

An Interview with APPLE Lecture Speaker Dr. Okim Kang

Daniel Eskin, Soo Joo, Kelly Katherine Frantz

INTRODUCTION

On April 18th, 2025, the *Studies in Applied Linguistics and TESOL (SALT)* journal interviewed Dr. Okim Kang, an invited speaker for the 2025 APPLE Lecture Series hosted by the Applied Linguistics and TESOL program at Teachers College, Columbia University.

Dr. Kang is a Professor of Applied Linguistics, Director of the Applied Linguistics Speech Lab, and Director of the Global Communication Center at Northern Arizona University. Her research interests are speech perception/production, language social psychology/attitudes, applied phonology/speech intelligibility, L2 oral assessment/testing, and automated speech recognition. She has published and co-authored eight books on the topics of spoken discourse, prosody, and communication success, including a handbook about contemporary English pronunciation. She has published over 110 academic articles and given 260 keynote and conference presentations. She is an associate editor for Applied Linguistics and serving on the editorial board for various prestigious journals (e.g., Language Testing, SSLA, Language Teaching, Language Assessment Quarterly, Journal of Second Language Pronunciation, and Languages). She obtained over 45 grants, including National Science Foundation and various testing agencies (e.g., Educational Testing Service, Duolingo, IELTS, the Cambridge English Assessment, and the British Council).

We thank Dr. Kang for her participation in this interview.

THE INTERVIEW

[Here](#) is the link to the full interview.

Background

1. You are well known for your research on second language pronunciation and intelligibility, linguistic stereotyping, and L2 speaking and listening assessment. Could you start by talking about what sparked your interest in these areas of research? Was there a particular moment or experience that drew you to this work? [[Q1 Video](#)]

Linguistic Stereotyping

2. I first came across your work on accent and speech perception through your NPR interview a couple of years ago. I found it fascinating—especially the concept of *linguistic stereotyping* and your classroom activity where students heard the same speech

while viewing photos of people of different races. It really highlighted how stereotypes can influence perception. You emphasized the importance of raising awareness about this issue. As a second language educator, one possible step might be incorporating a variety of accents into learning materials and assessments. However, I also read in your more recent work that learners often *prefer* the standard American accent in listening inputs.

From an assessment perspective, given that nowadays we are prioritizing localized assessment, do you think it would be effective to introduce the dominant local accent in those assessments? [[Q2 Video](#)]

Accents and AI

3. Bridging back to your recent work with AI models, a lot of AI systems are traditionally trained on the standard American accent. At the same time, I recently came across an AI-based app called *BoldVoice* that detects users' different accents. This does suggest that they are training these systems with different types of accents. However, I noticed that the app still markets itself as training for a standard American accent. This isn't unique to English. I'm originally from South Korea, and I've also heard of Korean AI models that struggle to recognize regional accents. So even if LLMs are *technically* trained on diverse accents, if their *practical* use still centers around the "standard," they seem to still be reinforcing linguistic stereotyping.

Many students in our program are working with these AI models. Do you have any suggestions or guidance for educators who want to integrate these AI models or develop their own systems? [[Q3 Video](#)]

Prosody and Pragmatic Meaning Conveyance

4. I appreciate the focus in your research on concepts like prosody, prominence, intonation, and tone choice where you are talking about more than simply pronunciation but how contextually appropriate meanings are conveyed in spoken English. Considering this research, do you have any recommendations for how ESL and EFL teachers can help students understand the role of these suprasegmental features in spoken interaction and pragmatics in English? [[Q4 Video](#)]
5. As a technological issue, you mentioned in your 2022 Duolingo talk (Kang, 2022, [see 48:43](#)) that acoustic models in the automatic speech recognition process, at the time, did not fully pick up on how suprasegmental features, such as intonation and prosody, convey pragmatic meaning. Based on your understanding, do you see a way forward in terms of how further training of these models could improve the automated recognition of these speech features? [[Q5 Video](#)]
6. In your 2018 publication on prosody and pragmatics (Kang & Kermad, 2018), one of the theoretical models on the relationship between prosody and emotion that you review was developed by Columbia computer scientist, Julia Hirshberg. Do you feel that her

framework characterizes how emotion is plotted onto the use of this speech feature sufficiently? [[Q6 Video](#)]

Automated L2 Speech Feedback

7. As someone who has developed automated tools for providing feedback for L2 learners' speech features, you've talked about the challenge of translating numeric values measuring these features into information that can be understood by teachers and learners and used to improve learners' speaking skills (Kang, 2022, [see 50:41](#)). In your opinion, do you feel that it is easier to provide automated feedback for certain speech features than others? [[Q7 Video](#)]
8. So, in your research on providing automated L2 speech feedback to learners, you first provided training to learners on how to properly interpret the feedback. Once learners received this training, did you and the other researchers see an effect on their speech features? [[Q8 Video](#)]

Intelligibility

9. On the topic of comprehensibility and intelligibility, I found it fascinating that there are different approaches to intelligibility, such as transcription accuracy, body language, and prosodic features. You mentioned prioritizing meaning and the contextual features. Are there any implications for that in terms of teaching and AI models? [[Q9 Video](#)]

Conclusion

10. You mentioned a forthcoming study about AI and bias. Would you be willing to tell us more about that? [[Q10 Video](#)]
11. In terms of your future research, it seems like your studies have found that it wasn't only the actual speech samples that affected the results of the output, but also these other multimodal features, such as the appearance of the person or the perception of what identity that speech sample plotted onto. Is that a future direction for training data? [[Q11 Video](#)]

REFERENCES

- Kang, O. (Dec, 2022). *Webinar: Understanding language for automated assessment of speaking* [Video]. YouTube. https://youtu.be/pYqIaC_XYrs?si=qVD_IsKFg9yu2WL1
- Kang, O., & Kermad, A. (2019). Prosody in L2 pragmatics research. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 78-92). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781351164085-6/prosody-l2-pragmatics-research-okim-kang-alyssa-kermad>