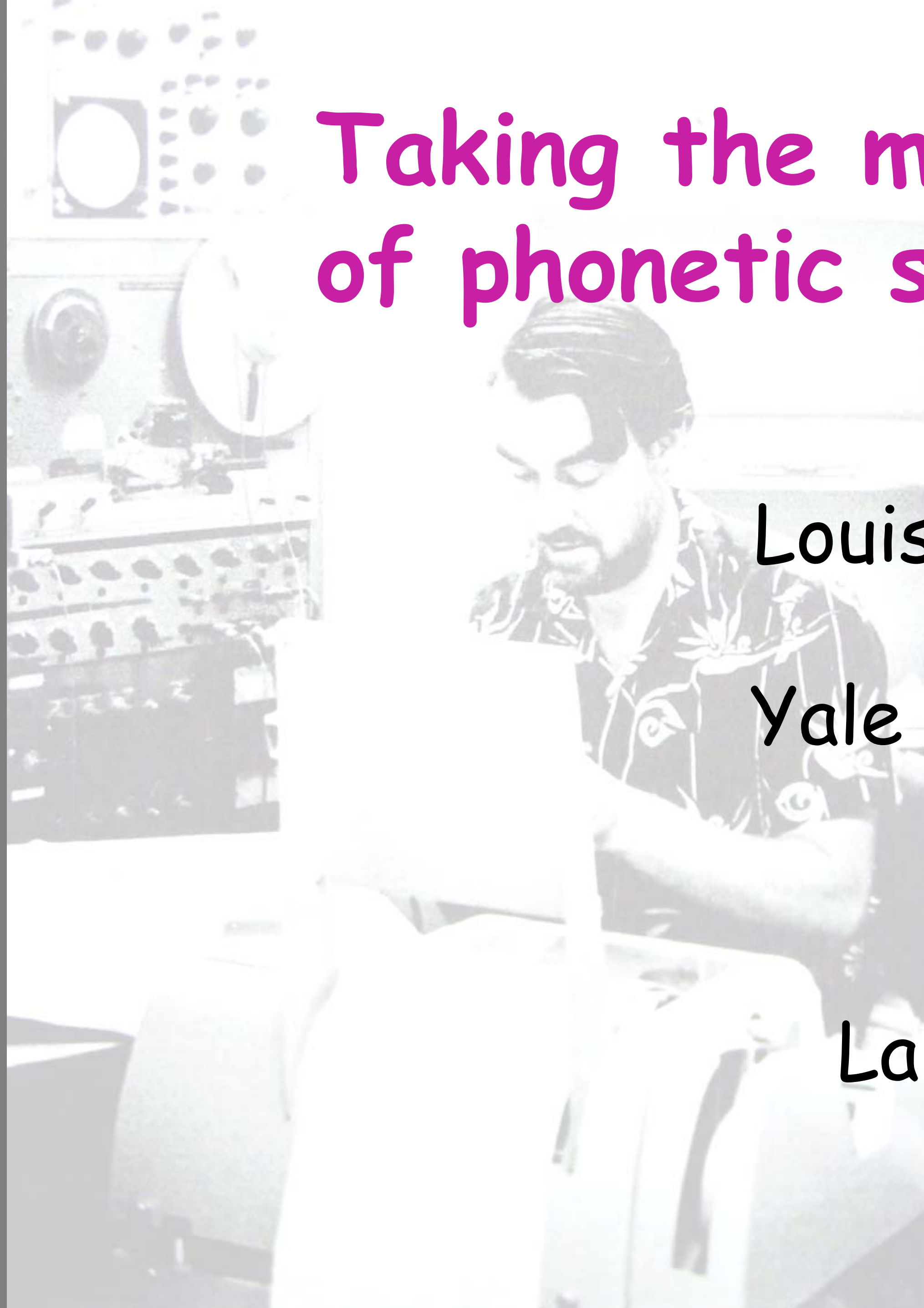


Taking the measure of phonetic structure

Louis Goldstein

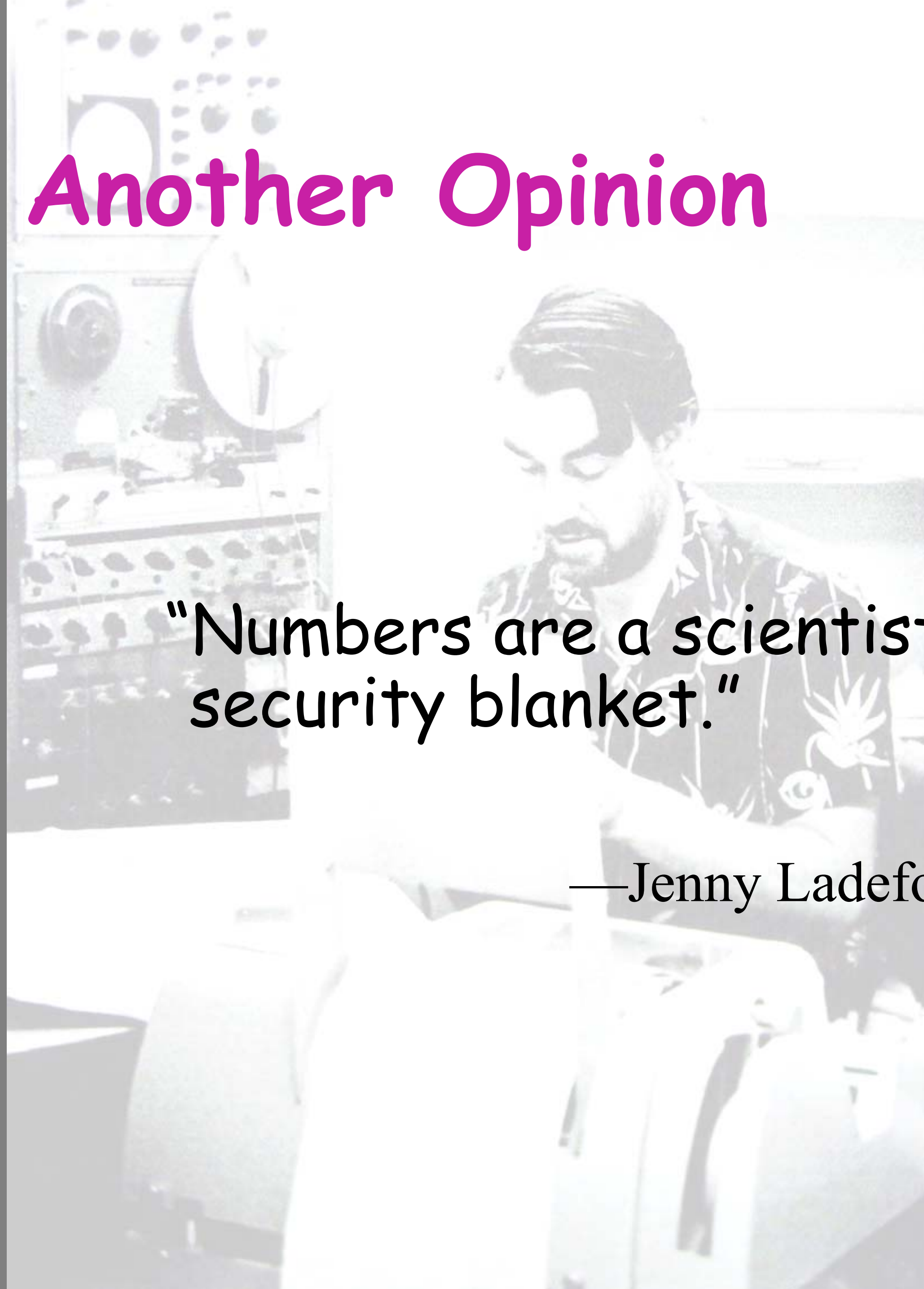
Yale University
and
Haskins
Laboratories

On Measurement:

"I often say that when you can measure what you are speaking about and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science."

—Lord Kelvin, quoted by Peter Ladefoged, ICPhS, Leeds, 1975

Another Opinion



"Numbers are a scientist's security blanket."

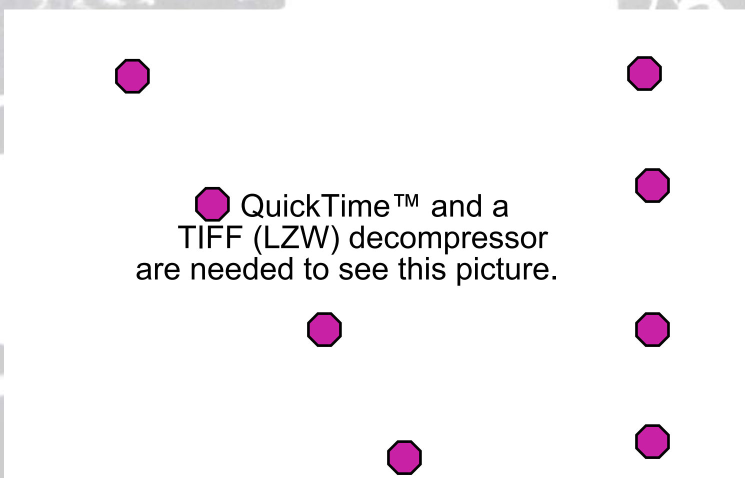
—Jenny Ladefoged

Describing the phonetic properties of languages

- They must be determined by “valid, reliable, significant” measurements.
  - measurement devices?
- This commitment has led to fundamental questions.
  - What are the appropriate reference frames within which to describe phonetic units?
  - Is there a set of universal phonetic categories?

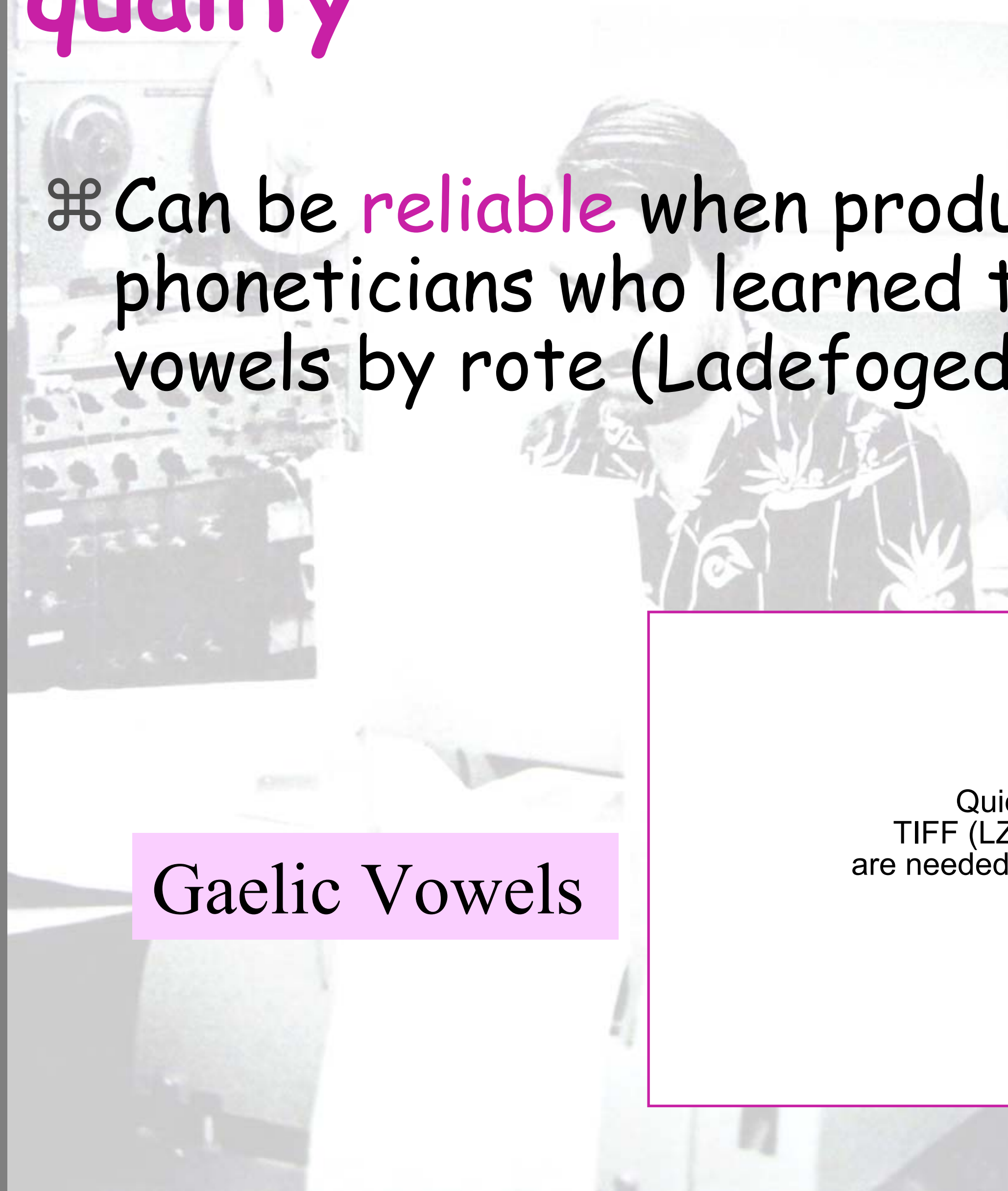
Reference frames for vowels

- Descriptions of *vowel quality* in terms of the “highest point of the tongue” are not valid.



S. Jones (1929)

Auditory judgments of vowel quality



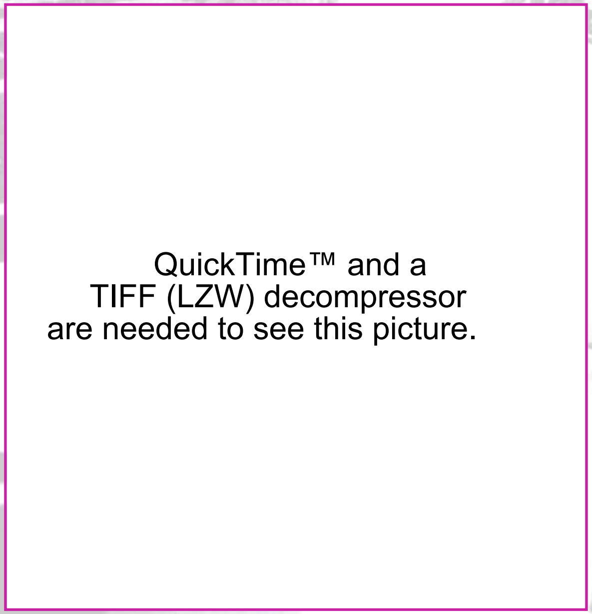
- Can be reliable when produced by phoneticians who learned the cardinal vowels by rote (Ladefoged, 1960)

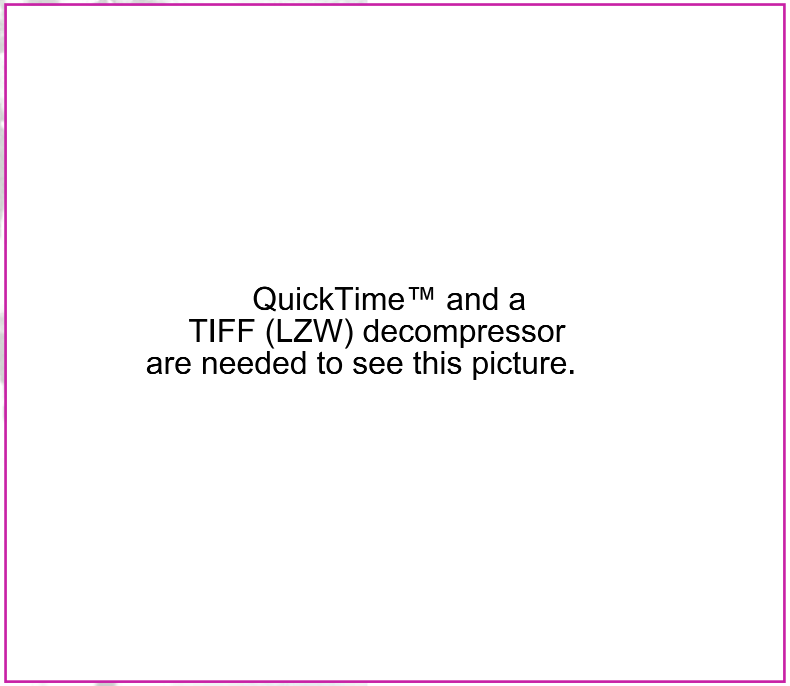
Gaelic Vowels

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

Formant frequency measurements

- Can be valid measures of vowel quality (Ladefoged, 1975)





Danish Vowels

Factor Analysis of Tongue Shapes

- Valid low-dimensional parameterization
- Compute entire tongue shape from 2 numbers (Harshman, Ladefoged & Goldstein, 1977)

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

Reference frame comparison

- Tongue factors for vowels can be computed from formant frequencies. (Ladefoged et al. 1978)
- Different reference frames for different purposes?
  - Phonetic specification of lexical items
    - Articulatory
  - Phonological patterning
    - Acoustic/Auditory
  - Speech production goals?

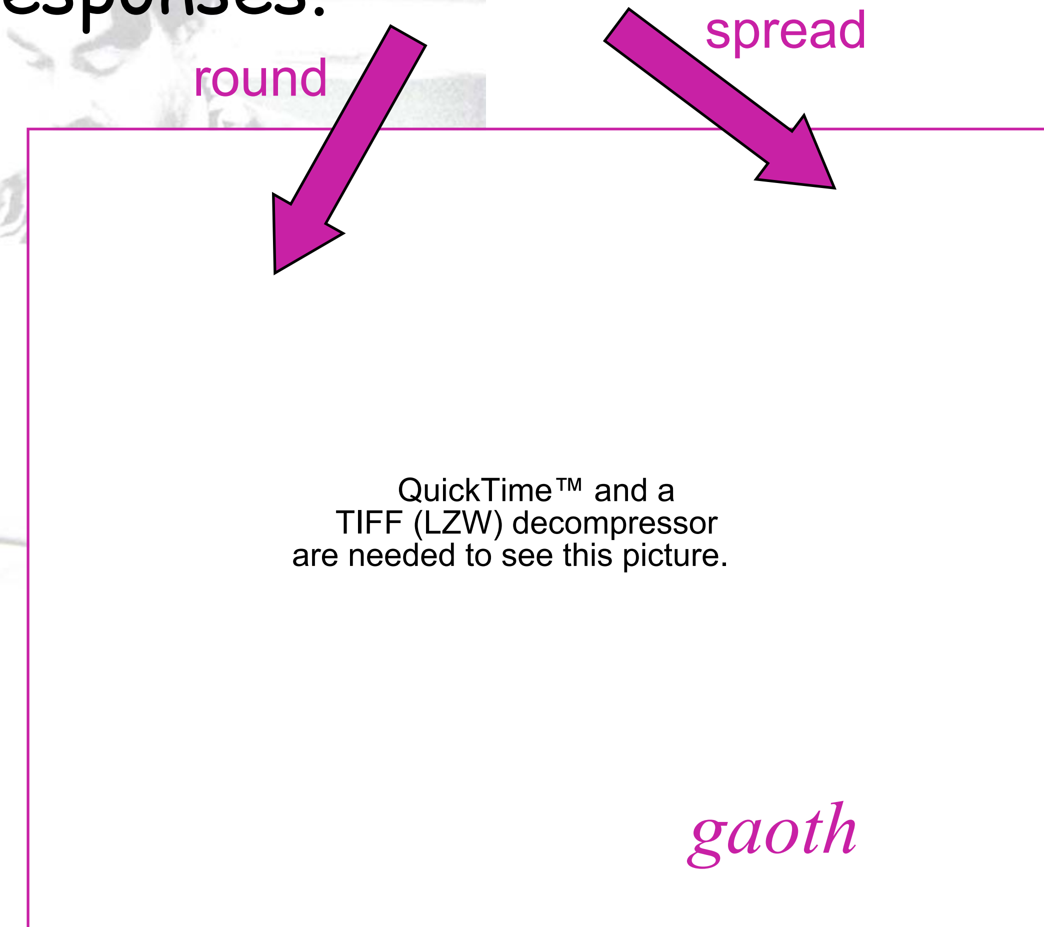
Continuing Debate: Acoustic vs. constriction goals

- Since tongue shapes and formants for vowels are inter-convertible, difficult to address for vowels.
  - Such a relation holds when the tongue produces a single constriction.
- Cross-speaker variability in tongue shapes (Johnson, Ladefoged & Lindau, 1993)
  - More variability than in auditory properties?
- Current debate about /r/
  - The relation between articulation and formants is more complex (in part because of multiple constrictions).

But wait… Ladefoged's (1960) experiment has more to say

One of the Gaelic vowels produced very inconsistent responses.

- Correlation of backness and rounding judgments
- Effect of rounding on F2



Implications for acoustic goals for vowels?

- Since front-rounded and back-unrounded vowels are so auditorily similar that skilled phoneticians confuse them, we would expect that, if goals were purely acoustic, or auditory, there would be languages in which individual speakers vary as to which of these types they produce.
- This doesn't appear to be the case.

Further Implications

- Ladefoged has argued (at various points) for a mixed specification for vowel goals:
  - Rounding is specified articulatorily;
  - Front-back, high-low are specified auditorily.
- But front-back judgments seem to be dependent on state of lips.
- McGurk experiment with phoneticians would probably have yielded different front-back judgments depending on lip display.
  - But then in what sense is front-back strictly an auditory (or acoustic) property?

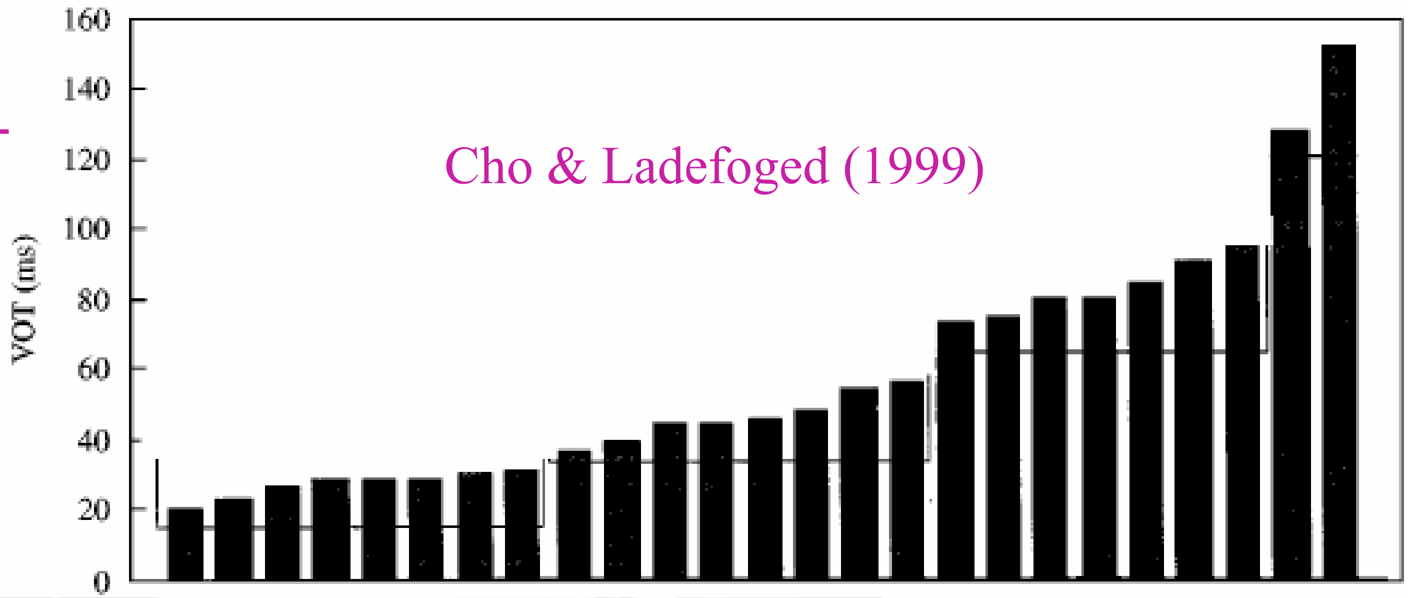
Universal phonetic categories?

- Careful measurement of segments across languages, initiated by Ladefoged, reveals more distinct types than could contrast in a single language
  - e.g. 8 types of coronal sibilants (Ladefoged, 2005)
- If phonetic categories (or features) are universal (part of universal grammar), more of them are required than are necessary for lexical contrasts and natural class specification.
- If phonetic categories are language-specific, then commonalities across languages are not formally captured.

How many distinct types?

- In some cases, it is not clear it is even possible to identify discrete potential categories.

VOT

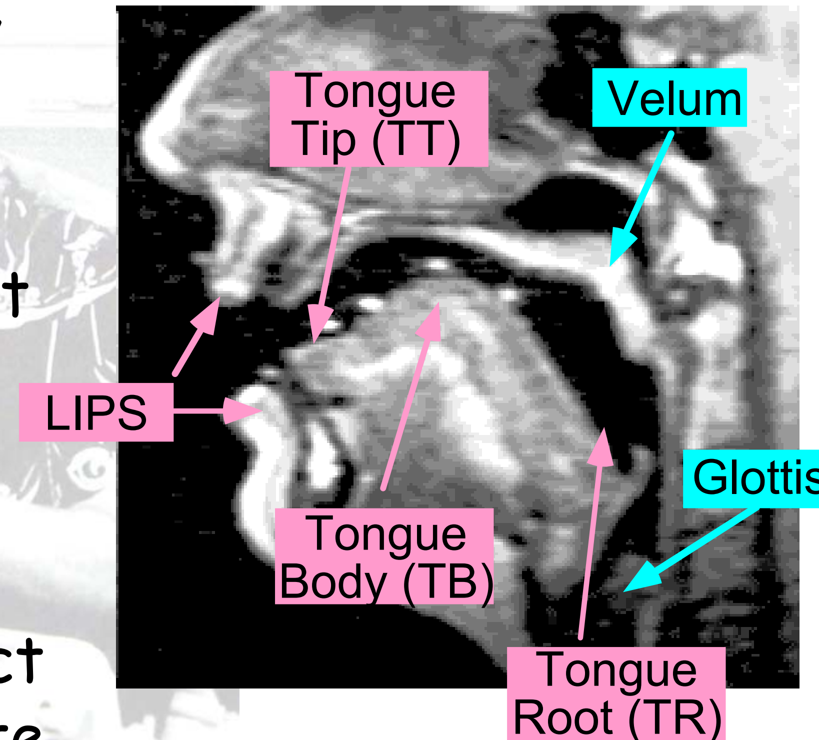


Articulatory Phonology

- Some categories are universal and others are language-specific.
- This follows from the nature of the constricting actions of the vocal tract and the sounds that they produce.
- Universal Grammar is not required to account for universal categories.

Gestures and constricting devices

- Fundamental units of phonology are gestures, vocal tract constriction actions.
- Gestures control functionally independent constricting devices, or organs.



- Constrictions of distinct organs count as discrete, potentially contrastive differences.

Universal constriction organs

- All speakers possess the same constricting organs.
- For a communication system to work, gestural actions must be shared by the members of the community (parity).
- Work on facial mimicry (Meltzoff & Moore, 1997) shows that humans can (very early) identify equivalences between the oro-facial organs of the self and others.
- Organs as the informational basis of a communication system satisfy parity.
- Use of one or another organ affords a universal category, while the actions performed are measurable and may differ from lg. to lg.

Primacy of between-organ contrasts: Adult phonology

- Of course, not all contrasting categories differ in organ employed. However...
- Between-organ contrasts are common and occur in nearly all languages. While not all within-organ contrasts are.

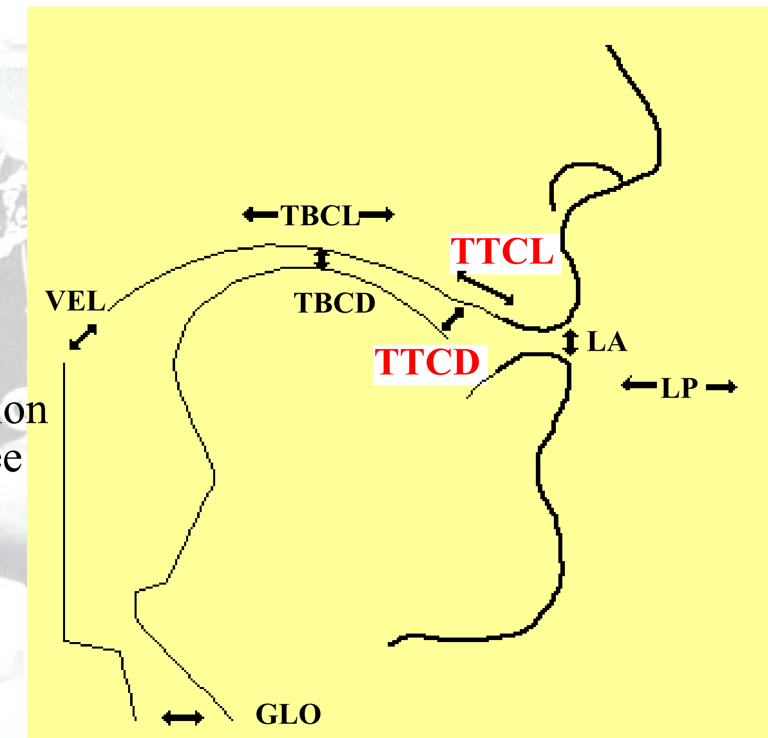
Within-organ differentiation

- Constriction gestures of a given organ can be distinguished by the degree and location of the constriction goal.

tick
sick
thick

Differ in

LP	lip protrusion
LA	lip aperture
TTCL	tongue tp constrict location
TTCD	tongue tip constrict degree
TBCL	tongue body constrict location
TBCD	tongue body constrict degree
VEL	velic aperture
GLO	glottal aperture



- These parameters are continua.
  How are they partitioned into categories?

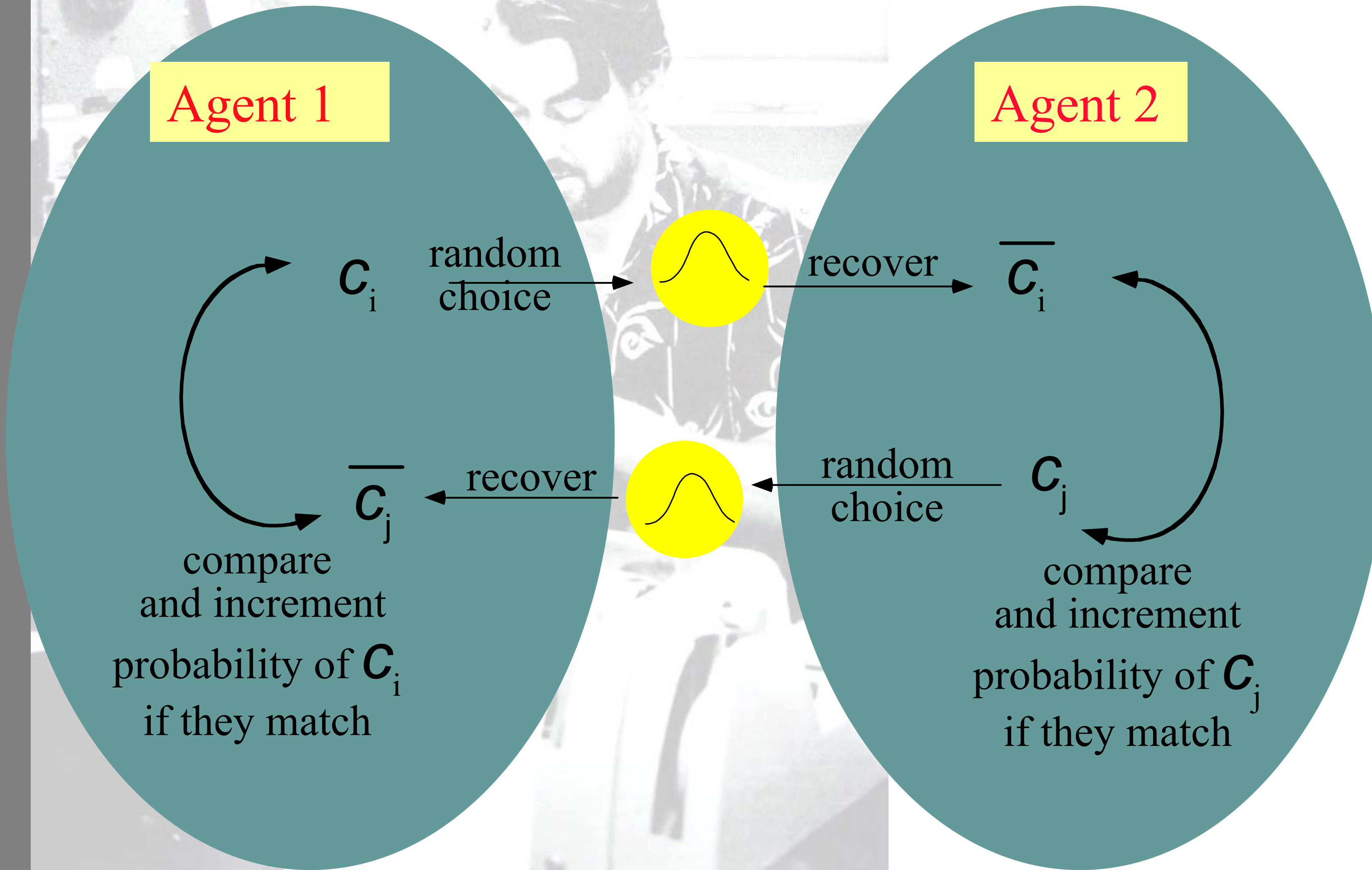
Within-organ categories

- Some within-organ categories are universal or nearly so.
  - e.g., constriction degree:
    - stop-fricative-approximant
  - Same categories are employed with multiple organs.
    - Stevens' "articulator-free" features
    - [continuant], [sonorant]
- Other within-organ categories are language-specific
  - e.g., Ladefoged's 8 phonetic categories for sibilants.

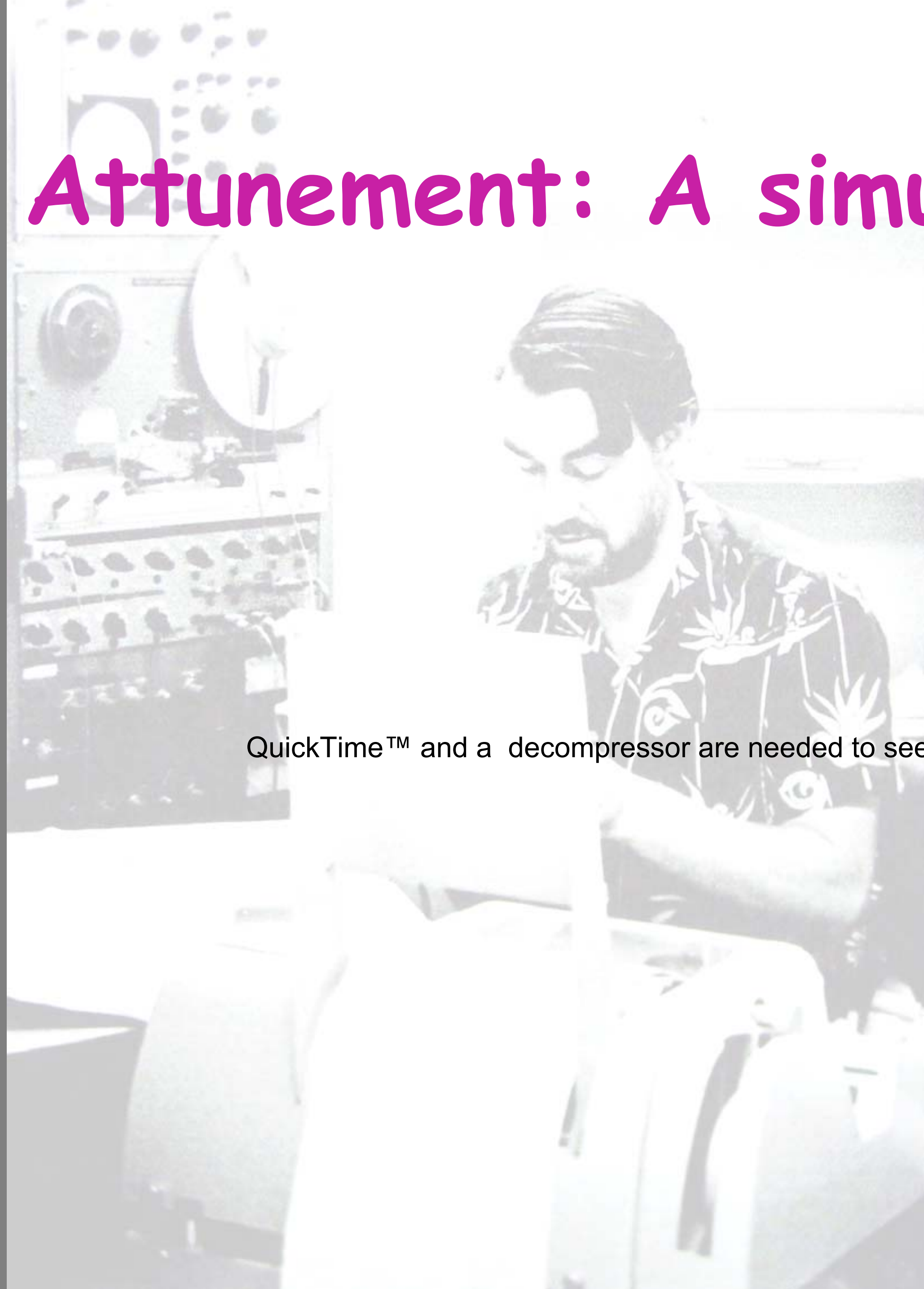
Emergence of within-organ categories through attunement

- Members of a community attune their actions to one another.
- Hypothesis: Shared narrow regions of a constriction continuum emerge as a consequence of attunement, thus satisfying parity.
  - Self-organization of phonological units
    - deBoer, 2000
    - Oudeyer, 2002
    - Goldstein, 2003

Simulation of attunement with agents



Attunement: A simulation



Agent 1

QuickTime™ and a decompressor are needed to see this picture.

Agent 2

Attunement & multiple modes

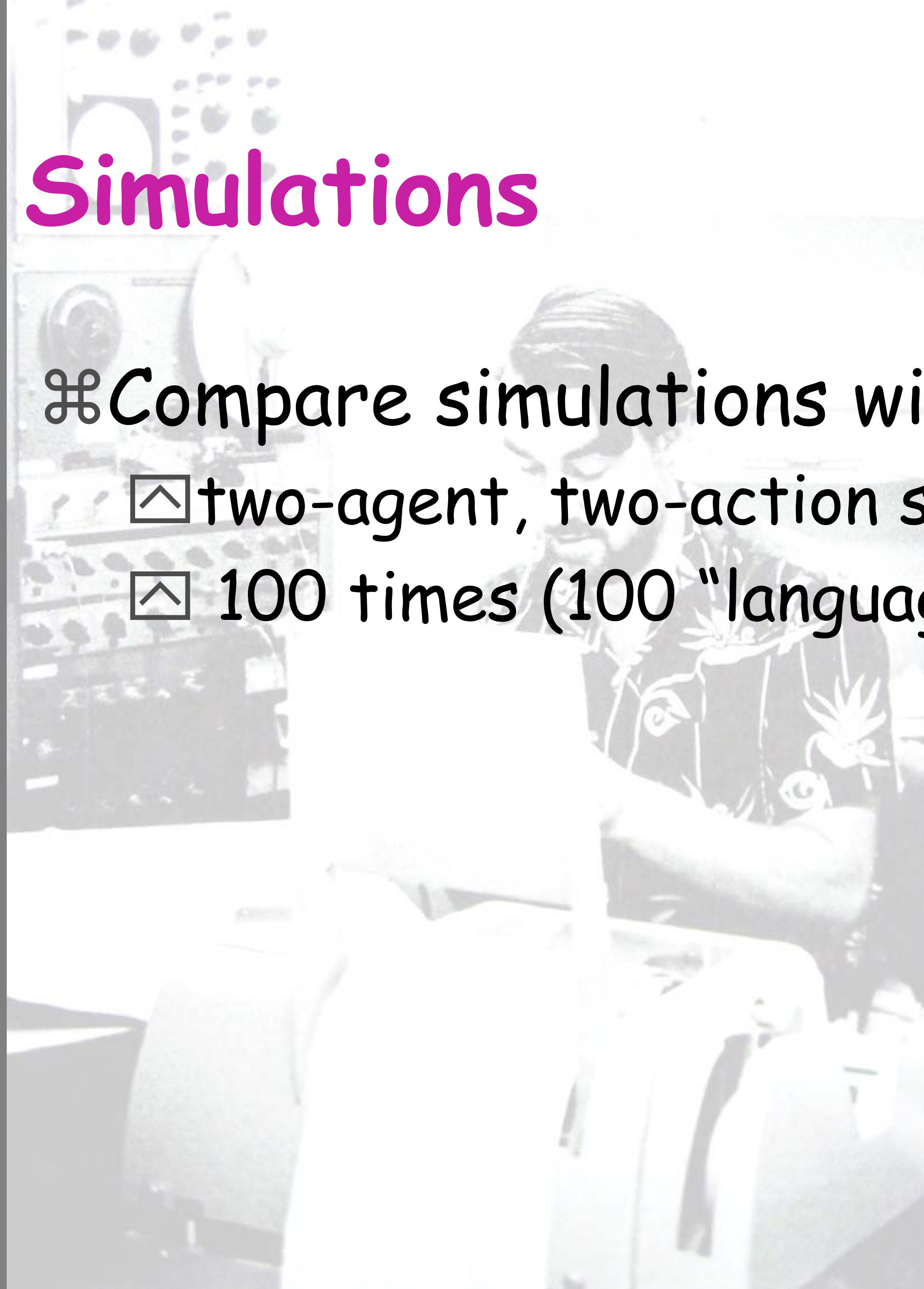
- Attunement produces convergence to a narrow range (shared by both agents).
- Multiple modes along the continuum (potentially contrasting values) can emerge in a similar fashion.
- Are the modes consistent across repeated simulations ("languages")?
  - Answer depends on the mapping from constriction parameter to acoustics.
  - Agents must recover constriction parameters from acoustics.

Constriction-acoustics maps

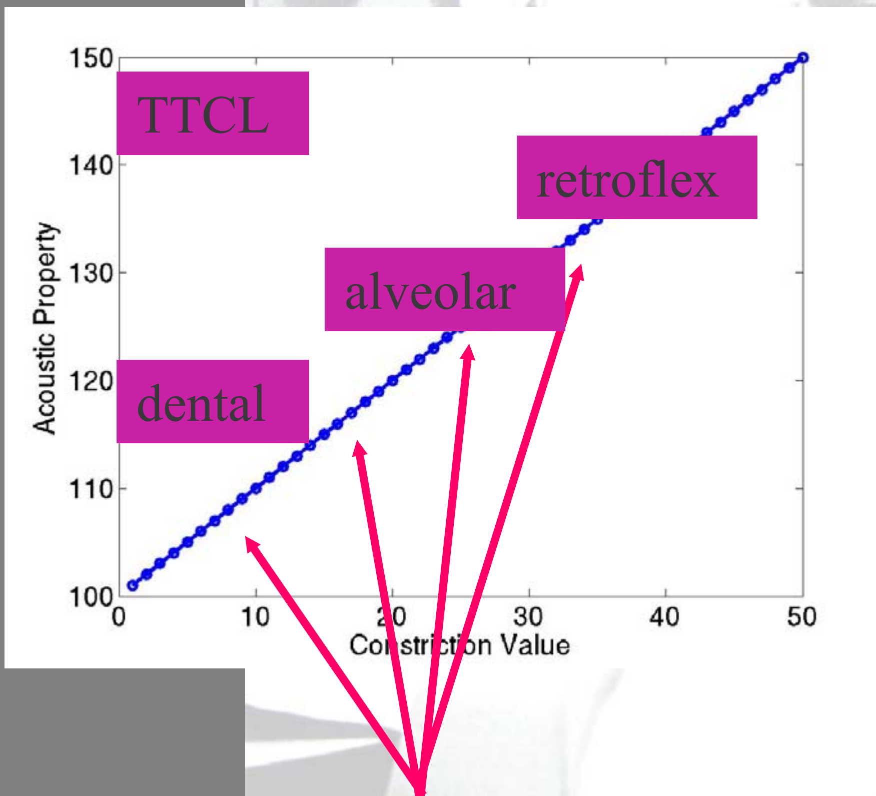
- Nature of mapping from constriction parameter to acoustics affects the consistency of modes obtained in simulation.
- Nonlinear Map (e.g. Stevens, 1989)
  - stable and unstable regions
  - Agents partition relatively consistently.
  - possible Model of Constriction Degree (e.g., TTCD)
- Linear Map
  - more variability in partitioning
  - possible Model of Constriction Location (e.g., TTCL) coronal sibilants

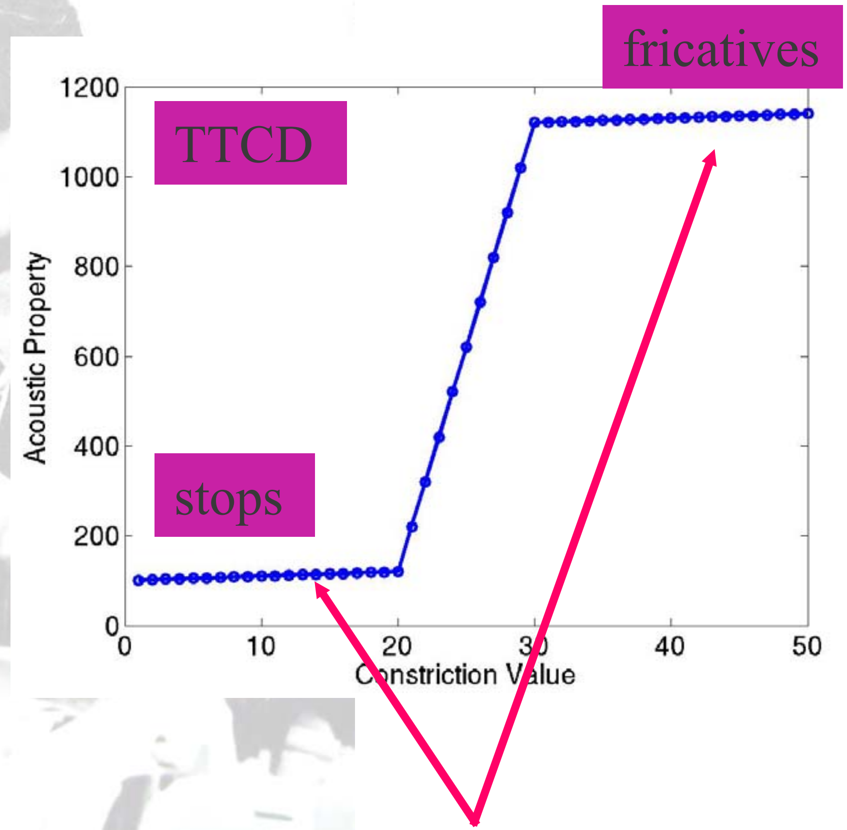
Simulations

- Compare simulations with these maps
  - two-agent, two-action simulations
  - 100 times (100 “languages”)



Results





languages' contrasting actions distributed over entire range

77% of languages contrast actions that span discontinuity

Organ hypothesis: phonological development

- Between-organ differences
  - Since neonates can already match organ selection with that of a model, we expect children's early words to match adult forms in organ employed.
- Within-organ differences
  - Since these require attunement and therefore specific experience, we expect that children's early words will not match the adult forms.

Experiment: children's early words (Goldstein 2003)

- Materials
  - Recordings of children's words by Bernstein-Ratner (1984) from CHILDES database
  - Data from 6 children (age range 1:1 - 1:9).
- Words with known adult targets were played to judges who classified initial consonants as English consonants.
- Based on judges' responses, child forms were compared to adult forms in organs employed and within-organ parameter values (CD).

Results

- Oral constriction organ (Lips, TT, TB)
  - For all 6 children, organ in child's production matched the adult target with > chance frequency.
- Glottis and Velum
  - Some children show significant matching with adult targets, some do not.
- Constriction Degree (stop, fricative, glide)
  - No children showed matching with > chance frequency.

Evidence from infant speech perception

- Young infants
  - may not be able to distinguish all adult within-organ categories
  - English /da/-/Da/ (Polka, 2001)
- Older infants
  - Classic decline in perception of non-native contrasts decline around 10 months of age involve within-organ contrasts
    - retroflex - dental
    - velar - uvular
  - Between-organ contrasts may not decline in the same way (Best & McRoberts, 2004).

The measure of Peter's contribution to phonetics

- Not just the vast amount of knowledge he created or inspired
- But also what he taught the field of linguistic phonetics about rigor.
  - measurement of data
  - modeling: measurable (testable) consequences of representational hypotheses