Toward phonetically grounded distinctive features.
Part II: Experimental evidence for blade features

Mark Pennington
mwpennin@indiana.edu

Abstract

In the first paper of this report "Toward phonetically grounded distinctive features. Part I: Acoustic-articulatory correlations in a four-region model of the vocal tract", the formant frequencies F1–F4 and the quality (or gain) factors Q1–Q4 are correlated with the positions, areas, or area ratios formed by the four active articulators: tongue root, tongue body, blade, lips. Among the findings, it was determined that: 1) when the blade position (location of smallest constriction) moves toward the lips, F3 frequency shifts higher; 2) blade aperture (blade area normalized by lip area) is directly correlated with Q3. The present paper applies these two blade relations to actual coronal speech sounds. To this end, an auditorily-based estimator of Q3 is developed: the peak energy factor PE3. The asymptotic ERB (equivalent rectangular bandwidth) of the auditory filter is about one-sixth octave wide. Hence one-sixth octave is adopted as the unit of formant frequency resolution. Vowel F3 frequencies are observed to span one octave or six one-sixth octaves. The six F3 distinctions are classified by the primary and secondary features of blade position [anterior posterior] and [AB RB], where AB and RB are advanced blade and retracted blade. Dentalveolars are [+anterior –posterior]; postalveolars are [–anterior +posterior]. Blade aperture is captured by the feature pair [elevated depressed]. Laminals are [+elevated –depressed]; apicals are [–elevated –depressed]. When the blade aperture increases from a small value (laminal) through a medium value (apical) to a large value (depressed), PE3 also increases. The coronal fricatives of American English, Toda, and Ubykh are examined as well as the coronal stops, nasals, and liquids of Central Arrernte. Both the palatographic evidence and the PE3 measures consistently show the laminality of [s] and the apicality of [ʃ ʂ]. Furthermore, the [s ʃ] sounds are always [+anterior]. In American English, for example, there is no statistically significant difference in F3 frequency between laminal [s] and apical [ʃ], which indicates very similar blade positions.

1. Introduction

In a previous paper [Pennington, this volume], a 27-tube frequency-domain vocal tract model (FDVT) was developed to calculate eight acoustic parameters: the first four formant frequencies F1–F4 and quality factors Q1–Q4. The quality or amplification factor Q is defined as the formant frequency F divided by its bandwidth B. Four articulator regions are delimited in the FDVT model, each characterized by an active articulator: the 8-tube tongue root region, the 9-tube tongue body region corresponding to a quarter wavelength at the second formant frequency, the 6-tube blade region corresponding to a quarter wavelength at the third formant frequency, and the 4-tube lip region. The vowel area functions of ten speakers were obtained from seven X-ray and MRI investigations and fit to the 27 equal-length tubes using cubic spline interpolation. Correlation matrices between the acoustic and articulatory parameters are calculated for the vowel system of each speaker. The coefficients of the parameter pairs are then averaged across the ten speakers. The results of the study gave the seven acoustic-articulatory relations:

1. Tongue root aperture (tongue root area normalized by lip area) is inversely correlated with F1 frequency.
2. As the tongue body position (location of smallest constriction) moves toward the lips, the F2 frequency also shifts higher.
3. Tongue body aperture (tongue body area normalized by lip area) is directly correlated with Q2.
4. As the blade position (location of smallest constriction) moves toward the lips, the F3 frequency also shifts higher.
5. Blade aperture (blade area normalized by lip area) is directly correlated with Q3.
6. Lip position (sum of tube lengths in lip region) displays an inverse correlation with F4 frequency.
7. Lip aperture (lip area) has a moderate direct correlation with F1 frequency and a weaker inverse correlation with Q4.

It was also shown that the dominant mechanism governing the quality factor of the higher formants (Q2–Q4) is radiation damping.

The purpose of the present study is to establish coronal features based on the relations of blade position and blade aperture:

4. As the blade position moves toward the lips, the F3 frequency also shifts higher.
5. Blade aperture (lip-normalized blade area) is directly correlated with Q3.

In order to use relations 4. and 5. as the foundation for a set of distinctive features with measurable acoustic correlates, it is first necessary to estimate the frequency (F3) and quality factor (Q3) of the third formant. Toward this end, approximate auditory filtering is performed and an estimator of Q3 is developed which has Q-like properties: the peak energy factor PE3 (Section 3).

To capture the coronal contrasts implemented by blade position, Chomsky and Halle [1968] proposed the binary feature [anterior]. They originally defined the anterior-nonanterior contrast as follows [p. 304]:

> "The consonants that in traditional terminology are described as palato-alveolar, retroflex, palatal, velar, uvular, or pharyngeal are therefore nonanterior, whereas labials, dentals, and alveolars are anterior."

Later phonological work restricted the anterior-nonanterior opposition to the blade [Hall, 1997, pp. 144–146]. Thus palato-alveolar and retroflex consonants are now considered to be [–anterior] whereas dental and alveolar consonants are [+anterior]. The binary feature [anterior] allows for a maximum of two distinctions in blade position. However in addition to the [+anterior] dentalveolar fricatives /θ s̪/, Toda also displays a phonemic contrast between the [–anterior] postalveolars /s̢/ and /s̢̠/, which differ only in the greater retraction of the /s̢̠/ (Section 4.4). Hence the binary feature [anterior] is unable to capture all the phonologically relevant distinctions in Toda blade position. Clearly, the opposition between [+anterior] dentalveolars and [–anterior] postalveolars needs to be supplemented by a finer-grained feature analysis of blade position.

The asymptotic ERB (equivalent rectangular bandwidth) of the auditory filter is almost equal to one-sixth octave [Glasberg and Moore, 1990]. Therefore one-sixth octave is selected as the unit of formant frequency resolution. Vowel formant measurements show that F3 frequencies range over one octave or six one-sixth octaves [Peterson and Barney, 1952; Hillenbrand et al., 1995]. Consequently, it seems reasonable to assume that there are likewise six phonetic distinctions in blade position given the correlation between blade position and F3 formant frequency. The six distinctions in F3 frequency are categorized in a two-by-three fashion by two equipollent feature pairs of blade position (Section 4.1). The primary and secondary feature pairs are respectively [anterior posterior] and [AB RB], where AB and RB designate advanced blade and retracted blade. Dentalveolars are [+anterior –posterior] while postalveolars are [–anterior +posterior]. Dentals and alveolars are respectively [+anterior –posterior, +AB] and [+anterior –posterior, –AB].

To capture the coronal contrasts implemented by blade aperture, Chomsky and Halle [p. 312] posited the binary feature [distributed]:

> "Distributed sounds are produced with a constriction that extends for a considerable distance along the direction of the air flow; nondistributed sounds are produced with a constriction that extends only for a short distance in this direction."

Laminal and nonretroflex sounds are then classed as [+distributed] while apical and retroflex sounds are [–distributed]. When the blade constriction is lengthened or shortened, the blade area decreases or increases accordingly. Therefore the binary feature [distributed] correctly conveys the opposition between a reduced (+) or an enlarged (–) blade aperture (Section 2). However in an electropalatographic study of Hindi, the groove-length measure could not distinguish [s] from [ʃ] whereas measures of groove width and contacted electrodes did so reliably [Dixit and Hoffman, 2004]. Because overall blade elevation appears to be a better indicator of blade area than groove length, the equipollent feature pair of blade aperture [elevated depressed] is adopted instead (Section 4.1). The laminal value [+elevated –depressed] indicates a small blade aperture while the apical value [–elevated –depressed] signals a medium one. The blade aperture (lip-normalized blade area) is directly correlated with Q3 and, by extension, its auditorily-based estimator PE3. Thus when the blade aperture increases from a small value (laminal) through a medium value (apical) to a large value (depressed), the PE3 should also increase.

To evaluate the performance of the acoustic correlates of blade position (F3) and blade aperture (PE3), known coronal contrasts are analyzed in American English (Section 4.3), Toda (Section 4.4), Ubykh (Section 4.5), and Central Arrernte (Section 4.6). With the exception of American English the test languages are characterized by very large coronal inventories. The acoustic analyses together with typological data permit some preliminary generalizations about coronal sounds (Section 5).

2. The apical-laminal contrast

In an important antecedent to current views on the apical-laminal distinction, Sweet [1877, p. 40] described [s z] as 'blade' sounds and [ʃ ʒ] as 'blade-point' sounds. He observed that the change from English alveolar [s] to palato-alveolar [ʃ] is made by “retracting the tongue somewhat from the (s) position, and pointing it more upwards, which brings the tip more into play.” Hence in Sweet’s description, there are two characteristics that separate palato-alveolar from alveolar strident fricatives: 1) tongue blade retraction and 2) an upward curving of the tongue tip or apicality. The IPA charts after the 1989 Kiel convention classify [ʃ ʒ] as *postalveolar* [International Phonetic Association, 1989; Handbook of the IPA, 1999], thereby emphasizing blade retraction over apicality as a characteristic trait of these sounds. However in the acoustic analysis of two American English speakers reported below, the [ʃ ʒ] set is kept distinct from [s z] only through apicality, not blade retraction. Therefore it seems preferable to refer to [ʃ ʒ] by means of the older and less restrictive IPA term *palato-alveolar*.

Chomsky and Halle [1968] proposed the binary feature [distributed], where [+distributed] and [–distributed] indicate respectively a long blade constriction (laminal) and a short blade constriction (apical). If the blade constriction is lengthened or shortened, then the blade aperture (blade area divided by lip area) likewise decreases or increases provided that the lip area remains constant. Hence, in principle, the feature [distributed] captures the contrast between a small (+) or a medium (–) blade aperture. Nevertheless, it seems more straightforward from an articulatory standpoint to express the apical-laminal distinction as two degrees of overall blade elevation, with an elevated blade implementing the laminal configuration (small blade area) and a non-elevated blade the apical configuration (medium blade area) [for a literature survey on the apical-laminal distinction, see Dart, 1991, Chapter 1].

The contrast between [s] and [ʃ ~ s̺] has been examined by palatography of the blade region. Fletcher and Newman [1991] recorded the electropalatographic contact patterns of two English speakers. The grand means of the contacted electrodes are [s] 47.7 and [ʃ] 28.9 (n = 96). The grand means of the groove widths are [s] 6.2 mm and [ʃ] 10.7 mm. Using a comparable electropalatographic technique, Dixit and Hoffman [2004] examined the [s ʃ] sounds of a Hindi speaker. The means of the contacted electrodes are [s] 41.7 and [ʃ] 30.2 (n = 96). The groove-width means of [s] and [ʃ] are 6 mm and 11.3 mm, whereas the groove-length means of [s] and [ʃ] are identical: 3.3 mm. Dart [1991, pp. 41–44] applied static palatography to the O'odham fricatives /s s̺/. The cross-subject means of the groove widths are /s/ 6.3 mm and /s̺/ 9.1 mm while those of the groove lengths are /s/ 5.4 mm and /s̺/ 3.1 mm. Despite differences in method, the groove-width means are remarkably similar in English, Hindi, and O'odham: about 6 mm for [s] and 10 mm for [ʃ ~ s̺]. The average number of contacted electrodes in the English and Hindi electropalatographic studies likewise displays a good deal of consistency, roughly 45 for [s] and 30 for [ʃ]. The groove-length measure, on the other hand, performs more poorly given the identical means of Hindi [s] and [ʃ]. Thus the apical-laminal distinction is better captured by overall blade elevation (mean blade area) than by groove length alone, as was suggested in the preceding paragraph. In a palatographic investigation of *sip*-*ship*, Ladefoged [1957, p. 773] found that each of the 164 speakers had a narrower channel for [s] than for [ʃ]. Gafos [1999, pp. 160–161] proposed a phonetic scale in which the cross-sectional area of the [s]-channel is always smaller than that of the [ʃ]-channel.

As was stated in the introduction, the blade aperture (lip-normalized blade area) is directly correlated with Q3. Because the correlation was obtained through vowel modeling, the question arises as to whether this relationship also holds for obstruents (stop transients and fricatives). When the vocal tract is excited at the glottis, the system response exhibits the natural (or pole) frequencies and no zeros. If the excitation occurs at another point along the vocal tract, then the system response displays zeros as well as poles [Flanagan 1972, p. 72–73). To evaluate the effects of the zeros on F3 and Q3, the area function of [t] [Story et al., 1996] was fit to 27 equal-length tubes of the FVDT model described previously [Pennington, this volume, Section 5]. The area of the tube corresponding to the stop closure was varied from 0.05 to 0.4 cm^2, thereby simulating a stop transient or fricative [cf. the 0.05–0.20 cm^2 fricative range of Stevens, 1998, p. 33]. The driving series pressure of this tube was set to unity. The results show that supraglottal excitation brings about negligible changes in the F3 formant frequency. The Q3 difference between the supraglottal and glottal excitation $\log Q3(sgl) - \log Q3(gl)$ is related to the driving tube area (cm^2) in the following way:

cm^2	dB
0.05	–1.37
0.10	–0.98
0.20	–0.84
0.40	–0.74.

The Q3 differences indicate that the supraglottal source leads to a quality factor only about 1 dB less than the glottal source, all else being the same. Thus the correlation between blade aperture and Q3 established for vowels appears valid for obstruents as well.

3. Auditorily-based spectral analysis

The palatographic evidence presented above shows that [s] has a small blade area (extensive blade contact) whereas [ʃ] has a medium blade area (limited blade contact). Hence [s] should produce a lower Q3 than [ʃ], given that the blade aperture correlates directly with the Q3 quality factor. In order to estimate the quality factor Q from actual speech, approximate auditory filtering is first carried out. Then a Q-like measure—the peak energy factor PE—is calculated. The peak energy factors PE1–PE4 correspond respectively to Q1–Q4.

Bell [1867, p. 16] proposed a vowel classification consisting of three primary height distinctions and three primary frontness distinctions. He also put forward three secondary height distinctions and three secondary frontness distinctions for a total of "nine degrees of vertical and nine of horizontal measurement." Because perceived vowel height is inversely related to F1 frequency and perceived vowel frontness is directly

	Calculated limits		Peterson & Barney		Hillenbrand et al.	
	Lowest F1	Highest F1	Lowest F1	Highest F1	Lowest F1	Highest F1
Man	287.4	812.7	270	730	342	768
Woman	341.7	966.5	310	860	437	936
Child	406.4	1149.4	370	1030	452	1002
	Lowest F2	Highest F2	Lowest F2	Highest F2	Lowest F2	Highest F2
Man	812.7	2298.8	840	2290	910	2322
Woman	966.5	2733.8	920	2790	1035	2761
Child	1149.4	3251.0	1060	3200	1137	3081

Table 1. The calculated F1 and F2 frequency limits of men, women, and children compared to the lowest and highest mean formant frequencies of the vowels in Peterson & Barney [1952] and Hillenbrand et al. [1995]. The F1 and F2 spans each consist of nine one-sixth octaves (= 1.5 octaves). The lowest first formant (287.4 Hz) is set one-sixth octave higher than the base frequency of 256 Hz. The man-to-woman and the woman-to-child scale factors are both 1/4 octave (18.9%) [Fant, 1966].

related to F2 frequency, Bell's proposal should entail nine phonetic distinctions in F1 and F2 formant frequency [for an overview of vowel height, see Pennington, this volume, Section 9]. Nine F1 and nine F2 distinctions do not appear excessive when formant frequency resolution is considered. Kewley-Port and Watson [1994] presented the F1 and F2 difference limens ($\Delta F / F \times 100$) of five previous studies. The mean difference limens of F1 and F2 are 4.58% (s.d. 2.88) and 4.56% (s.d. 2.70). These values fall somewhat below one-twelfth octave or a semitone, that is, 5.95%. In a later paper, Kewley-Port and Zheng [1999] determined the overall difference limen to be 0.28 Bark under ordinary listening conditions. To provide a helpful comparison, one-twelfth and one-sixth octaves are equivalent to 0.39 and 0.77 Bark when averaged over a similar frequency range [287–2734 Hz, see Equation 6 in Traunmüller, 1990 for Hertz-to-Bark conversion]. Thus the formant frequency difference limen lies below a semitone in ordinary listening conditions as well. To ensure unambiguous identification, formant frequencies should be separated from each other by a distance significantly larger than the difference limen. The equivalent rectangular bandwidth (ERB) of the auditory filter appears to be a good candidate since 1 ERB is on the order of the discrimination threshold of individual partials in complex tones [Moore and Ohgushi, 1993; cf. also Weitzman, 1992 who found that a formant difference of about 1 Bark was needed for reliable discrimination of synthetic vowels]. Glasberg and Moore [1990] measured auditory filter bandwidths using a notched noise method where a curve function is fitted to signal-to-masker thresholds at varying notch widths. They obtained the function

$$ERB = 24.7(4.37F + 1),$$

where *ERB* is the equivalent rectangular bandwidth in Hz, and *F* the center frequency in kHz. The equation may be rewritten as

$$ERB = \frac{f_c}{9.26} + 24.7,$$

where f_c is the center frequency in Hz, and 9.26 the asymptotic quality factor at higher frequencies. Since the inverse of a quality factor is the same as the bandwidth normalized by its center frequency, the asymptotic ERB is 10.8% ($= 1/9.26 \times 100$) or almost one-sixth octave (12.25%). Assuming that the first and second vowel formants each cover nine intervals of one-sixth octave (= 1.5 octaves), then the resulting frequency limits can be compared with those of actual vowels.

The calculated 1.5-octave limits of men, women, and children are given in Table 1, together with the lowest and highest vowel formant frequencies of two comprehensive studies of American English [Peterson and Barney, 1952; Hillenbrand et al., 1995]. The lowest first formant (287.4 Hz) is fixed at one-sixth octave above the base frequency of 256 Hz. The man-to-woman and the woman-to-child scale factors are both set to 1/4 octave (18.9%), following Fant [1966] who determined the cross-vowel average to be 18% for the former and 20% for the latter. Hillenbrand and Clark [2009] obtained man-to-woman factors of 18% (F1) and 17% (F2) from formant data in Hillenbrand et al. [1995]; Ménard et al. [2002] estimated the man-to-child factor to be about 40% or 1/2 octave (41.4%) on the basis of formant data in Lee et al. [1999]. Although the measured formant extrema show wide variability, the adopted 1.5-octave range spans them fairly well. Consequently, nine phonetic distinctions along the F1 and F2 vowel scales appear to be compatible with known formant frequency resolution.

The auditory filters consist of second-order digital resonators whose 3 dB bandwidths B_{3dB} are obtained from the relation $B_{3dB} = ERB \times (2/\pi)$ and the Glasberg and Moore function $ERB = 24.7(4.37F + 1)$ [cf. Hartmann, 1998, pp. 262–263 for the relation $B_{3dB} = ERB \times (2/\pi)$]. There are 48 filters per octave separated by the frequency ratio of $2^{1/48}$. Accordingly, there are 8 filters per one-sixth octave. The energy of each auditory filter E_i is calculated by squaring and then averaging its output over a 39.37 ms sliding rectangular window [for an extensive discussion of the frequency analysis method, see Pennington, 2005, Chapter 4].

Examination of the energy values of the auditory filters shows that speech spectra are too smoothed for the half-power points to be of use in assessing bandwidth. A measure is therefore needed that—like Q—is 1) dimensionless, 2) invariant with respect to multiplicative frequency shifts, and 3) an indicator of amplification at resonance. The peak energy value PE_i at frequency i meets the three criteria:

$$PE_i = \frac{E_i}{\sqrt{E_{i-8}E_{i+8}}}.$$

Because there are 1/48 octave steps, i–8 indicates a frequency one-sixth octave below i and i+8 a frequency one-sixth octave above i. Recall from the discussion above that a separation of at least one-sixth octave (≈ 1 ERB) is required for reliable discrimination of formant frequencies. The denominator is the geometric average of the energies of E_{i-8} and E_{i+8}. The four formants are evaluated over a total span of 28 one-sixth octaves, which is shifted according to the man's (×1), woman's (×1/4 octave), or child's (×1/2 octave) scale factors. There are five one-sixth octaves below the vowel F1 ranges presented in Table 1, eighteen one-sixth octaves comprising both F1 and F2, and five one-sixth octaves above the F2 ranges. The lowest calculated F1 frequency is 287.4 Hz. The frequency five one-sixth octaves below 287.4 Hz is 161.3 Hz. This value compares favorably with 170 Hz, the lowest F1 frequency of the closed vocal tract measured by Fant et al. [1976, p. 18]. Within each of the 28 one-sixth octaves, the maximum PE_i of

	Calculated limits		Peterson & Barney		Hillenbrand et al.	
	Lowest F3	Highest F3	Lowest F3	Highest F3	Lowest F3	Highest F3
Man	1625.5	3251.0	1690	3010	1710	3000
Woman	1933.1	3866.1	1960	3310	1929	3372
Child	2298.8	4597.6	2160	3730	2143	3702
	Lowest F4	Highest F4			Lowest F4	Highest F4
Man	2896.3	4096.0			3334	3687
Woman	3444.3	4871.0			3914	4352
Child	4096.0	5792.6			3788	4575

Table 2. The calculated F3 and F4 frequency limits of men, women, and children compared to the lowest and highest mean formant frequencies of the vowels in Peterson & Barney [1952] and Hillenbrand et al. [1995]. The F3 span consists of six one-sixth octaves (= 1 octave), the F4 span three one-six octaves (= 0.5 octave). The base frequency and scale factors are the same as those in Table 1.

the 8 filters is found and converted to decibels: $\log \max PE_i$ (= $10 \log_{10} \max PE_i$). Inside a given formant range, the peak formant frequency (F1–F4) is located at the largest value of $\log \max PE_i$ while the peak energy factor (PE1–PE4) is the largest value itself. The four candidate formants are cross-checked against the spectrographic and spectral-slice displays of a waveform editor.

To determine how well the peak energy factor functions as an estimator of amplification at resonance, two source signals were generated, one a 130 Hz sawtooth, the other a white noise waveform. Both signals were fed into a second-order digital resonator with a 500 Hz center frequency and Q values varying from 0 to 30 dB in 3 dB steps. The synthetic waveforms are analyzed by the auditory filter bank and the peak energy factor PE is measured at each Q value. A very strong output-input correlation is observed between PE and Q for both the periodic and aperiodic waveforms (r = 0.964), demonstrating that the peak energy factor performs comparably to the quality factor as an estimator of amplification at resonance [see Pennington, 2005, Section 5.6 for additional details].

4.1 Blade features: [anterior posterior, AB RB] and [elevated depressed]

In Table 1 the measured lowest and highest frequencies of F1 and F2 were found to coincide fairly well with the calculated 1.5-octave limits. In Table 2 the lowest and highest frequencies of F3 and F4 are provided [Peterson and Barney, 1952; Hillenbrand et al., 1995]. The base frequency of 256 Hz and the scale factors are the same as those of the first and second formants. Instead of the range of nine one-sixth octaves, the vowel F3 frequencies appear to span one octave or six one-sixth octaves. The calculated men's F3 limits are, for example, 1625.5 and 3251.0 Hz. Remark that the half-octave from 1625.5 to 2298.8 Hz also corresponds to the male F2 frequencies typical of front vowels. Consequently, there is an overlap of three one-sixth octaves between the second and third formants. The highest men's vowel F3 of about 3000 Hz lies somewhat below the calculated upper limit of 3251.0 Hz. Furthermore, the women's and children's F3 frequencies fall appreciably short of their calculated upper F3 bounds. However the mismatch seems to involve only the American English vowels since the F3 of a male

Blade Position	Coronal Subplace	anterior	posterior	AB	RB
advanced anterior	dental	+	–	+	–
plain anterior	plain alveolar	+	–	–	–
retracted anterior	retracted alveolar	+	–	–	+
advanced posterior	advanced postalveolar	–	+	+	–
plain posterior	plain postalveolar	–	+	–	–
retracted posterior	retracted postalveolar	–	+	–	+

Table 3. The primary and secondary feature pairs of blade position [anterior posterior] [AB RB] classifying the six coronal subplaces in a two-by-three fashion. AB designates Advanced Blade, RB Retracted Blade.

Toda dental /s̪/ reaches 3228 Hz (see Table 7 below). The calculated F4 frequencies span three intervals of one-sixth octave (= 0.5 octave). The highest men's vowel F4 of 3687 Hz is notably lower than the calculated upper limit of 4096.0 Hz. Yet the discrepancy concerns only the vowels again since the F4 frequency of a male Ubykh /s/ attains 4067 Hz (see Table 8 below). Given that the F4 frequency of the Toda postalveolar /s̠/ (2338 Hz) extends well below the F3 frequency of the Toda dental /s̪/ (3228 Hz), an overlap of one-sixth octave between the third and fourth formants is adopted. In all, the calculated F3 and F4 ranges consist respectively of six one-sixth octaves and three one-sixth octaves. Nevertheless, there are only five one-sixth octaves above the calculated F2 ranges in Table 1 because a) F3 overlaps F2 by three one-sixth octaves and b) F4 overlaps F3 by one-sixth octave.

In the introduction it was noted that Chomsky and Halle [1968] posited a binary division of blade position: [+anterior] and [–anterior]. Recall also from the introduction that blade position is best correlated with F3 frequency. As the calculated F3 frequencies range over six one-sixth octaves, a binary division of blade position thus yields an interval of three one-sixth octaves for each sign of [anterior]. In Table 3 the primary and secondary feature pairs of blade position [anterior posterior] [AB RB] classify the six coronal subplaces in a two-by-three fashion. The primary feature pair [anterior posterior] performs the same binary division of blade position as [±anterior] but is recast as an equipollent opposition. The secondary equipollent feature pair [AB RB] stands for [Advanced Blade Retracted Blade].

The dentalveolars in Table 3 are [+anterior –posterior] while the postalveolars are [–anterior +posterior]. The anterior blade region comprises both the dental and alveolar subplaces. Keating [1991, p. 31] considers the length of the anterior blade region to be on the order of 1.5 to 2 cm, half the estimate of 3 to 4 cm for the entire blade region. The alveolars [t d s z] occur more often than the dentals [t̪ d̪ s̪ z̪] according to Maddieson's typological survey [1984]. In the language descriptions that distinguish between them phonetically, there are 167 alveolar stops [t d] vs. 125 dental stops [t̪ d̪] [p. 35] and 138 alveolar fricatives [s z] vs. 44 dental fricatives [s̪ z̪] [p. 45]. Because there are very roughly twice as many alveolars (305) as dentals (169) overall, the observed distribution may simply reflect an equiprobable partition of the anterior blade region—with two

Blade Aperture (lip-normalized blade area)		PE3 (dB)	elevated	depressed
small	laminal	small	+	–
medium	apical	medium	–	–
large	depressed	large	–	+

Table 4. The equipollent feature pair of blade aperture [elevated depressed].

alveolar subplaces (plain alveolar, retracted alveolar) opposed to one dental subplace. The postalveolar stops and fricatives [ʈ ɖ ʂ ʐ] are of rather infrequent occurrence. Although there are 316 languages with dental or alveolar stops, only 36 languages have postalveolar stops [ʈ ɖ] [p. 32]. Similarly, the database contains only 20 postalveolar fricatives [ʂ ʐ] [p. 45].

As reviewed in Section 2, the apical-laminal contrast depends on overall blade elevation, with laminal and apical sounds corresponding respectively to small and medium mean blade areas. Recall also that the blade aperture (lip-normalized blade area) is positively correlated with Q3. The quality factor Q3 can not be directly estimated from auditory-filtered speech spectra. Consequently, another estimator of the amplification at resonance—the peak energy factor PE3—is developed in Section 3. The equipollent feature pair of blade aperture [elevated depressed] is presented in Table 4. As the blade aperture increases from a small value (laminal) through a medium value (apical) to a large value (depressed), the PE3 should likewise increase. When the blade is the primary articulator, the blade aperture can take only two values: laminal [+elevated –depressed] or apical [–elevated –depressed]. Coronal sounds are therefore featurally [–depressed].

4.2 Blade features: introduction to the evidence

To test how well the features of blade position (acoustic correlate: F3) and blade aperture (acoustic correlate: PE3) can capture phonological distinctions across languages, known coronal contrasts are examined in American English, Toda (Dravidian), Ubykh (Northwest Caucasian), and Central Arrernte (Australian). Both Toda and Ubykh possess an exceptionally large number of coronal fricatives [Ladefoged and Maddieson, 1996, pp. 156–163]. Central Arrernte like many other Australian languages has four coronal stops, nasals, and lateral approximants, of which two are laminal and two are apical [cf. Ladefoged and Maddieson, 1996, pp. 28–30].

Before turning to the individual languages, it is useful to present some background about the mapping of feature categories onto the acoustic continuum [see Repp, 1984 for a survey of categorical perception in speech]. There are two types of mapping: relational and absolute [Fant, 1986]. Relational invariance arises when the order of the delimited feature category C_i in the acoustic continuum is alone sufficient to bring about categorial identity. Absolute invariance occurs when there is an added condition that the feature categories be realized in the same way across different contexts. Let us assume that | indicates a between-category boundary along the increasing real-valued acoustic dimension {D}:

Features of Blade Position		Men’s F3	Women’s F3	Children’s F3
		—3251.0—	—3866.1—	—4597.6—
[+anterior –posterior]	[+AB –RB]			
		—2896.3—	—3444.3—	—4096.0—
[+anterior –posterior]	[–AB –RB]			
		—2580.3—	—3068.5—	—3649.1—
[+anterior –posterior]	[–AB +RB]			
		—2298.8—	—2733.8—	—3251.0—
[–anterior +posterior]	[+AB –RB]			
		—2048.0—	—2435.5—	—2896.3—
[–anterior +posterior]	[–AB –RB]			
		—1824.6—	—2169.8—	—2580.3—
[–anterior +posterior]	[–AB +RB]			
		—1625.5—	—1933.1—	—2298.8—

Table 5. The primary and secondary features of blade position [anterior posterior] [AB RB] classifying the six one-sixth octaves of men’s, women’s, and children’s F3 frequency ranges (Hz). The F3 frequency limits are the same as those given in Table 2. As before, the base frequency is 256 Hz, the man-to-woman and woman-to-child scale factors 1/4 octave (18.9%).

#1 {______ $|C_1|\ C_2\ |C_3|$ ______}
#2 {__ $|C_1|\ C_2\ |\ C_3\ |$ _________}

Because the feature category C_i obeys the ordering relation $C_1 < C_2 < C_3$, each C_i forms a relational equivalence class in #1 and #2. A good example of relational invariance is provided by the peak energy factor PE modeled in Section 3. When the Q increases from 0 to 30 dB, the measured PE of the sawtooth source ranges from 4.969 to 7.809 dB (roughly #1), that of the white noise source from 2.644 to 6.051 dB (#2). The feature categories follow the same order $C_1 < C_2 < C_3$ in both the sawtooth and white noise realizations. Furthermore, the voiced PE (sawtooth) is larger than the voiceless PE (white noise) at a given Q level. Hence the feature categories associated with the peak energy factors PE1–PE4 display relational invariance since they are a) ordered and b) realized differently according to the voicing context. To establish the contrastiveness of a relationally invariant speech parameter, the context must be held constant as in the method of minimal pairs.

The calculated F3 frequency limits of men, women, and children were furnished in Table 2. The lowest and highest F3 bounds comprise six intervals of one-sixth octaves: 1625.5 and 3251.0 Hz for men, 1933.1 and 3866.1 Hz for women, and 2298.8 and 4597.6 Hz for children. In Table 5, the primary and secondary features of blade position [anterior posterior] [AB RB] classify the six one-sixth octaves. As will be seen below, the blade positions specified by the calculated F3 frequencies generally correspond quite well to those in published descriptions regardless of the context. Because the feature categories are a) ordered along the continuum of F3 frequency and b) context-independent, the features of blade position show absolute invariance.

Man	F1 Hz	PE1 dB	F2 Hz	PE2 dB	F3 Hz	PE3 dB	F4 Hz	PE4 dB
s	432	3.410	1681	4.469	2622	0.737	3765	2.144
ʃ	415	3.582	1877	2.154	2628	3.999	3202	1.630
p-value	0.615	0.775	0.067	0.012	0.938	0.000	0.003	0.605
z	343	6.544	1608	4.200	2583	1.457	3665	2.968
ʒ	269	6.241	1803	1.587	2589	4.773	3448	1.784
p-value	0.138	0.701	0.003	0.001	0.906	0.001	0.102	0.164

Table 6.1. Formant frequency and peak energy means of the male American English fricatives [s z ʃ ʒ] (N = 9). The p-values result from a pairwise comparison between the laminals [s z] and apicals [ʃ ʒ] using a one-way ANOVA.

Woman	F1 Hz	PE1 dB	F2 Hz	PE2 dB	F3 Hz	PE3 dB	F4 Hz	PE4 dB
s	434	2.677	1975	2.394	3138	1.536	4305	0.927
ʃ	525	2.603	2153	1.507	3039	4.793	3841	2.029
p-value	0.008	0.889	0.210	0.111	0.119	0.000	0.000	0.022
z	334	5.484	1894	2.031	3097	2.752	4198	0.310
ʒ	373	3.435	2115	2.546	2929	4.892	3796	1.564
p-value	0.144	0.008	0.002	0.388	0.115	0.000	0.024	0.011

Table 6.2. Formant frequency and peak energy means of the female American English fricatives [s z ʃ ʒ] (N = 9).

4.3 Blade features: American English

In Section 2, palatographic data were provided showing that the blade area of [s] is smaller than the blade areas of [ʃ] and [ʂ]. Hence in accordance with Table 4, the following correspondences should hold:

laminal [s z]	small blade aperture	small PE3	[+elevated –depressed]
apical [ʃ ʒ]	medium blade aperture	medium PE3	[–elevated –depressed]

To verify that the apicals [ʃ ʒ] correspond to a medium PE3, and the laminals [s z] to a small PE3, the American English sounds [s z ʃ ʒ] are analyzed with the signal processing methods outlined in Section 3. A male and female speaker produced nonsense utterances of the form VCV, where the fricative C is flanked by vowels of the same quality [ɑ i u] (the recordings were graciously supplied by the House Ear Institute; the male and female speakers are coded as M4_NW and W4_JW). Nine tokens of each fricative were taken from three repetitions across the three vowel environments. There are two kinds of phonetically homogeneous subsegments: transient and hold. For example, stops, nasals, and fricatives consist of at most one approach transient, one or more hold subsegments, and one release transient [Hardcastle, 1976, pp. 134–137]. Subsegment boundaries are determined according to the method developed in Pennington [2005, Section 4.6], but with two additional rules: 1) the duration of a transient subsegment is less than the analysis window (< 39.37 ms) and 2) the duration of a hold subsegment is less than three times the analysis window (< 118.11 ms). The 118.11 ms limit is of the same order as the smallest onset-time difference (≈ 100 ms) that elicits a reliable percept of separated and successive auditory events [Hirsh, 1974; Divenyi, 2004]. Among the hold subsegments of each fricative, the one with the maximum overall intensity was selected for analysis.

The four formant frequencies (F1–F4) and peak energy factors (PE1–PE4) are then measured. The means of the formant frequencies and peak energy factors (N = 9) are presented in Tables 6.1 (man) and 6.2 (woman). A one-way analysis of variance (ANOVA) is performed to decide if the formant frequencies and peak energy factors differ significantly between the [ʃ ʒ] and [s z] sounds. The p-values resulting from their pairwise comparison are therefore also given. The significance level is assumed to be 0.05. Taking into account the seven acoustic-articulatory relations presented in the introduction (Section 1), the male and female averages exhibit the following regularities:

1) The voiced fricatives are associated with a lower F1 frequency than the voiceless fricatives, suggesting a more advanced tongue root (cf. relation 1. tongue root aperture in Section 1). They also have a larger PE1 than the voiceless fricatives, signaling a smaller glottal opening [see Pennington, 2005, Section 6.1 for PE1 as a cue for glottal aperture].
2) The apicals [ʃ ʒ] tend to have a higher F2 frequency than the laminals [s z], showing a more advanced tongue body (cf. relation 2. tongue body position in Section 1). Hence the American English [ʃ ʒ] sounds appear to be most often palatalized.
3) The apicals [ʃ ʒ] (medium blade aperture) are characterized by a medium PE3, the laminals [s z] (small blade aperture) by a small PE3, thereby verifying the correspondences presented in Table 4. Of all the formant variables (F1–F4, PE1–PE4), the PE3 measure provides the most significant differences between the [ʃ ʒ] and [s z] sounds (p-values ≅ 0.000). Thus the apical-laminal contrast constitutes the defining opposition between [ʃ ʒ] and [s z].
4) The apicals [ʃ ʒ] tend to have a lower F4 frequency than the laminals [s z], which indicates greater lip protrusion (cf. relation 6. lip position in Section 1). Ladefoged and Maddieson [1996, p. 148] state that “the secondary articulation of lip rounding is a feature of ʃ in some languages, such as English and French, but it is not found in many other languages, such as Russian.”

Remark that there are no statistically significant F3 differences between the apicals [ʃ ʒ] and the laminals [s z], which indicates that their blade positions are approximately the same. In the male Table 6.1 the apical and laminal sounds have almost identical F3 frequencies, whereas in the female Table 6.2 the apicals [ʃ ʒ] have a lower F3 frequency than the laminals [s z] but not significantly so. In Table 5, the six one-sixth octaves of men’s and women’s F3 frequency ranges are classified by the primary and secondary features of blade position [anterior posterior] [AB RB]. When the male and female F3 measures are categorized according to this grid, then the following features of blade position are found:

[s z ʃ ʒ] male plain alveolar	[+anterior –posterior, –AB –RB]
[s z] female plain alveolar	[+anterior –posterior, –AB –RB]
[ʃ ʒ] female retracted alveolar	[+anterior –posterior, –AB +RB]

Since, however, the female F3 differences between [s z] and [ʃ ʒ] are not significant, the features of blade position are more generally:

[s z ʃ ʒ] alveolar	[+anterior –posterior, –AB]

As was pointed out earlier, Sweet [1877] identified two characteristics that separate palato-alveolar from alveolar strident fricatives: 1) blade retraction and 2) apicality. Yet the formant data show that only the apical-laminal contrast consistently distinguishes between the [ʃ ʒ] and [s z] sounds in this sample of American English. In sum, the features of blade position and aperture are as follows:

[s z] laminal alveolar	[+anterior, –AB] [+elevated –depressed]
[ʃ ʒ] apical alveolar	[+anterior, –AB] [–elevated –depressed]

To enable a comparison with previous work, a table in the Appendix presents Jassem's formant measures [1965] of the strident coronal fricatives in American English and Polish. The means of the formant frequencies (F2–F4) and the relative formant amplitudes (A2–A4) are given for the speaker of each language. The formant amplitudes are normalized relative to the strongest peak in the spectrum. The American English formant frequencies in the Appendix follow the pattern of those in Tables 6.1 and 6.2 rather well. Once again the apicals [ʃ ʒ] have a higher F2 (more advanced tongue body) and a lower F4 (greater lip protrusion) than the laminals [s z], whereas the F3 differences between [s z] and [ʃ ʒ] remain small (similar blade positions). The quality factor Q is equal to $\left|H(F)\right|$, the magnitude of the transfer function at resonance [Fant, 1960, p. 54]. Accordingly, the relative formant amplitude should be a plausible estimator of Q. However the overall slope of the driving source spectrum can vary somewhat depending on the particular speech sound [cf. Fant, 1960, pp. 202–203 for fricatives]. This makes the relative formant amplitude a potentially ambiguous measure when sounds with different spectral slopes are compared. The peak energy factor PE, on the other hand, is less influenced by the overall spectral slope because it is a local estimator of Q. Nonetheless, the relative A3 amplitude does appear to yield results similar to those obtained with PE3. For example, as expected, the A3 amplitudes of American English [s z] are manifestly smaller than those of [ʃ ʒ]:

1) laminal [s z] ≈ -13 dB
2) apical [ʃ ʒ] 0 dB

Stevens [1985, p. 248] also observed that the third-formant energy is weaker for [s] than for [ʃ] in English.

The A3 amplitudes of Polish in the Appendix further suggest an apical-laminal contrast between [s z] and [ʃ ʒ ɕ ʑ]:

1) laminal [s z] ≈ –22 dB
2) apical [ʃ ʒ] ≈ –4 dB
3) apical [ɕ ʑ] 0 dB

The alveolo-palatals [ɕ ʑ] are then distinguished from the palato-alveolars [ʃ ʒ] by a higher F2 frequency, that is, by palatalization [on the equivalence IPA ɕ ʑ = $ʃ^j$ $ʒ^j$, see Hall, 1997, Section 2.4 for references].

Two kinds of American English approximant /ɹ/ have long been recognized: ‘bunched’ and ‘retroflex’ [Delattre and Freeman, 1968]. The retroflex type is almost certainly the apical variant of /ɹ/ (medium blade aperture), the bunched type most likely the laminal variant (small blade aperture). Dalston [1975] measured the F3 frequencies of word-initial /ɹ/ and found averages of 1546 Hz, 2078 Hz, and 2491 Hz, for men, women, and children. When these F3 frequencies are categorized according to Table 5, the features of blade position become:

/ɻ/ retracted postalveolar [–anterior +posterior, –AB +RB]

Zhou et al. [2008] determined by an MRI technique the area functions of /ɻ/ for two male speakers of American English, one with the laminal variant [ɻ̻] (bunched), the other with the apical variant [ɻ̺] (retroflex). Following the method described in Pennington [this volume, Section 5], the two area functions are each fit to 27 equal-length tubes (Xinhui Zhou kindly sent me the bunched and retroflex area functions in tabular form). Then the F3 frequency, Q3 value, and the blade aperture (blade area divided by lip area) are computed by the FDVT model. The F3 frequencies of laminal [ɻ̻] and apical [ɻ̺] are respectively 1538 Hz and 1594 Hz, in the vicinity of Dalston’s male average of 1546 Hz. The Q3 values of laminal [ɻ̻] and apical [ɻ̺] are 37.29 dB and 39.44 dB; the blade apertures of laminal [ɻ̻] and apical [ɻ̺] are 1.228 and 1.366. Although the Q3 values and blade apertures of apical [ɻ̺] exceed those of laminal [ɻ̻] by only small margins, the results do seem to indicate that the ‘bunched’ and ‘retroflex’ types are in fact the laminal and apical variants of /ɻ/. The feature specifications of laminal [ɻ̻] and apical [ɻ̺] are consequently:

[ɻ̻] laminal retracted postalveolar [–anterior, +RB] [+elevated –depressed]

[ɻ̺] apical retracted postalveolar [–anterior, +RB] [–elevated –depressed]

4.4 Blade features: Toda

Shalev et al. [1994] used static palatography to analyze the four strident coronal fricatives of Toda: s̪ s̱ ʃ s̢. The s̱ sound is described as an apical alveolar, the ʃ as palatalized, while dental s̪ and postalveolar s̢ have the usual IPA values. Alternatively, Hamann [2003, Section 2.2.6], following Sakthivel [1977], considers s̱ to be an apical

	F1	PE1	F2	PE2	F3	PE3	F4	PE4	ZCR	relINT
	Hz	dB	Hz	dB	Hz	dB	Hz	dB	Hz	dB
θ {θ}	523	3.957	1592	1.495	3091	0.982	4028	1.124	1550	–20.7
s̪ {s̪}	542	2.738	1629	0.673	3228	0.822	3537	4.226	2286	–17.7
s̢ {s̠}	504	2.210	1269	1.765	2095	3.225	2865	1.915	1761	–16.9
s̢ʲ {ʃ}	526	1.541	1440	1.083	1976	2.321	2821	3.363	2142	–18.3
s̢̠ {s̢}	531	1.993	1238	4.029	1648	3.898	2338	1.366	1846	–17.2

Table 7. The formant frequency and peak energy means of the male Toda coronal fricatives, each uttered three times. The mean positive zero-crossing rates (Hz) and intensities relative to the syllable peak (dB) are designated by ZCR and relINT. The phonetic symbols between braces are the phonemic transcriptions used in the original source.

postalveolar and s̢ a subapical palatal. The four sounds, together with the nonstrident fricative θ, occur in these near-minimal contrasts from the UCLA Phonetics Lab Archive (tcx_word-list_1992_09.wav; http://archive.phonetics.ucla.edu/Language/TCX/tcx.html):

t̪oːθ 'powdery, soft' koːs̪ 'money'
poːs̠ 'milk' poːʃ 'language'
poːs̢ 'clan name'

To detect the differences between nonstrident [θ] and strident [s̪ s̠ ʃ s̢], two measures are employed:

1) the positive zero-crossing rate (ZCR) of the sum of the auditory filter outputs
2) the relative intensity (relINT) normalized to the syllable peak.

Since the ZCR grows with increasing levels of turbulent velocity fluctuation [Sreenivasan et al., 1983], it appears to be the most appropriate parameter for distinguishing among the three types of obstruents, particularly the unvoiced ones: stop, nonstrident fricative, strident fricative [Reddy, 1967; Ito and Donaldson, 1971]. The relative intensity relINT has been shown to perform fairly well as an acoustic cue for the sonority scale, of which the stops and fricatives are a part [Parker, 2002, Chapter 5].

The male Toda coronal fricatives are illustrated in Table 7. The mean values are averaged over three repetitions. To avoid confusion, the phonetic symbols between curly braces will represent the original phonemic transcriptions. When the F3 frequencies of Table 7 are categorized according to Table 5, the blade features of Toda are as follows:

/θ/ {θ} nonstrident laminal dental [+anterior, +AB] [+elevated –depressed]
/s̪/ {s̪} laminal dental [+anterior, +AB] [+elevated –depressed]
/s̢/ {s̠} apical postalveolar [–anterior, –RB] [–elevated –depressed]
/s̢ʲ/ {ʃ} palatalized apical postalveolar [–anterior, –RB] [–elevated –depressed]
/s̢̠/ {s̢} apical retracted postalveolar [–anterior, +RB] [–elevated –depressed]

As expected, the lower values of ZCR and relINT keep the nonstrident fricative {θ} distinct from the strident ones {s̪ s̠ ʃ s̢}. Also in view of the small PE3, both {θ} and {s̪} are laminal in comparison to apical {s̠ ʃ s̺}.The /s̺ʲ/ sound appears to be the palatalized counterpart of nonretracted postalveolar /s̺/ [for a similar view, see Hamann, 2003, Section 4.7]. For instance, a one-way ANOVA conducted between /s̺ʲ/ and /s̺/ across the three repetitions reveals a significant difference in F2 (p-value = 0.042), but a non-significant difference in F3 (p-value = 0.171). This analysis is additionally supported by phonological data. Sakthivel [1977, pp. 44–45) shows that the locative case marker /-s̺/ undergoes morphophonemic assimilation to palatalized /-s̺ʲ/ after [j], and to retracted /-s̺̠/ after a (retracted) postalveolar:

/kas̺ʲtal/ + /-s̺/ → /kas̺ʲtals̺/ ‘in the darkness’

/poːj/ + /-s̺/ → /poːjs̺ʲ/ ‘in the mouth’

/pax uʈ/ + /-s̺/ → /pax uʈs̺̠/ ‘in the midst of cloud’

Note that the morphophonemic rules apply only to the /s̺ s̺ʲ s̺̠/ sounds, which suggests that they form the natural class of postalveolars. Recall further that the /s̺ s̺ʲ s̺̠/ sounds are all classified acoustically as postalveolars since their F3 formant frequencies fall within the [–anterior +posterior] range of Table 5. In consequence, there is good agreement between the phonological and acoustic classifications.

4.5 Blade features: Ubykh

The phonemic inventory of Ubykh includes eight strident coronal fricatives, which Ladefoged and Maddieson transcribe as {s ŝ ɕ s̺ z ẑ ʑ z̺} [1996, pp. 162–163). The non-IPA symbols {ŝ ẑ} are used to denote ‘hissing-hushing’ fricatives, following Catford [1977, p. 290] who finds them similar to [ʃ ʒ] except that “the tip of the tongue rests against the alveoles of the lower teeth.” However in his chart of Ubykh consonants, Hewitt [2004] represents these sounds with the standard IPA symbols /ʃ ʒ/. Words containing the fricatives are taken from the UCLA Phonetics Lab Archive (http://archive.phonetics.ucla.edu/Language/UBY/uby.html):

saːba ‘why’	za ‘one’
ŝa ‘three’	ẑaẑa ‘kidney’
ɕaɕa ‘mother-in-law’	ʑawa ‘shadow’
s̺a ‘head’	z̺a ‘firewood’

The single-token measurements of the male Ubykh fricatives are given in Table 8. When the F3 frequencies are organized according to Table 5, the resulting blade features are:

/s z/ {s z} laminal alveolar	[+anterior, –AB] [+elevated –depressed]
/ʃ ʒ/ {ŝ ẑ} apical alveolar	[+anterior, –AB] [–elevated –depressed]
/ʃʲ ʒʲ/ {ɕ ʑ} palatalized apical alveolar	[+anterior, –AB] [–elevated –depressed]
/s̺ z̺/ {s̺ z̺} apical postalveolar	[–anterior, 0AB 0RB] [–elevated –depressed]

	F1 Hz	PE1 dB	F2 Hz	PE2 dB	F3 Hz	PE3 dB	F4 Hz	PE4 dB	ZCR Hz	relINT dB
s {s}	494	2.072	1637	3.650	2714	0.844	4067	6.305	1651	–26.5
ʃ {ŝ}	501	3.027	1459	0.675	2794	8.345	2917	5.434	2642	–10.8
ʃʲ {ɕ}	427	1.998	2282	3.121	2383	4.870	3571	0.164	1651	–16.9
s̢ {s̢}	480	4.242	1417	2.861	2154	8.539	2834	0.077	2083	–10.4
z {z}	473	4.304	1892	0.169	2875	0.325	4067	0.205	127	–6.1
ʒ {ẑ}	365	5.448	1685	0.421	2754	2.392	2917	1.329	178	–5.8
ʒʲ {ʑ}	427	0.280	2562	1.830	2599	1.866	3623	–0.12	102	–10.0
z̢ {z̢}	466	5.462	1614	1.085	1710	2.369	2834	0.007	254	–5.9

Table 8. The formant frequencies and peak energy factors of the male Ubykh coronal fricatives, each uttered once.

The small PE3 values of laminal {s z} are opposed to the medium PE3 values of apical {ŝ ẑ, ɕ ʑ, s̢ z̢}. Remark that the PE3 values of {ŝ s̢} are significantly larger than usual, possibly indicating the lowered blade (and thus the exceptionally wide blade aperture) that Catford observed for {ŝ ẑ}. Although the F3 frequencies of /ʃʲ ʒʲ/ are distinctly lower than those of /ʃ ʒ/, they still remain within the alveolar category [+anterior, –AB]. Furthermore, the much higher F2 frequencies of /ʃʲ ʒʲ/ clearly make these sounds the palatalized versions of /ʃ ʒ/. Hence the Ubykh alveolo-palatals {ɕ ʑ} are simply the palatalized apical alveolars /ʃʲ ʒʲ/ (cf. the brief discussion of Polish in Section 4.3). The F3 frequencies of the voiceless and voiced postalveolars /s̢ z̢/ display a good deal of variation, 2154 and 1710 Hz, respectively. Therefore it seems likely that secondary distinctions of blade position are neutralized when the fricative is postalveolar: /s̢ z̢/ → [0AB 0RB].

4.6 Blade features: Central Arrernte

In an IPA illustration of Central Arrernte, Breen and Dobson [2005] furnished a phonological sketch and supplementary audio files, each exemplifying a man's phoneme (Central-Arrernte.zip; http://journals.cambridge.org/action/displayJournal?jid=IPA).

The coronal stops, nasals, lateral approximants, and rhotics are classified in the following manner:

Laminal dental: {t̪ n̪ l̪}

Laminal alveo-palatal: {tʲ nʲ lʲ}

Apical alveolar: {t n l ɾ}

Apical postalveolar: {ʈ ɳ ɭ ɻ}

Because of the discrepancies between the original transcription and the formant measurements, in particular for the laterals, the near-minimal contrasts are arranged in the order of Table 9 to enhance clarity:

	F1	PE1	F2	PE2	F3	PE3	F4	PE4	ZCR	relINT
	Hz	dB	Hz	dB	Hz	dB	Hz	dB	Hz	dB
t̻ {t̪}	392	1.773	1735	3.557	2562	0.838	3228	2.310	356	–23.7
t̻ʲ {tʲ}	365	1.735	1975	1.376	2489	1.462	3003	2.061	813	–26.4
t̺ʲ {t}	466	1.062	1785	1.253	2453	2.525	3136	1.710	229	–26.7
t̺ {ʈ}	339	1.165	1417	2.664	2250	2.105	3136	1.246	330	–22.4
n̻ {n̪}	285	6.233	1523	2.681	2794	1.234	3571	1.307	406	–0.4
n̻ʲ {nʲ}	334	7.750	2093	2.940	2637	1.550	3571	1.307	686	–0.9
n̺ʲ {ɳ}	285	6.189	2004	3.524	2599	1.902	3322	0.330	406	–1.5
n̺ {n}	273	6.760	1614	3.197	2599	2.873	3623	–0.37	356	–1.6
l̻ {l̪}	403	6.504	1357	0.525	2714	0.088	3676	0.054	279	–5.7
l̻ʲ {ɭ}	409	4.154	2123	0.985	2834	1.370	3623	–0.17	254	–6.3
l̺ʲ {lʲ}	421	6.395	2186	1.463	3046	2.279	3275	0.685	330	–8.4
l̺ {l}	501	5.758	1357	0.763	2714	2.787	3469	–0.12	356	–6.5
r {ɾ}	258	4.202	1459	1.190	2383	1.196	3228	–0.33	356	–3.2
ɻ {ɻ}	427	7.085	1568	2.360	2004	0.449	3136	0.017	406	–3.9

Table 9. The formant frequencies and peak energy factors of the male Central Arrernte stops, nasals, lateral approximants, and rhotics, each uttered once.

/t̻/ {t̪} at̪ək ‘grind-PAST’
/t̻ʲ/ {tʲ} atʲək ‘awake’
/t̺ʲ/ {t} atək ‘burst-PAST’
/t̺/ {ʈ} aʈək cover-PAST’
/l̻/ {l̪} al̪ək ‘go-PAST’
/l̻ʲ/ {ɭ} aɭəp ‘prickly wattle (tree)’
/l̺ʲ/ {lʲ} alʲək ‘boomerang-DAT’
/l̺/ {l} aləp ‘firestick’

/n̻/ {n̪} an̪ək ‘wet-PAST’
/n̻ʲ/ {nʲ} anʲək ‘head louse-DAT’
/n̺ʲ/ {ɳ} aɳək ‘stick-DAT’
/n̺/ {n} anək ‘sit-PAST’
/r/ {ɾ} aɾəŋ ‘father’s father’
/ɻ/ {ɻ} aɻək ‘see-PAST’

In a previous paper [Pennington, this volume, Sections 8.2–8.3], it was shown that the Q2, Q3, and Q4 quality factors are controlled chiefly by radiation damping. Recall also the acoustic-articulatory relations presented in the introduction:

3. Tongue body aperture (tongue body area normalized by lip area) is directly correlated with Q2.
5. Blade aperture (blade area normalized by lip area) is directly correlated with Q3
7. Lip aperture (lip area) has a moderate direct correlation with F1 and a weaker inverse correlation with Q4

During the hold phase of oral stops and nasals, either the tongue body area, blade area, or the lip area approaches zero, which renders the corresponding quality factor inoperative as a cue. Hence for noncontinuants (stops and nasals), only the approach and release transients can supply meaningful measures of PE2, PE3, and PE4. Accordingly, the release subsegments of the Arrernte coronal stops and nasals are the ones analyzed.

The /r/ {ɾ} sound in aɾəŋ ‘father’s father’ is realized as the trill [r], which occasionally occurs in the citation form according to Breen and Dobson [p. 250]. Lindau [1985] observes that the trill [r] consists of a sequence of closures and openings, where the closing phase is very similar to a tap or stop transient while the opening phase resembles an approximant. Therefore the coronal tap [ɾ] and trill [r] may be defined respectively as a single stop transient [d̆] and a sequence of stop transient + approximant hold subsegments: [[d̆][ɹ][d̆]...]. The alternating subsegmental composition of the trill [r] is also supported by phonological data since /r/ patterns with the tap [ɾ] [Hall, 1997: 122–124] as well as with the approximant [ɹ] [Walsh Dickey 1997, Table 3.3]. Of the two approximant hold [ɹ]-subsegments in the Central Arrernte /r/ (= /[d̆][ɹ][d̆][ɹ][d̆]/), the one with the maximum intensity is chosen for analysis.

When the F3 frequencies of Table 9 are categorized in light of Table 5, the blade features of the stops, nasals, lateral approximants, and rhotics become:

/t̻ n̻ l̻/ {t̪ n̪ l̪} laminal alveolar	[+anterior, –AB] [+elevated]
/t̻ʲ n̻ʲ l̻ʲ/ {tʲ nʲ ȴ} palatalized laminal alveolar	[+anterior, –AB] [+elevated]
/t̺ʲ n̺ʲ l̺ʲ/ {t ɳ lʲ} palatalized apical alveolar	[+anterior, –AB] [–elevated]
/t̺ n̺ l̺/ {ʈ n l} apical alveolar	[+anterior, –AB] [–elevated]
/r/ {ɾ} laminal alveolar	[+anterior, –AB] [+elevated]
/ɻ/ {ɻ} laminal postalveolar	[–anterior, 0AB 0RB] [+elevated]

With the sole exceptions of /t̺/ and /l̺ʲ/, the F3 frequencies of the stops, nasals, and laterals stay within the bounds of the alveolar category [+anterior, –AB]. Note, however, that the stops are retracted alveolars [–AB +RB] whereas the nasals and laterals are plain alveolars [–AB –RB]. The apicals /t̺ n̺ l̺ t̺ʲ n̺ʲ l̺ʲ/ have larger PE3 values than the laminals /t̻ n̻ l̻ t̻ʲ n̻ʲ l̻ʲ/, as would be expected. Ladefoged and Maddieson [1996, p. 30] mention that the spectra of the two apical stops in Eastern Arrernte exhibit strong mid-frequency peaks when compared to the spectra of the two laminal stops. As illustrated by their Figure 2.13, the peaks lie in the third formant range between 2.3 and 3.1 kHz—in concordance with the larger PE3 values of the Central Arrernte apicals.

Because the blade positions are categorized as alveolar [+anterior, –AB], the anterior-nonanterior distinction is not relevant for the stops, nasals, and laterals. Only the apical-laminal and palatalized-nonpalatalized distinctions are contrastive [for alternations within the latter opposition in Australian languages, see Dixon, 2002, Sections 12.2–12.3]. On the other hand, the anterior-nonanterior distinction does appear to be contrastive for the Central Arrente rhotics. Given their small PE3 values, the alveolar trill /r/ and postalveolar approximant /ɻ/ are both laminal. Therefore the two rhotics are set apart only by the lower F3 frequency of postalveolar /ɻ/ and the opposition tap/trill vs. continuant. As in Ubykh, secondary distinctions of blade position are most likely neutralized when the rhotic is postalveolar: /ɻ/ → [0AB 0RB].

5. Preliminary generalizations concerning coronal sounds

The language data provided above illustrate how the features of blade position and aperture can combine in various ways to form coronal sounds. Yet the feature combinations are constrained as evidenced by the regularities in coronal patterning. The following is a set of generalizations based on the present language data and Maddieson's typological survey of sound systems [1984]:

1) Coronals are most often [+anterior –posterior]. They are more rarely [–anterior +posterior] because the postalveolars [ʈ ɖ ʂ ʐ ɳ ɭ ɽ] are of rather infrequent occurrence [cf. Section 4.1 and Maddieson, 1984, p. 60, 77, 81]. Toda with three phonemic postalveolar fricatives /ʂ ʂ^j ʂ̠/ constitutes an obvious counter-example. Observe also the subphonemic distinction in American English between the two postalveolar rhotic approximants: laminal (bunched) [ɻ̻] and apical (retroflex) [ɻ̺].
2) The [s z] sounds of the tested languages are laminal without exception: [+elevated –depressed]. In a parallel manner, the [ʃ ʒ] sounds are always apical: [–elevated –depressed]. These results are in agreement with the palatographic evidence reviewed in Section 2. The Toda and Ubykh postalveolar fricatives are all apical. However the Central Arrente postalveolar /ɻ/ as well as the American English postalveolar 'bunched' [ɻ̻] are both laminal, thus demonstrating that postalveolars are not necessarily apical.
3) The coronal fricatives transcribed as [θ s z ʃ ʒ] are consistently dentalveolar: [+anterior –posterior]. American English shows no statistically significant F3 differences between laminal [s z] and apical [ʃ ʒ], an indication of very similar blade positions.
4) There can be no more than two contrasts in coronal subplace within the [+anterior] dentalveolars or the [+posterior] postalveolars, all else being the same [Hall, 1997, p. 93–94]. Toda, for instance, makes no more than a two-way distinction among nonpalatalized postalveolar fricatives (nonretracted /ʂ/ vs. retracted /ʂ̠/).
5) There appears to be a general preference for more apical fricatives than laminal fricatives within each voicing type. For example, Polish /s ʃ ʃ^j/ has one laminal and two apical fricatives (Section 4.3) whereas Ubykh /s ʃ ʃ^j ʂ/ and Toda /s ʂ ʂ^j ʂ̠/ have one laminal and three apical fricatives. The apical preference probably results from the fact that apical fricatives give rise to a more intense third formant (larger PE3) than laminal fricatives, thereby facilitating perception of differences in blade position when the language has a large number of coronals. On the other hand, the stop, nasal, and lateral series of Central Arrernte each consist of two laminals and two apicals as in many other Australian languages.

6. Summary and concluding remarks

The seven acoustic-articulatory relations given in Section 1 demonstrate that the first four formant frequencies (F1–F4) and quality factors (Q1–Q4) are critical acoustic parameters in speech. Hence in order to analyze actual speech signals, estimates of both the formant frequency F and quality factor Q are required. Toward this end, approximate auditory filtering is carried out using the Glasberg and Moore ERB function. Because the auditory speech spectra are too smoothed for the half-power points to be of use in assessing bandwidth, a Q-like measure is introduced: the peak energy factor PE. The peak energy factors PE1–PE4 are therefore auditorily-based estimators of Q1–Q4. The asymptotic ERB of the auditory filter is nearly one-sixth octave wide. Since a minimum separation of one ERB is necessary for reliable discrimination of formant frequencies, one-sixth octave is adopted as the unit of formant frequency resolution. Both the first and second formants are shown to cover nine one-sixth octaves, whereas the third and fourth formants span respectively six and three one-sixth octaves. The third-formant range overlaps that of the second by three one-sixth octaves; the fourth-formant range overlaps that of the third by one one-sixth octave. The formant-frequency scale factors of men, women, and children are ×1, ×1/4 octave, and ×1/2 octave (Section 3).

The six distinctions in F3 formant frequency are classified in a two-by-three fashion by two equipollent feature pairs of blade position. The primary and secondary feature pairs are [anterior posterior] and [AB RB], where AB and RB designate advanced blade and retracted blade. Dentalveolar sounds are [+anterior –posterior] while postalveolar sounds are [–anterior +posterior]. Dentals are [+anterior –posterior, +AB]; alveolars are [+anterior –posterior, –AB].

Distinctions in blade aperture (blade area normalized by lip area) are captured by the equipollent feature pair [elevated depressed]. Laminal sounds are [+elevated –depressed]; apical sounds are [–elevated –depressed]. As the blade aperture increases from a small value (laminal) through a medium value (apical) to a large value (depressed), PE3 also increases. When the blade is the primary articulator, the blade aperture can assume only two values: laminal [+elevated –depressed] or apical [–elevated –depressed]. Thus coronals are featurally [–depressed].

The coronal fricatives of American English, Toda, and Ubykh are analyzed as well as the coronal stops, nasals, and liquids of Central Arrernte. The blade positions specified by the calculated F3 frequencies in Table 5 generally correspond quite well to those expected from the phonetic descriptions. For example, Toda /s̪/ and /θ/ are categorized as dentals in Table 5 because of their very high F3 frequencies (3228 and 3091 Hz). Both sounds are equally assessed as dentals based on the transcription and palatographic data of Shalev et al. [1994].

As was mentioned in the introduction, the palato-alveolars [ʃ ʒ] are currently considered to be [–anterior] postalveolars. This assumption became established in the 1960s [Catford, 1968; Chomsky and Halle, 1968]. However earlier views on the place of articulation of the palato-alveolars were more nuanced (cf. Sweet's description in Section 2). For example, Heffner wrote in a standard textbook of phonetics [1949, p. 156]:

> "The exact point of the constriction is relatively unimportant for the sound [ʃ], except that it may not be against the upper teeth themselves, for in that event an [s] sound results, but the [ʃ] may be gingival, alveolar, or as the International Phonetic Association describes it, palatoalveolar."

Heffner's term gingival refers to the region near the gum line of the upper teeth.

The usual binary feature analysis of the strident coronal fricatives in the Chomsky and Halle 1968 framework is as follows [Keating, 1988, p. 6; Hall, 1997, p. 98):

s̻ laminal dental	[+anterior] [+distributed]
s apical alveolar	[+anterior] [–distributed]
ʃ laminal postalveolar	[–anterior] [+distributed]
s̺ apical postalveolor	[–anterior] [–distributed]

The features [+distributed] and [–distributed] correspond respectively to laminal and apical coronals as discussed in Section 2. Since the postalveolar fricatives of Toda and Ubykh are apical, the [s̺] is assumed to be apical as well. The equipollent feature analysis of these sounds is:

s̻ laminal dental	[+anterior –posterior, +AB] [+elevated –depressed]
s laminal alveolar	[+anterior –posterior, –AB] [+elevated –depressed]
ʃ apical dentalveolar	[+anterior –posterior] [–elevated –depressed]
s̺ apical postalveolar	[–anterior +posterior] [–elevated –depressed]

The [s̻ s] and [ʃ s̺] fricative sets form natural classes in each feature system (cf. French *sifflantes* for hissing s-like fricatives and *chuintantes* for hushing sh-like fricatives). In the binary analysis, the [s̻ s] set is dentalveolar while the [ʃ s̺] set is postalveolar. In the equipollent analysis, the [s̻ s] set is laminal (small blade aperture) while the [ʃ s̺] set is apical (medium blade aperture). The palatographic evidence as well as the PE3 measures of the examined languages consistently show the laminality of [s̻ s] and the apicality of [ʃ s̺]. Furthermore, the F3 frequencies of [s ʃ] span about the same range, signaling similar [+anterior] blade positions (see Section 4.3 for American English). Consequently, the equipollent feature analysis appears to be the correct one.

Appendix

	F2 (Hz)	A2 (dB)	F3 (Hz)	A3 (dB)	F4 (Hz)	A4 (dB)
			American English			
s	1600	–13	2620	–13	3950	0
ʃ	2020	–11	2610	0	3420	–9
z	1570	–6	2720	–12	3920	0
ʒ	2090	–4	2650	0	3350	–9
			Polish			
s	1900	–21	2920	–21	4170	–3
ʃ	1750	–16	2630	–7	3200	0
ɕ	2300	–11	2930	0	3540	–1
z	1770	–23	2870	–23	4210	–3
ʒ	1750	–13	2640	–1	3220	0
ʑ	2220	–9	2870	0	3580	0

American English and Polish formant measures. The means of the formant frequencies (F2–F4) and the relative formant amplitudes (A2–A4) of fricatives produced by an American English and a Polish speaker. The values are taken from Table 1 in Jassem [1965]. The formant amplitude is normalized relative to the most intense peak in the spectrum.

References

Bell, A.M.: Visible speech (Simkin, Marshall & Co., London 1867).

Breen, G.; Dobson, V.: Central Arrernte. J. Int. Phon. Ass. *35:* 249–254 (2005).

Catford, J.C.: The articulatory possibilities of man; in Malmberg, Manual of Phonetics, pp. 309–333 (North Holland, Amsterdam 1968).

Catford, J.C.: Mountain of tongues: the languages of the Caucasus. Ann. Rev. Anthropol. *6:* 283–314 (1977).

Chomsky, N.; Halle, M.: The sound pattern of English (Harper & Row, New York 1968).

Dart, S.N.: Articulatory and acoustic properties of apical and laminal articulations. UCLA Working Papers Phonet. *79* (1991).

Dalston, R.M.: Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. J. acoust. Soc. Am. *57:* 462–469 (1975).

Delattre, P.; Freeman, D.C.: A dialect study of American English r’s by X-ray motion picture. Linguistics *44:* 29–68 (1968).

Divenyi, P.L.: The times of Ira Hirsh: multiple ranges of auditory temporal perception. Semin. Hear. *25:* 229–239 (2004).

Dixit, R.P.; Hoffman, P.R.: Articulatory characteristics of fricatives and affricates in Hindi: an electropalatographic study. J. Int. Phon. Ass. *34:* 141–159 (2004).

Dixon, R.M.W.: Australian languages: their nature and development (Cambridge University Press, 2002).

Fant, G.: Acoustic theory of speech production (Mouton, The Hague 1960).

Fant, G.: A note on vocal tract size factors and non-uniform F-pattern scalings. Q. Prog. Status Rep., Speech Transm. Lab., R. Inst. Technol., Stockh., No. 7, pp. 22–30 (1966).

Fant, G.: Features: fiction and facts; in Perkell, Klatt, Invariance and variability in speech processes, pp. 480–492 (Lawrence Erlbaum Ass., Hillsdale 1986).

Fant, G.; Nord, L.; Branderud, P.: A note on the vocal tract wall impedance. Q. Prog. Status Rep., Speech Transm. Lab., R. Inst. Technol., Stockh., No. 17, pp. 13–20 (1976).

Flanagan, J.L.: Speech analysis, synthesis and perception (Springer Verlag, Berlin 1972).
Fletcher, S.G.; Newman, D.G.: [s] and [ʃ] as a function of linguapalatal contact place and sibilant groove width. J. acoust. Soc. Am. *89:* 850–858 (1991).
Gafos, A.I.: The articulatory basis of locality in phonology (Garland, New York 1999).
Glasberg, B.R.; Moore, B.C.J.: Derivation of auditory filter shapes from notched-noise data. Hearing Res. *47:* 103–138 (1990).
Hall, T.A.: The phonology of coronals (John Benjamins, Amsterdam 1997).
Hamann, S.R.: The phonetics and phonology of retroflexes; PhD diss. Utrecht University (2003).
Handbook of the International Phonetic Association (Cambridge University Press, 1999).
Hardcastle, W.J.: Physiology of speech production. An introduction for speech scientists (Academic Press, London 1976).
Hartmann, W.M.: Signals, sound, and sensation (Springer Verlag, 1998).
Heffner, R.M.S.: General phonetics (University of Wisconsin Press, Madison 1949).
Hewitt, G.: Introduction to the study of the languages of the Caucasus (Lincom Europa, Munich 2004).
Hillenbrand, J.; Clark, M.J.: The role of f_0 and formant frequencies in distinguishing the voices of men and women. Atten. Percept. Psychophys. *71:* 1150–1166 (2009).
Hillenbrand, J.; Getty, L.A.; Clark, M.J.; Wheeler, K.: Acoustic characteristics of American English vowels. J. acoust. Soc. Am. *97:* 3099–3111 (1995).
Hirsh, I.J.: Temporal order and auditory perception; in Moskowitz, Scharf, Stevens, Sensation and measurement: papers in honor of S. S. Stevens, pp. 251–258 (D. Reidel, Dordrecht 1974).
International Phonetic Association: Report on the 1989 Kiel convention. J. Int. Phon. Ass. *19:* 67–80 (1989).
Ito, M.R.; Donaldson, R.W.: Zero-crossing measurements for analysis and recognition of speech sounds. IEEE Trans. Audio Electroacoust. *19:* 235–242 (1971).
Jassem, W.: The formants of fricative consonants. Lang. Speech *8:* 1–16 (1965).
Keating, P.: A survey of phonological features (IULC, Bloomington 1988).
Keating, P.: Coronal places of articulation; in Paradis, Prunet, Phonetics and Phonology 2. The special status of coronals: internal and external evidence, pp. 29–48 (Academic Press, San Diego 1991).
Kewley-Port, D.; Watson, C.S.: Formant frequency discrimination for isolated English vowels. J. acoust. Soc. Am. *95:* 485–496 (1994).
Kewley-Port, D.; Zheng, Y.: Vowel formant discrimination: towards more ordinary listening conditions. J. acoust. Soc. Am. *106:* 2945–2958 (1999).
Ladefoged, P.: Use of palatography. J. Speech Hear. Disord. *22:* 764–774 (1957).
Ladefoged, P.; Maddieson, I.: The sounds of the world's languages (Blackwell Publishers, Oxford 1996).
Lee, S.; Potamianos, A.; Narayanan, S.: Acoustics of children's speech: developmental changes of temporal and spectral parameters. J. acoust. Soc. Am. *105:* 1455–1468 (1999).
Lindau, M.: The story of /r/; in Fromkin, Phonetic linguistics, pp. 157–168 (Academic Press, Orlando 1985).
Maddieson, I.: Patterns of sounds (Cambridge University Press, 1984).
Ménard, L.; Schwartz, J.L.; Boë, L.J.; Kandel, S.; Vallée, N.: Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. J. acoust. Soc. Am. *111:* 1892–1905 (2002).
Moore, B.C.J.; Ohgushi, K.: Audibility of partials in inharmonic complex tones. J. acoust. Soc. Am. *93:* 452–461 (1993).
Parker, S.G.: Quantifying the sonority hierarchy; PhD diss. University of Massachusetts (2002).

Pennington, M.: The phonetics and phonology of glottal manner features; PhD diss. Indiana University (2005).
Pennington, M.: Toward phonetically grounded distinctive features. Part I: Acoustic-articulatory correlations in a four-region model of the vocal tract (in this volume).
Peterson, G.E.; Barney, H.L.: Control methods used in a study of vowels. J. acoust. Soc. Am. *24:* 175–184 (1952).
Reddy, D.R.: Phoneme grouping for speech recognition. J. acoust. Soc. Am. *41:* 1295–1300 (1967).
Repp, B.H.: Categorical perception: issues, methods, and findings; in Lass, Speech and language: advances in basic research and practice (vol. 10), pp. 243–335 (Academic Press, 1984).
Sakthivel, S: A grammar of the Toda language (Annamalai University, Annamalainagar 1977).
Shalev, M.; Ladefoged, P.; Bhaskararao, P.: Phonetics of Toda. PILC J. Dravidic Studies *4:* 19–56 (1994).
Sreenivasan, K.R.; Prabhu, A.; Narasimha, R.: Zero-crossings in turbulent signals. J. Fluid Mech. *137:* 251–272 (1983).
Stevens, K.N.: Evidence for the role of acoustic boundaries in the perception of speech sounds; in Fromkin, Phonetic linguistics, pp. 243–255 (Academic Press, Orlando 1985).
Stevens, K.N.: Acoustic phonetics (MIT Press, 1998).
Story, B.H.; Titze, I.R.; Hoffman, E.A.: Vocal tract area functions from magnetic resonance imaging. J. acoust. Soc. Am. *100:* 537–554 (1996).
Sweet, H.: A handbook of phonetics (Clarendon Press, Oxford 1877).
Traunmüller, H.: Analytical expressions for the tonotopic sensory scale. J. acoust. Soc. Am. *88:* 97–100 (1990).
Walsh Dickey, L.: The phonology of liquids; PhD diss. University of Massachusetts (1997).
Weitzman, R.S.: Vowel categorization and the critical band. Lang. Speech *35:* 115–125 (1992).
Zhou, X.; Espy-Wilson, C.Y.; Boyce, S.; Tiede, M.; Holland, C.; Choe, A.: A magnetic resonance imaging-based articulatory and acoustic study of “retroflex” and “bunched” American English /r/. J. acoust. Soc. Am. *123:* 4466–4481 (2008).