10

Visual Representation of Speech and Sounds

Paul Fletcher and James Elkins

Spectrograms reveal the acoustic nature of sound in two dimensions. The horizontal axis records time. The dark spectrogram is the song of the canyon wren; it is a series of rapid rising notes. As the song goes on — really only four seconds — the little rising notes fall, and to a human ear the result is a falling melody. The green spectrogram is part of the lark sparrow's song. It is much more complex, and over the course of three seconds it makes many rhythmic and tonal changes. To an unaided human ear, it sounds like lovely chirping, but with the help of the sonogram it is possible to hear much more — the visual dissection helps the ear to discern things that would otherwise be too rapid to catch.

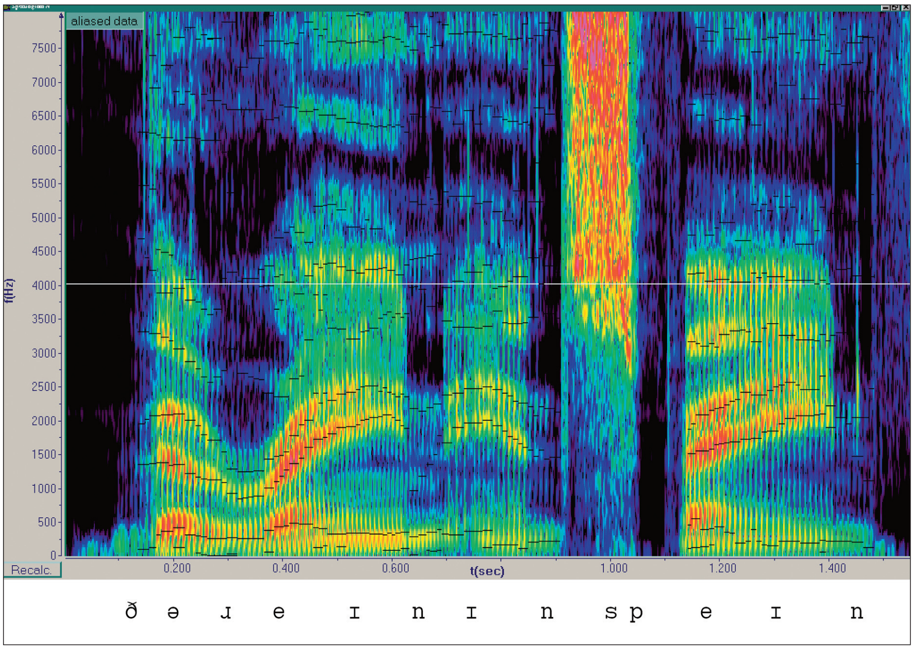
Bird sonograms are increasingly popular, and several software packages make them accessible to amateurs. The Bird sonograms at the beginning and the end of this Chapter were made using software called Raven and Amadeus; the former is used by professional ornithologists.

Human speech spectrograms

In the spectrogram on the next page, someone is saying "the rain in Spain"; it takes less than 1 1/2 seconds.

The second dimension of a spectrogram, after time, is frequency, in this case measured from 0 to 8000 cycles per second (cps, or Hertz). Frequency is labelled on the y axis. The human ear can detect frequencies from 15 Hz to 20,000 Hz, and the voice includes frequencies higher than those shown here.

The third dimension is intensity (loudness). Here color represents loudness: black corresponds to the lowest volume, and louder sounds move through violet, blue, yellow, and orange, to red.



The continuous acoustic signal

The acoustic signal is continuous. Because the linguistic representation that we are most familiar with is the written form, we tend to think that words have spaces between them. As you can see, this is not the case. There is some black on the spectrogram, which signals silence, but the most obvious instance comes within a word, when the "p" sound in "Spain" is being pronounced. This is because in producing this sound you close your lips very briefly (for less than one-tenth of a second) and no sound comes out.

Vowels

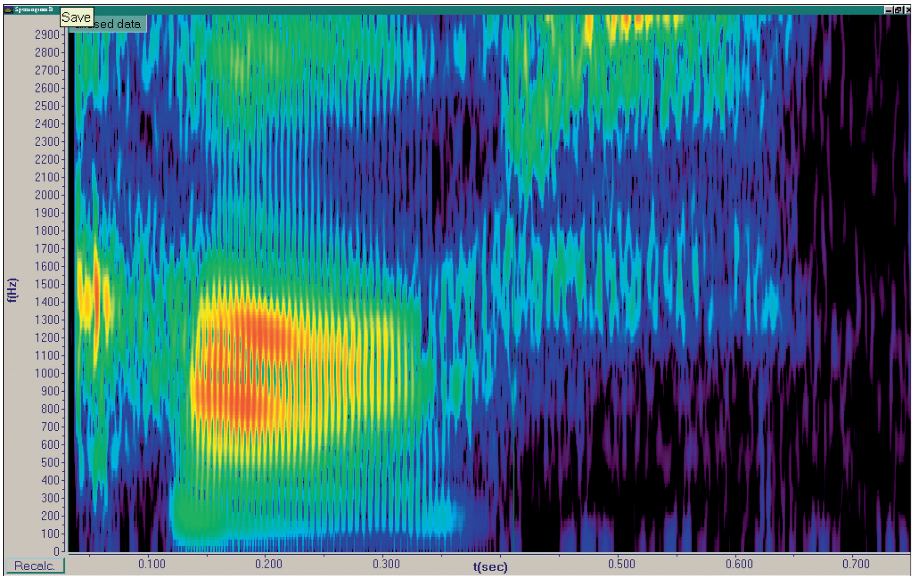
You will see concentrations of acoustic energy, some colored red, in the lower half of the frequency range, between 500 and 2000 Hz. These are characteristic of vowels, and each of the vowels has three obvious arc-shaped concentrations, called 'formants'. It is the shape of these *formants*, their relationship to one another, and their duration, that enables us to identify different vowels.

Formants are a little like harmonics in music, but not as predictable. They are resonance frequences overlaid on the fundamental frequency by the structure of the vocal tract. We hear the vowels in "rain" and "Spain" as the same. In the spectrogram, their formants have the same shape, are in the same spatial orientation to one another, and are the same length. The individual sounds are identified below the spectrogram using symbols from the International Phonetic Alphabet (IPA). This is used because English orthography can be misleading, especially as to the sounds of vowels.

Accents

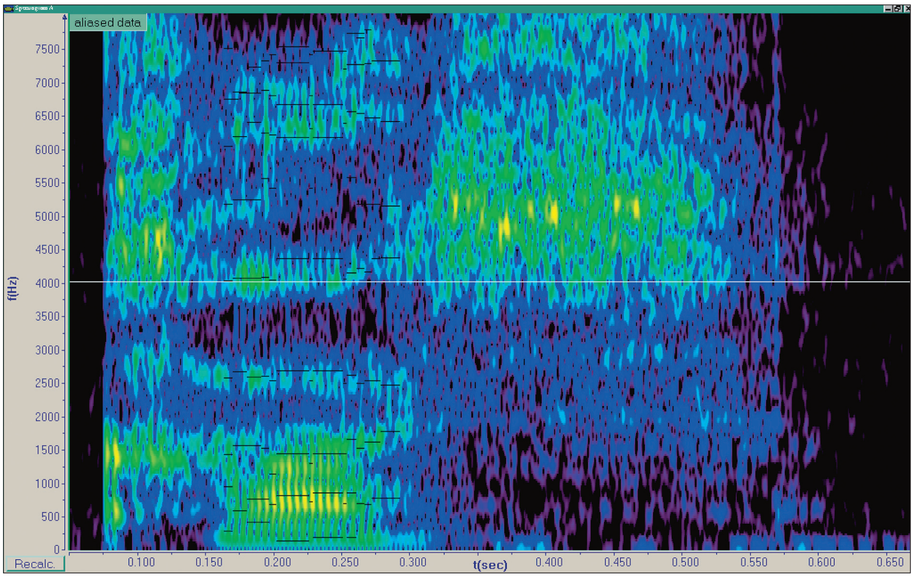
The quality of vowels is very important in distinguishing accents of English. Here are two spectrograms comparing the word "caught" spoken by a native of Tralee in Ireland, and by a native of Bandon, both female and of similar ages. Once again IPA symbols under the spectrogram identifies the different sounds in the words. (This is a funny comparison to people in Ireland: Tralee is a tourist destination, and home of a popular beauty contest; Bandon, to put it politely, isn't a tourist destination.)

Most noticeable is the difference in the length of the vowels. The Bandon vowel is shorter and sounds more like the vowel in "cot," while the Tralee pronunciation is longer. The shape of the formants differs also. Most relevant to our identification of accentual differences are the first three formants. Formants higher up the range (easier to see in the Tralee vowel) are important in identifying individual differences between speakers.



kʰ ɑː tʰ

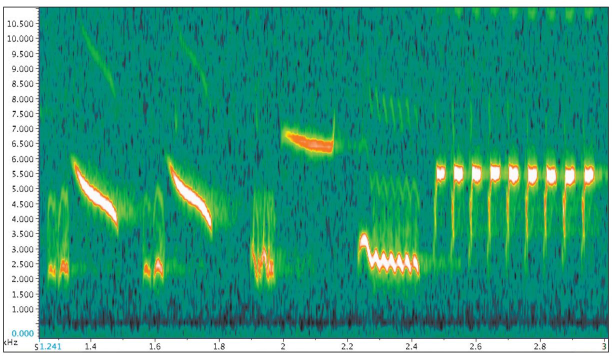
Tralee Informant pronouncing the word "caught"



Individual speakers

Phonograms are common in the mass media when they are shown as "voice-prints": ways of identifying individual speakers on audiotapes. Such differences are very subtle. The Bandon speaker's voice looks "rougher": the corrugations are glottal (laryngeal) pulses, possibly traits of that individual speaker.

Phoneticians distinguish between speaker verification and elimination (which is relatively easy) and speaker identification (which is prone to error). In the first cases one compares a spectrogram from speaker A with a previously stored spectrogram from speaker A, or another speaker. Speaker identification (the voiceprint case) is much trickier, and informed research would suggest that it cannot yet be done reliably.



Spectrograms as pictures

Spectrograms are a good example of one of the themes proposed in the Introduction: they look like naturalistic images, but they aren't. (There have already been several examples, in Chapters 1 and 6.) Somehow, their image-like configuration makes them more useful than they might be if their information were presented in tabular form (as in the Nagios software discussed in Chapter 7). Somehow, they help us to discern what we hear more accurately. Birdwatchers slow down spectrograms so they can pry apart the rapid changes in birdsong, and there are ways of exaggerating the vertical scale and the color schema. But even without those enhancements, the picturelike quality of the spectrograms lets them speak — or sing — more clearly.

For further reading

A. Hughes and P. Trudgill, *English Accents and Dialects* (London: Arnold, 1996); R. Kent and C. Read, *Acoustic Analysis of Speech,* second edition (Albany NY: Singular, 2002); P. Ladefoged, *Vowels and Consonants* (Oxford: Blackwell Publishers, 2001); David Rothenberg, *Why Birds Sing: A Journey Through the Mystery of Bird Song* (New York: Basic Books, 2005).

11

Matching Shades of Crowns

Francis Burke and Catherine Gorman

It may seem that teeth just range from sparkly white to tea-stain brown: that would be a simple change in hue, and a lowering of value. But in fact some teeth have a wide color range, and they also vary in chroma. (Color is best quantified according to its three dimensions: hue, for example blue, red, or green; the chroma, which is the intensity or saturation of color; and the value, the level of brightness of the color.)

The challenge for dentists is that the differences in the hue, value, and chroma of individual teeth are very subtle but crucial to providing the patient with a crown that matches perfectly. (Artists have it easy by comparison!)

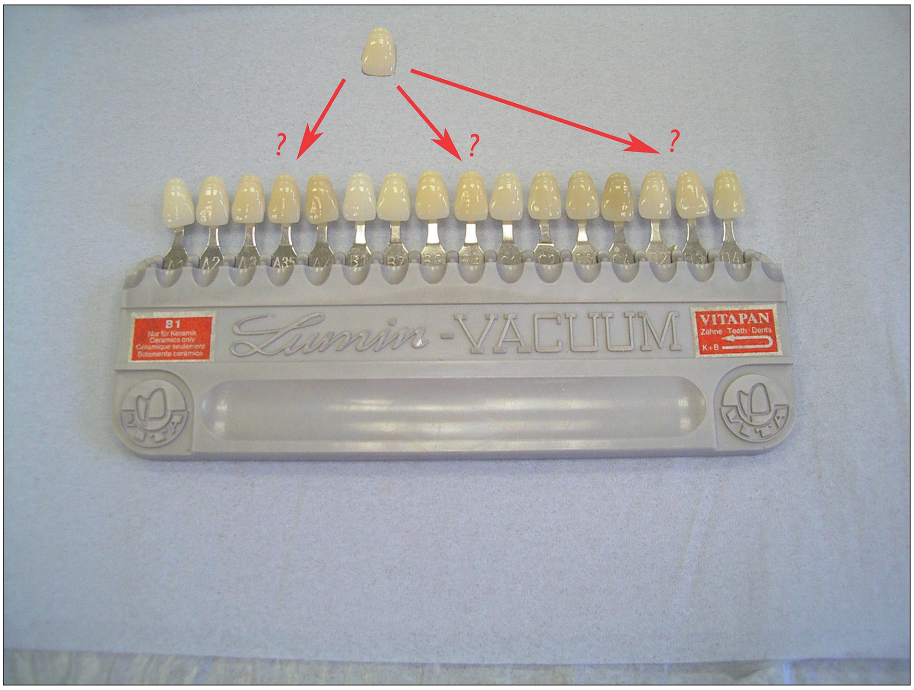
Vita's first method

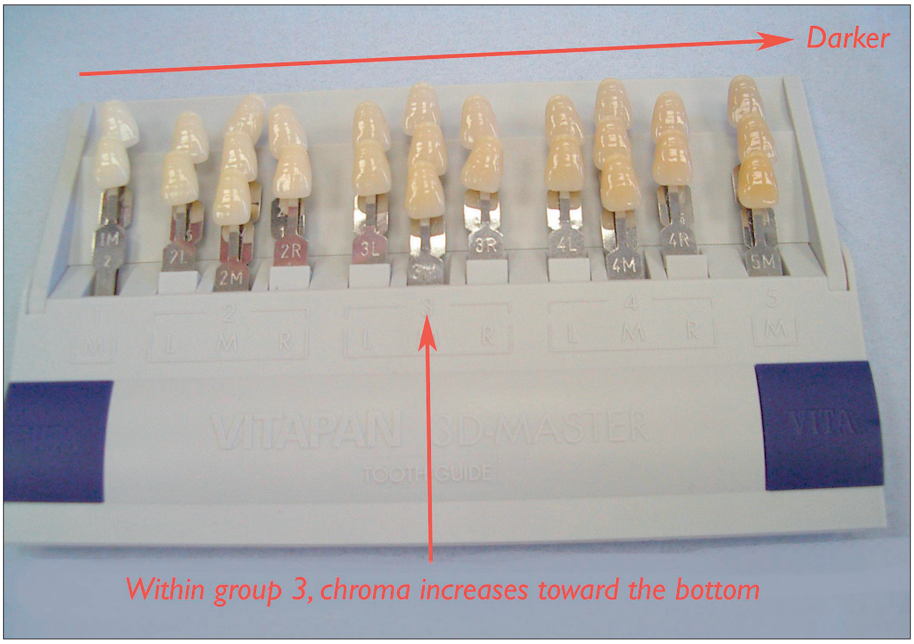
Vita Zahnfabrik in Bad Sackingen, Germany, is a large dental company. They offer two arrangements.

In the first (top photo on the next page), porcelain teeth are grouped into four categories: A, B, C, and D. Notice the small letters at the base of the teeth.) Each represents a different hue. Each category is subdivided into numbers: A1, A2, A3, A3.5, A4. Here the number represents the chroma (intensity). The teeth can be arranged from lightest to darkest for comparison of value to the natural teeth.

Vita's second method

In the 3-D Master Toothguide, the shade samples are grouped in 5 lightness levels (values). (See the bottom photo on the next page.) This arrangement incorporates all three dimensions of color perception; the manufacturers claim this results in increased accuracy when selecting a shade. Value is selected first; in this example the dentist selects the three teeth labelled 3M.


?
?
?
Lumin-VACUUM
VITAPAN


Darker
Within group 3, chroma increases toward the bottom

The three teeth here are labelled 3M1, 3M2, and 3M3. The differences between them are higher and lower chroma (intensity). And finally, the hue is decided (L,M,R).

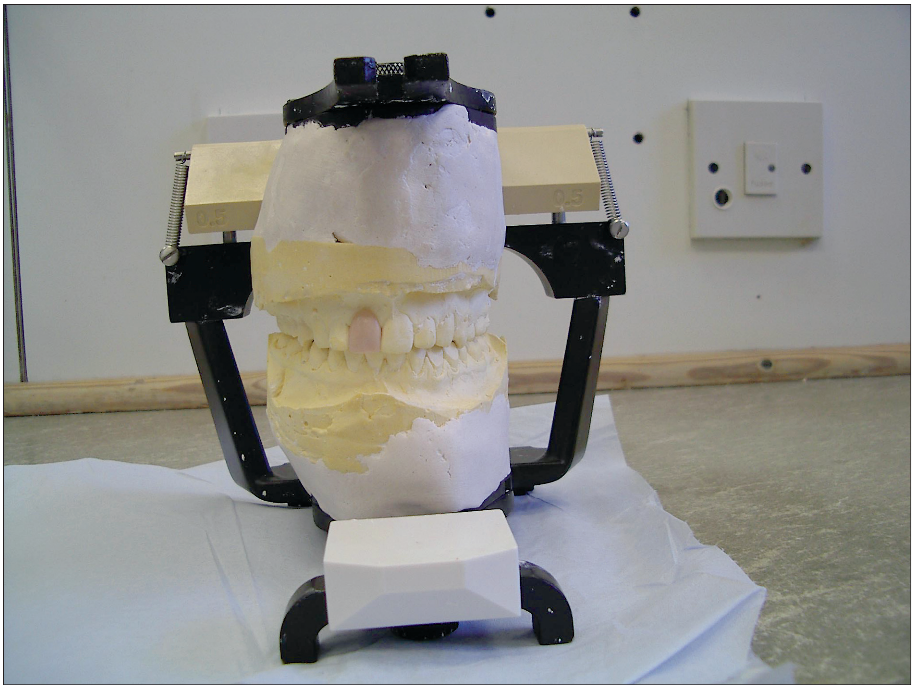
Color wheels

Notice how subtle this is, in comparison to ordinary color wheels. Can you see, for example, that the three teeth in the vertical row 3M differ in chroma (intensity) and *not* in value (darkness)? That the teeth in the third grouping differ in hue? The ones on the right are more red, and the ones on the left are more yellow.

This level of subtlety is beyond most instruction in art academies, where Josef Albers's *Interaction of Color* is generally considered the most detailed source. (The colors in his book are more obvious than these.)

Color sensitivity

In the original exhibition on which this book is based, we offered a model of teeth , a mirror, and a color-corrected lamp. Visitors were encouraged to compare the porcelain samples with their teeth, and see if they could name the color classifications of each of their teeth.



The makers of porcelain teeth recommend that dentists only compare tooth colors for five to seven seconds at a stretch, because the colors are so close in value, hue, and chroma that the eye's sensitivity begins to trail off. Very little in the art world compares with this degree of sensitivity: the closest parallel would be professional printers, who have to be very accurate in matching print runs and in approximating the colors of original prints when they are reproduced; but there is software and hardware for those tasks. In art education, Joseph Albers's *Interaction of Color* (1963) remains a standard textbook. It is used in conjunction with colored papers, and one of Albers's purposes was to increase artists' sensitivity to minute changes in hue and chroma. But even that text — in its many abridged editions and computer implementations — does not come anywhere near the precision required in matching porcelain teeth. In terms of visual subtlety and sheer disciminatory accuracy, the techniques described in this section surpass even art and history of art.

For further reading

Shimya Hobo et al., editors, *Fundamentals of Fixed Prosthodontics*, third edition (Chicago IL: Quintessence, 1997); N. Ray, "Color and Color Matching" (unpublished manuscript, 1998, University College Cork, available at www.ucc.ie/ucc/depts/restorative/downloads.htm); and R.W. Dykema, C.J. Goodacre, and R.W. *Phillips Johnston's Modern Practice in Fixed Prosthodontics,* fourth edition (Philadelphia PA: W.B. Saunders, 2001); Elkins, *How to Use Your Eyes* (New York: Routledge, 2000), chapter 27, "How to Look at Color."