(1) Learn as much quantitative methodology as you can. If you cannot evaluate quantitative studies, if you cannot perform item and reliability analyses, you are crippling yourself professionally. Quantitative methods are crucial things to have in your toolbox. Don't be a lumberjack who only knows how to use an axe, but not a saw.

(2) Learn some qualitative methods as well. As someone who prefers quantitative approaches myself, I ask you: If you can only do quantitative research, how do you know what is worth looking at, what constructs you should try to quantify? Also, your interpretations of quantitative results will be much richer with some qualitative findings, some rich description, to help contextualize them and clarify their interpretation. Don't be a lumberjack who only uses saws, not axes.

(3) Get some teaching experience before you graduate, if you possibly can. Most people with doctoral degrees will become professors, and you can lose your job if you do not know how to teach well. University students can be very demanding—ask

yourself what you and your classmates would expect from a professor! If you can hit the ground running as far as teaching goes, that's one less thing for you to worry about.

(4) Learn to work well with others, share the work, and compromise. As a faculty member, you are going to be serving on a lot of committees and sitting in a lot of meetings. Learn to get along and compromise, and to push for what you and your program want, without being abrasive and alienating.

(5) Do not give up. Do. Not. Give. Up. That may be the most important thing to getting through graduate school, getting a job, getting published, and getting tenure. Self-doubt is a normal part of graduate school, and even of being a junior faculty member. During my doctoral program, I was repeatedly convinced that I was academically doomed, but I resolved early on that I was not going to give up, I would keep fighting, and if I failed, at least it would not be because I quit. Later, as a junior faculty member I was frequently discouraged and/or alarmed by obstacles, but I stubbornly put my head down and kept pushing. It worked. Determination is not the only thing you need to be professionally successful, but you definitely need it.

"Do not give up. Do. Not. Give. Up. That may be the most important thing to getting through graduate school, getting a job, getting published, and getting tenure."

Special Feature 1

Messages from AALA Executive Board Members

Jirada Wudthayagorn

Assistant Professor,

Head, Language Assessment and Evaluation Track, EIL Program

Chulalongkorn University Language Institute, Thailand

Hello AALA student members! I am Jirada Wudthayagorn. I am very pleased to be part of AALA and honored to share my story in the Special Feature in this Issue.

I have diverse education backgrounds. I earned a bachelor's degree in Education with honors and a master's degree in Educational Research from Chulalongkorn University. Then, I won a Royal Thai Government Scholarship to pursue my Ph.D. in the US. In 1994, I studied theoretical linguistics at the University of Delaware. Here, I got an M.A. in Linguistics. However, I knew that I could not move forward with Linguistics. It is too theoretical for me. I made a decision to move to Georgetown to study Sociolinguistics. Life was not that smooth. My advisor moved to a new university. Although I got lost here, I managed myself to earn an M.S. in Sociolinguistics at Georgetown University. Then, I transferred my Georgetown credits to the University of Pittsburgh. Here I focused on how American kids learned Japanese in a Foreign Language in Elementary School (FLES) program. This is part of the project funded State Department directed by Professor Richard Donato and Professor G Richard Tucker (Carnegie Mellon University). As a research assistant, I was responsible for assessing kids' attitude and motivation in relation to their achievement. With this hand-on experience, I understood more about assessing language and related variables that contributed to language achievement.

Right after my Ph.D. graduation in 2000, I started my teaching career at Maejo University in Chiang Mai, Thailand. I taught English Foundation courses which seemed not to relevant to my education background. Fortunately, Professor Tucker introduced me to ETS to serve as a committee member of TOEFL Board Grants and Awards (2003-2008). This opportunity has widened my horizon and my network in language assessment.

I moved to Chulalongkorn University in 2012. Here I have two main roles--(1) Lecturer of English at Chulalongkorn University Language Institute and (2) Head of Language Assessment and Evaluation Track of the English as an International Language (EIL) Program, Graduate School. I can say that, in Thailand, Chulalongkorn University is the only place that offers the study of Language Assessment and Evaluation at a graduate level. Here, I have more opportunity to do more research in language assessment. My on-going research is the link between language policy and the university exit exam, and mapping the university test with the CEFR. Our program gained more reputation when we hosted the 2nd AALA conference in Bangkok in 2015.

Language assessment and evaluation in Thailand is continuously growing up. More and more graduate students are interested in this field as they know that working in this area is challenging yet rewarding to the improvement of English education. It is very fortunate that we have the Asian Association of Language Assessment so that all students in this field, especially in Asia, can learn more from each other and help to promote the growth of language assessment. Good Luck with my best wishes ☺

Special Feature 2

2016 AALA Conference Student Award Winners

The AALA Best Student Paper Award and AALA Best Student Poster Award are officially established in 2016. Both are merit-based awards that support the attendance of the winners at the AALA annual conference.

Best Student Paper:

Paul Wicking (1st place winner)
Meijo University, Japan
Understanding teacher belief and practice as a foundation for developing learning-oriented assessment

Bjorn Lennart Norrbom (2nd place winner)
National Center for Assessment in Higher Education, Saudi Arabia/University of Jyväskylä, Finland,
Taking test validation a step further through a two-level structure of Assessment Use Arguments

Best Student Poster:

Roxanne Wong
City University of Hong Kong/University of Jyväskylä
Design and implementation of an in-house computerized feedback system

AALA Founding Members (in alphabetical order)

He, Lianzhen (Zhejiang University, China)
Imao, Yasu (Osaka University, Japan)
Jin, Yan (Shanghai Jiao Tong University, China)
Kunnan, Antony John (University of Macau, Macau)
Lee, Yong-Won (Seoul National University, South Korea)
Lee, Young-Shik (Hannam University, South Korea)
Qian, David (Hong Kong Polytechnic University, Hong Kong)
Sawaki, Yasuyo (Waseda University, Japan)
Wu, Jessica Row-Whei (Language Training and Testing Center, Taiwan)

AALA Executive Board (2016-2017)

President
David Qian (Hong Kong Polytechnic University, Hong Kong)
Co-President
Jessica Wu (Language Training and Testing Center, Taiwan)
1st Vice President
Yan Jin (Shanghai Jiao Tong University, China)
2nd Vice President
Jirada Wudthayagorn (Chulalongkorn University, Thailand)
Immediate Past President
Antony John Kunnan (University of Macau, Macau)
Secretary/Treasurer
Yasuyo Sawaki (Waseda University, Japan)
Communications Officer
Yasuhiro Imao (Osaka University, Japan)
Members at Large
Lianzhen He (Zhejiang University, China)
Young Shik Lee (Hannam University, South Korea)

AALA Student Committee (2015-2016)

Advisor
Yan Jin (Shanghai Jiao Tong University, China)
Co-chairs
Rongchan Lin (Columbia University, USA)
Sha Liu (University of Bristol, UK)
Secretary
Carol Dabarera (Nanyang Technological University, Singapore)
Treasurer
Xiaohua Liu (The University of Auckland, New Zealand)
Public Relations Officer
Lin Lin (The University of Hong Kong, Hong Kong)
Committee Members
Junghee Byun (Seoul National University, South Korea),
Matthew Wallace (Nanyang Technological University, Singapore)

Committee Contact Information

AALA Website:
http://www.aalawebsite.com/

Email: aalastudent@hotmail.com

Facebook Page:
https://www.facebook.com/aalastudentcommittee/

Facebook Group:
https://www.facebook.com/groups/AALAStudentGroup/

LinkedIn:
https://www.linkedin.com/groups/6941774/profile

AALA Student Committee In Conversation Series

Issue 2 (Fall/ Winter 2016)

Editors: Carol Dabarera & Sha Liu