

# Speech Technology and its Social Impact

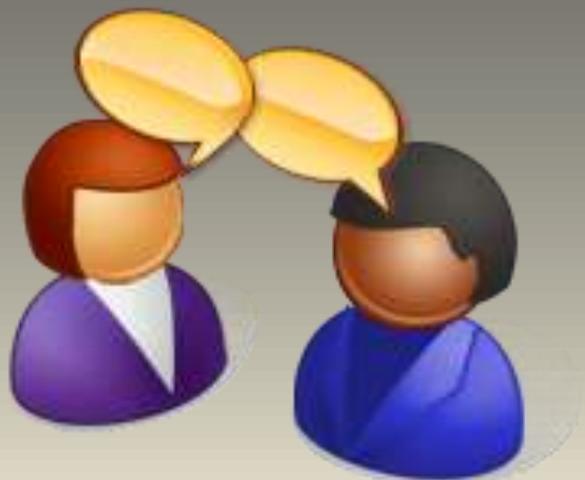
Patti Price, PPRICE Speech and Language Technology

## ➤ **Speech as a social technology**

- The social aspects of speech make it challenging
- We don't model those aspects well at present
- People vs. automatic speech recognition
- Opportunities
- Summary

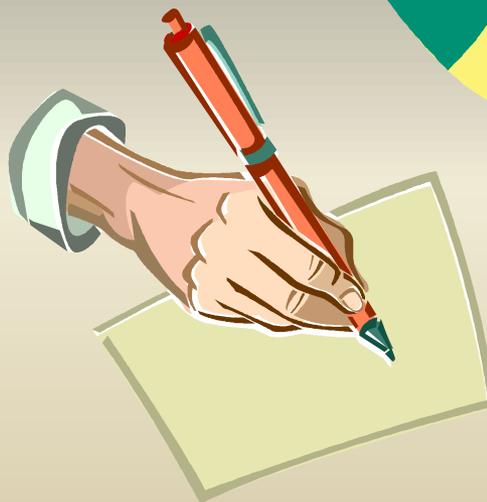


**Pre-language:** Information access by direct experience

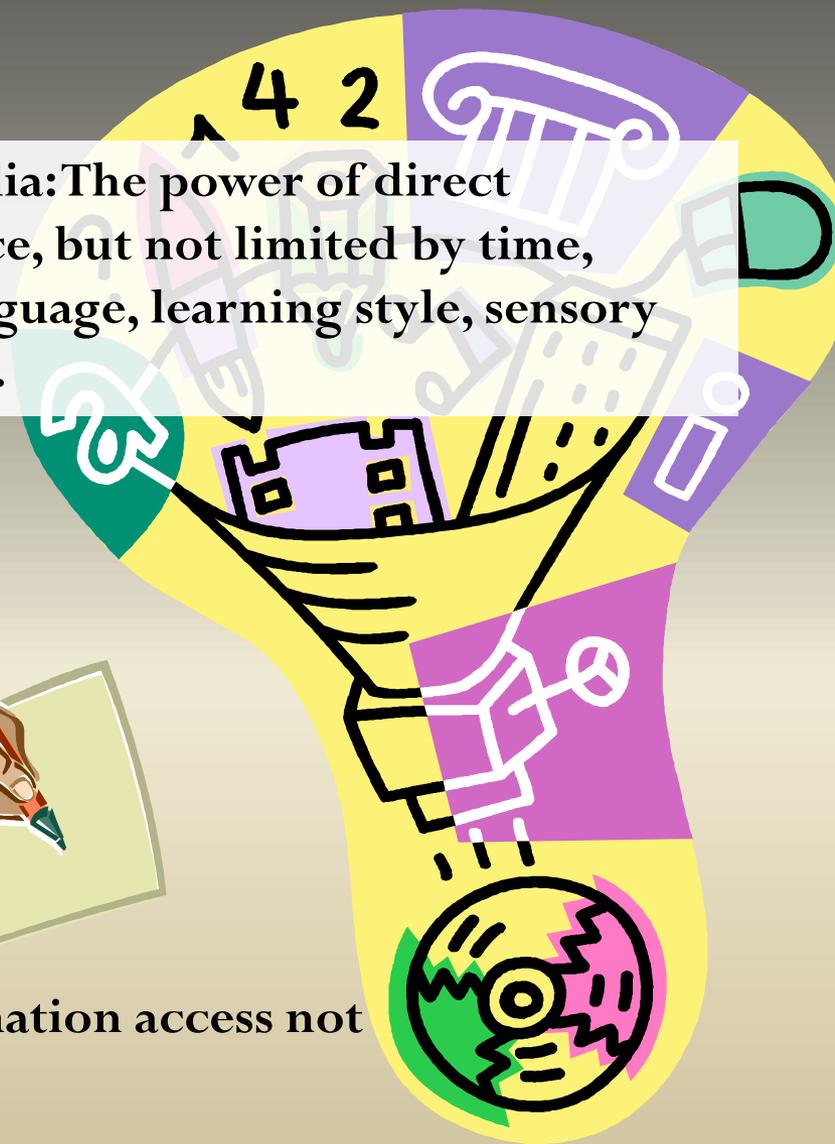


**Oral language:** Information access also from someone in same time and place with same language

**Multimedia:** The power of direct experience, but not limited by time, place, language, learning style, sensory abilities...

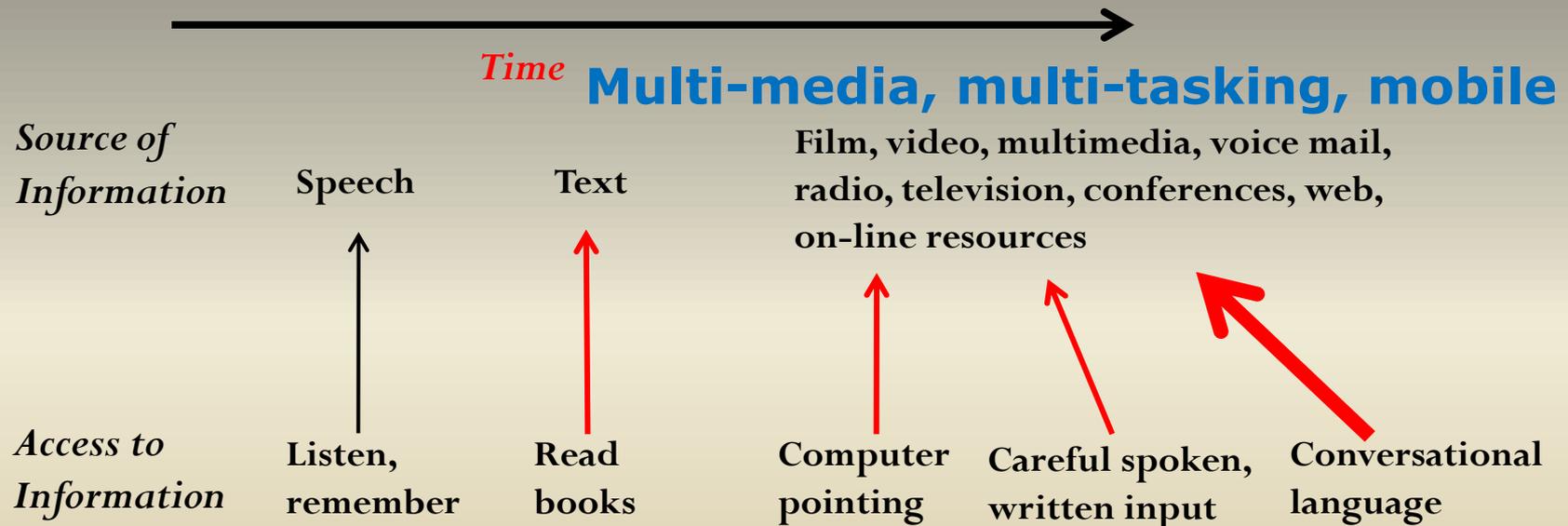


**Written language:** Information access not limited by time and place



# Speech in the Information Age

- Speech & text were revolutionary because of information access
- New media and connectivity yield information overload
- Can speech technology help?



**Speech is social in ways our technology is not. Can it become a complementary partner in with humans?**

# We wanted the knowledge navigator

- Instead, we got  
geth... on  
And... lie

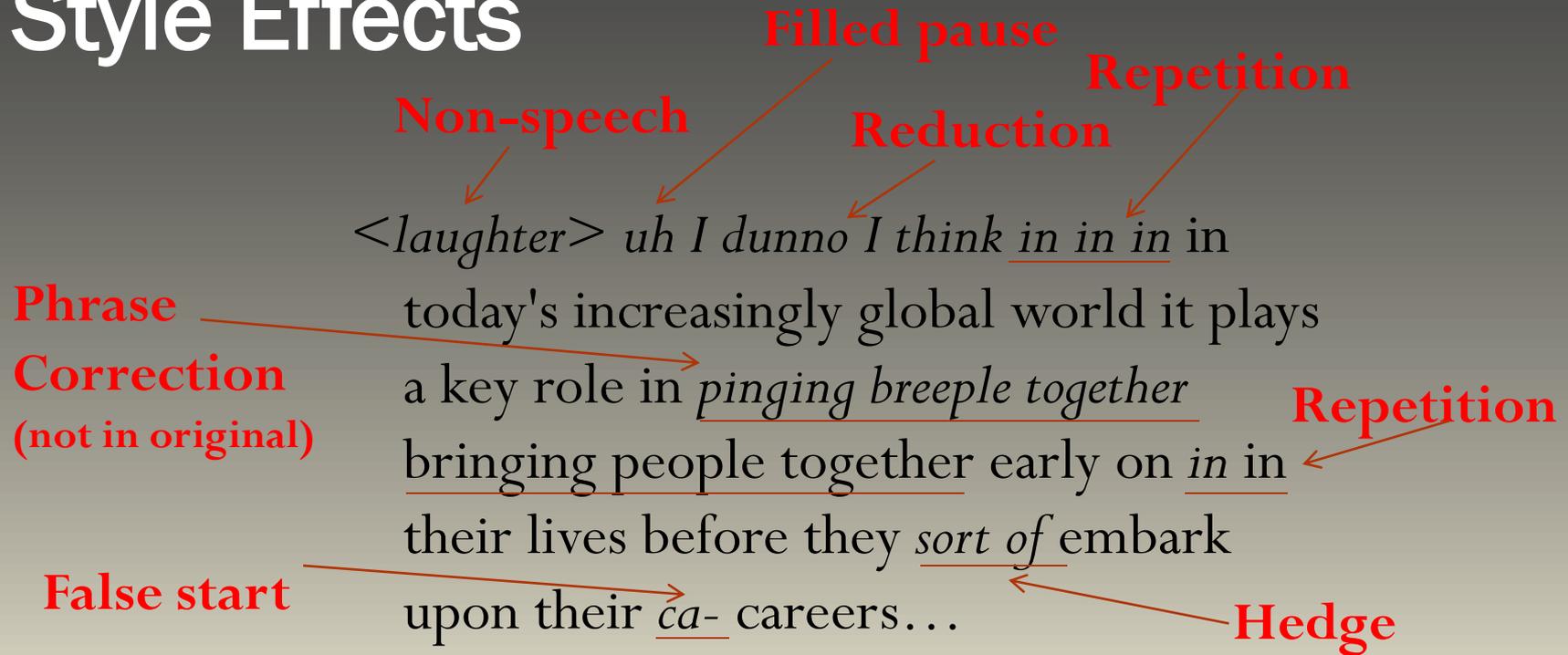
**We need to do better**

From Apple 1987 visionary video



From Saturday Night Live, April 2006

# Style Effects



In today's increasingly global world it plays a key role in bringing people together early on in their lives before they embark upon their careers.

**Overlapping speech not illustrated**

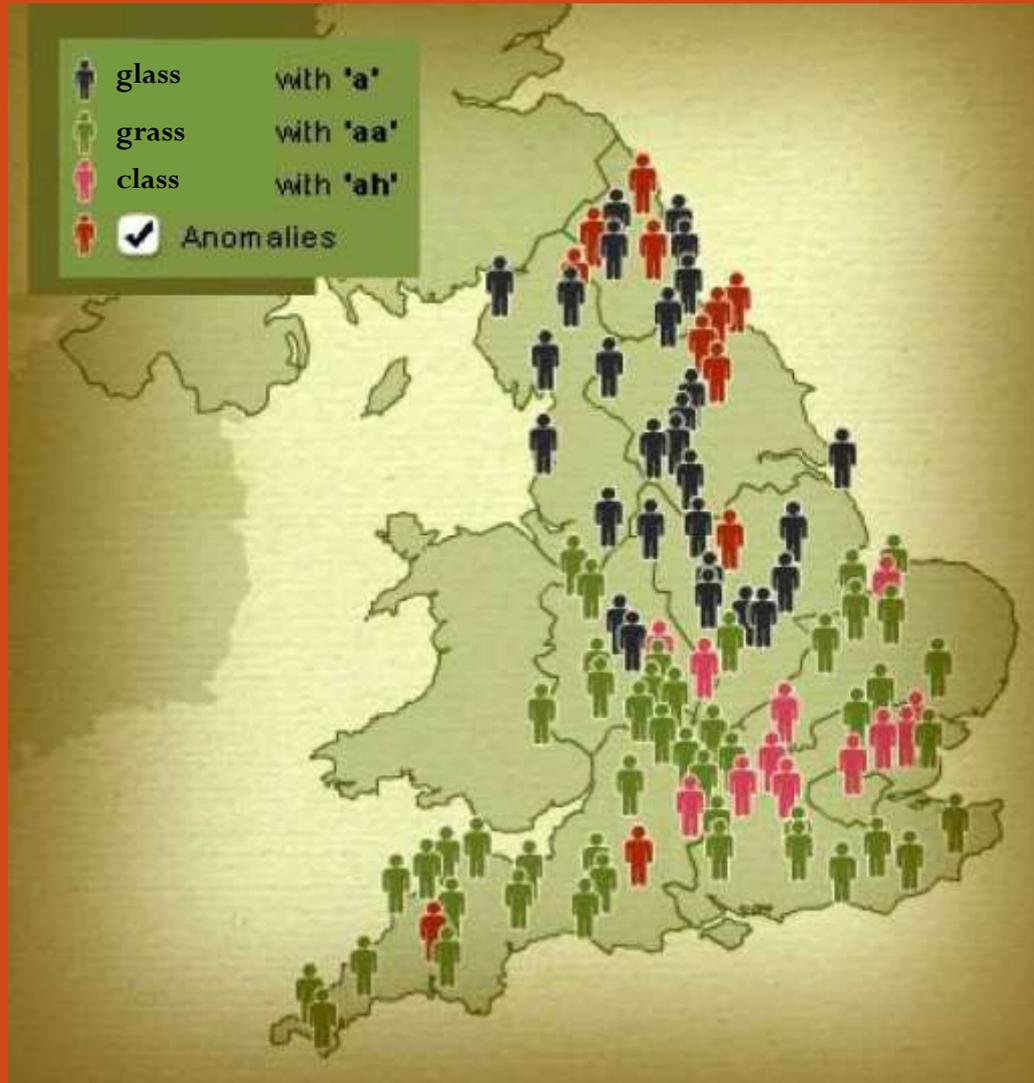
**When disfluencies are removed, spontaneous speech had same recognition error rates as read speech.**

Butzberger et al. 1992



# Pronunciation Varies with Geography

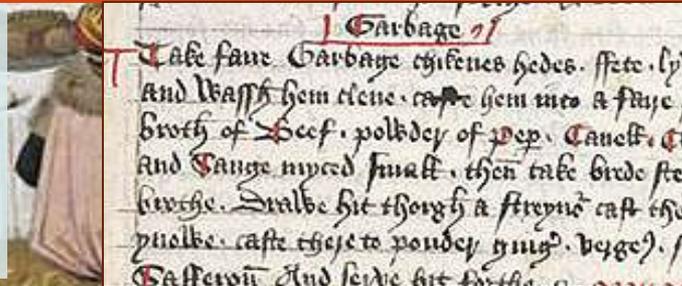
[www.bl.uk/learning/langlit/sounds/changing-voices/phonological-change/](http://www.bl.uk/learning/langlit/sounds/changing-voices/phonological-change/)



# Language Varies with Time

## Example from British Library web page:

Take faire Garbage chikenes hedes, ffete, lyvers, And gysers and wassh hem clene. Caste hem into a faire potte. And caste fressh broth of Beef, poudrer of Pep, Canell, Clowes, ...



Medieval

...where lifting first the lid off one pie, out skips some Frogs, which makes the Ladies to skip and shreek; next after the other Pie, whence comes out the Birds...

*Triumphs and Trophies in Cookery,  
to be used at Festival Times,  
as Twelfth-Day, &c.*

1600's

The breast is the only part which is looked on as fine in a turkey, the legs being very seldom cut off and eaten at table... to appear only in a form which seems to have a special attraction at a bachelor's supper-table...

**M**ake the likenes of a Ship in Paffe board, with Flags and Screemers, the Guns belonging to it of Kikkles, binde them about with packthred, and cover them with course paffe proportionable to the fashion of a ship, and place them in places convenient, as

### ROAST TURKEY.

1005. A noble dish is a turkey, roast or boiled. A Christmas dinner with the middle classes of this empire, would scarcely be a Christmas dinner without its turkey; and we can hardly imagine an object of greater envy than is presented by a respected portly pate familiar carving, at the season devoted to good cheer and genial charity, his own turkey, and carving it well. The only art consists, as in the carving of a goose,



ROAST TURKEY.

getting from the breast as many fine slices as possible; and all may have remarked the very great difference in the large number of people whom a good carver will find slices for, and the comparatively few

1800's

# Dialect Varies Demographically

- Age, sex, social class,
- Grammar, lexicon, pronunciation



RP, young, female: “I guess the first time I kind of went abroad really by myself was straight after A-Levels and I went to Paris”

*Seldom Seen:*

(Harry Belafonte

in *Kansas City*

By Robert Altman)

No subject

pronoun

1 Got to say this for you, you got guts.

Regular *is* no brains.

Habitual *be* mean nothin' 2.

Guts is 3 che *be* 4 F\*\*\*in' pig got guts.

Blue here 5 be sittin' up all night long,  
suckin' on some pig guts.

Negative  
concord

From *Peter L Patrick*,  
University of Essex

**An accent is not just something someone else has.**

**Everyone has an accent. Everyone has a dialect.**

**Some are more prestigious than others.**

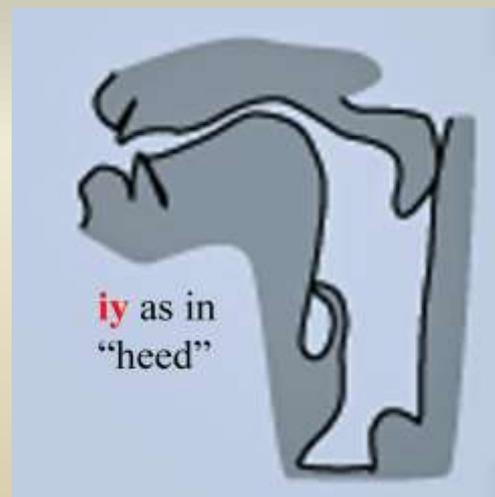
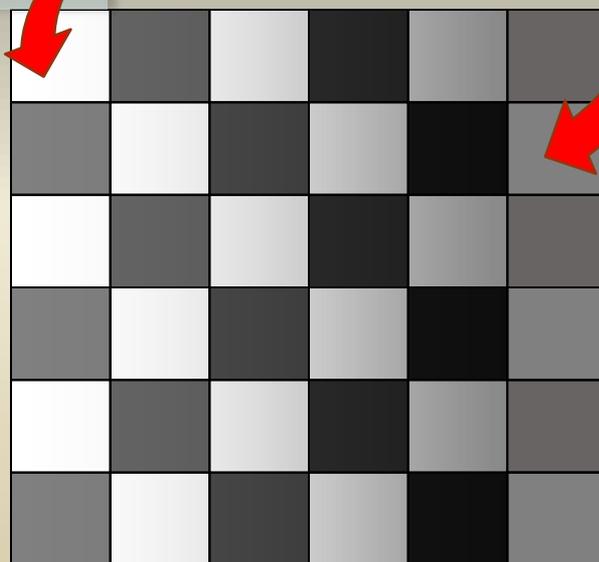
# Dialect Modeling in ASR

Dialect is a system

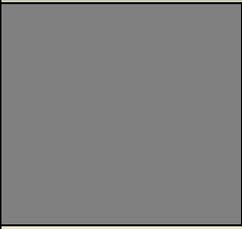
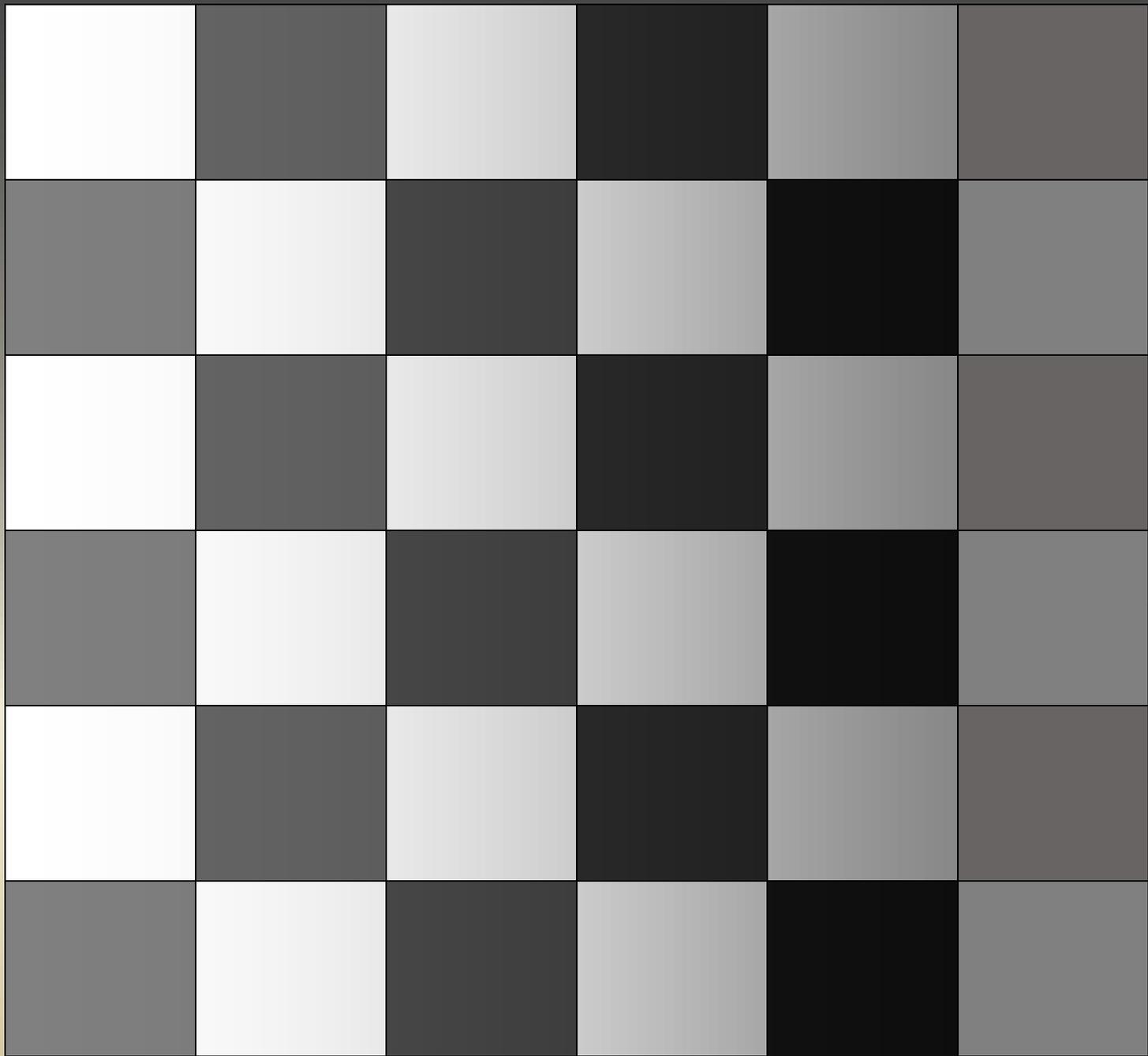
The system is not well modeled in ASR

These are  
being merged

When the point is the  
Contrast system

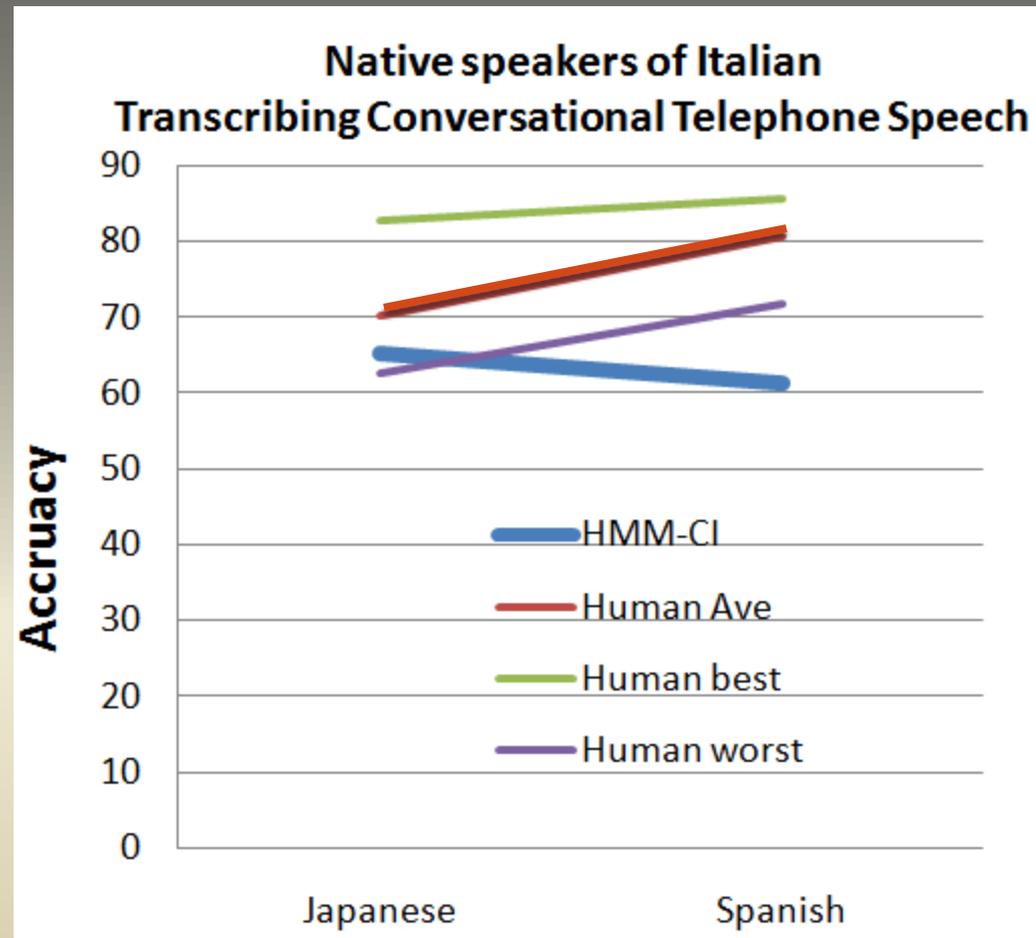


Head sections adapted from Gunnar Fant's *Acoustic Theory of Speech Production*



# Human Recognition vs. ASR

- Remove 'language model' but still use natural speech
- Phonetic inventories are similar for Italian, Japanese and Spanish
- Simple ASR is about the same as the worst of the 15 Italian transcribers
- (Spanish and Italian are close in phonotactics)



Data from Shen et al., Interspeech 2008

# ASR is sometimes better than people...

Large vocabulary tasks where people may not know the vocabulary

(e.g., thousands of names of companies in stock trading)

Small vocabulary tasks where memory plays a role

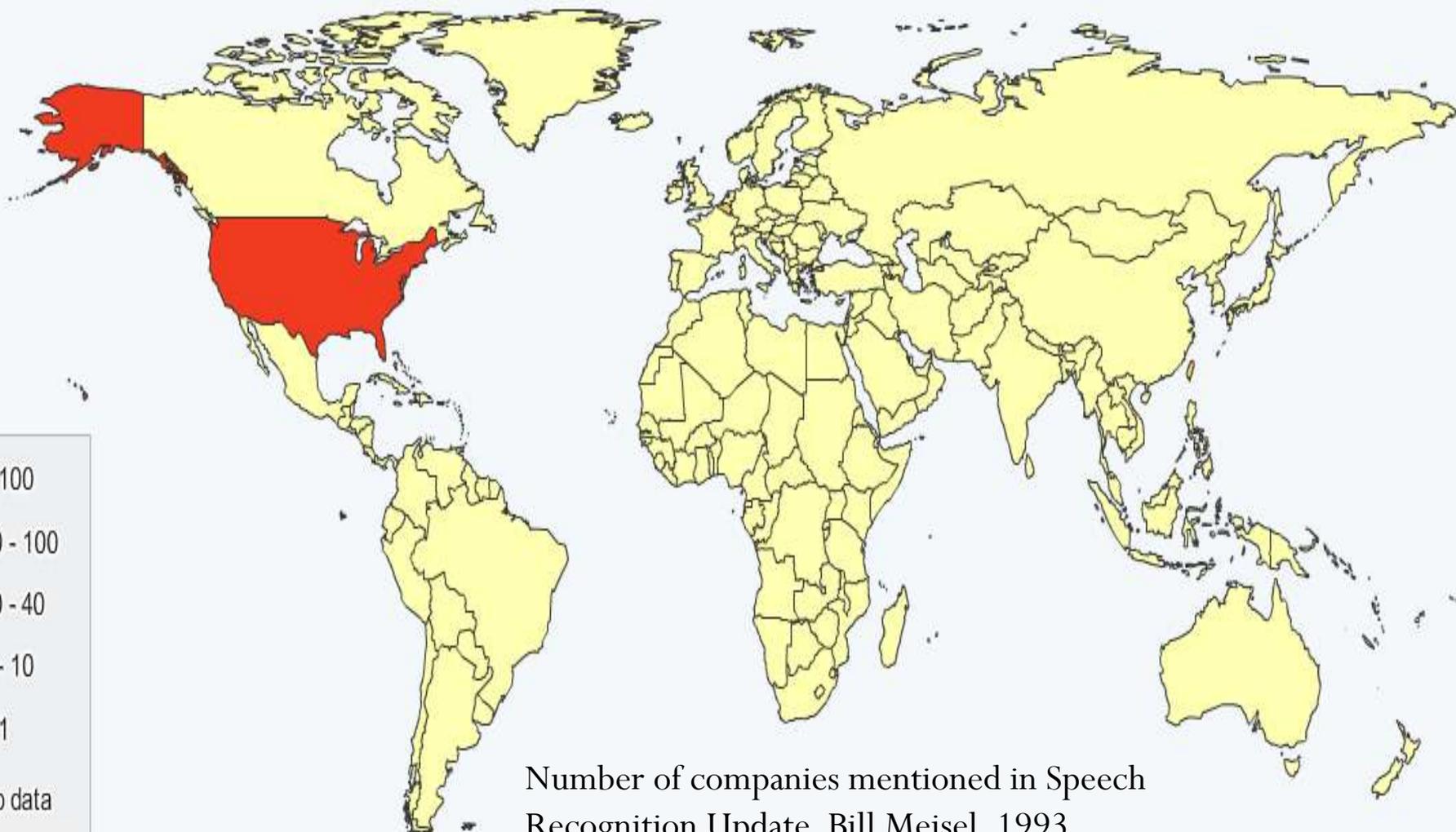
(e.g., transcribing sequence of 12 digits tracking numbers)

Artifacts of poorly designed experiments

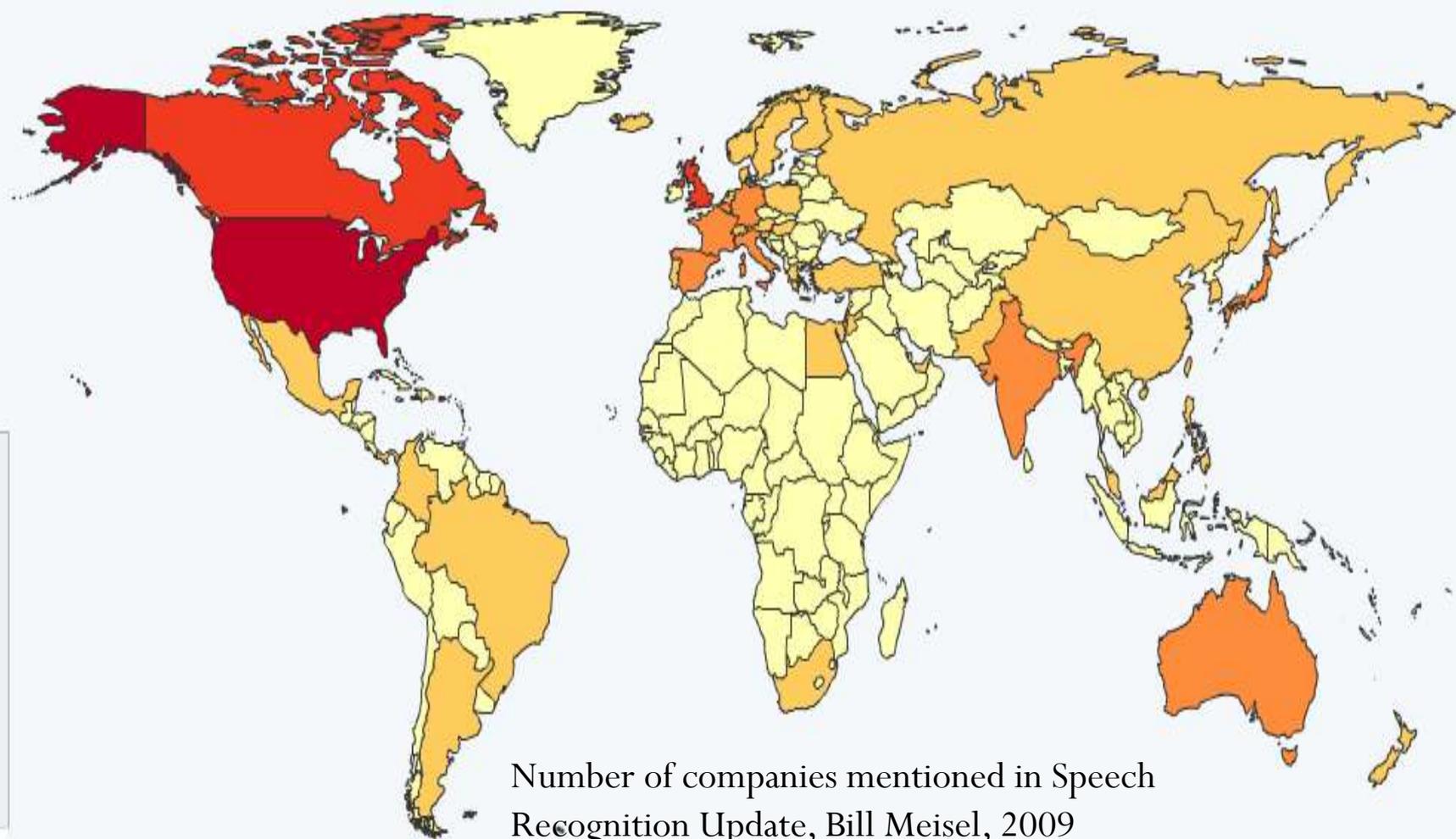
(e.g., testing on training data, correlational data that helps...)

**But generally people are more robust, flexible, adaptable... to situations that are normal human variability**

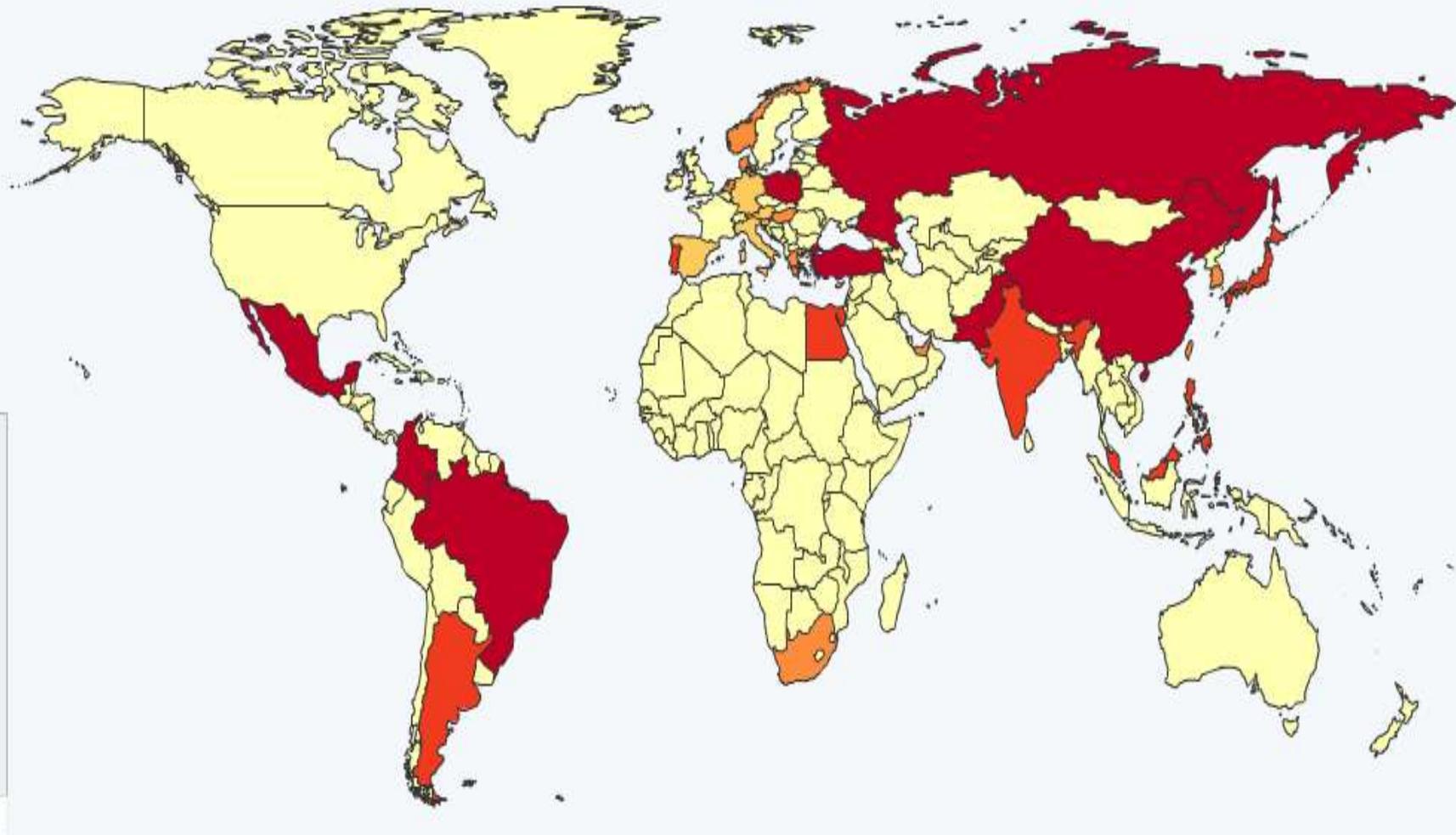
# Companies 1993



# Companies 2009



# Companies 2009, GDP/Company



# Trends

Speech as Access

What are the words?



Speech as Source

What does it mean?



Information as Partner

Here's what you need.

Can our technology to help with information overload?

## **Limitations to speech technology arise from the evolution of speech as a social construct**

- Constrained by evolutionary history, production, perception, cognition
- Balancing needs of both speaker and hearer
- At the least, an existence proof, at best, a model we can improve on
- Speech technology lacks social skills; what do we do?