LANGUAGE TESTING, WORLD ENGLISHES AND ENGLISH AS A LINGUA FRANCA

The case for evidence-based change

Luke Harding

Lancaster University (UK)

Outline

1. Overview of the challenges of the World Englishes (WE) and English as a Lingua Franca (ELF) paradigms for language testing

2. A case study where empirical evidence was collected to address a particular challenge relating to speaker accents in listening assessment

3. Some future directions for language testing in meeting the challenges of WE and ELF
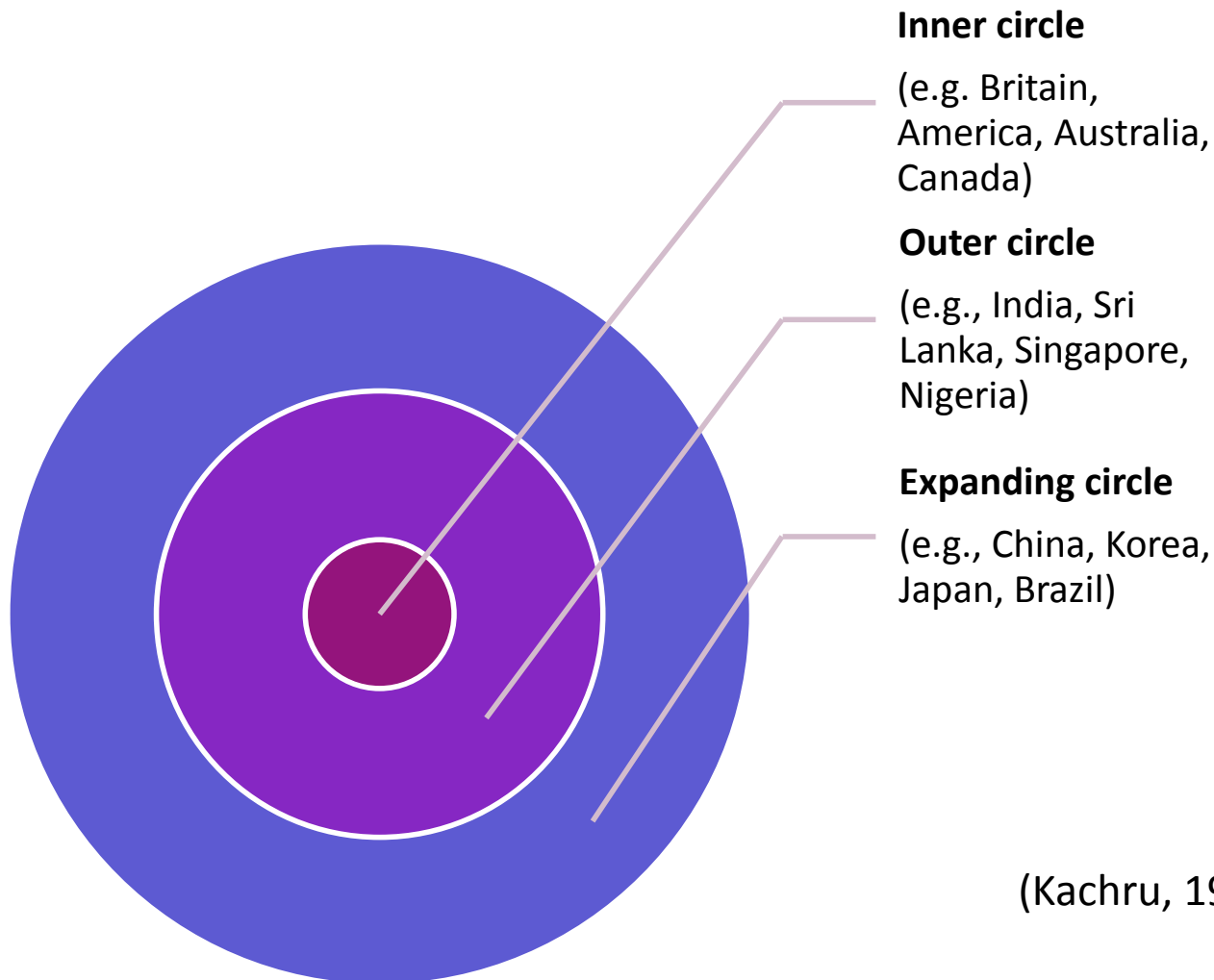
Challenge #1: Whose norms?

- Which norms should take primacy in assessment standards?

> … in language testing, an implicit (and frequently explicit) assumption has long been that the criteria for measuring proficiency in English round the world should be candidates' use of particular features of English which are used and accepted as norms by highly educated native speakers of English.

> (Lowenberg, 1993, p.95)

The World Englishes perspective

Inner circle

(e.g. Britain, America, Australia, Canada)

Outer circle

(e.g., India, Sri Lanka, Singapore, Nigeria)

Expanding circle

(e.g., China, Korea, Japan, Brazil)

(Kachru, 1985)

New and emerging varieties

- Hong Kong English (Hung, 2000)
- China English (Deterding, 2006; He & Li, 2009)
- Korean English (Shim, 1999)
- Japanese English (Morrow, 1987)

Applying native speaker norms (I)

- 1.   You must .... a lot of friends at school.

        A) be having
        B) have
        C) will have
        D) has

Applying native speaker norms (II)

"First, you have to beat the egg" (Singapore English)

[bit]

Or …

To what degree is this speaker's pronunciation influenced by an L1?

Implications of the spread of English

- Problematises notion of "correctness"

- As a result (in certain contexts), test designers need to re-think:
  - How they design criteria for speaking and writing assessment
  - How they create answer keys for grammar or vocabulary tests
  - How they train raters to be aware of which forms will be acceptable
  - How they select input for reading and listening tests

Challenge #2: What type of proficiency?

- Are current models of proficiency enough?
- What sort of proficiency should we be testing?

- English as a Lingua Franca (ELF)
  - What is it?
  - Why is it important to language testing?

English as a Lingua Franca (ELF)

- Communication between two (or more) non-native speakers (sometimes NNS – NS communication)
- English now spoken by 25% of world population (Source: British Council)
- More ELF conversations daily than conversations between native speakers
- Empirical work has begun, e.g.:
  - Jenkins' (2000) "Lingua Franca Core" for ELF pronunciation
  - Seidlhofer's (2001) Vienna-Oxford International Corpus of English
  - Mauranen's (2003) Corpus of Academic English

Proficiency in ELF?

- Successful ELF communication requires different types of competence:

    In a context where we have to constantly shuttle between different varieties and communities, proficiency becomes complex. To be really proficient in English today, one has to be multidialectal. This does not mean that one needs production skills in all the varieties of English. One needs the capacity to negotiate diverse varieties to facilitate communication. The passive competence to understand new varieties is part of this multidialectal competence ... Proficiency means, then, the ability to shuttle between different varieties of English and different speech communities.

    (Canagarajah, 2006, p.233)

Example – Aviation English

- Kim (2009; forthcoming)
- Context: Introduction of ICAO proficiency requirements
- Investigated whether lack of proficiency in English = communication breakdowns
- Findings showed that miscommunication caused by:
  - Failure to accommodate, negotiate and repair
- Should construct be broadened to include ELF competence?

Reasons for resistance in language testing

Reason	Explanation
Ideological conservatism	Testers as "gatekeepers"
Concerns for stability	Test-takers need to know what to expect
Concerns for fairness	Will all test-takers' varieties be represented equally
Concerns for acceptability	What norms do learners themselves want to aspire to? (see Taylor, 2006)

Dealing with the challenges

- Language testers now thinking about challenges in test design
- Much opinion, little "hard" research so far
- Language testing is high-stakes – scores can determine:
    - Entrance to a degree program
    - Professional qualification
    - Visa
    - Citizenship
- Need for careful, evidence-based change (see Elder & Harding, 2008)
- Some research has begun to address these issues …

A case study

- Project: *Accent and listening assessment: A validation study of the use of speakers with L2 accents on an academic English listening test*

- Focus: test-takers experiences of accent-related difficulty and the nature of "online" attitudes towards speakers with L2 accents.

Language Testing and Evaluation 21

Luke Harding

Accent and Listening Assessment

PETER LANG
Internationaler Verlag der Wissenschaften

Context

- Current testing practice takes an "orthodox" approach with only native-speaker varieties represented in listening test input (e.g., IELTS, TOEFL)
- Strong theoretical rationale for L2 accents:
    - Authenticity
    - Construct representation
- Orthodox approach driven by concerns over:
    - Intelligibility
    - Acceptability
    - Construct validity
    - Fairness
- Little empirical research to support or contest the orthodox approach

Aims and approach

Aims:

- To investigate how, and to what extent, the introduction of L2 accents affects performance on a listening test
- To investigate the nature of test-takers' "online" attitudes towards L2 speakers

Approach:

- Verbal report methodology:

  "Gathering data by asking individuals to vocalise what is going through their minds as they are solving a problem or performing a task" (Gass & Mackey, 2000, p.13)

- Conducted as immediate retrospective verbal reports (e.g., Buck, 1990; Wu, 1998)

Instrument

- University Test of English as a Second Language (UTESL)

- **Listening sub-test:**
  - 1 lecture (approx. 25 minutes)
  - Divided into 4 sections
  - Range of fixed-choice and open-ended tasks
  - 30-40 items per test

- **Topics:**
  - Sleep
  - The Oldest Old

Speakers

- **Japanese English accent**
  - Recorded input for the Sleep test

- **Mandarin Chinese English accent**
  - Recorded input for the Oldest Old test

Data collection

- **Participants:**
  - 8 NESB university students
  - 4 Japanese L1 and 4 Mandarin Chinese L1

- **Procedure:**
  - Pauses inserted at regular points throughout test
  - At each pause, retrospective verbal report was elicited
  - Response on test paper was used as a stimulus
  - Data collected for every item on the test + general comments at the end of each major section
  - Data were transcribed and Speaker, Pronunciation or Accent Related Episodes (SPAREs) were identified

Example of a SPARE

(LI-SL.GC1)

Li: <u>I just don't like the speaker</u> (laughs)

Researcher: *Okay … um can you explain a bit more?*

Li: <u>Yeah because her accent,</u> so just not familiar with the word, it just take me a long uh um maybe a few second to (get) used to it and to think about "oh is that the word?"
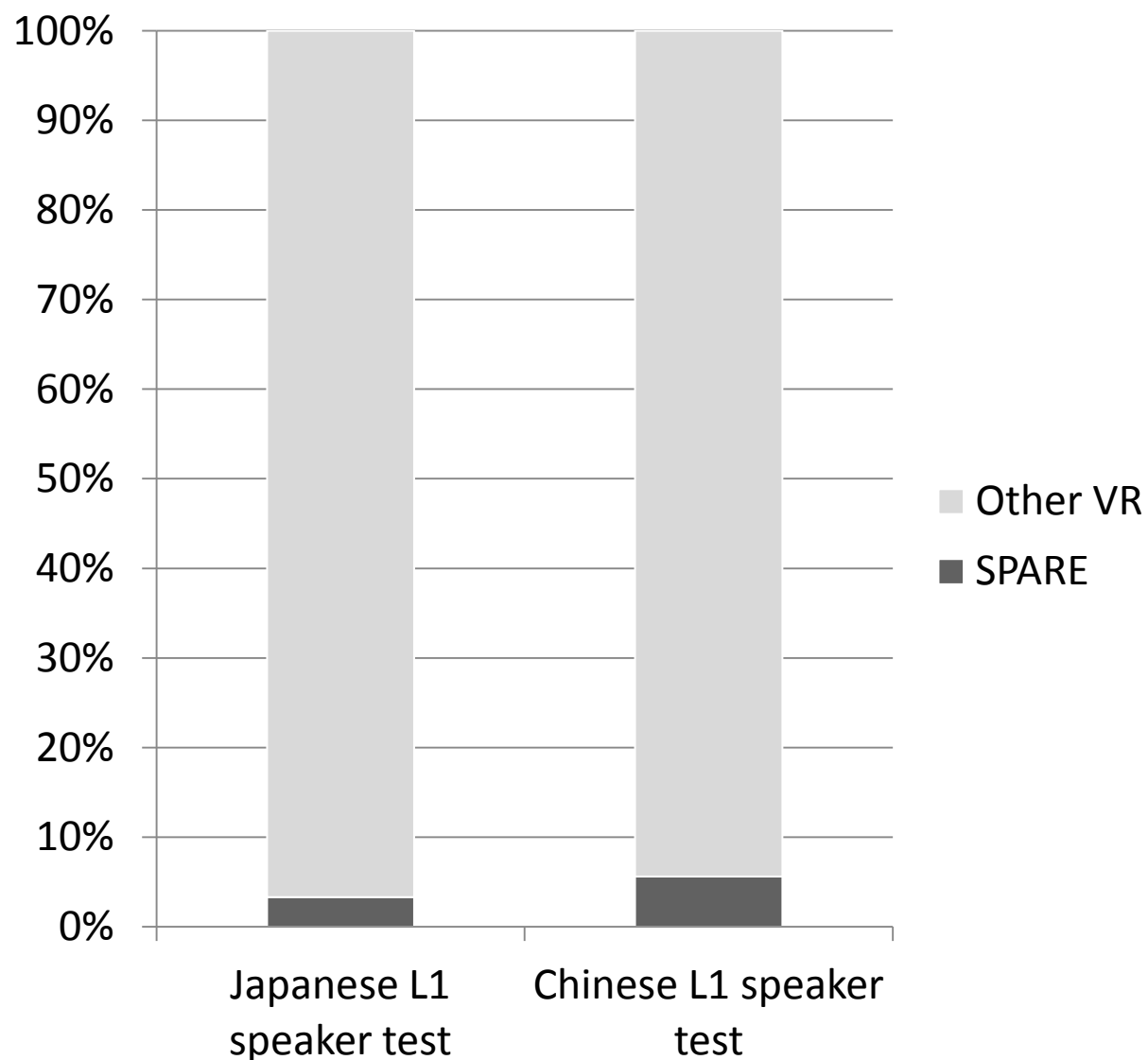
Researcher: *Okay …*

Li: Yeah it's just cost time you know … but if we cost time and we can't concentrate on it we just missed the information, it's not good

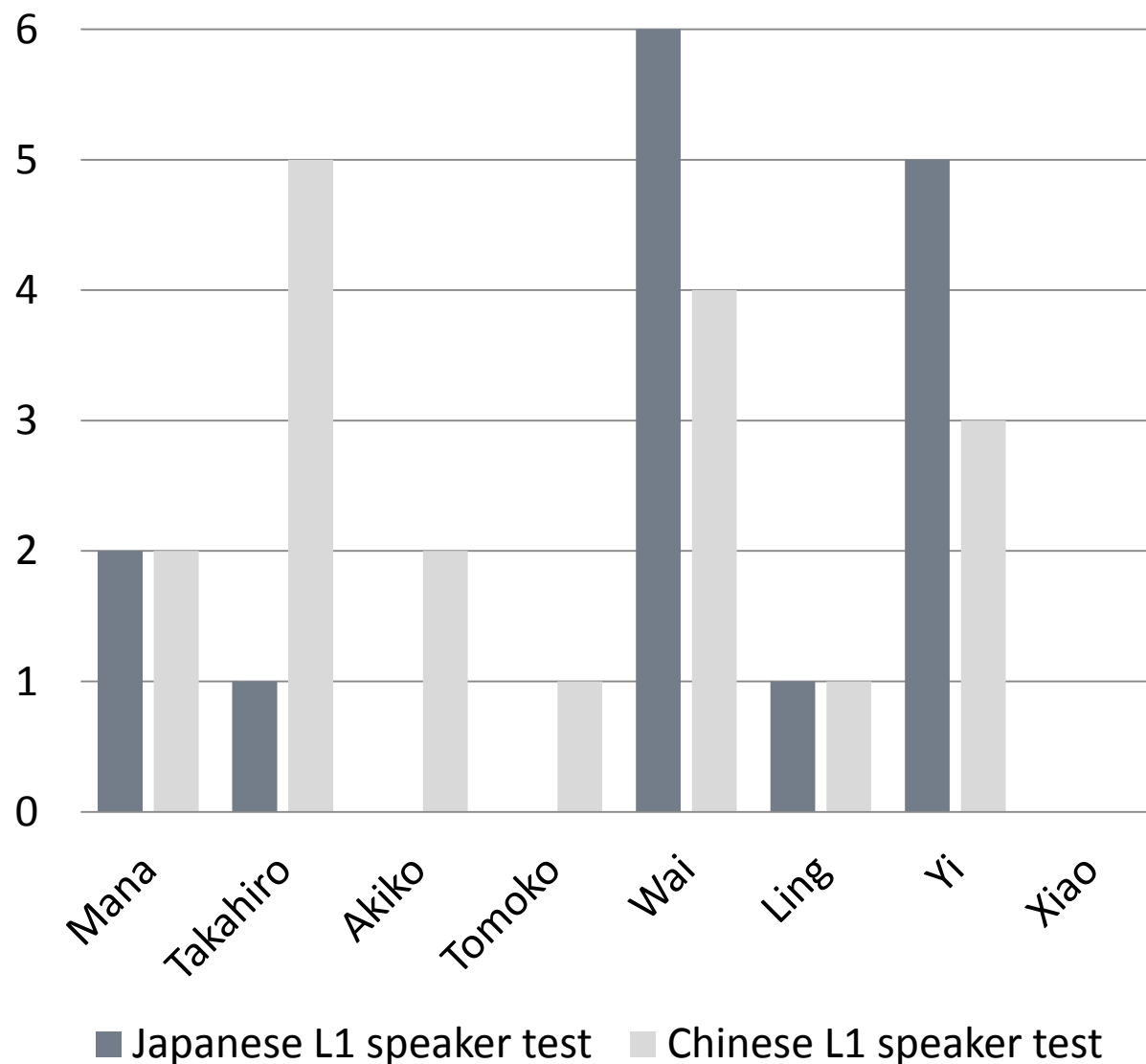**Frequency of speaker, pronunciation or accent related-episodes**

3.3% of total VR instances in test featuring Japanese L1 speaker

5.6% of total VR instances in test featuring Chinese L1 speaker

Frequency of speaker, pronunciation or accent related-episodes

- Bulk of spares came from 3 listeners

- 1 listener did not problematise speaker on either test

Coding of SPAREs

- Coding of all SPAREs revealed the following categories:

- Accent-related difficulty
  - Word recognition/perception
  - Processing cost

- Attitudinal response
  - Negative evaluation
  - Positive evaluation

Word recognition/perception

(MA-OO.12)

Researcher: *Um what were you thinking during that question?*

Mana: Um for first blank I thought um she would tell that place … and with I don't know … I really don't know with … but I've never heard of place before … <u>she said "central" … something</u> … so … I didn't really get where it is

Researcher: *So what difficulty do you think you were having there?*

Mana: Mm: … <u>unfamiliar word … and … her pronunciation I guess</u> (laughs) …

Inferencing

(TA-OO.23)

Researcher: *All right let's go back to the first ah answer you had there …*

Takahiro: Ah "environmental factors"

Researcher: *Okay what were you thinking while you were answering that one?*

Takahiro: Mm … the, <u>I thought that was a totally different word, j(ust) it sounded totally different to "environmental" … she has, I think she have has a bit of accent from I don't know where</u> so: …

Researcher: *What led you to that answer?*

Takahiro: Um <u>because of the cons, in context um … I just guessed and browsed my very thin dictionary (laughs) and I thought this is the answer I should put</u> mm

Tuning-in to an accent

(LI-OO.22)

Li: With this one ah of course we don't know what kind of things we should, we will hear(d) hear so just missed it, <u>just write "gyms" but it's not right? I don't think that's right</u>

Researcher: *What do you recall hearing?*

Li: <u>A just remember "gyms" … because she say "gyms" I say "how come why why she want to say 'gyms'?" so … suddenly missed the rest thing</u>

(LI-OO.25)

Li: <u>Ah: it's "genes" … I think do some sports like "gyms" (laughs) yeah I find the answer here</u>

Researcher: *Okay …*

Li: So the oldest old have low number of E-4 …

Researcher: *Okay what were you thinking there?*

Li: <u>Mm: this sentence is good, yeah it's remind me again so I can fix the, this one [points to item 22]</u>

Negative evaluation of speaker

- **Four types of negative evaluative response:**

(1) Expressions of personal dislike of the speaker's accent

(2) Judgments of speaker competence – both as an English user and as a professional

(3) Comments on the "normality" or alternatively the "weirdness" of certain pronunciation features

(4) Expressions of feelings of discomfort

Examples

(LI-SL.GC5)

Researcher: *Any general comments?*

Li: This one is easier than the first one [the Oldest Old] … yeah but <u>the reader is not good enough</u> (laughs)

Researcher: *Okay okay can you explain a bit more?*

Li: Um just because <u>her pronunciation is very weird</u> so just need yeah try to understand what she is saying then you can gots the information from it …

(MA-SL.36)

Researcher: *What what uh what were you thinking during that section?*

Mana: Mm: … <u>to be honest I don't like her English</u>

Researcher: *Okay*

Mana: (laughs) um … <u>like personally I like native English speaker … so mm … sounds like uncomfortable</u>

Difficulty → negative attitude

(LI-SL.GC1)

Researcher: *Okay um any general comments at this point?*

Li: <u>I just don't like the speaker</u> (laughs)

Researcher: *Okay … um can you explain a bit more?*

Li: <u>Yeah because her accent, so just not familiar with the word, it just take me a long uh um maybe a few second to (get) used to it and to think about "oh is that the word?"</u>

Negative attitude → difficulty (?)

(WA-SL.3)

Researcher: *Ah okay, and what were you thinking while you were trying to answer that that question?*

Wai: Um … <u>I've heard she saying something like "growth level" or … yeah I was lost</u> (laughs)

Researcher: *Um you say you were lost, what difficulties do you think you had there?*

Wai: Um … <u>cause when I like heard she um speaking I put bit more attention on … she pronounce "s" instead of "th" so … it sort of like draw me uh draw my attention on some, somewhere else</u>

Wai's negative attitudes

(WA-SL.GC2)

Researcher: *Do you have any general comments on that section?*

Wai: Um … still the … um … <u>the accent sort of make me feel uncomfortable</u>

Researcher: *What ah … what features of the accent or wh[y] why is the accent making you feel uncomfortable?*

Wai: Um … cause it's just um … it's not very, <u>it's not ver[y] English</u> … and uh … yeah it's just like, <u>it just feel like abnormal stuff so you feel it's weird</u> …

Researcher: *Okay, let's continue …*

Wai: Oh and one thing … um … like <u>if the accent is a bit weird like those other country's accent make me feel like not very professional like the information she give to me I will not believe like trust her that … much as some, someone sounds more professional</u> …

Summary

- Study provides evidence that:
    - L2 accents caused relatively few problems for listeners across two academic listening tests
    - Individual listeners varied widely in the degree to which they problematised L2 speech
- When accent/pronunciation was identified as a source of difficulty:
    - Key-word recognition was the most common difficulty reported
    - But … compensatory strategies were drawn on to deal with problems at the perceptual level
- When attitudinal responses were reported:
    - These tended to be negative; discourse of "correctness" and "normality"
    - But … a narrow type of ELF/intercultural communicative competence may be under test: ability to tolerate accents that are difficult/not preferred

Implication

- By attempting to integrate World English varieties on a listening test, some ELF competencies appear to be activated = a richer construct

Future directions

- Assessing English as a Lingua Franca?

> It seems to me that there is a great need for re-thinking the construct involved in language assessment. I think we are at a moment of very significant change, the sort of change that only comes along once in a generation or longer – the challenge that is emerging in our developing understanding of what is involved in ELF communication.
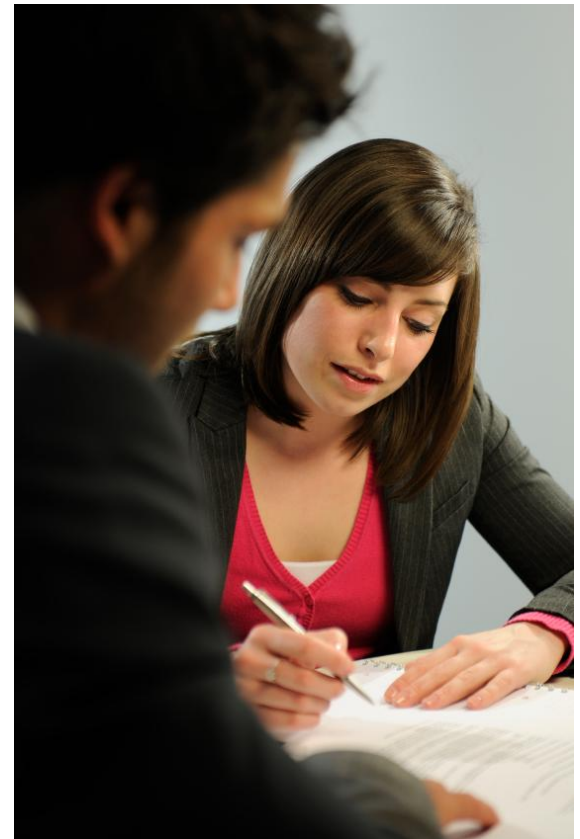
<div align="right">(McNamara, 2011)</div>

What would an ELF test look like?

- The ability to tolerate and comprehend different varieties of English: different accents, different syntactic forms and different discourse styles
- The ability to negotiate meaning when meaning is ambiguous
- The ability to use those phonological features which are crucial for intelligibility across speakers of different L1 backgrounds
- An awareness of appropriate pragmatics (e.g., awareness of politeness in cross-cultural situations)
- The ability to accommodate your interlocutor, to make yourself understandable to whomever you are speaking with
- The ability to notice and repair breakdowns in communication
- Others?

Is it already happening?

- Paired speaking tests
- NNS – NNS communication
- Test-takers need to:
  - Negotiate
  - Deal with variation
  - Accommodate
  - Repair
- Rating scales still focus on traditional communicative competences

What's changing?

TOEFL: ANNOUNCEMENT: TOEFL COE 2012 RESEARCH PROGRAM - Windows Internet Explorer

http://www.ets.org/toefl/grants/coe_2013_research_program

File   Edit   View   Favorites   Tools   Help

Favorites   Serviced Apartments in Hon...   Web Slice Gallery ▼   Suggested Sites ▼

TOEFL: ANNOUNCEMENT...   How to Capture Screen Shot...   Page ▼ Safety ▼ Tools ▼

immigration), (b) determining placement and promotion, or (c) informing pedagogy.

5. INNOVATION AND DEVELOPMENT

5.1 Conduct needs analysis of language use in diverse English-medium higher education contexts that might reveal new constructs not covered by the TOEFL iBT test (e.g., technology-mediated communication, lingua franca English instruction contexts).

5.2 Operationalize and investigate the measurement of intelligibility of different varieties of English (i.e., measures that define the thresholds of intelligible speech from distinct perspectives) to inform the design of listening assessment.

5.3 Investigate innovative ways of assessing academic English proficiency (focusing on one or more of the following: delivery modes, task/item types, psychometric models, scoring approaches).

TOEFL iBT dataset and research instrument available to COE-funded researchers

The dataset and research instrument described below will be made available to principal investigators funded by the TOEFL COE research program if needed.

TOEFL iBT Public Use Dataset

Three separate datasets from each of two test forms are available and include demographic information about test takers (age, gender, native language, etc.) and copies of relevant test material:

Done   Internet   200%

start   The Challenge of Engl...   Test-taker attitudes t...   Sent Items - Microsof...   LTRC 2008   TOEFL: ANNOUNCEM...   EN   11:16

What next?

- Tests of ELF may be on the horizon
- Raises issues of fairness – who should take these tests? Native speakers too?
- More research required to understand what is important in lingua franca communication
- World Englishes perspective needs to be incorporated in good test design – for validity's sake
- Views of stakeholders sought at every step of the process
- As English changes, so must assessment practices