

Accent and Variability of L2 Speech

Okim Kang

Northern Arizona University

AAAL Webinar, 2020

February 7, 2020



NORTHERN ARIZONA
UNIVERSITY

Outline

- Accent
- L2 speech constructs
- Learner attitude (e.g., Kang, 2015)
- Speaker variability (e.g., Kermad & Kang, 2018)
- Listener variability (e.g., Kang & Rubin, 2009)
- Listener training (e.g., Kang et al., 2015 and others)
- Accent varieties in global contexts
(Kang et al., 2019; Kang & Liu, 2018)
- Implications



What is accent exactly?

- Most people are interested in accents.
- People can tell whether someone is speaking with an accent almost immediately (2-3 seconds).
- Even if what they are listening to is played backwards!
- We use accent to make judgments about other people.



Disney Movies

Jafar (*Aladdin*)



Scar (*Lion King*)



Clocksworth and
Candle (*Beauty
and the Best*)



If we look at children's movies, accents can serve one of the two purposes: Either an evil genius or an funny character.

Disney Movies

Aladdin



Simba &
Mufasa



Belle



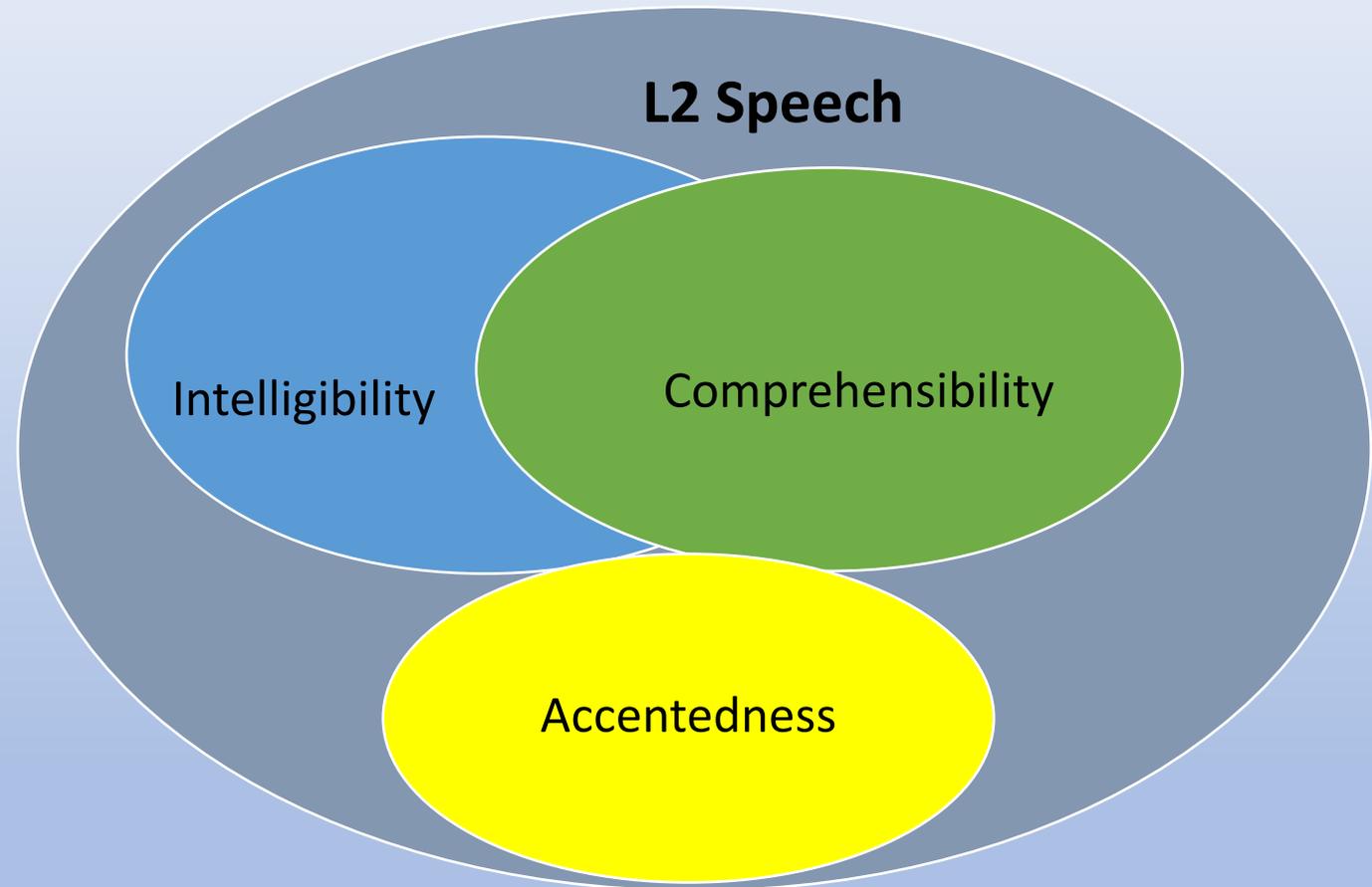
- Never the main characters (especially, no prince or princess).
- Children are growing up with this expectation and stereotype that good people should have no accent.

Accents

- What is accent?
 - A distinctive mode of pronunciation of a language, especially one associated with a particular nation, locality, or social class.
 - Phonological difference from a local norm – we all have accents.

Major speech constructs (Derwing & Munro, 2005)

- In L2 speech, we often talk about the following 3 constructs.
- **Comprehensibility:** Perception of how easy it is to understand an utterance.
- **Accentedness:** Perception of how different a speaker's accent is from that of L1 community.
- **Intelligibility:** how much the listener actually understands of the intended message.



Although intelligibility and comprehensibility are closely related, it is not a one to one correlation and accent is partially independent of either of the other dimensions.

3 Dimensions of Speech

Accent

Phonological
difference

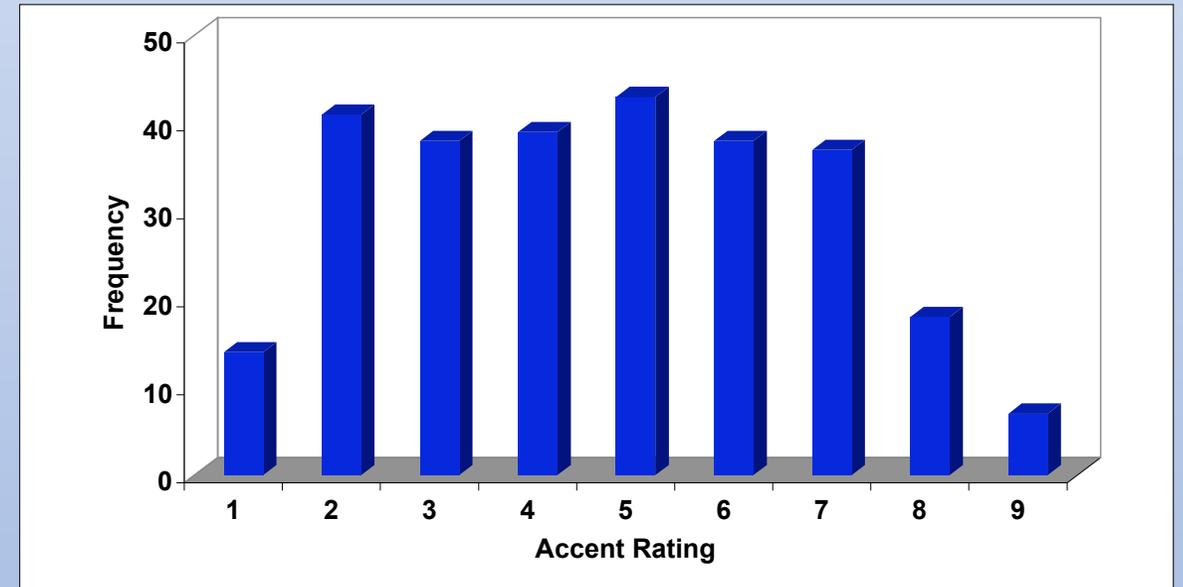
Comprehensibility

Listener effort

Intelligibility

The extent to which
a listener
understands an
utterance

**Relations between Accent and
Intelligibility (Munro & Derwing, 1995)**



Ratings of utterances with 100% intelligibility by transcription, but the accent ratings for these utterances ranged from 1 to 9. L2 speech can be completely intelligible, but can be heavily accented.

Learners' Perceptions Toward Accent

- Many learners still prefer to model native speakers from inner-circle countries such as the UK or from North America (e.g., Bayard, Gallois, Ray, Weatherall, & Sullivan, 2002)
- The majority of learners consider speaking with perfectly native pronunciation to be a desirable goal (Derwing, 2003; Kang, 2014; Scales, Wennerstrom, Richard, and Wu, 2006)
- Students are dissatisfied with their current curriculum of learning pronunciation due to misunderstanding of various models and accents made available to them (Kang, 2014).

Accent/Pronunciation Instructional Goal

- Pronunciation instruction should be
 - Moving toward setting up an international version of English
 - Helping learners to build a realistic goal for their pronunciation (Derwing & Munro, 2005; 2017)
 - Developing their own intelligible accent (Kang et al., 2019)

- English Learners...
- often relying on native speakers' judgments on their accent
- *"I am the one who has the accent or it's my English that is the problem"*.



- The actual problem is...
(in many cases), not the speakers (i.e., the learners)
.....but the listeners.

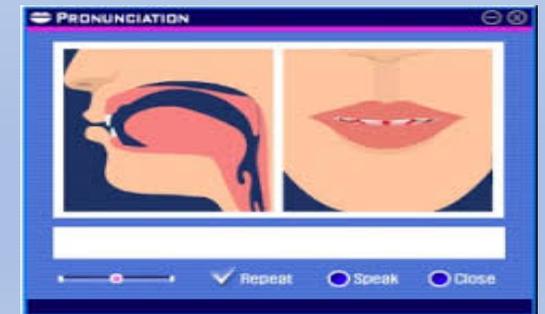
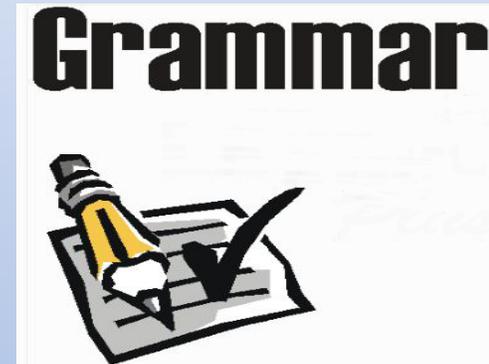
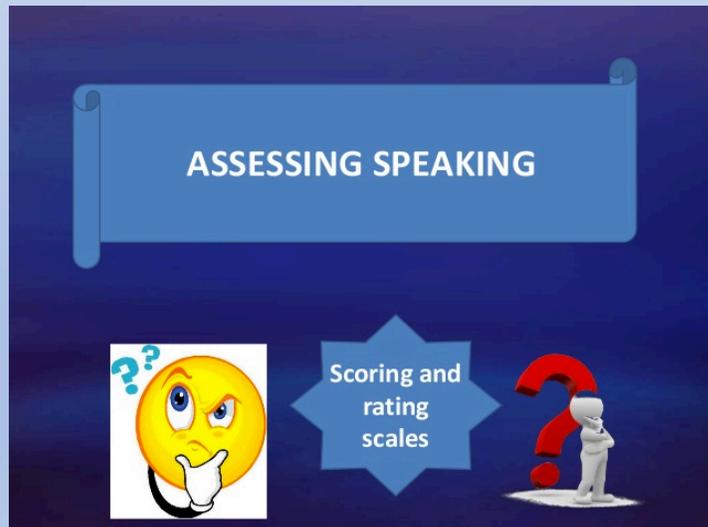


Speaker Variability



L2 Speech Ratings

- How accurately can we assess a non-native speaker's speech?
- On what basis do we make the decision?



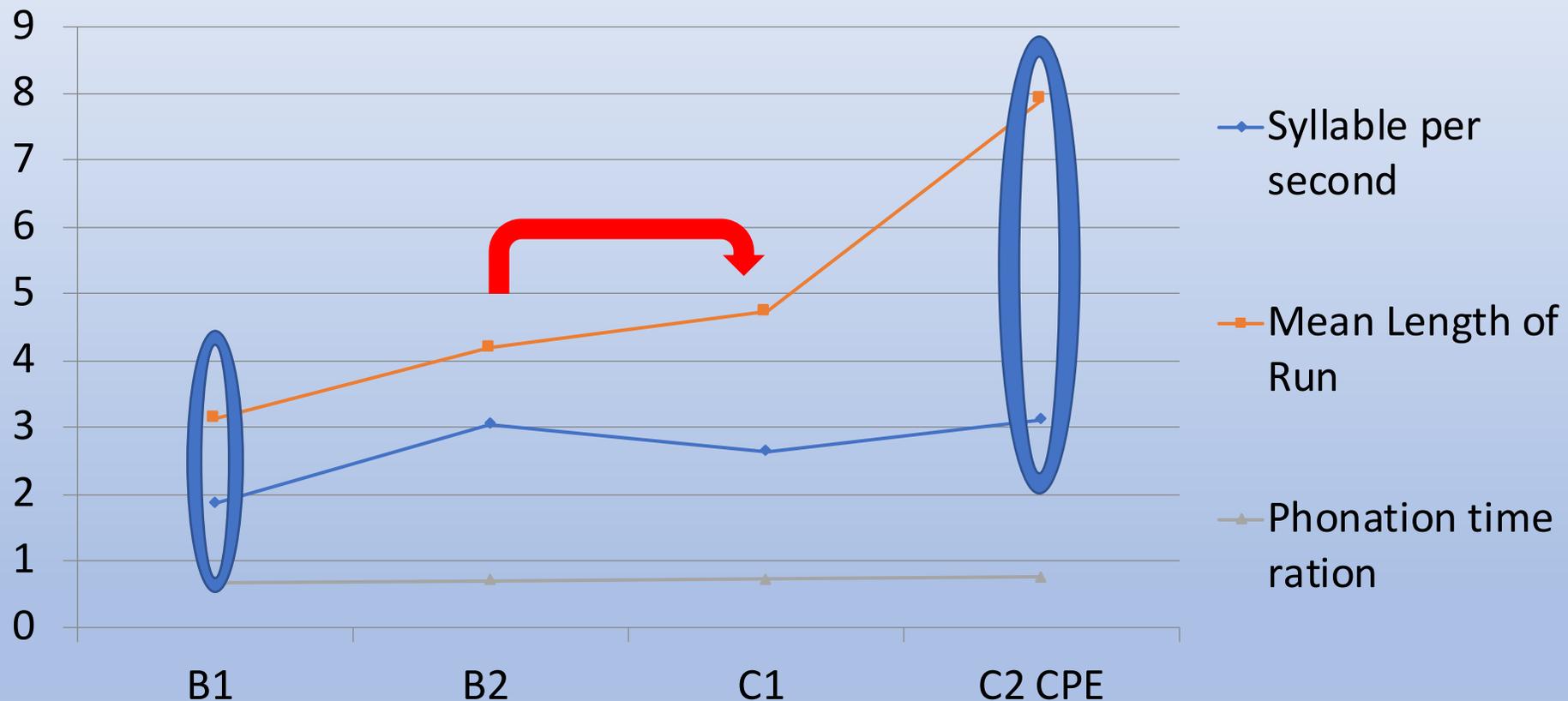
Rating Criteria and Speech Properties

- **Relationships between ratings and speech features** (Iwashita et al., 2008; Kang et al., 2010)
- The rating features of speaking skills (e.g., pronunciation) have not been clearly identified.
- The lack of consistent differences between adjacent groups or levels
 - No linear relationship in the development of linguistic features



The lack of consistent differences between adjacent levels

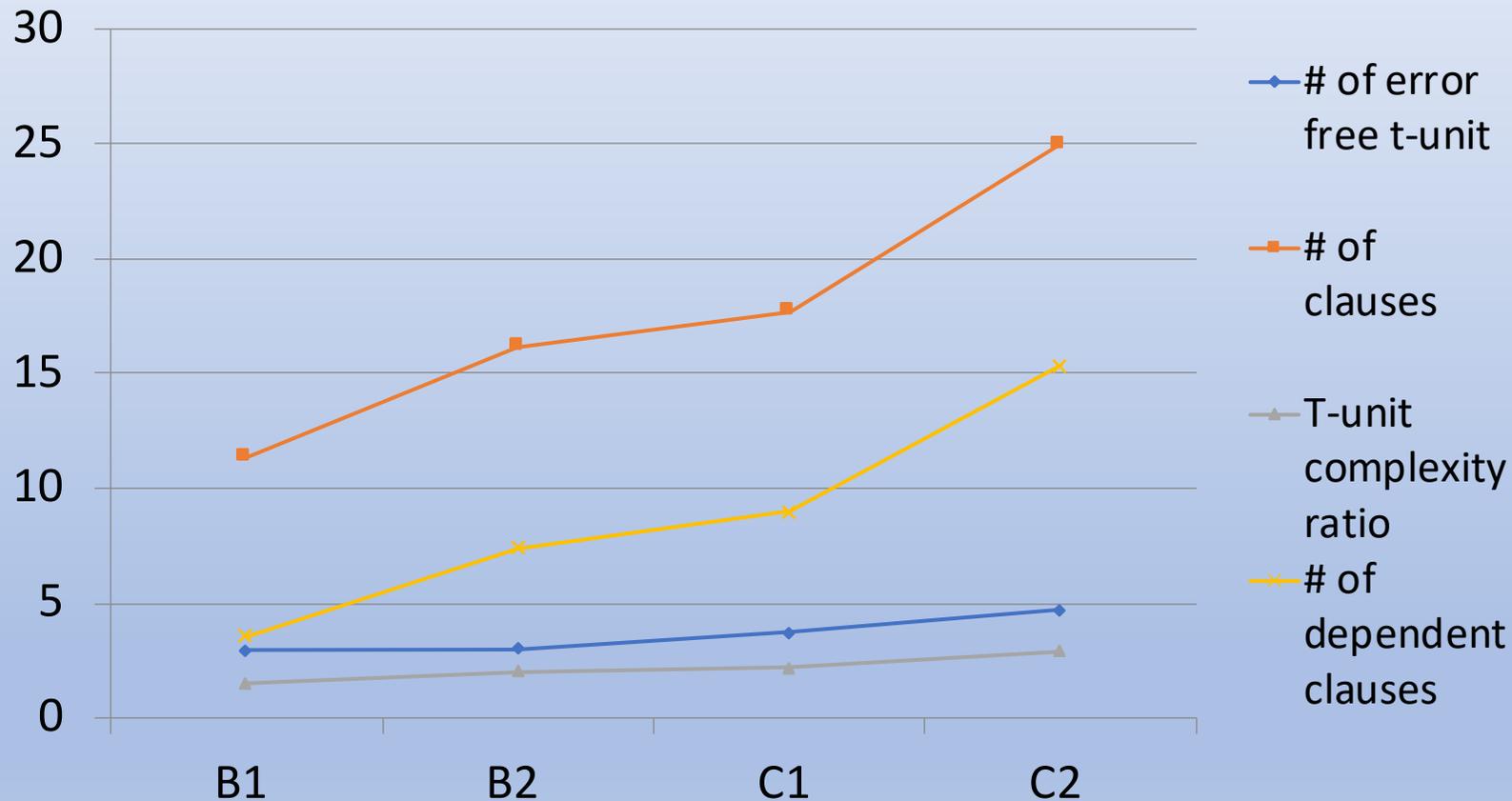
Fluency Features Across Levels in Cambridge English Assessment (Kang, 2014), analyzing 120 files (30 files from each level)



The CEFR: common European framework of reference for languages

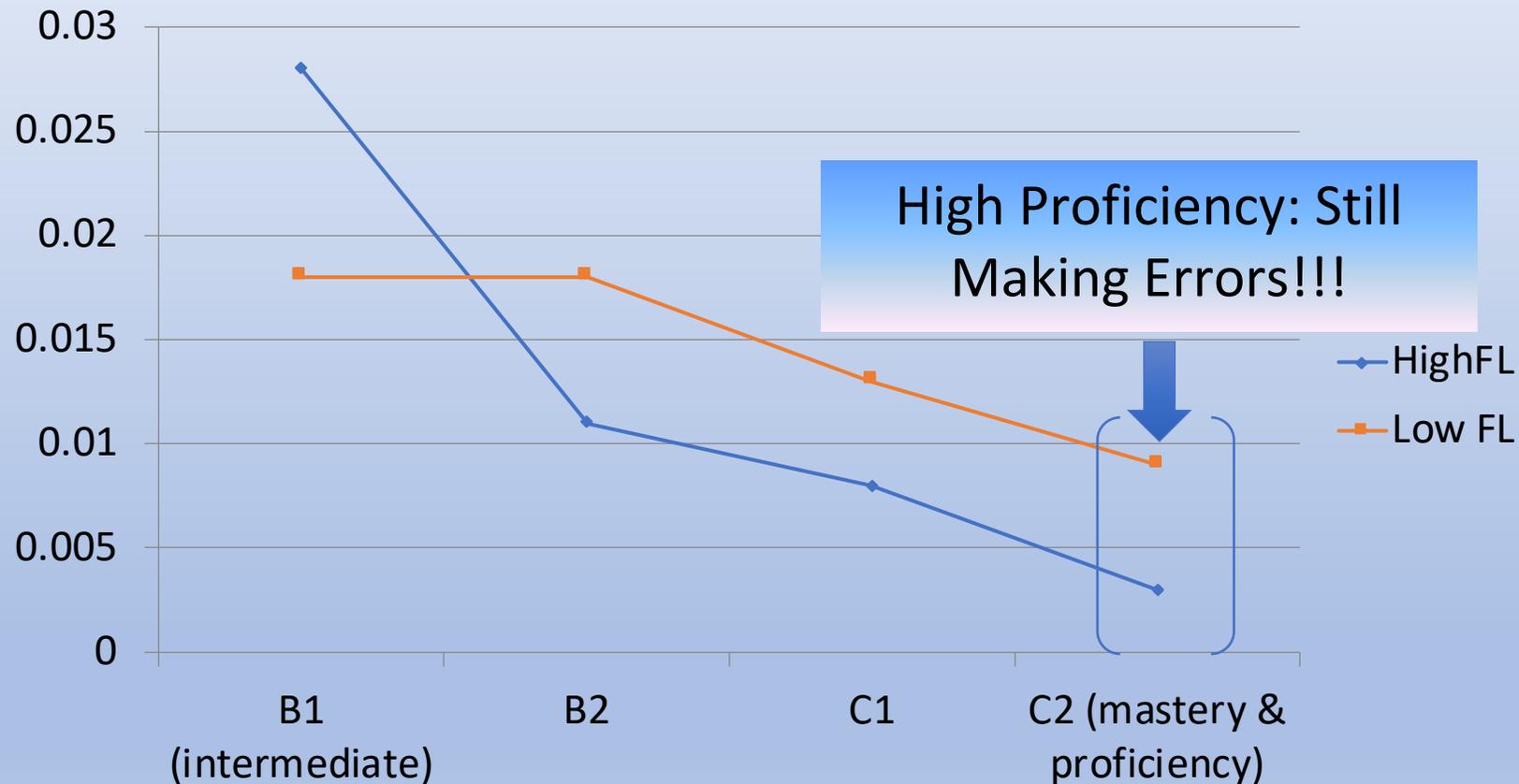
The lack of consistent differences between adjacent levels

Grammatical Complexity Features Across Levels in Cambridge English Assessment (Kang, 2014), analyzing 120 files (30 files from each level)



The CEFR: common European framework of reference for languages

Functional Load Segmental Errors (Kang & Moran, 2014): -Analyzing Cambridge Language Assessment 120 sound files-



The CEFR: common European framework of reference for languages

Phonological Analyses (Kang, Rubin, & Pickering, 2010)

- **TOEFL iBT spoken responses**
- **Analyzing 112 sound files**
- **29 detailed acoustic analysis**
- **Rate**
 - Mean length of run
 - Syllables per second
 - Phonation time, etc.
- **Pausing**
 - Total pause time
 - Mean length of pauses, etc.
- **Pitch range**
- **Intonation** (Brazil, 1997)
 - Falling + rise-falling = falling
 - Rising + fall-rising = rising

Multiple
Regression

50% of
the
variance

DV1

Accentedness ratings (1-7)

DV2

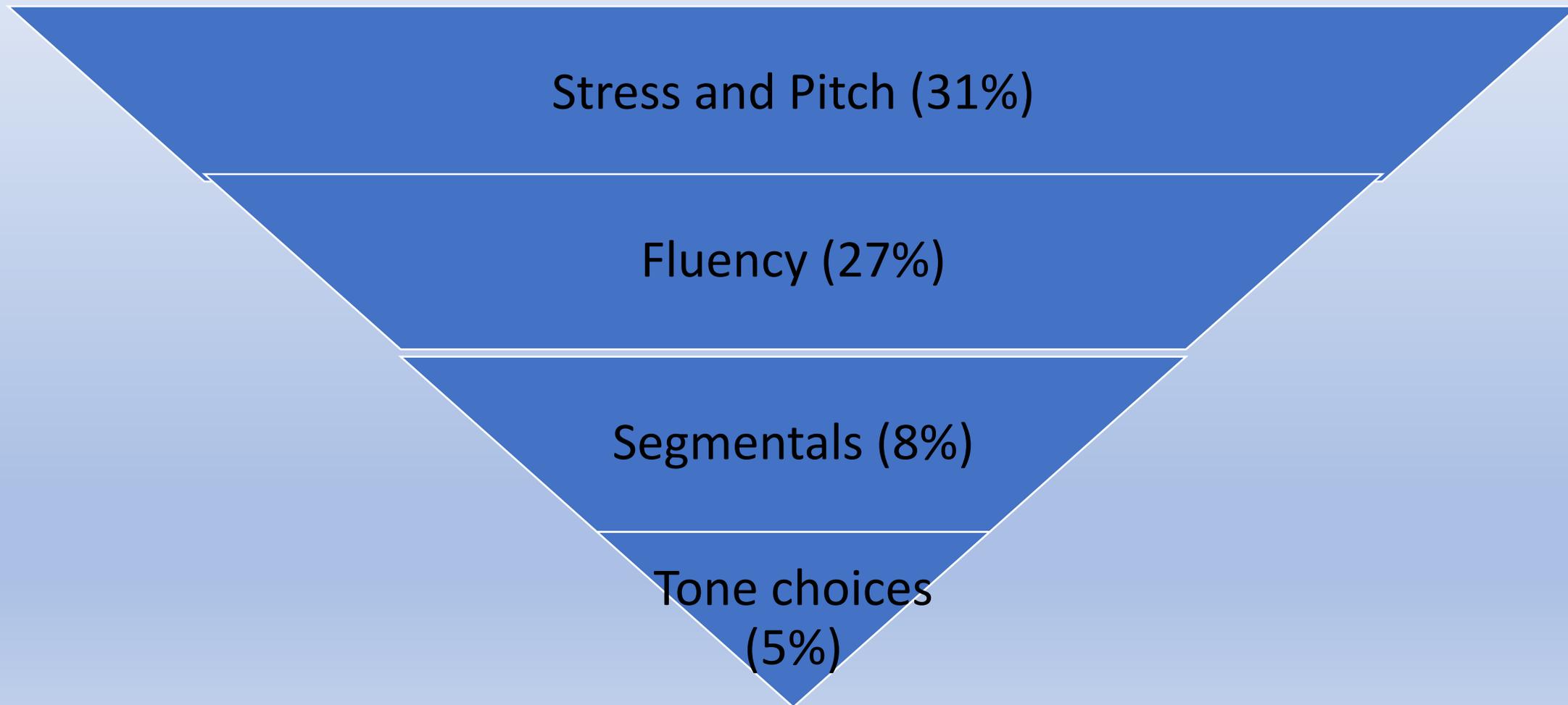
• Holistic ratings (0-4)

DV3

• Comprehensibility ratings (1-7)

Hierarchical Priority in Pronunciation Features

(Kang, 2013)



Factors that Affect L2 Speech

What factors influence L2 speech?

As we know...

- Age (Flege, 1987, 1992, 1995)
- Proficiency (Isaacs et al., 2015; Saito et al., 2016)
- L1 (Swan & Smith, 2001)
- Length of residence (Piske et al., 2001)

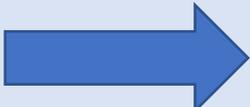
Also, individual differences (Kermad & Kang, under review) : 20 speakers + 60 undergraduate listeners

- Motivation
- Aptitude
- Anxiety

Kermad & Kang (under review)

	Comprehensibility	Intelligibility
R^2	16%	12%
Motivation	(+)	(+)
Aptitude	(+)	(+)
Anxiety	(+)	(+)

Effect of IDs on L2 Pronunciation

Significant Individual Differences	Phonological Feature	Statistics for Final Model Step
Motivation(-) Aptitude (-) Anxiety (+) 	Segmental Deviations	✓ $F(9, 10)=6.809, p=.003$ ✓ $R^2 =.860$
Age (-) Motivation (+) Anxiety (-) 	Fluency	✓ $F(10, 9)=15.402, p=.011$ ✓ $R^2 =.851$

Listener Variability



Rater Error

Speech Ratings



**How much error
is attributable to
listener
attitudes/
backgrounds?**

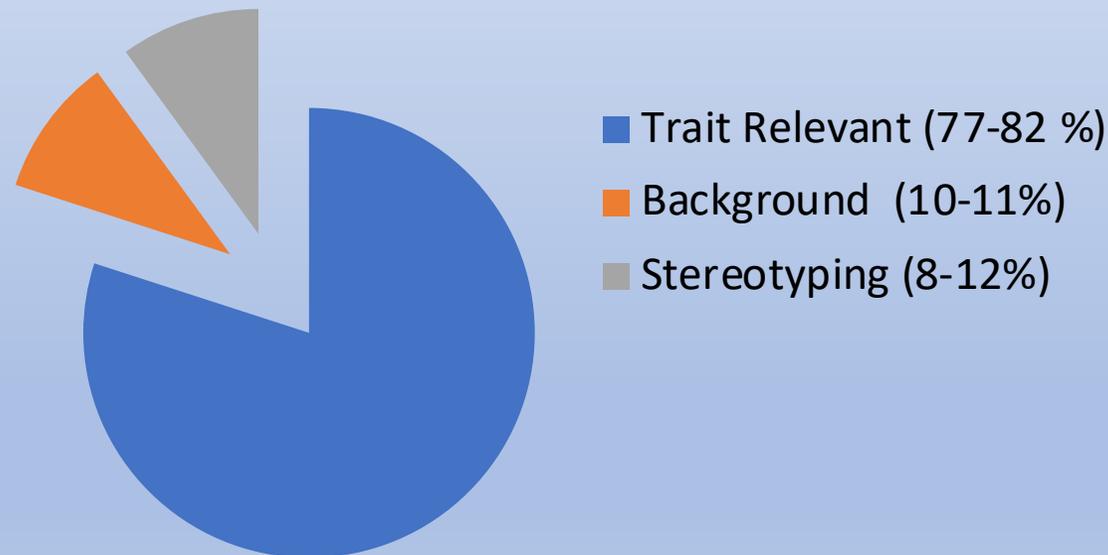
Subjectivity of Human Raters

- Rater's attitudinal and background factors
 - 18-23 % of variance (Kang & Rubin, 2009)

Listener's background factors (Kang, 2012)

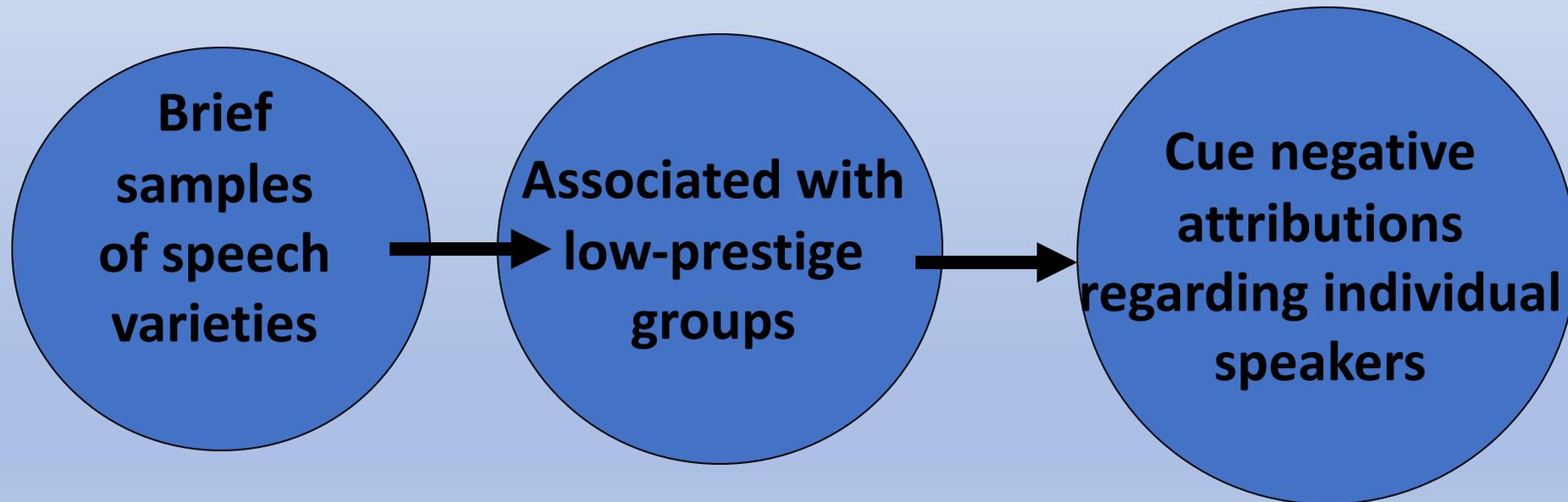
- Native speaker status
- Previous teaching experience
- Negative experience with NNSs

Listener Judgements



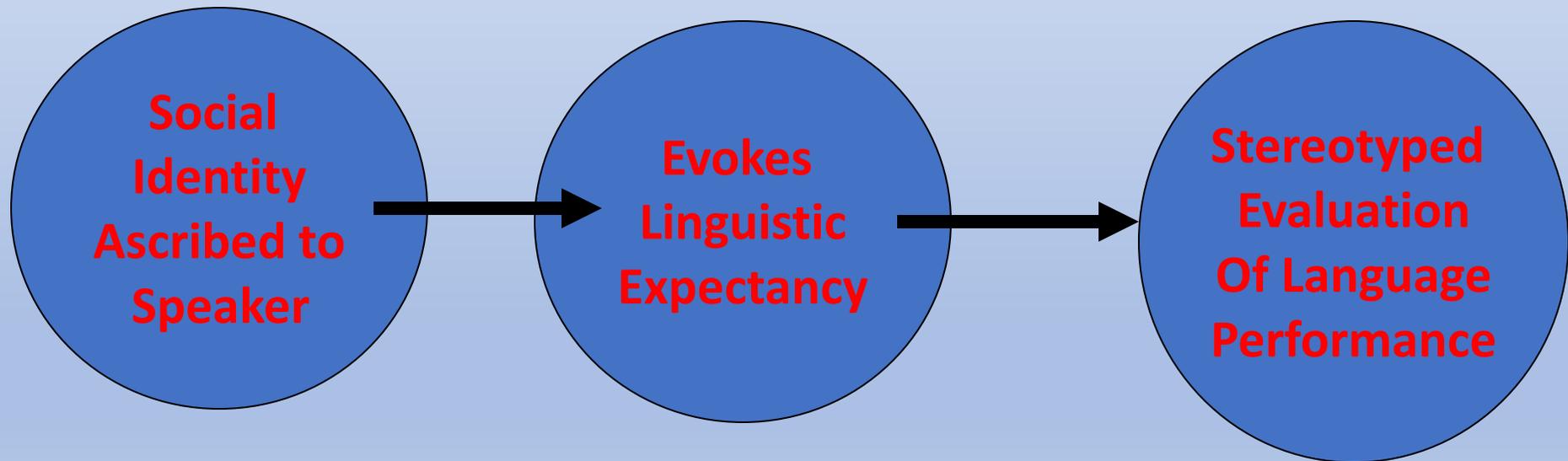
The Linguistic Stereotyping Paradigm

- Listeners ascribe to speakers stereotyped traits on the basis of their speech (Lambert et al., 1960)



The **Reverse** Linguistic Stereotyping Paradigm (RLS)

- Listeners ascribe stereotyped characteristics to speech based on information supplied about the speaker's social identity: Kang & Rubin (2009); Rubin (1992)



US Listener Bias (Kang & Rubin, 2009)

- Listener's own expectation and stereotyping
 - Attributions of a group membership
 - Perceived ethnicity of the speaker



Asian Face

**Listening
Comprehension**



Caucasian Face

ESL Student Bias (Ghanem & Kang, accepted)

- Understanding of the Lecture on Conditionals



Asian Female Teacher



Caucasian Female Teacher

Accent & Credibility (Lev-Ari & Keysar, 2010)

- Accent can determine the credibility of information.
- Accented speech is less believable than non-accented.
- Examples of those trivia statements:
 - *There are approximately 20,000 feathers on an eagle. (T or I don't know)*
 - *A mosquito has 2 teeth. (T or I don't know)*



Accent & Grammaticality (Ghanem & Kang, 2015):



NS-ungrammatical

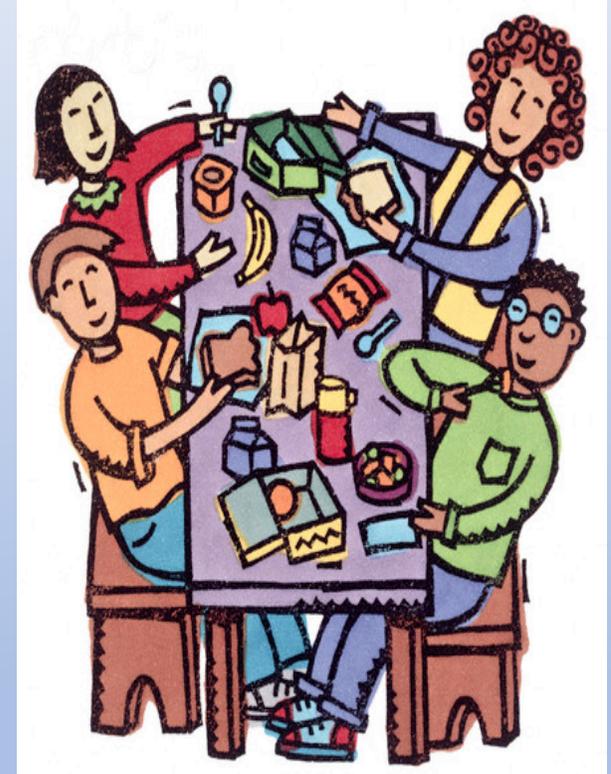


NNS-grammatical

- Accented speech makes English sentences less grammatical than non-accented speech?
- When listeners hear a sentence with a NNS accent, they find it ungrammatical, even though the sentence is grammatical.
- In contrast, when listeners hear a sentence with a NS accent, they find it grammatical, even though it is ungrammatical.

How to Overcome Bias/Stereotyping

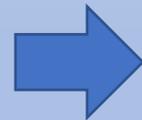
- Using Allport (1954) contact hypothesis...
- Short inter-group contact intervention (Kang, Rubin, & Lindemann, 2015)
 - Interacted informally with ITAs while sharing pizzas and drinks for an hour
- Institutionally supported contact activities (Staples, Kang & Wittner, 2014)
 - Participated in a language partner program once a week throughout the semester
- Course embedded contact activities (Kang & Moran, 2015)
 - Students enrolled in *Anthology 103* participated in culture exchange programs with international students three times a semester



Inter-group Contact

- After the *intercultural* contact activities, US undergraduate students showed positive changes in their attitudes toward international teaching assistants (ITAs) or accented speech in general.
- They evaluated the ITAs' speech as more comprehensible and thought the ITAs were instructionally more competent.

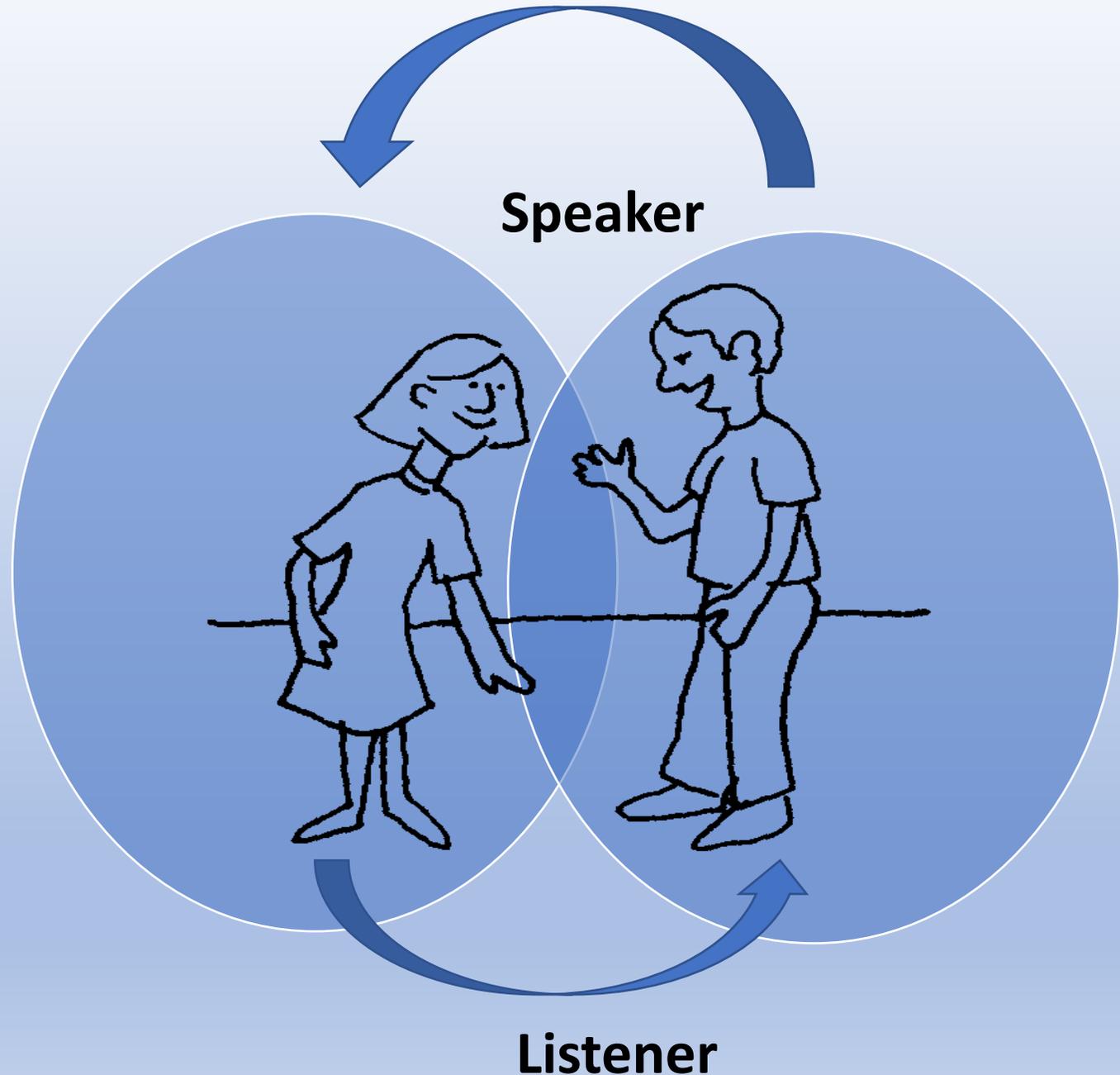
Positive change
in undergraduate
attitudes toward
ITAs



Influenced students'
perceptions of ITA
speech
performances

L2 Speech

- Conversation is a two-way street.
- The responsibility for effective communication between NNSs and NSs lies not only with the former as speakers, but also with the latter as active, responsive, and unprejudiced listeners



Training in Assessment Contexts

- Effect of training and rater individual differences on oral proficiency assessment. (Kang, Rubin, & Kermad, *Language Testing*, 2019).
- To what degree can a brief session of rater training neutralize the impact of potentially biasing rater individual differences on speaking scores associated with the iBT TOEFL®?
 - Individual differences includes
 - Teaching experience,
 - Native speaker status
 - Linguistic training
 - NNS weekly contact
 - Stereotyping measured by the RLS



Methods

Speakers

- 112 NNSs/examinees' sound files
- speech response to four iBT TOEFL® speaking **integrated** tasks
- equal number of speakers from four nationalities (Spanish, Korean, Arabic, and Chinese)
- range of possible holistic scores (0-4) determined by prior ETS ratings

Raters

- $n = 82$ (48 NSs and 34 NNSs)
- Naïve raters
- Undergraduate or graduate students from Southeastern university and/or residents of surrounding community

Methods

First Rating

All 82 naïve
raters

Rater Training

40 raters (20 severe and 20 lenient)

Training took place at least a month following
the completion of untrained ratings

ETS-prepared website for training iBT TOEFL®
raters (6-8 hours)

Second Rating

40 trained raters re-
rated the speech files

Listener individual differences on Ratings

	Pre-training*			Post-training		
Rating Outcomes	F	p	R ²	F	p	R ²
Holistic scores	4.34	.002	.529	.608	.745	.121

Training status	Holistic scores	
	M	s.d.
Naïve	3.03	.162
Trained	2.80	.278
Naïve	2.46	.220
Trained	2.65	.281



- The proportion of variance accounted for by these background variables was around 53% → down to 12 % after training.
- Dramatic effect of rater training
 - The impact of the extraneous rater variables was reduced by about 41%.
 - Both the previously severe and the previously lenient groups moved toward a more moderate stance.

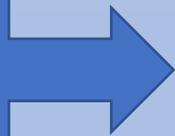
New Perspectives Toward World Englishes

The role of
English as an
international
language

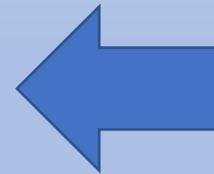


Different
perspectives to the
field of Applied
Linguistics

Ecological
Validity



Use of
Authentic
Materials in
Assessment



Different Varieties of English in Listening Comprehension

- The effects of international accents and shared L1 on listening comprehension tests (Kang, Thomson, & Moran, *TESOL Quarterly*, 2019)
- To what extent do different varieties of English have an impact on listeners' comprehension scores in the TOEFL listening test?



Participants: Speakers

- **18 speakers** were carefully chosen
 - 3 from 6 different countries
 - Inner Circle: The U.S. and England
 - Outer Circle: India and South Africa
 - Expanding Circle: Mexico and China
- Range of intelligibility (high, mid, and low), as determined by trained (expert) raters' scalar judgments + 47 listeners
- Professors or graduate students currently teaching at the university level.



Participants: Listeners & Materials

- 60 listeners from the same 6 countries as the speakers (**10 per country**)
- Senior undergraduate or early graduate students with a TOEFL score of **>100 (for NNSs)**
- TOEFL passages were controlled for:
 - style (professor's monologic speech)
 - Length (500-800 words), speech rate (3.2-3.6 syl/sec)
 - Number of questions (6)
 - Vocabulary range confirmed
 - Item difficulty controlled



The Effect of Speakers' Accent and a Shared L1 Effect with All 18 Speakers

Estimates of Fixed Effects for the Interaction of the Speaker's L1 and the Shared-L1 Influence for All 18 Speakers (Linear Mixed-Effects Model Analysis)

Parameter (interaction)	Estimates	Std. error	df	t	Sig.
[Indian speaker] X [accent-shared]	.700	.237	710	2.951	.003
[South African speaker] X [accent-shared]	.633	.237	710	2.670	.008

- The Indian listeners performed significantly better on lectures delivered by Indian speakers (Interlanguage Speech Intelligibility Benefit, Bent & Bradlow, 2003)
- The same pattern was found with the South African listeners.

The effect of speakers' accent and a shared-accent effect with the 10 highly comprehensible speakers

Listener Group	Mean	
American	5.516	
British	5.441	
Indian	5.475	
South African	5.500	.960
Mexican	5.550	.597
Chinese	5.600	.632
Total	5.500	.739

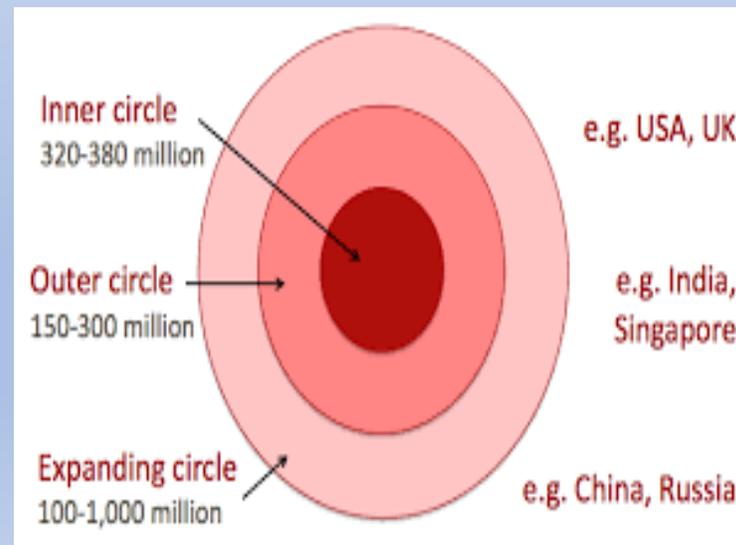
There was no significant difference among these 10 highly intelligible speakers

Proficiency Effect in Different Varieties of Accent (Kang, Moran, Ahn, & Park, *SSLA*, 2019)

- Proficiency as a mediating variable of intelligibility for different varieties of accents
 - 30 beginners + 30 intermediate + 30 advanced Korean learners of English
 - **highly proficient listeners** did not score any lower for any of the accented varieties than they did for the NES varieties (British and American English).
 - The **low proficiency listeners** also didn't show a difference, as they performed uniformly lowly for all speakers.
 - **However, the intermediate proficiency listeners' scores were significantly different ($p < .05$) across different varieties of speakers.**

WE instruction in EFL Classrooms Kang & Liu (2019; under review)

- To what extent can incorporating different varieties of World Englishes in the EFL classrooms affect students' listening comprehension?



Methods

- **Students Participants**
- 110 Chinese university students in China
 - 51: Experimental group
 - 59: Control group
- 81 Females, 29 Males
- Age from 18–20

6 speakers – 1 American, 1 British, 1 South African, 1 Indian, 1 Mexican, 1 Chinese

- Represented Inner, Outer, and Expanding Circle (Kachru, 1992)
- Speakers were screened and validated as part of a larger project (Kang et al., 2018)

WE-based Instruction

- 8 weeks of instruction

Experimental Group (51)

- Focused instruction on WE intelligibility
- Models with 6 different varieties of English (British, American, Indian, South African, Spanish, and Chinese).

Control Group (59)

- A traditional English-test preparation curriculum
- American accent only as a model

Methods – Data Collection

Students took listening
comprehension/perception **pre-tests**



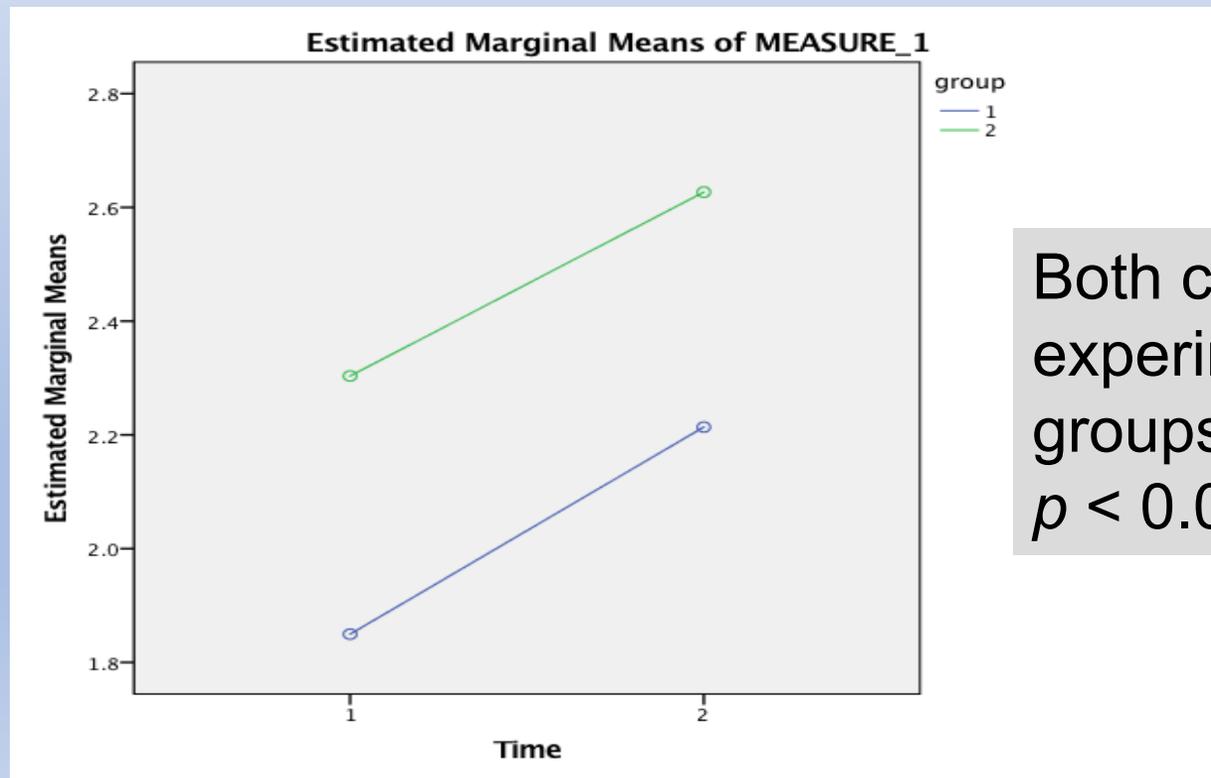
8 weeks' lessons



Students took listening
comprehension/perception **post-tests**

Overall Listening Comprehension Scores with All 6 speakers

Group	Pre-test Mean (SD)	Post-test Mean (SD)
Control	1.85 (1.43)	2.21 (1.42)
Experimental	2.30 (1.36)	2.62 (1.53)

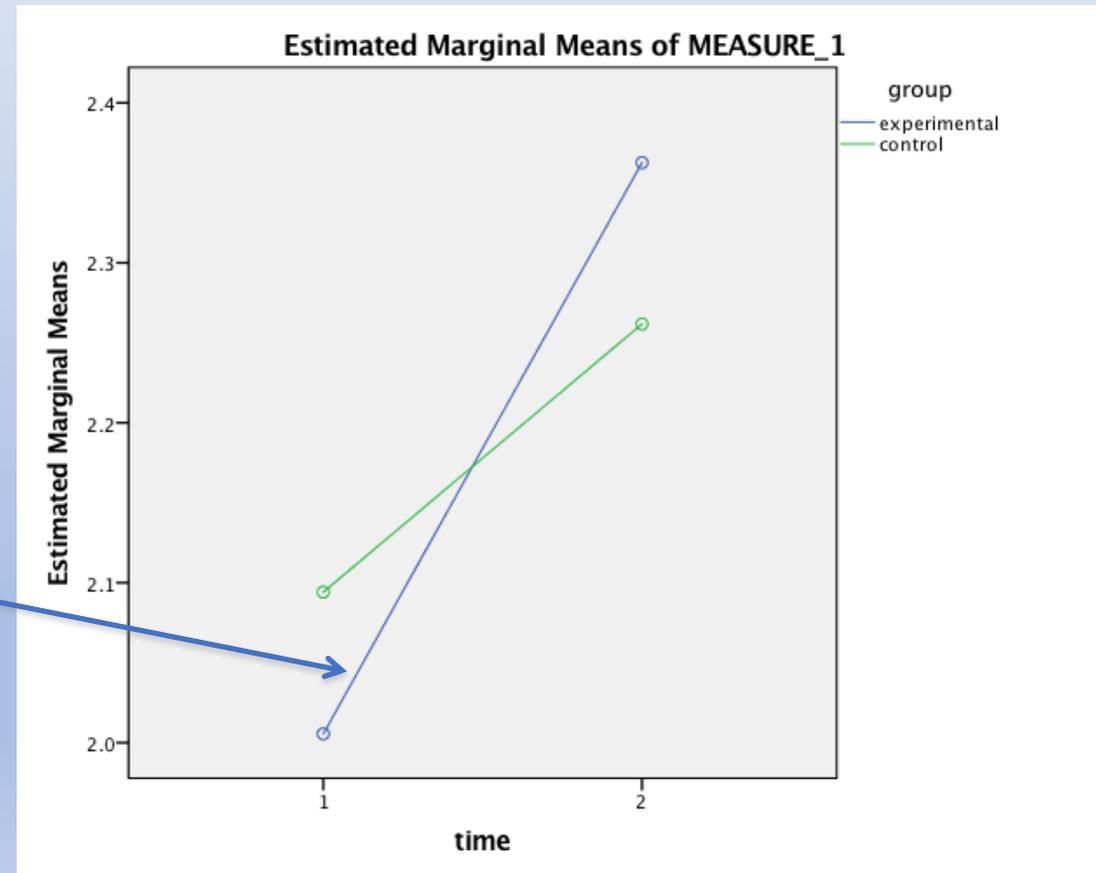


Both control and experimental groups improved. $p < 0.01$

Listening Comprehension Scores with 4 Non-inner Circle Speakers

Group	Pre-test Mean (SD)	Post-test Mean (SD)
Con.	2.01 (1.20)	2.26 (1.30)
Exp.	2.00 (1.46)	2.36 (1.20)

Only the experimental group improved. ($p = 0.005$)



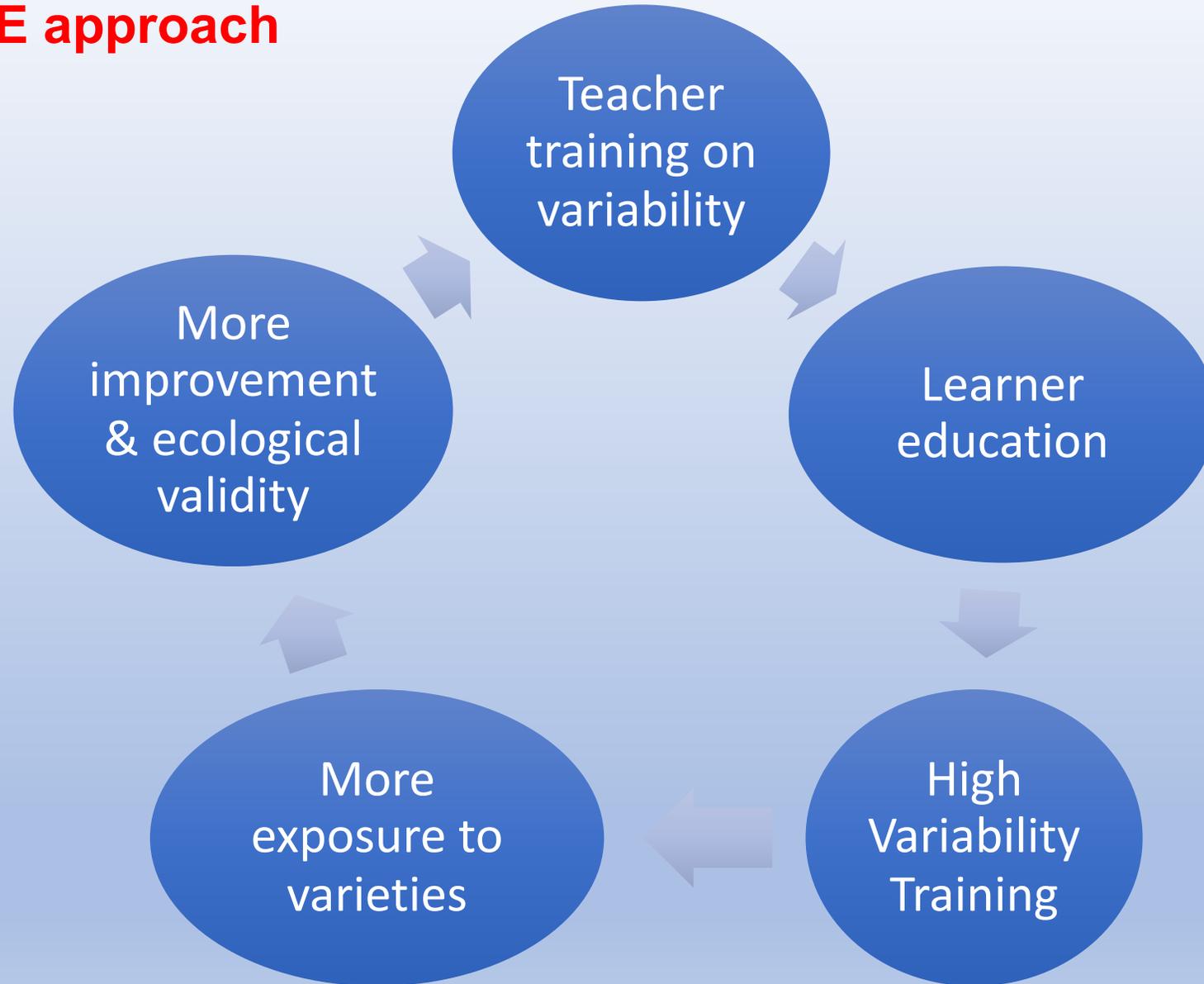
Implication & Future Research

- How does the incorporation of WE varieties, with highly intelligible speakers, affect students listening comprehension?



- More exposure, more familiar
- High variability phonetic training (Bradlow, 2008; Barriuso & Hayes-Harb, 2018)
- Ecological validity and fairness in language assessment
- EFL teachers are not comfortable with NNS models (Jenkins, 2005; Kang, Dalman, Larson, Wheeler, & Yaw, 2019)

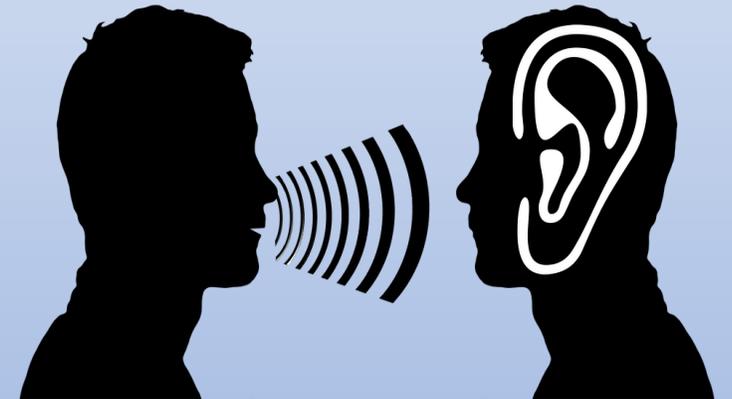
WE approach



Accent and Variability of L2 Speech

How can we better understand the sources of variance from both speakers and listeners in the global communication contexts?

- We all have accents.
- Communication is a two-way street.
- Speaker related factors (e.g., motivation, aptitude) and speech properties
- Listener backgrounds, biases, and stereotyping
- WE varieties in listening comprehension
- Training in different varieties of English
- Others....



My talk can inform your practice in language learning, teaching, and assessment.

Thank You

Okim.Kang@nau.edu