On the Importance of Standardized Word List Data in Dialectology Research

Douglas S. Bigham

University of Texas at Austin

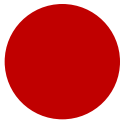
douglas.s.bigham@gmail.com

Methods In Dialectology XIII – Leeds, United Kingdom

4th August, 2008

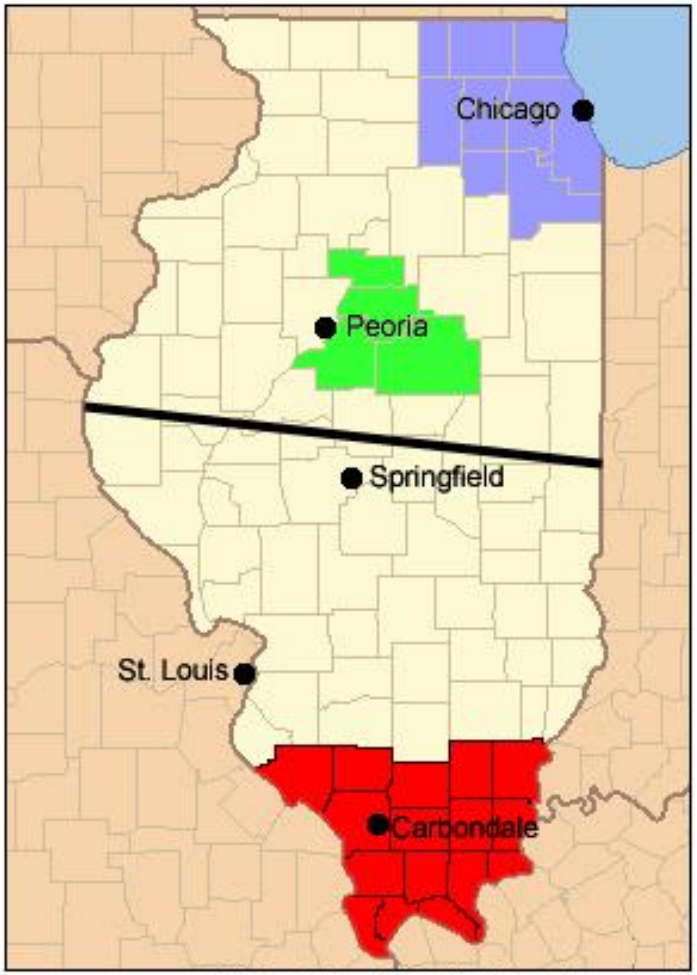
Overview

- Dialect Contact
  - What happens when speakers of two different dialects interact with each other?
  - Trudgill, 1986; 2004
- University Students
  - Southern Illinois University – Carbondale (SIUC)
  - Close, persistent, intimate contact
  - Transient, dynamic, “anchored” populations
- Emerging Adulthood
  - Period between High School and a Career
  - Roughly age 18-26
  - Marked by exploration, self-discovery, and transience
  - J. Arnett, 2001
- Vowels of Interest:
  - LOT, THOUGHT, and TRAP



Dialect Geography for this Study

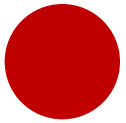
The position of Illinois in the United States



Dialect Regions in Illinois

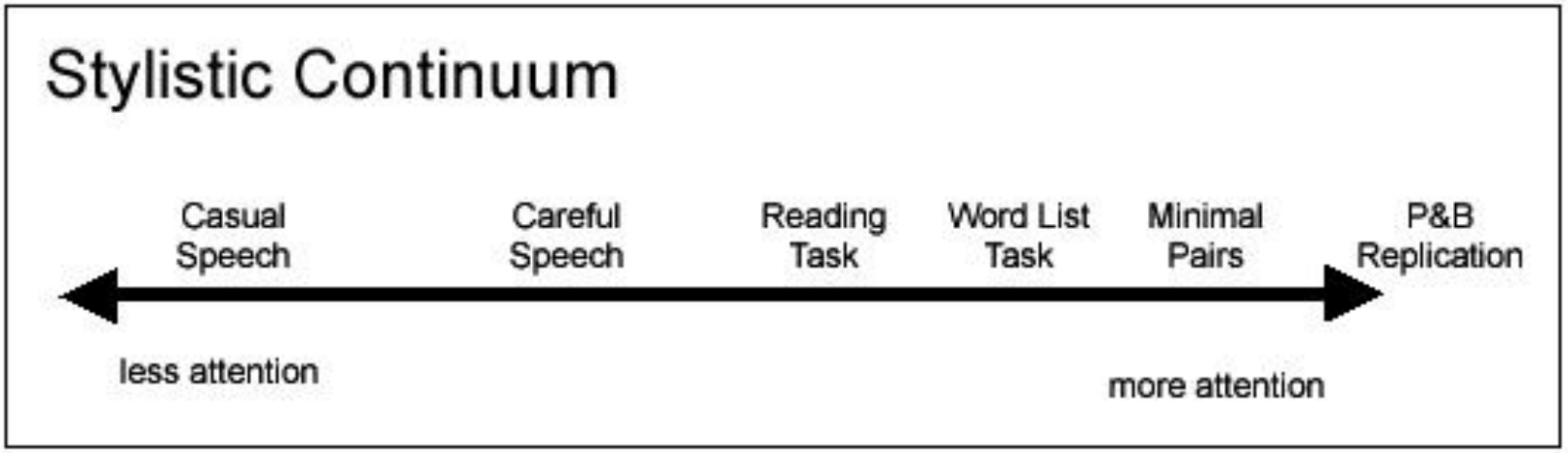
Data to be Presented

- LOT, THOUGHT, and TRAP vowels
  - Wells' Key Words (Wells, 1982)
- Eight speakers sampled (from 126 speaker corpus)
  - 4 Chicagoland / 4 Southern Illinois
- "Natural" speech data
  - Tokens taken from casual interviews
  - Interviews averaged 1.5 hours per speaker
- "Word List" recitation data
  - Peterson & Barney (1952) repetition (*P&B-task*)
  - 11 vowels of American English in *h_d* & *b_t* contexts
  - Each reading repeated 5 times



The "Stylistic Continuum" and Problems with "Natural" Speech

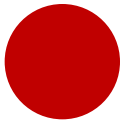
- The Stylistic Continuum
  - Capturing non-self-aware speech



- "Style" and "Audience Design" (Bell, 1984)
- "Awareness" and "Formality" (Schilling-Estes, 1998)
- Interviews as situated speech (Fuller, 2000)
- Variation as "Indexical" (Eckert, 2000)

Virtues of P&B-type Data

- P&B-type data support interview-based data
  - Hillenbrand, et al. (1995) – “NCS” features in Michigan
  - Hagiwara (1997) – “Californian” features in California
  - Clopper, et al. (2005) – regional features confirmed for multiple American English dialects
- Practical Considerations
  - Easier to collect
  - Easier to analyze
  - Provide a uniform comparison against a known “benchmark”
  - Provide uniform comparability across tokens & speakers



Comparability of Data: Overall Token Frequency

- Comparable token counts
  - Often lacking in “interview speech”
  - e.g., FOOT is especially uncommon in my data

	LOT	TRAP	DRESS	THOUGHT	GOAT	GOOSE	FOOT	STRUT
Interview	17%	16%	14%	19%	7%	12%	9%	6%
WordList	12%	14%	14%	12%	12%	13%	13%	10%

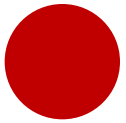
- Word List Data, while not perfect, show a much greater comparability across tokens
- This problem is exacerbated when consonantal contexts is considered

Comparability of Speakers' Data: Following Consonantal Context

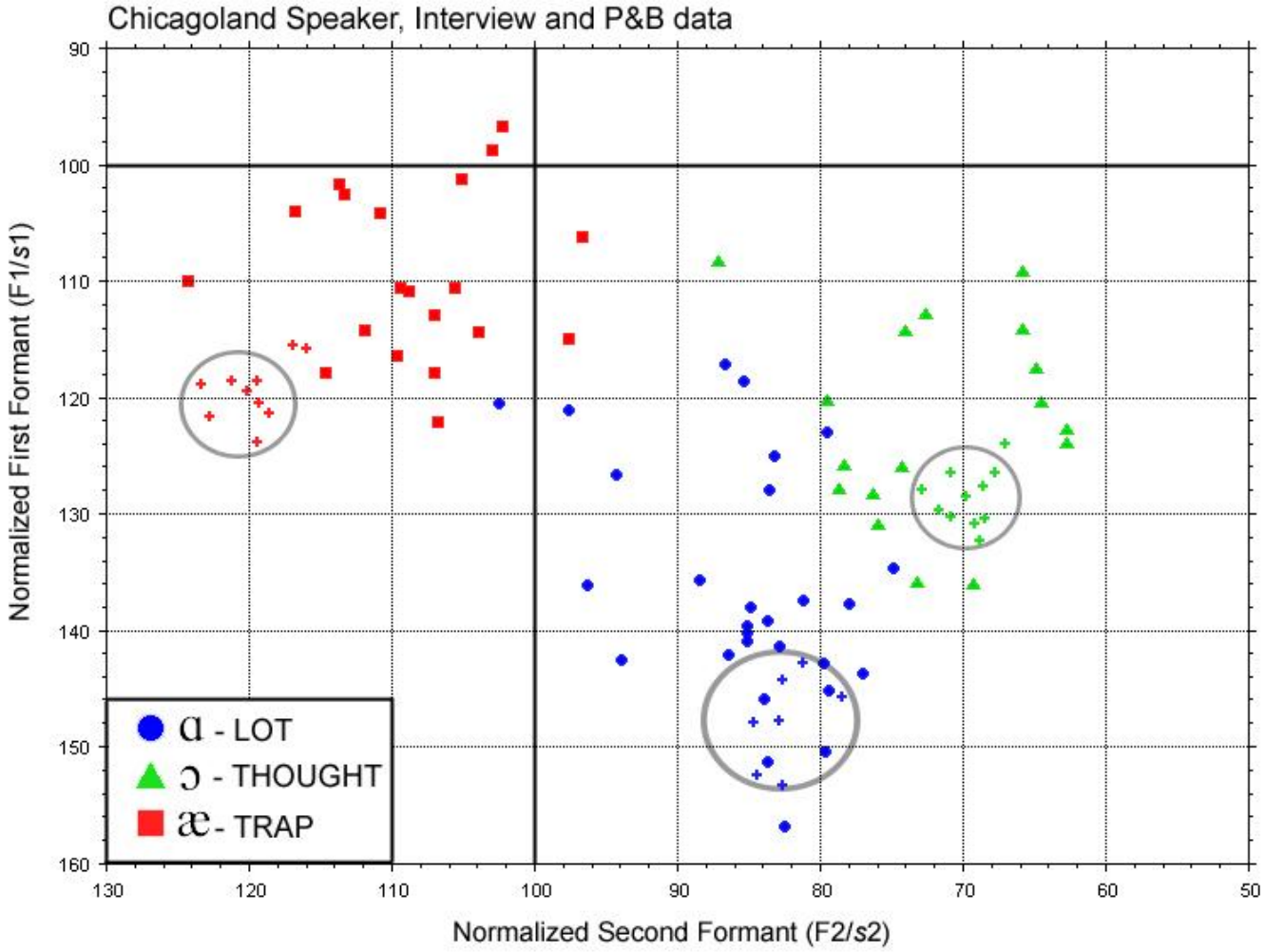
	LOT	TRAP	DRESS	THOUGHT	GOAT	GOOSE	FOOT	STRUT
TOTAL FREQUENCY	**17%**	**16%**	**14%**	**19%**	**7%**	**12%**	**9%**	**6%**
LABIAL	40%	8%	15%	26%	12%	30%		40%
ALVEOLAR	36%	71%	79%	57%	35%	43%	76%	53%
VELAR	24%	21%	6%	11%	6%		24%	7%
WORD FINAL				6%	47%	27%		
VOICED	64%	39%	21%	52%	44%	91%	57%	47%
VOICELESS	36%	61%	79%	48%	63%	9%	43%	53%
PLOSIVE	67%	45%	64%	15%	29%	20%	95%	40%
FRICATIVE	2%	34%	36%	32%	24%	13%		33%
LIQUID/GLIDE/ APPROXIMANT	17%			23%		23%	5%	
NASAL	14%	21%		23%		17%		
AFFRICATE								27%

Phonemic vs. Phonetic Findings

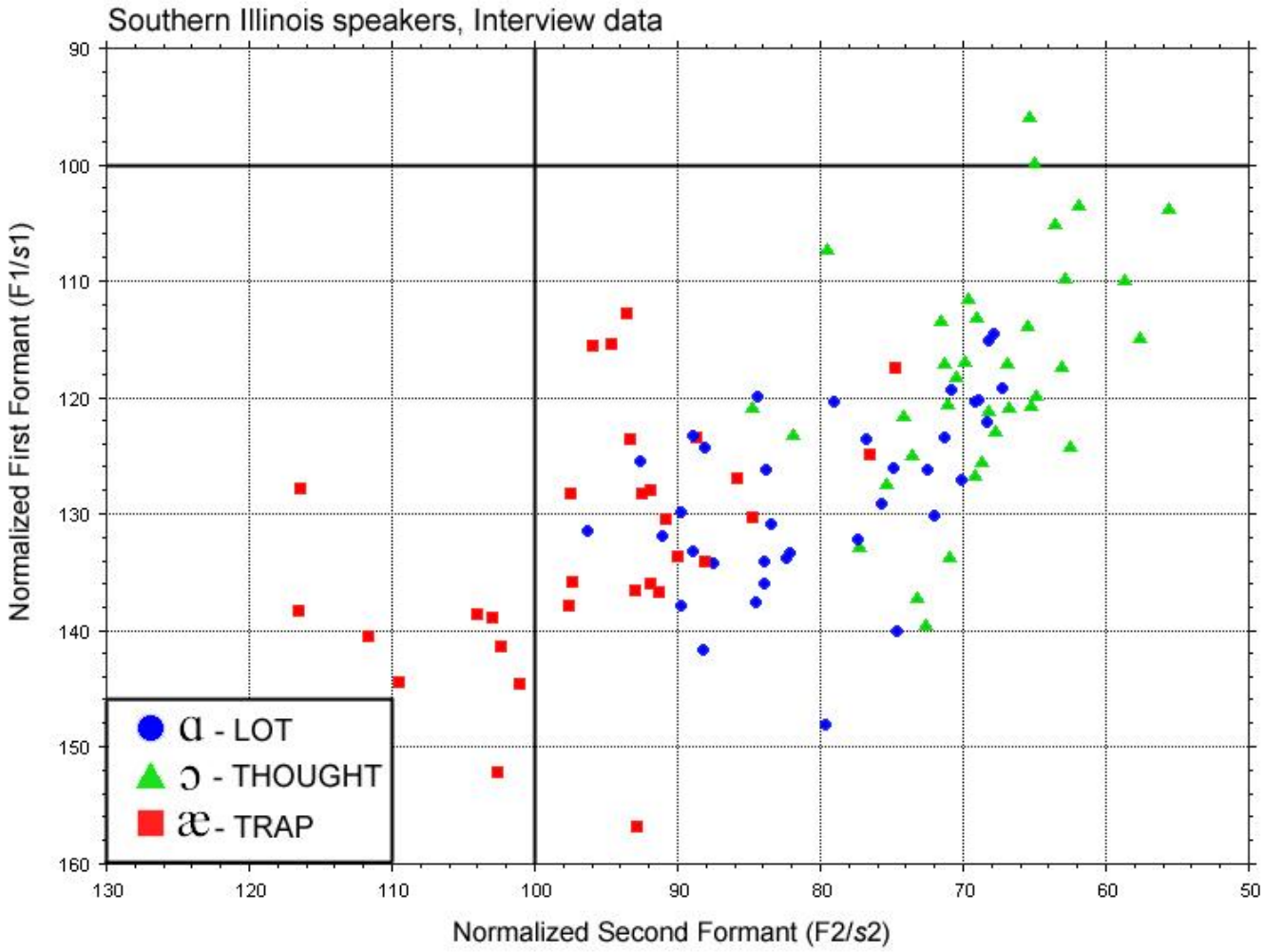
- LOT, THOUGHT, the Low-Back Vowel Merger
  - LOT~THOUGHT merged for most U.S. Midlands and Western speakers
  - Unmerged among Inland North and Core South speakers
  - Areas near the Midlands east of the Mississippi River show a variety of transitional and near-merged forms
  - The low-back vowel merger is a progressive merger... younger speakers are more likely to merge than older speakers in the same region



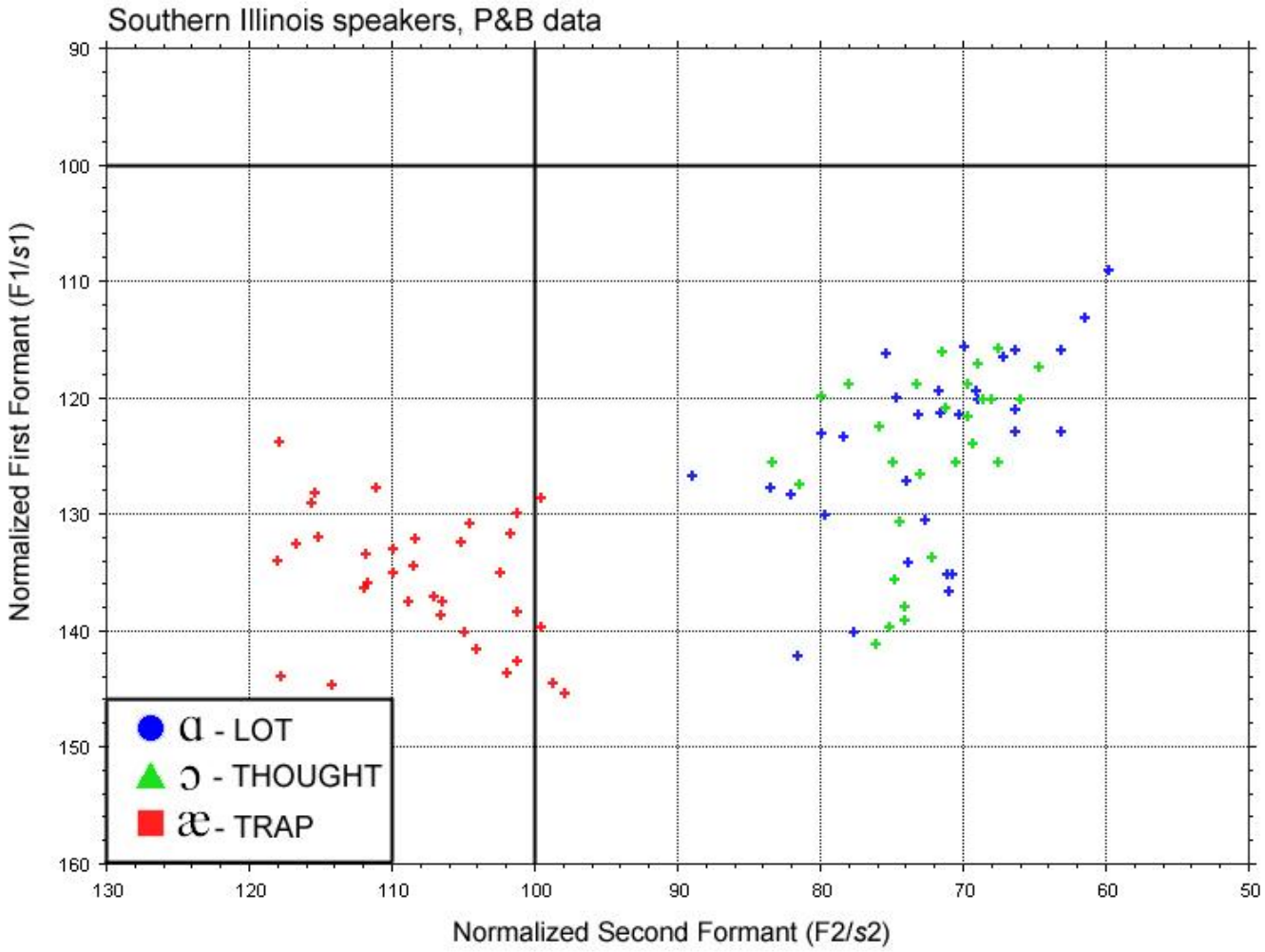
LOT, THOUGHT, & TRAP: UNMERGED CHICAGOLAND SPEAKER



LOT, THOUGHT, & TRAP (INTERVIEW DATA): MERGED SOUTHERN ILLINOIS SPEAKERS



LOT, THOUGHT, & TRAP (P&B DATA): MERGED SOUTHERN ILLINOIS SPEAKERS

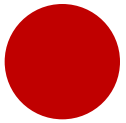


LOT, THOUGHT, & TRAP: MERGED OR NOT?

- Cartesian Distances
  - Baranowski, 2006
- Problem
  - Where is the merger?

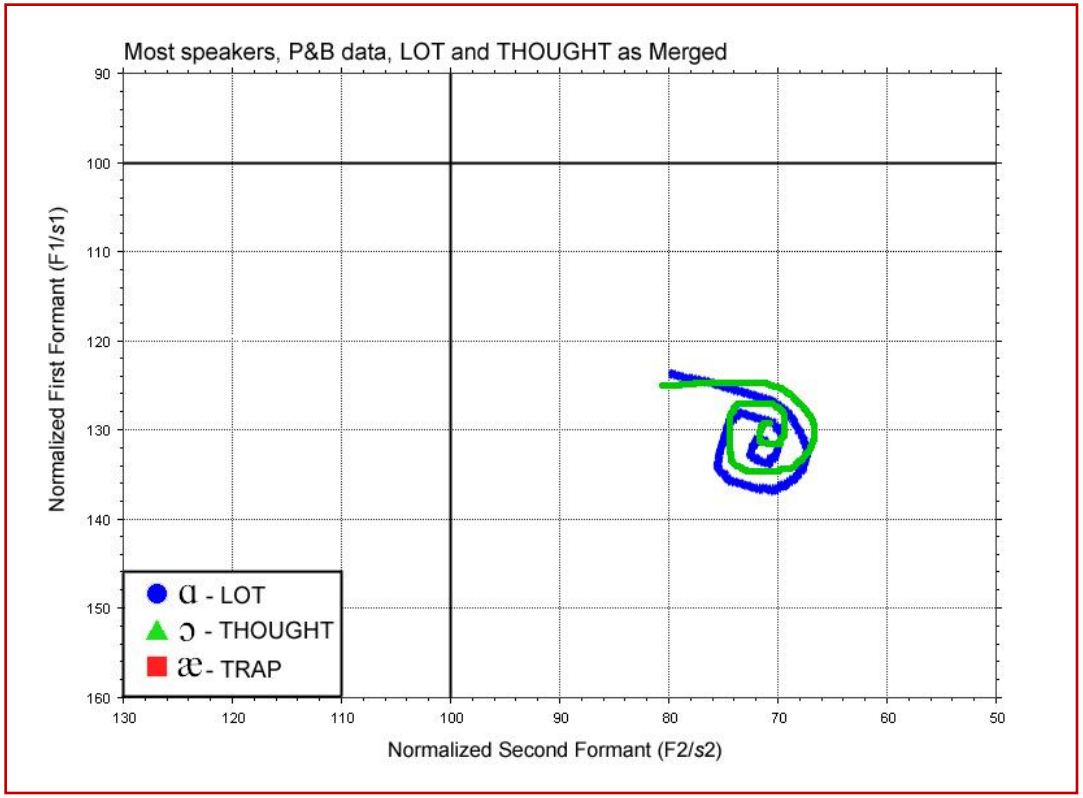
	Distance in Normalized "Units"		
	LOT ~ THOUGHT	LOT ~ TRAP	TRAP ~ THOUGHT
Females, Interview Data	**23**	**26**	**49**
Females, P&B Data	**4**	**48**	**44**
Males, Interview Data	**24**	**15**	**35**
Males, P&B Data	**1**	**40**	**41**

- Target Undershoot (Lindblom, 1963)
  - Rapid speech produced more centralized vowels
  - Interview speech is more rapid than word list reading
- Word Lists = More Attentive Speech
  - "More attentive" speech more accurately reflects phonological categories

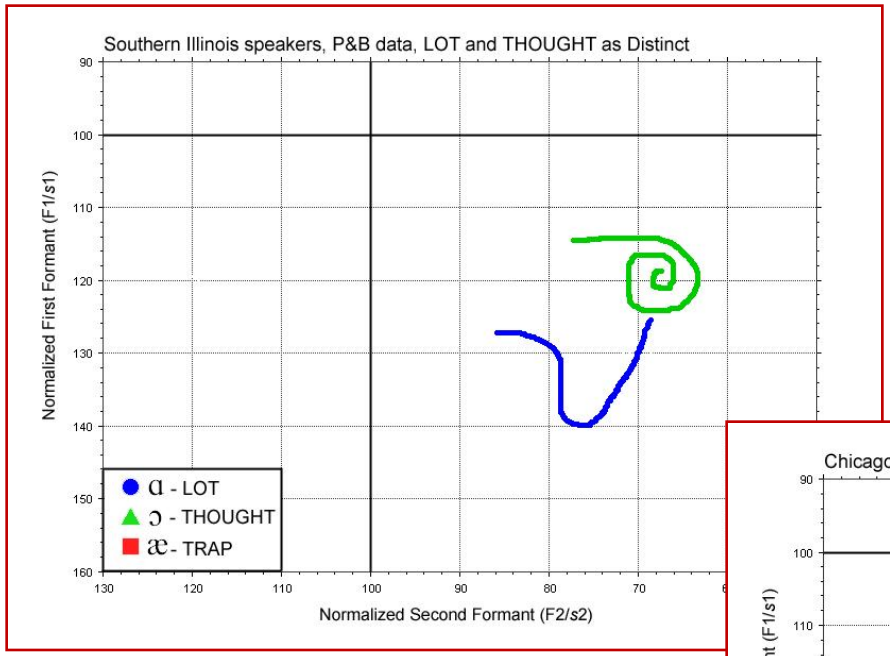


Phonetic Detail: Vowel Trajectories

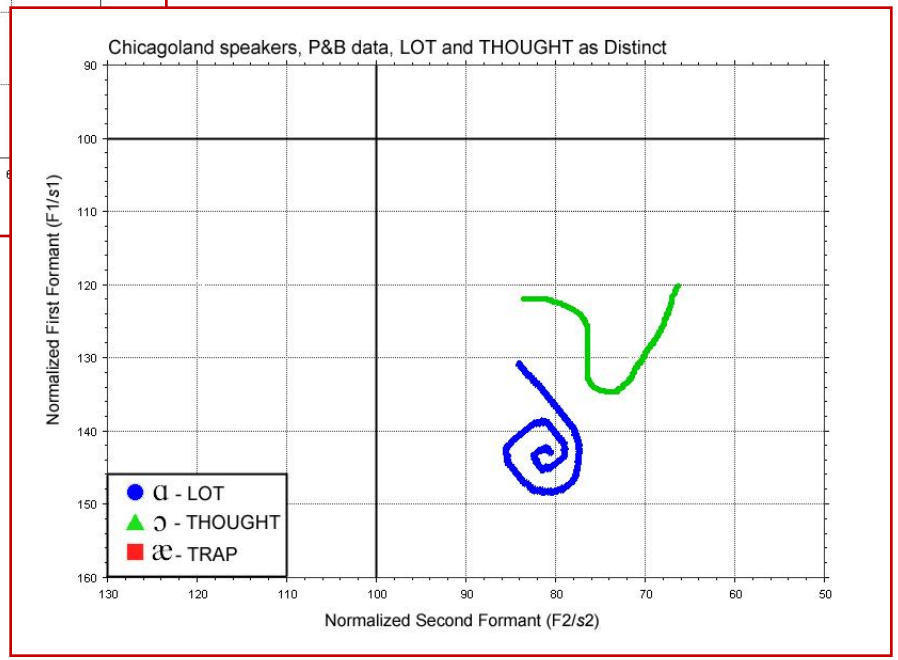
- P&B tokens are longer in duration
  - Longer duration = better phonetic detail
  - ... such as comparing vowel trajectories
- MERGED
  - most speakers
  - all regions



Phonetic Detail: Vowel Trajectories



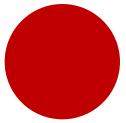
- Southern Illinois
  - Un-merged
  - THOUGHT = stable
  - LOT = moving



- Chicagoland
  - Un-merged
  - THOUGHT = moving
  - LOT = stable

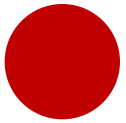
Conclusion:
In Defense of Word List Data

- Comparability
  - Across disciplines / researchers / subjects
  - Across token types
  - Across consonantal & linguistic contexts
- Higher Level Phonological Access
  - More attentive speech provides more accurate phonological targets
  - Word list speech not subject to “supra-phonetic interference”
- Finer-grained Phonetic Detail
  - tokens are longer in duration
  - vowel trajectories provide compelling information



Problems, Issues, Assumptions

- Is the “attention paid to speech” model accurate?
- Does “more attentive” speech fall closer to the “phonological target”?
- Where does attentive (“unnatural”) speech fit in a model of language change?
- How can these two kinds of data be reconciled?
- Why are some speakers’ interview data and P&B data more closely “matched” than others?



Thank you!

Contact Info for References, Further Questions... job offers...

Douglas S. Bigham, Ph.D.

University of Texas at Austin

douglas.s.bigham@gmail.com

References (Partial)

- Arnett, J. (2000). Emerging adulthood : A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469-480.
- Bell, A. (1984). Language style as audience design. *Language in Society, 13(2),* 145-204.
- Chambers, J. & P. Trudgill. (1980). *Dialectology*. London: Cambridge University Press.
- Fuller, J.M. (2000). Changing perspectives on data: Interviews as situated speech. *American Speech, 75(4),* 388-390.
- Gordon, M. J. (2001). *Small-town values and big-city vowels: A study of the Northern Cities Shift in Michigan*. Publications of the American Dialect Society. *PADS 84*. Durham, NC: Duke University Press.
- Habick, T. (1980). "Sound Change in Farmer City: A Sociolinguistic Study Based on Acoustic Data". Unpublished doctoral thesis. University of Illinois at Urbana-Champaign.
- Kerswill, P. (1994). *Dialects converging: Rural speech in urban Norway.* Oxford: Clarendon Press.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (2001). *Principles of linguistic change: Vol. 2. Social factors*. Oxford: Blackwell.
- Labov, W., S. Ash, & C. Boberg. (2006). *Atlas of North American English.* Berlin: Mouton.
- Peterson, G. E. & H. L. Barney. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24,* 175-184.
- Schilling-Estes, N. (1998). Investigating “self-conscious” speech: The performance register in Ocracoke English. *Language In Society, 27*, 53-83.
- Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.
- Wells, J. (1982). *Accents of English*. Vol. 1. London: Cambridge University Press.