The Fractal Structure of Language:
Phonetic Measurements from the American South

William A. Kretzschmar, Jr., University of Georgia

In previous study of the Linguistic Atlas data from the Middle and South Atlantic States (e.g.

Kretzschmar 2009, 2015; these books provide extensive background for language as a complex

system), I have shown that the frequency profiles of variant lexical responses to the same cue are

all patterned in nonlinear asymptotic hyperbolic curves, which I call A-curves. Phonetic

transcriptions from the Atlas data set in both detailed and simplified IPA have also been shown

to have the same A-curve frequency profiles. Moreover, these frequency profiles are scale-free,

in that the same A-curve patterns occur at every level of scale.

Figure 1: Frequency profiles of four subsets of pronunciations of *one*

Figure 1 shows pronunciations for the word *one* (available online, www.lap.uga.edu), which

have a similar curve for each subset of the data. I have argued that the nonlinear A-curve

distribution arises because human speech operates as a complex system, but that is not my point today. What you should see in what follows is that human speech has a fractal structure, the same frequency curve at every level of scale. Digital assistance and Big Data allow us to see these curves, and once we know they are there, they constitute a fundamental fact about speech that we must recognize in order to improve speech recognition and other computer-mediated processing of language.

The father of fractals is the mathematician Benoit Mandelbrot (1977, 1982). He showed how fractals were common in the world around us, as in the measurement of coastlines, or in the operation of economic markets. Their crucial property is that they are self-similar at every level of scale, as illustrated in Figure 2: the big curves have identical smaller curves at the ends and edges, and these smaller curves have even smaller curves at their own ends and edges.

Figure 2: Fractal design (cover of Mandelbrot 1982)


THE FRACTAL GEOMETRY OF NATURE

Benoit B. Mandelbrot
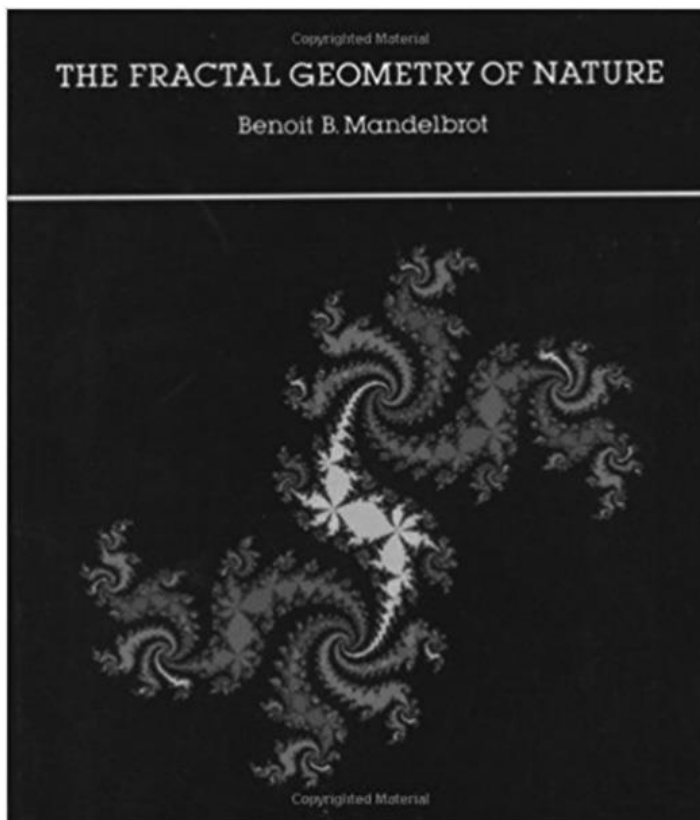
Mandelbrot described mathematical equations for figures like this one, and the equations all relied on exponents to produce the self-similar effects. This is not the place for detailed exposition of the mathematics; it is enough here to note the self-similarity.

My language data here comes from our in-progress, grant-funded automatic harvesting of vowel data from the interviews with sixty-four speakers across the American South (Kretzschmar and Renwick 2016-2018), 380 hours of speech by 34 men and 30 women, born from 1886–1965, across a mixture of ethnicities, social classes, education levels, and ages. Included below are F1/F2 plots, as phoneticians usually describe formants of the speech signal, of the vowels we have harvested so far from 50 speakers. Each measurement is taken at the midpoint of the vowel (our methods are described in Miller, Olsen, Renwick, and Kretzschmar 2017). I will be focusing only on the vowel in GOAT, in all environments. The digital tool used to make the visualizations is a Shiny app in R, made by Joey Stanley of the Atlas team.

Figure 3 shows the picture at the largest level of scale, 50 speakers. We have 16,323 measurements shown here, as tiny dots so that you can still see the grid and shading (boxes must have at least one token for the label to appear). In order to show the distribution of the measurements in F1/F2 space I have applied a 7 x 9 grid, a common technique called point-pattern analysis in spatial science. There are 63 possible boxes and only 8 boxes have no points in them, so the GOAT vowel clearly appears all over the chart. I have also applied shading by density of the points, so boxes with high density are darkest and boxes with lower density appear in lighter shades. Box G5 stands out as the densest one, and we might say that that is how Southerners usually say the vowel in GOAT. But they say the vowel in lots of other ways, too.

Figure 3: F1/F2 plot of realizations of the GOAT vowel (50 speakers)
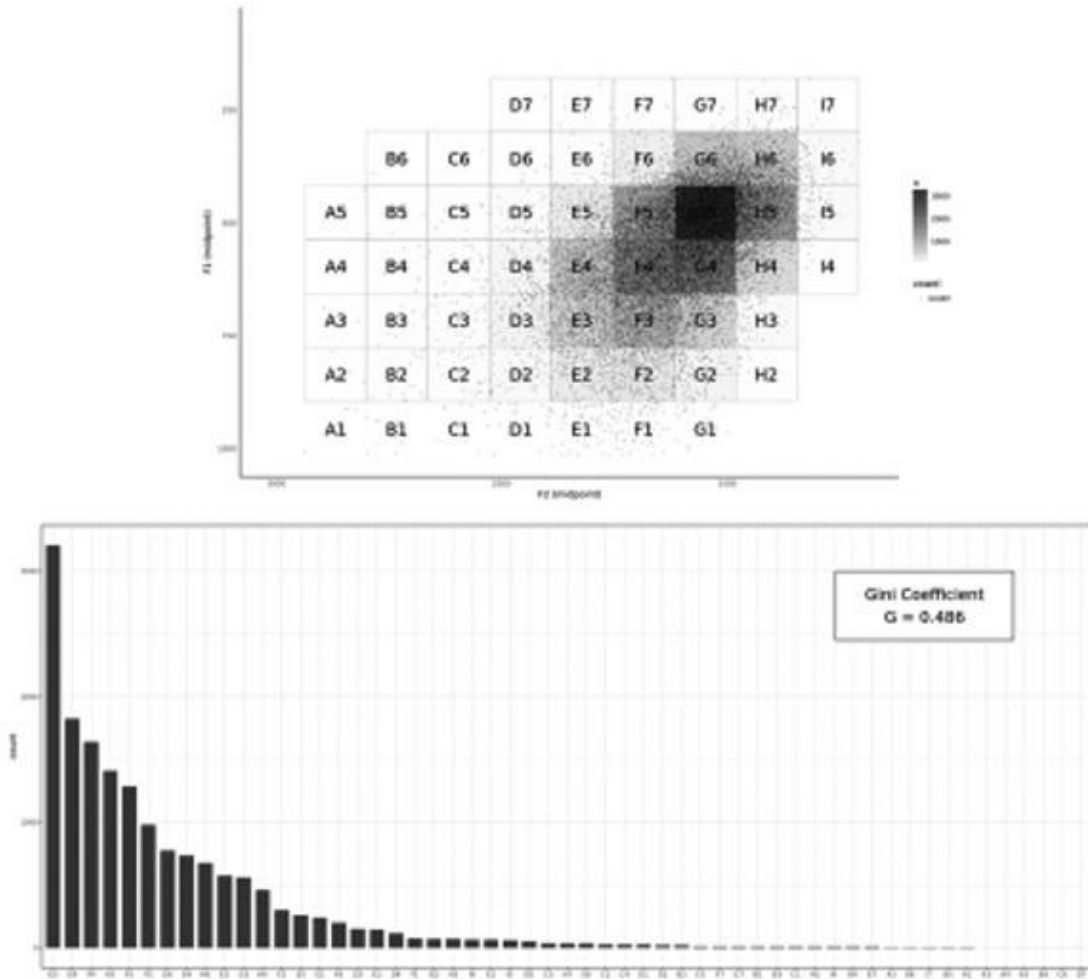
Below the F1/F2 chart you see the A-curve, for the boxes ranked by frequency. We clearly get the same kind of curve from this point-pattern analysis of F1/F2 space as we were getting from Atlas IPA transcriptions in Figure 1, now with over 16000 tokens. The Gini Coefficient shown above the A-curve is a measure of the depth of the nonlinear curve. As I have reported elsewhere (Kretzschmar 2015: Ch. 7), normal distributions always have a Gini Coefficient below 0.2, and the Gini Coefficient here is certainly nonlinear instead of normal, at 0.486. This means that the distribution of vowel realizations does not have a central tendency as in a Gaussian model.

4

Now, the scaling property as shown in subsamples of the same dataset. Figure 4 shows just the men in the sample so far (25 of them), with 10,141 measurements. The densest box is still G5, and the A-curve is still present, now with a slightly higher Gini Coefficient of 0.494.

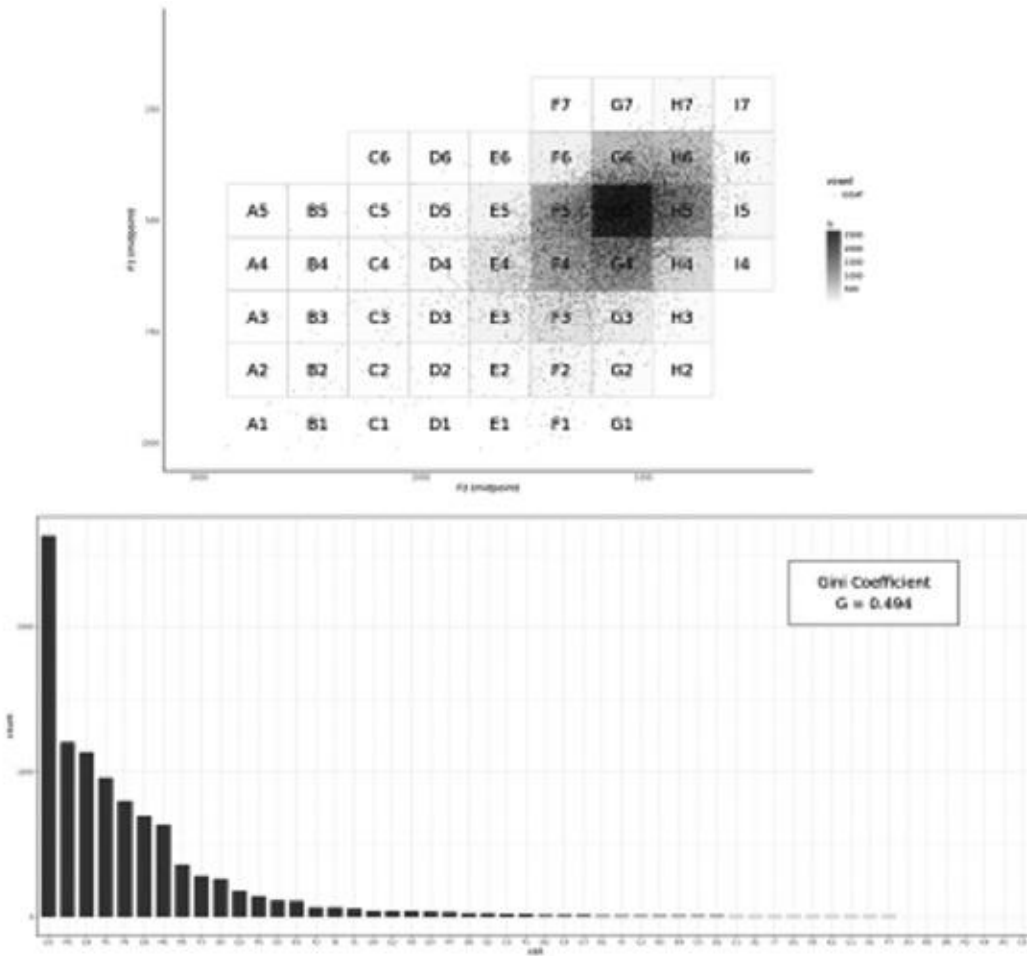Figure 4: F1/F2 plot of realizations of the GOAT vowel (25 male speakers)

Figure 5 shows just the women (25 of them), with 6182 measurements. Now the densest box is G4, with a slightly lower Gini Coefficient of 0.463. What made the Gini change is that women have higher densities of points in more boxes, notably F4 and F3 but also G3, G5, and H5. The women have a somewhat broader distributional pattern than the men.

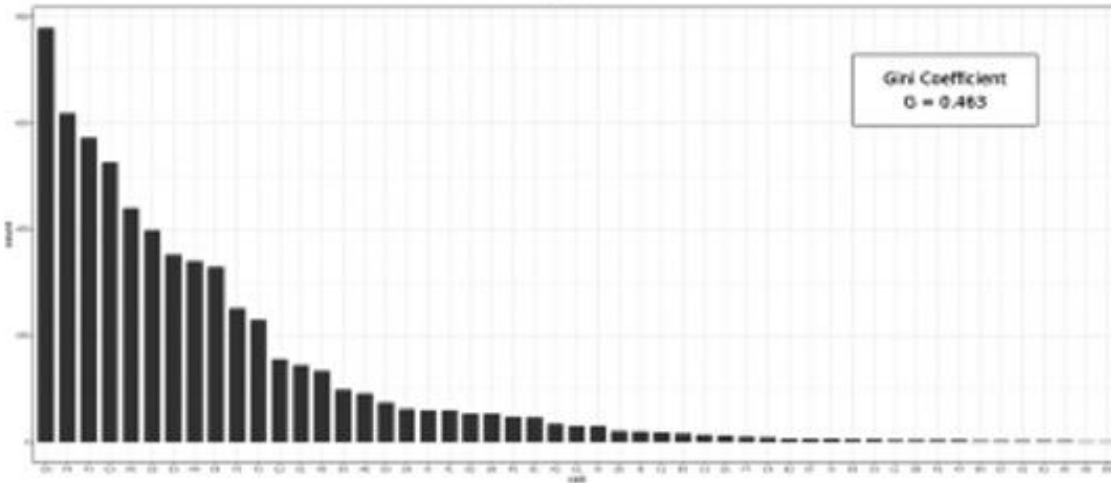Figure 5: F1/F2 plot of realizations of the GOAT vowel (25 female speakers)
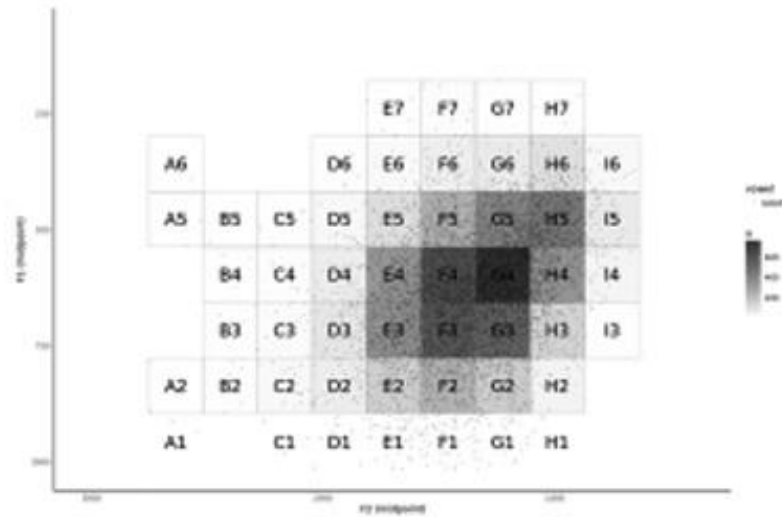

Figure 6 shows just the Black speakers (10 of them), with 3905 measurements. The densest box is again G5, and the Gini Coefficient is back up again to 0.497. The African American speakers are thus not so different from other Southerners in the vowel of GOAT. However, if we look at just the four African American women in Figure 7, with 1271 measurements, we see that the densest box is now H6, higher and backer in articulation, and the Gini Coefficient is up to 0.5.

Figure 6: F1/F2 plot of realizations of the GOAT vowel (10 Black speakers)
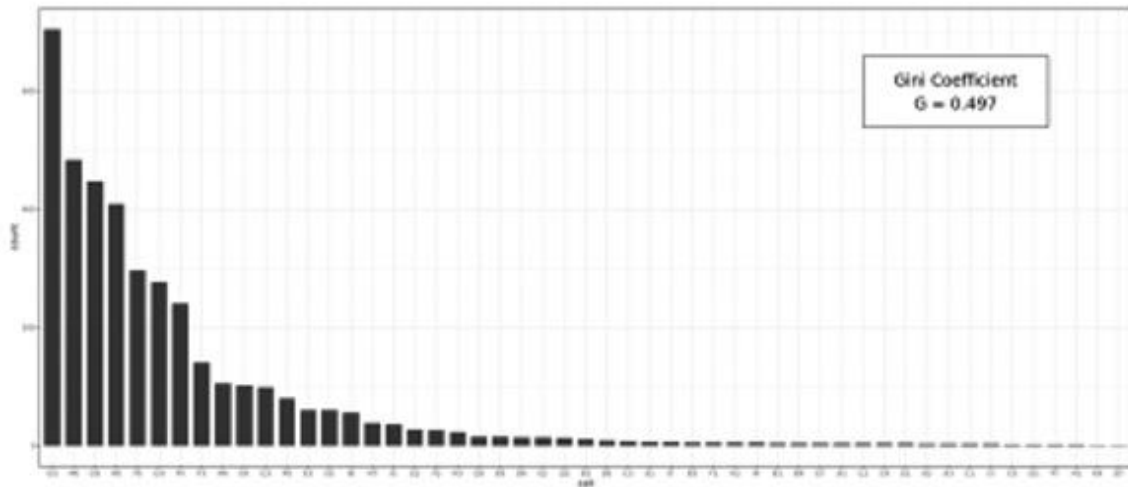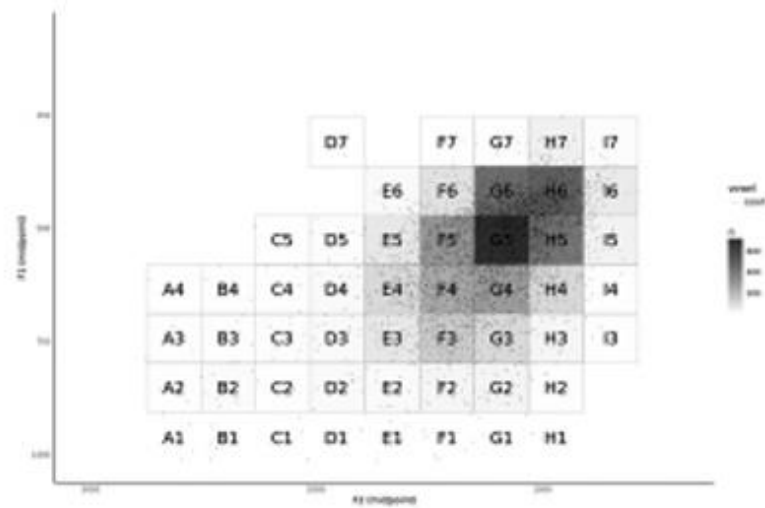
The African American men and the African American women are different, just as the men and women overall were different, but when the Black women's measurements are aggregated with the African American men, they make an A-curve, just as measurements aggregated for the men and women overall make the A-curve for the dataset overall.

Figure 7: F1/F2 plot of realizations of the GOAT vowel (4 female Black speakers)

We are looking at self-similarity, the same shape, at different levels of scale, even when the particulars of the distributional pattern are different. If we observe even smaller groups, down to individual speakers, we see that the individual scale of analysis is the least predictable, but the self-similarity of their overall pattern of realization of the GOAT vowel is the same: the A-curve. In short, human speech as exemplified here in data for the GOAT vowel is self-similar.

So, at the end of the day, what does this mean when we try to pursue our goals in linguistics? We now know, from digital tools and Big Data, that the empirical basis for any

8

generalization about language is self-similarity at different levels of scale. There will not be a

Gaussian central tendency with errors distributed around it, because all distributions are

nonlinear. Speech recognition needs to model many different subgroups of a population, rather

than assuming that one size fits all. It is also the case that variation is always wider than we had

assumed, but always patterned, so that the best speech recognition algorithms will include

interaction that allows speakers to produce a recognizable frequent realization after using

infrequent ones. Phonetic data is never normally distributed, so we need to make phonologies in

ways that respect the self-similarity of speech as people use it.

References

Kretzschmar, William A., Jr. 2009. *The Linguistics of Speech*. Cambridge: Cambridge University
        Press.
Kretzschmar, William A., Jr. 2015. *Language and Complex Systems*. Cambridge: Cambridge
        University Press.
Kretzschmar, William A., Jr., and Margaret Renwick. 2016-2018. NSF BCS- 1625680,
        "Automated Large-Scale Phonetic Analysis: DASS Pilot."
Mandelbrot, Benoit. 1977. *Fractals: Form, Chance, and Dimension.* San Francisco: Freeman.
Mandelbrot, Benoit. 1982. *The Fractal Geometry of Nature.* San Francisco: Freeman.
Miller, Rachel, Michael Olsen, Margaret Renwick, and William A. Kretzschmar, Jr. 2017.
        Methods for Transcription and Forced Alignment of a Legacy Speech Corpus.
        Proceedings of Meetings on Acoustics. Jul 2017: 1-13.