Acoustic-articulatory correlations in a four-region model of the vocal tract: Experimental evidence for blade features

MARK PENNINGTON

Indiana University

Abstract

In the first part of this report, 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. Vowel area functions of ten speakers were taken from seven X-ray and MRI studies and fit to 27 equal-length tubes by means of cubic spline interpolation. 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. In the second part of the report, evidence for these two blade relations is provided using acoustic recordings of 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. Measured F3 frequencies are observed to span six one-sixth octaves (one octave). 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 [s̻ s] and the apicality of [ʃ s̺]. Furthermore, the [s ʃ] sounds are both found to be [+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.

---

*This paper is an abridged and updated version of a two-part report that had previously appeared in Vol. 11 of the IULC Working Papers. I wish to thank Jennifer Cole for suggesting a more thorough introduction to the concept of the distinctive feature than was presented earlier. Geoffrey Stewart Morrison pointed out that the supraglottal excitation of obstruents may lead to results different from those obtained with the original vowel glottal-source model. He also called attention to the possible presence of a sublingual cavity during sibilant production. Both concerns are now addressed in Sections 2.7 and 3.1, respectively. In 2013 Mark Tiede completed his Yale dissertation "An *MRI-based morphological approach to vocal tract area function estimation*" in which he measured the vowel area functions of 12 speakers – on the same order as the 10 analyzed here. A future paper will examine the acoustic-articulatory correlations of the four articulator regions using his data set.

1. Introduction

The concept of the distinctive feature was introduced in 1928 by the Prague School phonologists R. Jakobson, S. Karcevsky and N. Trubetzkoy. They pointed out that in phonological correlations there is a common principle which may be thought of independently of each couple of opposing terms. In the series of English binary oppositions dill:till :: gill:kill :: bill:pill, for example, the presence or absence of voice (the common substrate) can be extracted as the two terms of comparison. Later Jakobson, Fant and Halle (1952) firmly established the featural notation +F vs. –F for the opposing terms, where F represents a common dimension like voicing and the positive or negative sign indicates the value of the opposing term. Thus at its origin, distinctive feature theory assumed that (a) minimal segmental contrasts and their corresponding acoustic-motor dimensions are determined on the basis of a particular language; (b) opposing feature values can be used to specify disjoint natural classes since [+voice] characterizes the class of voiced stops and [–voice] the class of voiceless stops in the English series above (see Mielke 2011 for the additional use of distinctive features to describe phonological alternations, Baltaxe 1978 for their early history, as well as Clements 2009 and Cohn 2011 for recent discussion).

The goal of this study is to evaluate a set of blade features by acoustically analyzing the minimal segmental contrasts of four genetically diverse languages: American English, Toda, Ubykh and Central Arrernte. In comparison to American English the other languages have very large coronal inventories. Maddieson's typological data (1984) on coronal phonemes are also taken into consideration. The development of distinctive feature theory has been hindered by persistent uncertainty about the relation between speech gesture and acoustic result. Jakobson, Fant and Halle (1952) identified correlates of production and perception for the twelve binary oppositions that they detected in the world's languages. However subsequent research has generally abandoned the concern with co-varying motor and acoustic parameters. Most emphasis has been devoted to the motor dimension (e.g. Chomsky and Halle 1968; Clements and Hume 1995) or exceptionally to the acoustic dimension (Flemming 2002). Before 1990, in any case, the assessment of acoustic-articulatory correlations would have been problematic due to the scarcity of vocal tract area functions in the literature.

This report consists of two parts. The first part (Section 2) finds the acoustic-articulatory correlations for the four active articulators: tongue root, tongue body, blade, and lips (cf. Halle, Vaux and Wolfe 2000 for the same set of oral articulators). The second part (Section 3) assembles palatographic and especially acoustic evidence in support of the proposed blade features. It is important to give a sufficiently detailed account of the acoustic-articulatory correlations in the first part because they provide the theoretical foundation for the empirical acoustic measurements in the second part. Although the present focus is on the blade, acoustic-articulatory correlations are also performed for the other three articulators since coronal sounds may be accompanied by secondary articulations like labialization (lip protrusion) or palatalization (tongue body advancement).

To obtain the acoustic-articulatory correlations, a 27-tube frequency-domain vocal tract model (FDVT) calculates 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 taken from seven previously published 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 yield 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.

The relations 4. blade position and 5. blade aperture establish the evidence base for a set of coronal features with the acoustic correlates F3 and Q3. Accordingly, in order to identify the features of actual coronal speech sounds, one must first find their third-formant frequencies and quality factors. Toward this end, approximate auditory filtering is applied to the speech waveform and an estimator of Q3 is developed which has Q-like properties: the peak energy factor PE3.

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

> "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: 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 s̪̺/, Toda also displays a phonemic contrast between the [–anterior] postalveolars /s̺̠/ and /s̠/, which differ only in the greater retraction of the /s̠/. 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. Formant measurements show that F3 frequencies range over about six one-sixth octaves (one octave). 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. 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 (1968: 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. 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. 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 3.4.1), Toda (Section 3.4.2), Ubykh (Section 3.4.3), and Central Arrernte (Section 3.4.4). The acoustic analyses together with typological data permit some preliminary generalizations about coronal sounds (Section 3.5).

2. Acoustic-articulatory correlations in a four-region model of the vocal tract

2.1 *Description and validation of the Frequency-Domain Vocal Tract model (FDVT)*

A frequency-domain vocal tract model (FDVT) was developed by the author using Fortran 77. The vocal tract frequency response is computed by a transmission line consisting of 27 single-tube T-sections (Fant 1960: 36–38). Each T-section is made up of two series circuits and one shunt circuit. Flanagan (1972: 28–35) reviews the analogous acoustic elements of the T-section and provides the corresponding numerical values. The wall impedance of the relaxed cheek is implemented by a mass-compliance-viscous loss shunt (Ishizaka, French and Flanagan 1975). A circular piston mounted in an infinite plane baffle models the radiation impedance (Aarts and

Janssen 2003), which involves the computation of first-order Bessel and Struve functions within the main Fortran program (Zhang and Jin 1996: 134–136, 347–348). The output volume velocity is calculated in 1 Hz steps from 12 to 6502 Hz; the overall power PWR is defined as the sum of the squares of the output volume velocity. The peak formant frequencies (F1–F4) and 3 dB bandwidths (B1–B4) are then determined interactively. Once the bandwidth of a formant is found, the logarithmic quality factor is calculated $\log Q = 20\log_{10}(F/B)$. The dimensionless log Q is a decibel measure of amplification at resonance (Kinsler and Frey 1962: 195).

To test the validity of the FDVT model, the formant frequencies and bandwidths of a 4 cm^2 and a 1 cm^2 uniform pipe 18 cm in length are compared with those calculated by the Matlab-based VTAR program (Zhang and Espy-Wilson 2004). The VTAR constants are set identical to the ones used in the FDVT program. VTAR provides an on-off switch for the radiation impedance; however, the method of computing the radiation impedance is not documented. The simulations yield typical differences of less than 1% between the FDVT and VTAR acoustic parameters. Thus the validity of the FDVT model is confirmed satisfactorily. The radiation bandwidth is found to grow progressively larger as the formant frequency increases, especially for the higher formants (F2–F4). Furthermore, the radiation bandwidth is larger for the 4 cm^2 pipe than for the 1 cm^2 pipe. These observations are in conformity with the analytical results for the open-closed uniform pipe, which indicate that the radiation bandwidth is proportional to both f^2 and the pipe area A (Fant 1960: 307; Stevens 1998: 155).

2.2.1 *Articulator regions in a uniform pipe model of the vocal tract*

The volume velocity and the pressure both vary sinusoidally along the axis of a uniform pipe open at one end and closed at the other. The spatial volume velocity is a reciprocal function of the spatial pressure; hence the volume velocity minimum corresponds to a pressure maximum and vice versa. Chiba and Kajiyama (1958: 151) discovered that there was a systematic relationship between the location of a constriction in the vocal tract and the resulting changes in formant frequency:

> "When part of a pipe is constricted, its resonant frequency becomes low or high according as the constricted part is near the maximum point of the volume current or of the excess pressure."

For example, a contraction or an expansion at the lips (volume velocity maximum) lowers or raises F1 respectively whereas a contraction or an expansion at the glottis (pressure maximum) raises or lowers F1 respectively (see Fant 1975 for a detailed examination of perturbation analysis and Story 2006 for further references).

Figure 1 displays the spatial volume velocity magnitudes of the second and third formants in a lossless open-closed uniform pipe. The second formant (F2) shows a volume velocity maximum at one-third the vocal tract length from the glottis and a volume velocity minimum at two-thirds the overall length from the glottis. Likewise, the third formant (F3) shows a volume velocity maximum at three-fifths the vocal tract length from the glottis and a volume velocity minimum at four-fifths the overall length from the glottis. As the constriction locations of F2 and F3 move toward the lips from the volume velocity maximum (pressure minimum) to the volume velocity minimum (pressure maximum), the formant frequencies should shift from their lowest values to their highest values in agreement with perturbation analysis. An examination of Fant's nomograms (1960: Figures 1.4-9 & 1.4-11) reveals that F2

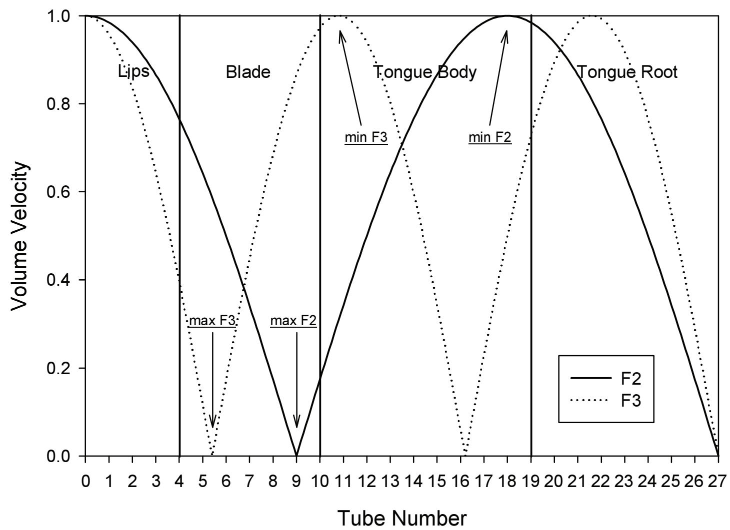

Figure 1. *Articulator regions and spatial volume velocity magnitudes of the second and third formants (F2, F3) in a lossless open-closed uniform pipe. The arrows indicate the constriction locations of the highest and lowest second formant frequencies max F2, min F2 as well as the highest and lowest third formant frequencies max F3, min F3.*

and F3 are raised as expected when the constrictions are moved forward through their respective one-quarter wavelengths.

The vocal tract model of 27 equal-length tubes is partitioned into four regions, each of which corresponds to an active articulator: the lips, the blade, the tongue body, and the tongue root. The lip region consists of 4 tubes (No. 1–4), the blade region 6 tubes (No. 5–10), the tongue body region 9 tubes (No. 11–19), and the tongue root region 8 tubes (No. 20–27). It is useful for the ensuing discussion to introduce a convenient reference length of 18 cm for the 27-tube vocal tract. Each tube is therefore 0.67 cm long. For the third formant, the one-quarter wavelength between three-fifths and four-fifths 18 cm from the glottis is 3.60 cm (= $1/5 \times 18$ cm). This F3 quarter wavelength is modeled by a six-tube blade region 4 cm long (= $6/27 \times 18$ cm). If instead the interval were approximated by five tubes 3.33 cm long (= $5/27 \times 18$ cm), the blade would most likely be too short. Keating (1991), for example, gives an upper bound of 3 to 4 cm for blade length. The entire six-tube blade region is then advanced from the F3 volume velocity maximum in order to avoid excessive overlap with the F2 volume velocity minimum. As a result, the lip region consists of four tubes and is 2.67 cm long (= $4/27 \times 18$ cm). For the second formant, the one-quarter wavelength between one-third and two-thirds 18 cm from the glottis is 6 cm (= $1/3 \times 18$ cm). The F2 quarter wavelength is modeled by the exact interval – a nine-tube tongue body region 6 cm in length (= $9/27 \times 18$ cm). However the entire tongue body region is shifted one tube back from the F2 volume velocity minimum so that the posterior blade remains close to the F3 volume velocity maximum. Hence the tongue root region is made up of eight tubes rather than nine and is 5.33 cm long (= $8/27 \times 18$ cm).

2.2.2 *Articulator regions in a non-uniform pipe model of the vocal tract*

The above partition of the vocal tract into four articulator regions presupposes that the locations of the volume velocity maxima and minima of speech sounds do not deviate markedly from those of the lossless uniform pipe open at one end and closed at the glottis. The closed glottis condition holds when the glottal area is appreciably smaller than the area of the open end, thereby sustaining the odd resonance modes at frequencies on the order of three (F2), five (F3), and seven (F4) times the fundamental resonance frequency (F1) of the vocal tract. A peak glottal area from 0.05 to 0.2 cm^2 is typical of adult voiced sounds whereas glottal areas between 0.1 and 0.4 cm^2 are characteristic of voiceless sounds (Stevens 1998: 35). Using an analysis-by-synthesis procedure, Badin (1989) found the resonance frequencies of [ʃ] to be in the neighborhood of 430 Hz (F1), 1750 Hz (F2), 2680 Hz (F3), and 3200 Hz (F4) with the glottal areas set at both 0.1 and 0.25 cm^2. Clearly, the pattern of odd resonance frequencies is preserved even for the relatively large glottal opening of 0.25 cm^2. On the basis of X-ray area functions, Mrayati and Carré (1976) computed the damped volume velocities of F1, F2, and F3 of eleven French vowels modeled as non-uniform pipes. Although the volume velocity amplitudes along the vocal tract often differed considerably among the synthetic vowels, the locations of the volume velocity maxima and minima shifted little with respect to their locations in the uniform pipe. In sum, the partition of the vocal tract into the four articulator regions appears to be valid for speech sounds because (a) the closed glottis condition is nearly always met and (b) the locations of the volume velocity maxima and minima of a lossy non-uniform pipe largely coincide with those of the uniform pipe in Figure 1.

2.3 *Vowel area functions*

The vowel area functions of ten speakers are taken from seven X-ray and MRI studies (Fant 1960; Mrayati and Guérin 1976; Baer et al. 1991; Yang and Kasuya 1994; Story, Titze and Hoffman 1996, 1998; Takemoto et al. 2006). The area functions were fit to 27 equal-length tubes by means of cubic spline interpolation. The spectra of the original and the interpolated area functions were then compared graphically using the VTAR program. For the first and second formants, the original and the interpolated area functions always showed negligible frequency differences. On the other hand, the third and fourth formants often displayed a good deal of sensitivity to deviations from the original area function.

2.4 *Articulatory and acoustic parameters*

The FVDT program calculates the following ten articulatory parameters (see Figure 1 for the tube numbers):

(1) The lip length parameter $L(lip)$ is the sum of tube lengths in the lip region from tube 1 to 4 (lip protrusion).

(2) The minimum area index $I_{\min A}$ is the integer index of the tube with the smallest constriction in a given articulator region. Therefore the blade and tongue body positions are: $6 \le I_{\min A}(blade) \le 1$ (tubes 5–10), $9 \le I_{\min A}(body) \le 1$ (tubes 11–19).

Table 1.1. *First formant and power correlations. The acoustic-articulatory correlations of each speaker's vowel system are calculated, then the cross-speaker mean and its standard deviation (s.d.) is found and ranked (N = 10).*

First Formant and Power Correlations: ranked means and standard deviations				
Acoustic Parameter	Rank	Articulatory Parameter	mean	s.d.
log *F1*	1	$\log \overline{A}(root)/\overline{A}(lip)$	–0.935	0.063
log *F1*	2	$\log \overline{A}(root)$	–0.915	0.071
log *F1*	3	$\log \overline{A}(lip)$	0.723	0.224
log *Q1*	1	$\log \overline{A}(root)/\overline{A}(lip)$	–0.849	0.121
log *Q1*	2	$\log \overline{A}(root)$	–0.841	0.110
PWR	1	$\log \overline{A}(root)/\overline{A}(lip)$	–0.844	0.097
PWR	2	$\log \overline{A}(root)$	–0.768	0.113

(3) The mean area $\overline{A}$ is the length-weighted mean of the tube areas $A(n)$ in a given articulator region $\overline{A} = \sum_n^m l(n)A(n) / \sum_n^m l(n)$, where the tube indices n, m range from 1 to 4 for the lip area $\overline{A}(lip)$, 5 to 10 for the blade area $\overline{A}(blade)$, 11 to 19 for the tongue body area $\overline{A}(body)$, 20 to 27 for the tongue root area $\overline{A}(root)$. When the tube length l remains constant, the mean area $\overline{A}$ is simply the sum of the tube areas $A(n)$ in the articulator region divided by the number of tubes. The mean area $\overline{A}$ is converted to its base-2 logarithm $\log_2 \overline{A}$.

(4) The three area ratios are the mean areas of the blade, tongue body, and tongue root regions normalized by the mean lip area: $\overline{A}(blade)/\overline{A}(lip)$, $\overline{A}(body)/\overline{A}(lip)$, and $\overline{A}(root)/\overline{A}(lip)$. These area ratios are likewise transformed into their base-2 logarithms.

There are thirteen acoustic parameters. The first four formant frequencies (F1–F4), bandwidths (B1–B4), quality factors (Q1–Q4), and the overall power (PWR) are determined as set forth in Section 2.1. The formant frequency F is converted to $\log_2 F$ in order to account for the log frequency (octave) scales of pitch and formant perception (Miller 1989). As mentioned earlier, log Q ($= 20\log_{10} Q$) is a decibel measure of amplification at resonance.

2.5 *Acoustic-articulatory correlations*

2.5.1 *First formant and power correlations*

To estimate the strength of association between the articulatory and acoustic parameters, Pearson correlation matrices are calculated for the vowel system of each speaker. Then the coefficients of the parameter pairs are averaged across the ten speakers.

In Table 1.1 the frequency parameter log F1 exhibits the largest correlation coefficient with the lip-normalized area ratio $\log \overline{A}(root)/\overline{A}(lip)$ (r = –0.935), followed by $\log \overline{A}(root)$ (r = –0.915) and then by $\log \overline{A}(lip)$ (r = 0.723). The tongue root area $\log \overline{A}(root)$ manifestly

Table 1.2. *Second formant correlations.*

Second Formant Correlations: ranked means and standard deviations				
Acoustic Parameter	Rank	Articulatory Parameter	mean	s.d.
log *F2*	1	$I_{\min A}(body)$	0.617	0.280
log *F2*	2	$L(lip)$	–0.531	0.462
log *Q2*	1	$\log \overline{A}(body)/\overline{A}(lip)$	0.607	0.294
log *Q2*	2	$\log \overline{A}(lip)$	–0.580	0.235

dominates the area ratio $\log \overline{A}(root)/\overline{A}(lip)$. The trading relation between the lip and tongue root areas – as indicated by the ratio $\log \overline{A}(root)/\overline{A}(lip)$ – conforms to perturbation analysis since a wider lip or a narrower tongue root area increases F1 while a narrower lip or a wider tongue root area decreases F1.

The quality factor log Q1 patterns after log F1, the largest coefficient being associated with $\log \overline{A}(root)/\overline{A}(lip)$, and the next largest with $\log \overline{A}(root)$. The close correspondence between log Q1 and log F1 is consistent with the magnitude of the transfer function at resonance $|H(F)| = Q = F/B$ which predicts a linear relation between Q and F on the assumption that B remains constant (Fant 1960: 54). To detect a possible linear increase of log Q1 with log F1, a linear regression analysis is applied to each vowel system. The mean cross-speaker slope is found to be 7.62 dB/octave, not greatly different from the linear slope of 6 dB/octave. Because wall losses diminish with increasing frequency, B1 becomes smaller as F1 increases (Flanagan 1972: 69). Thus the observed mean slope grows somewhat faster than 6 dB/octave.

The overall power PWR also behaves similarly to log F1, with the largest coefficient corresponding to $\log \overline{A}(root)/\overline{A}(lip)$ and the second largest to $\log \overline{A}(root)$. The clear parallel between PWR and log F1 suggests that most of the intrinsic power of vowels is concentrated in the first formant and is therefore nearly equal to $\log|H(F1)|$ $(= 20\log_{10}|H(F1)|)$, or equivalently $\log Q1$ $(= 20\log_{10} Q1)$. A regression analysis of PWR and log F1 yields a mean slope of 8.17 dB/octave, which is reasonably comparable to the slope of log Q1 above.

2.5.2 *Second and third formant correlations*

The mean correlation coefficients for the second and third formants are presented in Tables 1.2 and 1.3. Recall from Section 2.4 that $I_{\min A}(body)$ and $I_{\min A}(blade)$ are the tongue body and blade indices of the tube with the smallest constriction in the articulator region. The tongue body and blade ranges are therefore $9 \le I_{\min A}(body) \le 1$ and $6 \le I_{\min A}(blade) \le 1$, with the larger integer near the lips and the smaller near the glottis. Summarizing the known correspondences between formant frequencies and articulator positions, Fant (1960: 26) stated that a very low or high F2 formant frequency indicates either a retracted articulation or a palatal position of the tongue. As expected, there is a positive correlation (r = 0.617) between log F2 and the tongue body position $I_{\min A}(body)$. He also observed that a very low or high F3 formant frequency signals either a retroflex modification or a prepalatal/dental articulation. Accordingly, there is a

Table 1.3. *Third formant correlations.*

Third Formant Correlations: ranked means and standard deviations				
Acoustic Parameter	Rank	Articulatory Parameter	mean	s.d.
log *F3*	1	$L(lip)$	–0.439	0.438
log *F3*	2	$I_{\min A}(blade)$	0.400	0.314
log *Q3*	1	$\log \overline{A}(blade)/\overline{A}(lip)$	0.611	0.231
log *Q3*	2	$\log \overline{A}(lip)$	–0.499	0.324

positive correlation (r = 0.400) between log F3 and the blade position $I_{\min A}(blade)$. Thus when $I_{\min A}(body)$ and $I_{\min A}(blade)$ increase from 1 to 9 and from 1 to 6, so do the F2 and F3 frequencies in agreement with the perturbation analysis of Section 2.2.1. Note, however, that log F3 and lip length $L(lip)$ achieve a correlation coefficient similar in magnitude (r = –0.439) to that between log F3 and blade position, indicating that lip protrusion may lower F3 as much as blade advancement raises it. Lip protrusion likewise lowers F2, though to a lesser degree on account of the larger correlation between log F2 and tongue body position $I_{\min A}(body)$. Therefore in order to achieve their F3 and F2 targets, the blade and tongue body positions must compensate for changes in the lip length $L(lip)$.

The quality factors log Q2 and log Q3 display the largest correlations with the respective lip-normalized area ratios $\log \overline{A}(body)/\overline{A}(lip)$ and $\log \overline{A}(blade)/\overline{A}(lip)$. To determine how log Q2 and log Q3 change with the area ratios, regression analyses are performed. For log Q2 the mean cross-speaker slope is 2.54 dB per doubling of $\overline{A}(body)/\overline{A}(lip)$ whereas for log Q3 the mean slope is 3.85 dB per doubling of $\overline{A}(blade)/\overline{A}(lip)$.

The validity of these observations is tested by means of a simple two-pipe model. The areas of the tongue body or blade region are varied in an otherwise uniform pipe 4 cm^2 in area and 18 cm long. The six areas of the tongue body pipe (tubes 11–19) or the blade pipe (tubes 5–10) are the following: 8 cm^2, 4 cm^2, 2 cm^2, 1 cm^2, 0.5 cm^2, and 0.25 cm^2. The frequency responses resulting from the area variations of the blade pipe is plotted in Figure 2. As the area of the lip region remains fixed at 4 cm^2, the spectra demonstrate the expected lowering of the formant intensity $\log|H(F3)|$ with increasingly smaller blade areas. Note that the attenuation of the third formant is accompanied by strikingly few changes in the other formants even when the blade area is only 0.25 cm^2. To investigate the two-pipe model further, regression slopes for the second and third formants are calculated. The log Q2 slope is 1.62 dB per doubling of $\overline{A}(body)/\overline{A}(lip)$ and the log Q3 slope is 4.45 dB per doubling of $\overline{A}(blade)/\overline{A}(lip)$. Hence the log Q3 slope of the two-pipe model does not differ substantially from the mean Q3 slope of 3.85 dB per doubling of $\overline{A}(blade)/\overline{A}(lip)$. As pointed out in Section 2.1, radiation bandwidth is proportional to the area of a uniform pipe. This suggests that radiation loss is a plausible mechanism for the behavior of Q2 and Q3. To determine whether the slopes of log Q2 and log Q3 are mainly due to radiation damping, the analysis of the two-pipe model is conducted without radiation loss. The log Q2 slope is then 0.68 dB per doubling of $\overline{A}(body)/\overline{A}(lip)$ while the log

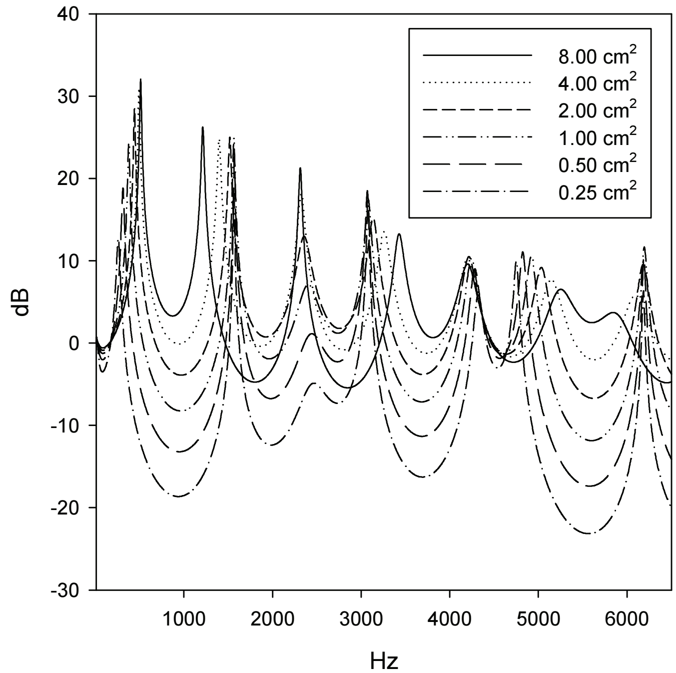

Figure 2. *Spectra of a 4 cm² uniform pipe 18 cm in length, with the blade region (tubes 5–10) varying from 8 cm² to 0.25 cm².*

Q3 slope is 0.19 dB per doubling of $\overline{A}(blade)/\overline{A}(lip)$. In view of these negligible values, both slopes must be governed by radiation damping.

2.5.3 *Fourth formant correlations*

The mean correlation coefficients for the fourth formant are given in Table 1.4. The frequency parameter log F4 and lip length $L(lip)$ yield the correlation coefficient with the largest magnitude (r = –0.388), the negative sign indicating that F4 frequency decreases as lip length increases and vice versa. Changes in lip length also cause inversely proportional shifts of the other formant frequencies, F2 and F3 in particular (cf. Tables 1.2 and 1.3). However, unlike them, the F4 frequency is nearly free from the confounding effects of the lingual articulators. For example, the correlation coefficients of log F4 with the positions of the tongue body $I_{\min A}(body)$ and blade $I_{\min A}(blade)$ are only 0.165 and 0.01 (both s.d. > 0.4). As a consequence, the F4 frequency furnishes the optimal cue for the lip length parameter $L(lip)$. Although log F4 and log F3 attain strong relative correlations with lip length $L(lip)$ and blade position $I_{\min A}(blade)$, respectively, the absolute values are modest in themselves ($r \approx |0.4|$). These moderate correlations are unanticipated because (a) F4, like the other formant frequencies, is inversely proportional to the length of a uniform pipe; (b) F3 and blade position should show the same degree of association as F2 and tongue body position in light of perturbation analysis. In Section 2.3 it was pointed out that the F3 and F4 frequencies often exhibit high sensitivity to small

Table 1.4. *Fourth formant correlations.*

Fourth Formant Correlations: ranked means and standard deviations				
Acoustic Parameter	Rank	Articulatory Parameter	mean	s.d.
log *F4*	1	$L(lip)$	–0.388	0.293
log *F4*	2	$I_{\min A}(body)$	0.165	0.409
log *Q4*	1	$\log \bar{A}(lip)$	–0.619	0.252
log *Q4*	2	$\log \bar{A}(lip)/\bar{A}(root)$	–0.584	0.259

deviations of the vowel area function whereas the F1 and F2 frequencies do not. As a result, the large-scale changes of the area function needed to control F1 and F2 may also give rise to unpredictable shifts in F3 and F4. Extrinsic noise sources include the original measurement error and the cubic spline interpolation of the present study.

The range of lip protrusion can be tentatively estimated by subtracting the mean of the shortest vocal tracts (vowels with drawn lips) from the mean of the longest vocal tracts (vowels with protruded lips) for the seven adult male speakers. The mean of the shortest vocal tracts is 16.43 cm, that of the longest vocal tracts is 18.64 cm. Thus the adult male range of presumed lip protrusion is on average 2.21 cm (s.d. 0.82). However the simplifying assumption that lip protrusion is responsible for all changes in vocal tract length is not warranted. Raising and lowering the larynx can also shorten and lengthen the vocal tract (Ewan and Krones 1974). Hoole and Kroos (1998) point out that there can be substantial interspeaker variability in larynx height. Two of their three male subjects showed a maximum difference of 0.7–1.0 cm between German protruded and drawn vowels, yet the other displayed only a 0.2 cm difference. Remark that the correlation between log F4 and $L(lip)$ (the sum of tube lengths in the lip region) is exactly the same as the correlation between log F4 and the summed tube lengths of any other region – including the entire length of the vocal tract. This reflects the fact that the correlation coefficient is invariant under linear transformations such as proportional length changes. Because F4 frequency is identically correlated with both lip length $L(lip)$ and total vocal tract length, larynx height must also have an acoustic effect. Nevertheless, lip protrusion will be considered the main determinant of vocal tract length since the lips constitute the most visible active articulator (see Rosenblum 2008 for a review of the lips in visual speech perception).

The quality factor log Q4 reveals an inverse correlation (r = –0.619) with the lip area $\log \bar{A}(lip)$. The mean of the log Q4 regression slope is –5.67 dB per doubling of $\bar{A}(lip)$. This value approaches the slope of –6 dB per doubling of area that results from the strict proportionality between radiation bandwidth and the area of a uniform pipe. Hence the slope of log Q4 is governed by radiation damping like the quality factors Q2 and Q3 in the previous section.

2.6 *Acoustic-articulatory correlations: summary*

It is convenient to characterize an articulatory configuration by its motor dimension (position or aperture) as opposed to rather unwieldy areas or area ratios. The seven motor dimensions given below bear some similarity to the oral tract variables suggested by Browman and Goldstein

(1989: 210; see also Goldstein, Byrd and Saltzman 2006). Paired with the corresponding motor dimension, the most strongly correlated acoustic and articulatory parameters are the following:

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

2.7 The *acoustic-articulatory correlations applied to consonants*

The acoustic-articulatory correlations illustrated above were obtained from the vowel area functions of ten speakers. In order to use the same method for noncontinuants (stops, nasals) and continuants (fricatives, liquids, glides), a comparable number of area functions would be needed to produce similarly stable correlation averages. However there are only a few consonant area functions presented as tables in the literature (Narayanan 1995; Story, Titze and Hoffman 1996, 1998), which precludes the direct extension of the method. Nevertheless, according to Section 2.2.2, the vowel acoustic-articulatory correlations can still be applied to consonants on the condition that the vocal tract approximates a uniform pipe open at the mouth and closed at the glottis, the odd resonance modes being on the order of three (F2), five (F3), and seven (F4) times that of the fundamental resonance frequency (F1). When the open-closed vocal tract is excited at the glottis (glottal source), the system response shows odd resonance modes and no zeros (spectral minima). If the excitation occurs at another point along the vocal tract (supraglottal source), then the system response displays zeros as well as the odd resonance modes (Flanagan 1972: 72–73). Fant (1960: 169) summarizes the acoustic consequences of the open-closed vocal tract and a glottal or supraglottal source:

"Acoustically, the common denominator of all sounds produced from a resonator system of a prescribed configuration is the particular set of formant frequencies of the vocal tract, i.e., the F-pattern. The differences in location of the source and the spectrum envelope of the source will only influence the relative intensity levels of the formants."

Consonants are composed of at most an approach transient phase, a hold phase, and a release transient phase (Hardcastle 1976: 134–137). The transient phase of the voiced stop, signaled by an abrupt spectral discontinuity, contains critical perceptual cues for place of

Table 1.5.1 *The formant frequencies (F1-F4) and quality factors (Q1-Q4) from the area function of [t] (Story, Titze and Hoffman 1996). The second column gives the varying areas of tube 5 in the 27-tube FDVT model. The glottal source is tube 27, the supraglottal source tube 5.*

[t] Source	Area cm^2	F1 Hz	Q1 dB	F2 Hz	Q2 dB	F3 Hz	Q3 dB	F4 Hz	Q4 dB	PWR dB
glottal	0.0125	235	11.9	1674	38.0	2656	40.9	3281	18.0	31.5
supraglottal	0.0125	247	10.5	1681	15.3	2663	28.2	3282	20.3	–6.7
glottal	0.025	264	14.1	1680	38.1	2666	40.6	3335	20.3	36.6
supraglottal	0.025	270	13.8	1686	31.3	2672	36.7	3327	20.3	–2.2
glottal	0.05	303	17.4	1689	37.7	2680	39.3	3423	18.5	40.9
supraglottal	0.05	306	17.2	1693	35.3	2685	37.9	3401	20.6	2.5
glottal	0.10	347	20.7	1700	38.2	2694	38.2	3533	18.9	44.2
supraglottal	0.10	349	20.5	1703	36.7	2698	37.3	3495	21.3	7.1
glottal	0.20	387	23.1	1711	38.2	2701	36.8	3610	22.1	46.2
supraglottal	0.20	388	23.2	1713	37.1	2705	36.0	3569	22.7	10.7
glottal	0.40	418	25.2	1719	37.9	2695	35.7	3623	25.2	47.5
supraglottal	0.40	418	25.2	1721	37.5	2699	35.0	3588	24.1	13.1

Table 1.5.2 *The formant frequencies (F1-F4) and quality factors (Q1-Q4) from the area function of [k] (Story, Titze and Hoffman 1996). The second column gives the varying areas of tube 12 in the 27-tube FDVT model. The glottal source is tube 27, the supraglottal source tube 12.*

[k] Source	Area cm^2	F1 Hz	Q1 dB	F2 Hz	Q2 dB	F3 Hz	Q3 dB	F4 Hz	Q4 dB	PWR dB
glottal	0.05	327	16.5	1540	27.2	2205	37.6	3691	39.7	40.0
supraglottal	0.05	332	16.4	1538	27.3	2212	34.6	3706	28.3	–2.2
glottal	0.10	337	17.1	1544	27.4	2202	37.6	3670	39.5	40.6
supraglottal	0.10	341	17.0	1543	27.4	2208	38.0	3684	27.0	–1.3
glottal	0.20	343	17.5	1547	27.3	2190	37.3	3625	39.4	40.4
supraglottal	0.20	347	17.4	1545	27.4	2196	35.3	3639	25.7	–1.0

Table 1.5.3 *The formant frequencies (F1-F4) and quality factors (Q1-Q4) from the area function of [p] (Story, Titze and Hoffman 1996). The second column gives the varying areas of tube 1 in the 27-tube FDVT model. The glottal source is tube 27, the supraglottal source tube 1.*

[p] Source	Area cm^2	F1 Hz	Q1 dB	F2 Hz	Q2 dB	F3 Hz	Q3 dB	F4 Hz	Q4 dB	PWR dB
glottal	0.05	313	17.4	1237	35.4	2176	39.5	3259	41.6	35.0
supraglottal	0.05	316	17.3	1238	35.0	2182	35.4	3269	23.5	1.0
glottal	0.10	365	20.9	1281	35.3	2192	39.2	3265	41.3	39.7
supraglottal	0.10	366	20.9	1282	35.3	2195	37.9	3273	34.0	6.5
glottal	0.20	420	24.2	1346	35.7	2217	39.0	3274	41.4	43.7
supraglottal	0.20	421	24.2	1346	35.7	2220	38.6	3280	38.1	11.7

articulation (Blumstein and Stevens 1980). Since the voiceless stop has a silent hold phase, all the acoustic information within the stop segment must be attributed to its transient phases (Stevens and Blumstein 1981: 3). If the transient phase of stops can be shown to exhibit odd resonance modes like vowels, this would provide strong evidence that the acoustic-articulatory correlations remain valid. To test the premise, the area functions of [t], [k], and [p] (Story, Titze and Hoffman 1996) were fit to 27 equal-length tubes of the FVDT model. The areas of the tube corresponding to the stop closure are set to 0.05, 0.1 and 0.2 cm^2, thereby simulating a stop transient or fricative (cf. the 0.05–0.2 cm^2 fricative range of Stevens 1998: 33). The areas of the [t] closure are additionally set to 0.0125, 0.025 and 0.4 cm^2. To evaluate the effects of spectral zeros, each area function is calculated with a glottal or a supraglottal source. Similarly to the vowels earlier, the volume velocity source at the closed glottis (tube 27) is assigned the value of 1 for all frequencies. The series pressure sources of [t] (tube 5), [k] (tube 12) and [p] (tube 1) are also assigned the value of 1 at all frequencies (for the series connection of the supraglottal pressure source, see Fant 1960: 36; Flanagan 1972: 54; Stevens 1998: 102).

The results for [t], [k], and [p] are shown in Tables 1.5.1, 1.5.2, and 1.5.3. A brief examination of the data reveals that both the glottal and the supraglottal sources excite the odd resonance modes of the whole vocal tract, even when the area of the [t] closure is only 0.0125 cm^2. Thus regarding sibilant fricatives like [s ʃ], the following assumption by Toda, Maeda and Honda does not appear tenable (2010: 343): “since the generation of fricatives’ source requires a narrow constriction, the back cavity located behind the constriction tends to be acoustically inactive.” Moreover, the statement is incompatible with the fact that tongue body gestures in the back cavity determine F2 frequency, the main cue distinguishing palatalized and nonpalatalized sibilants (Shupljakov, Fant and de Serpa-Leitão 1968). Because telephone speech bandlimited to 4000 Hz is still intelligible, frequencies in the range of the first four formants are much more important than higher ones (cf. the 8000 Hz sampling rate of ITU-T Recommendation G.711).

As expected, formant frequency differences between the glottal and supraglottal source are slight. Quality factor differences between the two source locations are also minor, with two classes of exceptions. For [k] and [p], the supraglottal Q4 values are considerably less than the glottal ones; this is consistent with Fant’s noise spectrum slopes of –6 to –3 dB/octave for dorsal and labial obstruents but 0 db/octave for coronals (1960: 202–203). Recall from Section 2.6 that Q4 is a weaker correlate of lip aperture than F1 frequency. For the smallest [t] closures of 0.0125 and 0.025 cm^2, the supraglottal Q2 and Q3 values are also less than their glottal counterparts.

On the basis of pattern-playback and tape-splicing experiments, Fant (1960: 148) concluded that “the formant transitions in the oral sound intervals next to the nasal murmur are the most prominent auditory cues for differentiating the various nasal phonemes.” Therefore like the stop, the transient phase is a significantly better identifier of the nasal segment than the hold phase (see Ohde, Haley and Barnes 2006 for current evidence). On the other hand, the hold phase of nasals can carry contrastive tone in a number of African languages (Welmers 1973: 66–68). To analyze the transient phase of nasals, a side branch is placed in parallel with tube 20 of the FDVT model, the most anterior tube of the tongue root region. The nasal pipe, 19 elementary vocal tract (VT) tubes in length, consists of three tubes: the nasal port (variable area, 4 VT tubes long), the nasal cavity (3 cm^2, 11 VT tubes long), and the nostrils (1.5 cm^2, 4 VT tubes long); the shape factors of the port, cavity, and nostril tubes are fixed at 1, 4, and 2, respectively (Dang, Honda and Suzuki 1994). Without extra damping the calculated nasal resonance in the range of 550–1100 Hz would be too pronounced because nasal peaks exceeding the first formant are not often observed in nasalized vowel spectra (Chen 1997). To keep the level of the nasal peak

Table 1.5.4 *The formant frequencies (F1-F4) and quality factors (Q1-Q4) from the area functions of [t k p] (Story, Titze and Hoffman 1996), with a transient closure area of 0.05 cm², two nasal port areas (0.05 and 0.2 cm²), and a glottal source.*

Transient 0.05 cm²	Nasal Port cm²	F1 Hz	Q1 dB	F2 Hz	Q2 dB	F3 Hz	Q3 dB	F4 Hz	Q4 dB	PWR dB
nasal [t]	0.05	331	16.2	1692	37.7	2682	39.6	3424	18.3	40.5
	0.20	361	8.7	1699	36.6	2686	38.8	3426	17.7	39.7
nasal [k]	0.05	392	14.1	1540	27.4	2205	37.6	3689	39.5	39.6
	0.20	522	6.5	1539	27.4	2205	40.8	3684	37.7	39.7
nasal [p]	0.05	343	16.4	1244	35.5	2177	39.5	3260	41.3	33.1
	0.20	375	9.3	1266	33.1	2177	39.5	3263	40.2	31.7

generally below that of the first formant, a resistance of 30 cgs acoustic ohms is added in series to the nasal cavity tube. For easier comparison, the transient phase of nasals is modeled using (a) the area functions of [t k p] above, (b) an oral closure set to 0.05 cm^2 like the narrowest fricative, and (c) a glottal source. Two nasal port areas are investigated: (1) Stevens' estimate of 0.2 cm^2 for nasal consonants (1998: 487) and (2) a markedly smaller one of 0.05 cm^2 given the negligible difference between the first formant peaks of Central Arrernte nasals and laterals (cf. Table 8 below). The details of the nasal modeling are supplied in Table 1.5.4. With respect to the corresponding stops, the data sets indicate that Q1 decreases and F1 increases as the area of the nasal port grows larger while the other quality factors Q2–Q4 and formant frequencies F2–F4 remain almost unchanged. The results are in accord with those of House and Stevens (1956) as well as later nasalization studies (for a review, see Pruthi 2007: Chapter 2).

Thus for the transient phase of noncontinuants (stops, nasals) and the hold phase of fricative continuants, the FDVT simulations show that the odd resonance modes are excited along the entire vocal tract even when the smallest oral constriction is 0.05 cm^2 or less. For the hold phase of liquid [l ɹ] and glide [j w] continuants with a typically larger minimum oral constriction (Stevens 1998: 545), further simulations are not necessary since their vowel-like formant structure has already been demonstrated by both measurement and modeling work (Lehiste 1964; Story, Titze and Hoffman 1996, 1998). Hence, in principle, the vowel acoustic-articulatory correlations are equally applicable to all consonants.

3. Experimental evidence for blade features

3.1 *The apical-laminal contrast*

In an important antecedent to current views on the apical-laminal distinction, Sweet (1877: 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*.

As was discussed in the introduction, Chomsky and Halle (1968: 312) proposed the binary feature [distributed], where [+distributed] and [–distributed] designate a long and short constriction along the direction of the air flow. The authors then state that "phonetics books traditionally distinguish apical from laminal and retroflex from nonretroflex consonants. As a first approximation (to be further refined below), we class the former as [–distributed] and the latter as [+distributed]." When the length of the blade constriction increases or decreases, the blade aperture (blade area/lip area) likewise decreases or increases provided that the lip area remains constant. Hence the feature [distributed] captures the contrast between the small blade aperture (+) of laminal sounds and the medium (–) blade aperture of apical sounds. 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). As evidence for this, the tip and blade heights of [s z t d n l] are found to vary almost inversely in a cinefluorographic analysis of British English (Bladon and Nolan 1977: Figure 5). Because blade aperture and its acoustic correlate Q3 are continuous parameters, apical-laminal articulations form a phonetic gradient. For example, in the Bladon and Nolan study, one speaker produced more apicalized variants of regularly laminal [s z] than the other seven. Therefore only a phonological contrast (established using minimal pairs, for example) can define the phonetic boundary between apical and laminal sounds. As will be demonstrated in the rest of the paper, the palatographic and third-formant peak measures indicate that contrasting /s/ and /ʃ/ are respectively laminal and apical without exception (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: 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 or constriction length alone, as was suggested in the preceding paragraph. In a palatographic investigation of *sip-ship*, Ladefoged (1957: 773) found that each of the 164 speakers had a narrower channel for [s] than for [ʃ]. Gafos (1999: 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.

An MRI survey of American English voiceless and voiced fricatives was carried out by Narayanan (1995). Complete area functions of [s z] and [ʃ ʒ] were acquired from a man and woman (Tables 6.9–6.12). The eight area functions are adjusted to 27 equal-length tubes by cubic spline interpolation. Then the FVDT program computes the following parameters: F3 (Hz), Q3 (dB), mean blade area (cm^2), blade aperture (blade area/lip area). A glottal source is used since the formants of glottal and supraglottal sources differ little when the minimum blade constriction ≥ 0.05 cm^2 (cf. Table 1.5.1). The calculated F3 frequencies of male [s] and [ʃ] are respectively 2593 and 3369 Hz, the Q3 values 41.8 and 28.8 dB, the mean blade areas 1.73 and 0.81 cm^2, and the blade apertures 2.27 and 0.53. Note that the minimum constriction areas of [s] and [ʃ] are 0.18 and 0.16 cm^2 in the interpolated area function, 0.14 and 0.16 cm^2 in the original area function. There is a striking discrepancy in blade elevation between the MRI and the palatographic results. In the MRI data the mean blade area of [s] is more than twice that of [ʃ] whereas palatographic methods uniformly find that [s] is characterized by a narrower channel and more extensive contact than [ʃ]. The reversal of the expected results occurs throughout the data set. For instance, the mean blade areas of male [z] and [ʒ] are respectively 1.6 and 1.58 cm^2, those of female [s] and [ʃ] 1.0 and 0.54 cm^2, those of female [z] and [ʒ] 0.9 and 0.65 cm^2. The MRI data set is even more problematic in that all the Q3 values increase with mean blade area and blade aperture. This provides support for Q3 as the acoustic correlate of blade aperture, but also entails that [s] will have a stronger third-formant peak than [ʃ]. However Stevens (1985: 248) observed that the amplitude of third-formant energy is weaker for [s] than for [ʃ] in English (cf. similar findings in Section 3.4.1 below). In sum, the proposed area functions of [s z] and [ʃ ʒ] appear subject to serious measurement error.

Blade articulations can be accompanied by a sublingual cavity of variable size, but there is little information on its acoustic effect. Espy-Wilson et al. (2000) modeled the sublingual space of American English /r/ by joining a tube of the same volume to the front cavity, in either a series or a parallel connection. The volume included not only the sublingual cavity proper, but also any airspace along the sides of the tongue. Both connections lowered F3 by a similar amount (200–300 Hz), which suggests that the choice of connection is not crucial. Nevertheless, it is not clear how much the sublingual cavity proper contributed to the F3 lowering since the side airspace was tallied as part of the sublingual volume. Ladefoged and Maddieson (1996: Chapter 5) provide representative sagittal sections of strident coronal fricatives. Two kinds of coronal fricatives have negligible sublingual cavities, (1) those with the blade approaching or contacting the lower teeth (Figures 5.9, 5.11, 5.18) and (2) those with the blade almost perpendicular to the mandible (Figures 5.12, 5.14, 5.16). The two types make up nearly all of their sample. Only the Toda fricatives illustrated in Figure 5.13 (the laminal dental) and Figure 5.15 (the palatalized apical postalveolar) appear to have significant sublingual cavities. But as will be seen in Section 3.4.2, the two fricatives display somewhat weaker third-formant peaks than the other Toda laminals or apicals, a possible indication that their sublingual cavities play only a minor acoustic role.

3.2 *Auditorily-based spectral analysis*

The palatographic evidence presented earlier 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.

The auditorily-based analysis covers the range of formant frequencies generally observed for men, women and children. Bell (1867: 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 related to F2 frequency, Bell's proposal should entail nine phonetic distinctions in F1 and F2 formant frequency. 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 $ERB = f_c / 9.26 + 24.7$, 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

Table 2.1. *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).*

	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

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 2.1, 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). The man-to-woman factors are in substantial conformity with the 1.2 ratio of adult male (16.93 cm) and female (14.09 cm) vocal tract lengths (Goldstein 1980: 186), given the inverse proportionality between the length of an open-closed uniform pipe and its resonance frequencies. 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, Potamianos and Narayanan (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: 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: $PE_i = E_i / \sqrt{E_{i-8} E_{i+8}}$. 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

Table 2.2. *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-sixth octaves (= 0.5 octave). The base frequency and scale factors are the same as those in Table 2.1.*

	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

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 2.1, 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, Nord and Branderud (1976: 18). Within each of the 28 one-sixth octaves, the maximum PE_i of 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).

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

In Table 2.1 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 2.2 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

Table 3.1. *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.*

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	–	+	–	+

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 Toda dental /s̪/ reaches 3228 Hz (see Table 6 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 7 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 2.1 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 summary of a acoustic-articulatory correlations (Section 2.6) 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 3.1 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 3.1 are [+anterior –posterior] while the postalveolars are [–anterior +posterior]. The anterior blade region comprises both the dental and alveolar subplaces. Keating (1991: 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

Table 3.2. *The equipollent feature pair of blade aperture [elevated depressed].*

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

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

As reviewed in Section 3.1, 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 from the summary in Section 2.6 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 above. The equipollent feature pair of blade aperture [elevated depressed] is presented in Table 3.2. 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].

3.4 *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: 156–163). Central Arrernte like many other Australian languages has four coronal stops, nasals, and lateral continuants, of which two are laminal and two are apical (cf. Ladefoged and Maddieson 1996: 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

Table 4. *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 2.2. As before, the base frequency is 256 Hz, the man-to-woman and woman-to-child scale factors 1/4 octave (18.9%).*

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—

across different contexts. Let us assume that | indicates a between-category boundary along the increasing real-valued acoustic dimension {D}:

#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 3.2. 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 2.2. 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 4, 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.

3.4.1 *American English*

In Section 3.1, palatographic data were provided showing that the blade area of [s] is smaller than the blade areas of [ʃ] and [s̺]. Hence in accordance with Table 3.2, 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 3.2. 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 man and woman speakers are coded as M4_NW and W4_JW; in their description of the recordings Shannon et al. 1999 state that "talkers were chosen who had no noticeable regional accent – standard American Midwest dialect."). 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. Consonants are composed of at most one approach transient subsegment, one or more hold subsegments, and one release transient subsegment (cf. Section 2.7). 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 intensity is selected for analysis because it offers the best environmental signal-to-noise ratio.

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 5.1 (man) and 5.2 (woman). A one-way analysis of variance (ANOVA) is performed on the nine tokens 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 summary, 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 2.6).

Table 5.1. *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.*

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

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

(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 2.6). 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 3.2. 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 2.6). Ladefoged and Maddieson (1996: 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 5.1 the apical and laminal sounds have almost identical F3 frequencies, whereas in the female Table 5.2 the apicals [ʃ ʒ] have a lower F3 frequency than the laminals [s z] but not significantly so. In Table 4, 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, Table A1 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 Table A1 follow the pattern of those in Tables 5.1 and 5.2 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 $|H(F)|$, the magnitude of the transfer function at resonance (Fant 1960: 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, as discussed in Section 2.7. 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

The A3 amplitudes of Polish in Table A1 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 = ʃʲ ʒʲ, see Hall 1997: Section 2.4 for references).

Two kinds of American English continuant /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 4, 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). 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]

Table 6. *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.*

	F1 Hz	PE1 dB	F2 Hz	PE2 dB	F3 Hz	PE3 dB	F4 Hz	PE4 dB	ZCR Hz	relINT 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

3.4.2 *Toda*

Shalev, Ladefoged and Bhaskararao (1994) used static palatography to analyze the four strident coronal fricatives of Toda: s̪ s̠ ʃ s̢. The s̠ 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 s̠ to be an apical 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[1]:

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

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). Since the ZCR grows with increasing turbulent fluctuations (Sreenivasan, Prabhu and Narasimha 1983), it likewise appears to be a useful 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 male Toda coronal fricatives are illustrated in Table 6. 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 6 are categorized according to Table 4, the blade features of Toda are as follows:

/θ/ {θ} nonstrident laminal dental [+anterior, +AB] [+elevated –depressed]
/s̻/ {s̻} laminal dental [+anterior, +AB] [+elevated –depressed]
/s̺/ {s̺} apical postalveolar [–anterior, –RB] [–elevated –depressed]
/s̺ʲ/ {ʃ} palatalized apical postalveolar [–anterior, –RB] [–elevated –depressed]
/s̺̠/ {s̺} apical retracted postalveolar [–anterior, +RB] [–elevated –depressed]

The lower values of relINT and ZCR keep the nonstrident fricative {θ} distinct from the strident ones {s̻ s̺ ʃ s̺}. Also in view of the small PE3, both {θ} and {s̻} are laminal in comparison to apical {s̺ 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: 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 4. In consequence, there is good agreement between the phonological and acoustic classifications.

3.4.3 *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: 162–163). The non-IPA symbols {ŝ ẑ} are used to denote ‘hissing-hushing’ fricatives, following Catford (1977: 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[2]:

sa:ba ‘why’ za ‘one’
ŝa ‘three’ ẑaẑa ‘kidney’

Table 7. *The formant frequencies and peak energy factors of the male Ubykh coronal fricatives, each uttered once.*

	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
$ʃ^j$ {ɕ}	427	1.998	2282	3.121	2383	4.870	3571	0.164	1651	–16.9
ʂ {ʂ}	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
$ʒ^j$ {ʑ}	427	0.280	2562	1.830	2599	1.866	3623	–0.12	102	–10.0
ʐ {ʐ}	466	5.462	1614	1.085	1710	2.369	2834	0.007	254	–5.9

ɕaɕa 'mother-in-law' ʑawa 'shadow'

ʂa 'head' ʐa 'firewood'

The single-token measurements of the male Ubykh fricatives are given in Table 7. When the F3 frequencies are organized according to Table 4, 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]

/$ʃ^j$ $ʒ^j$/ {ɕ ʑ} palatalized apical alveolar [+anterior, –AB] [–elevated –depressed]

/ʂ ʐ/ {ʂ ʐ} apical postalveolar [–anterior, 0AB 0RB] [–elevated –depressed]

The small PE3 values of laminal {s z} are opposed to the medium PE3 values of apical {ŝ ẑ, ɕ ʑ, ʂ ʐ}. Remark that the PE3 values of {ŝ ʂ} 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 /$ʃ^j$ $ʒ^j$/ are distinctly lower than those of /ʃ ʒ/, they still remain within the alveolar category [+anterior, –AB]. Furthermore, the much higher F2 frequencies of /$ʃ^j$ $ʒ^j$/ clearly make these sounds the palatalized versions of /ʃ ʒ/. Hence the Ubykh alveolo-palatals {ɕ ʑ} are simply the palatalized apical alveolars /$ʃ^j$ $ʒ^j$/ (cf. the brief discussion of Polish in Section 3.4.1). The F3 frequencies of the voiceless and voiced postalveolars /ʂ ʐ/ 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: /ʂ ʐ/ → [0AB 0RB].

3.4.4 *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. The coronal stops, nasals, lateral continuants, 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 8 to enhance clarity:

/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 Sections 2.5.2 and 2.5.3, 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 Section 2.6:

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 reaches zero, which renders the corresponding quality factor inoperative as a cue. Hence for noncontinuants (stops, nasals), only the approach and release transients can supply meaningful measures of PE2, PE3, and PE4. The release subsegments of the Central Arrernte stops and nasals are the ones analyzed since they are all of greater intensity than the approach subsegments.

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

Table 8. *The formant frequencies and peak energy factors of the male Central Arrernte stops, nasals, lateral continuants, and rhotics, each uttered once.*

	F1 Hz	PE1 dB	F2 Hz	PE2 dB	F3 Hz	PE3 dB	F4 Hz	PE4 dB	ZCR Hz	relINT 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

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 + continuant 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 rhotic continuant [ɹ] (Walsh Dickey 1997: Table 3.3). Of the two 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 8 are categorized in light of Table 4, the blade features of the stops, nasals, lateral continuants, and rhotics become:

/t̻ n̻ l̻/ {t̪ n̪ l̪} laminal alveolar	[+anterior, –AB] [+elevated]
/t̻ʲ n̻ʲ l̻ʲ/ {tʲ nʲ ɭ} 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: 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. Anderson (2000: Chapter 3) gives palatographic data from Western Arrernte showing that midline contact length is shorter for the apical stops and nasals than for the laminal ones.

Because the blade positions are categorized as alveolar [+anterior, –AB], the anterior-nonanterior distinction is not relevant for the stops, nasals, and laterals (excepting perhaps /t̺/). Only the apical-laminal and palatalized-nonpalatalized distinctions are contrastive. A palatalized-nonpalatalized distinction between the traditionally termed apical alveolars /t̺ʲ n̺ʲ l̺ʲ/ and apical postalveolars /t̺ n̺ l̺/ has been reported in two other Australian languages. For Yanyuwa and especially Yindjibarndi, Tabain and Butcher (1999) found the apical alveolar stops to have higher vowel-onset F2 frequencies than the apical postalveolars although the difference is not robust. On the other hand, the anterior-nonanterior distinction does appear to be contrastive for the Central Arrernte rhotics. Given their small PE3 values, the alveolar trill /r/ and postalveolar /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. The pervasive neutralization that affects the anterior-nonanterior opposition in this Central Arrernte sample is apparently not attested in Western Arrernte. For example, the dental and postalveolar subplaces of stops and nasals are sharply distinguished by a palatographic measure of frontmost contact (Anderson 2000). In Wubuy the dentalveolar and postalveolar stops are kept distinct by a similarly unambiguous F3 difference at consonantal closure (Bundgaard-Nielsen et al. 2012: Table 9).

3.5 *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 (Maddieson 1984). Toda with three phonemic postalveolar fricatives /ʂ ʂʲ ʂ̠/ constitutes an obvious counter-example. Observe also the subphonemic distinction in American English between the two postalveolar rhotic continuants: 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 3.1. The Toda and Ubykh postalveolar fricatives are all apical. However the Central Arrernte

postalveolar /ɹ̻/ as well as the American English postalveolar 'bunched' [ɹ̻] 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: 93–94). Toda, for instance, makes no more than a two-way distinction among nonpalatalized postalveolar fricatives (nonretracted /s̺/ vs. retracted /s̠/).

(5) There appears to be a general preference for more apical fricatives than laminal fricatives within each voicing type. For example, Polish /s ʃ ʃʲ/ has one laminal and two apical fricatives (Section 3.4.1) whereas Ubykh /s ʃ ʃʲ s̠/ and Toda /s s̺ s̺ʲ s̠/ 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.

4. Concluding remarks

The acoustic-articulatory correlations obtained in the first part of this paper demonstrate that the four formant frequencies (F1–F4) and quality factors (Q1–Q4) are critical parameters in speech. The findings help explain why telephone speech bandlimited to 4000 Hz remains intelligible, seeing that the first two formants and the third formant of men and women fall within this range (cf. Section 2.7 and Tables 2.1-2.2). In the second part of the study the formant frequencies (F1–F4) and peak energy factors (PE1–PE4) are determined for the coronal fricatives of American English, Toda, and Ubykh as well as for the coronal stops, nasals, and liquids of Central Arrernte. The blade positions specified by the calculated third formant frequencies in Table 4 generally correspond quite well to those expected from the phonetic descriptions. For example, Toda /s̪/ and /θ/ are categorized as dentals because of their very high F3 measures (3228 and 3091 Hz, respectively). Both sounds are equally assessed as dentals based on the transcription and palatographic data of Shalev, Ladefoged and Bhaskararao (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 3.1). For example, Heffner wrote in a standard textbook of phonetics (1949: 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: 6; Hall 1997: 98):

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

The features [+distributed] and [–distributed] correspond respectively to laminal and apical coronals as discussed in Section 3.1. 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:

s̪ 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 [s̪ 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 [s̪ s] set is dentalveolar while the [ʃ s̺] set is postalveolar. In the equipollent analysis, the [s̪ 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 [s̪ 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 3.4.1 for American English). Consequently, the equipollent feature analysis appears to be the correct one.

Acknowledgements

I wish to thank Stuart Davis, Samuel Obeng, Robert Port and Charles Watson for reading the original report and giving me their valuable comments and criticisms. I would also like to express my gratitude to Kenneth de Jong who provided a great deal of helpful feedback, not only for this abridged and updated version of the paper, but for the earlier two-part report as well.

Correspondence e-mail address: mwpennin@indiana.edu

Notes

1. http://archive.phonetics.ucla.edu/Language/TCX/tcx.html: tcx_word-list_1992_09.wav.
2. http://archive.phonetics.ucla.edu/Language/UBY/uby.html.

Appendix

Table A1. *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.*

	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
ç	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

References

Aarts, Ronald M. & Augustus J. E. M. Janssen. 2003. Approximation of the Struve function H_1 occurring in impedance calculations. *Journal of the Acoustical Society of America* 113(5). 2635–2637.

Anderson, Victoria B. 2000. *Giving weight to phonetic principles: The case of place of articulation in Western Arrernte*. Ph.D. dissertation, University of California, Los Angeles.

Badin, Pierre. 1989. Acoustics of voiceless fricatives: Production theory and data. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *Royal Institute of Technology, Stockholm* 30(3). 33–55.

Baer, T., J. C. Gore, L. C. Gracco, & P. W. Nye. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *Journal of the Acoustical Society of America* 90(2). 799–828.

Baltaxe, Christiane A. M. 1978. *Foundations of distinctive feature theory*. Baltimore: University Park Press.

Bell, Alexander M. 1867. *Visible speech*. London: Simkin, Marshall & Co.

Bladon, R. A. W. & F. J. Nolan. 1977. A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics* 5. 185–193.

Blumstein, Sheila E. & Kenneth N. Stevens. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America* 67(2). 648–662.

Breen, Gaven & Veronica Dobson. 2005. Central Arrernte. *Journal of the International Phonetic Association* 35(2). 249–254.

Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6(2). 201–251.

Bundgaard-Nielsen, Rikke L., Brett J. Baker, Christian Kroos, Mark Harvey & Catherine T. Best. 2012. Vowel acoustics reliably differentiate three coronal stops of Wubuy across prosodic contexts. *Laboratory Phonology* 3. 133–161.

Catford, John C. 1968. The articulatory possibilities of man. In Bertil Malmberg (ed.), *Manual of phonetics*, 309–333. Amsterdam: North Holland.

Catford, John C. 1977. Mountain of tongues: The languages of the Caucasus. *Annual Review of Anthropology* 6. 283–314.

Chen, Marilyn Y. 1997. Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America* 102(4). 2360–2370.

Chiba, Tsutomu & Masato Kajiyama. 1958. *The vowel, its nature and structure*. Tokyo: Phonetic Society of Japan.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. 2009. The role of features in phonological inventories. In Eric Raimy & Charles E. Cairns (eds.), *Contemporary views on architecture and representations in phonology*, 19–68. Cambridge: MIT Press.

Clements, G. N., and Elizabeth V. Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 245–306. Cambridge: Blackwell.

Cohn, Abigail, C. 2011. Features, segments, and the sources of phonological primitives. In G. Nick Clements & Rachid Ridouane (eds.), *Where do phonological features come from. Cognitive, physical and developmental bases of distinctive speech categories*, 13–42. Amsterdam: John Benjamins.

Dalston, Rodger M. 1975. Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *Journal of the Acoustical Society of America* 57(2). 462–469.

Dang, Jianwu, Kiyoishi Honda & Hisayoshi Suzuki. 1994. Morphological and acoustical analysis of the nasal and the paranasal cavities. *Journal of the Acoustical Society of America* 96(4). 2088–2100.

Dart, Sarah N. 1991. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics* 79.

Delattre, Pierre & Donald C. Freeman. 1968. A dialect study of American English r's by X-ray motion picture. *Linguistics* 6(44). 29–68.

Divenyi, Pierre L. 2004. The times of Ira Hirsh: Multiple ranges of auditory temporal perception. *Seminars in Hearing* 25(3). 229–239.

Dixit, R. Prakash & Paul R. Hoffman. 2004. Articulatory characteristics of fricatives and affricates in Hindi: An electropalatographic study. *Journal of the International Phonetic Association* 34(2). 141–159.

Espy-Wilson, Carol Y., Suzanne E. Boyce, Michel Jackson, Shrikanth Narayanan & Abeer Alwan. 2000. Acoustic modeling of American English /r/. *Journal of the Acoustical Society of America* 108(1). 343–356.

Ewan, William G. & Robert Krones. 1974. Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics* 2. 327–335.

Fant, Gunnar. 1960. *Acoustic theory of speech production*. The Hague: Mouton.

Fant, Gunnar. 1966. A note on vocal tract size factors and non-uniform F-pattern scalings. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *Royal Institute of Technology, Stockholm* 7(4). 22–30.

Fant, Gunnar. 1975. Vocal-tract area and length perturbations. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *Royal Institute of Technology, Stockholm* 16(4). 1–14.

Fant, Gunnar. 1986. Features: fiction and facts. In Joseph S. Perkell & Dennis H. Klatt (eds.), *Invariance and variability in speech processes*, 480–492. Hillsdale, New Jersey: Lawrence Erlbaum.

Fant, Gunnar, L. Nord & P. Branderud. 1976. A note on the vocal tract wall impedance. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *Royal Institute of Technology, Stockholm* 17(4). 13–20.

Flanagan, James L. 1972. *Speech analysis, synthesis and perception*. Berlin: Springer Verlag.
Fletcher, Samuel G. & Dennis G. Newman. 1991. [s] and [ʃ] as a function of linguapalatal contact place and sibilant groove width. *Journal of the Acoustical Society of America* 89(2). 850–858.
Flemming, Edward S. 2002. *Auditory representations in phonology*. New York: Routledge.
Gafos, Adamantios I. 1999. *The articulatory basis of locality in phonology*. New York: Garland.
Glasberg, Brian R. & Brian C. J. Moore. 1990. Derivation of auditory filter shapes from notched-noise data. *Hearing Research* 47(1–2). 103–138.
Goldstein, Ursula G. 1980. *An articulatory model for the vocal tracts of growing children*. Ph.D. dissertation, Massachusetts Institute of Technology.
Goldstein, Louis, Dani Byrd, & Elliot Saltzman. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In Michael A. Arbib (ed.), *Action to language via the mirror neuron system*, 215–249. Cambridge: Cambridge University Press.
Hall, T. Alan. 1997. *The phonology of coronals*. Amsterdam: John Benjamins.
Halle, Morris, Bert Vaux. & Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31(3). 387–444.
Hamann, Silke R. 2003. *The phonetics and phonology of retroflexes*. Ph.D. dissertation, Utrecht University.
Handbook of the International Phonetic Association. 1999. Cambridge: Cambridge University Press.
Hardcastle, William J. 1976. *Physiology of speech production. An introduction for speech scientists*. London: Academic Press.
Hartmann, William M. 1998. *Signals, sound, and sensation*. New York: Springer.
Heffner, Roe-Merrill S. 1949. *General phonetics*. Madison: University of Wisconsin Press.
Hewitt, George. 2004. *Introduction to the study of the languages of the Caucasus*. Munich: Lincom Europa.
Hillenbrand, James M. & Michael J. Clark. 2009. The role of f_0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics* 71(5). 1150–1166.
Hillenbrand, James M., Laura A. Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5). 3099–3111.
Hirsh, Ira. J. 1974. Temporal order and auditory perception. In Howard R. Moskowitz, Bertram Scharf & Joseph C. Stevens (eds.), *Sensation and measurement: Papers in honor of S. S. Stevens*, 251–258. Dordrecht: D. Reidel.
Hoole, Philip & Christian Kroos. 1998. Control of larynx height in vowel production. *Fifth International Conference on Spoken Language Processing (ICSLP-1998)*. 531–534.
House, Arthur S. & Kenneth N. Stevens. 1956. Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders* 21(2). 218–232.
International Phonetic Association. 1989. Report on the 1989 Kiel convention. *Journal of the International Phonetic Association* 19(2). 67–80.
Ishizaka, Kenzo, J. C. French & James L. Flanagan. 1975. Direct determination of vocal tract wall impedance. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 23(4). 370–373.
Ito, M. Robert & Robert W. Donaldson. 1971. Zero-crossing measurements for analysis and recognition of speech sounds. *IEEE Transactions on Audio and Electroacoustics* 19(3). 235–242.
Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to speech analysis*. *The distinctive features and their correlates*. Cambridge: MIT Press.
Jakobson, Roman, S. Karcevsky & N. Trubetzkoy. 1928. Quelles sont les méthodes les mieux appropriées à un exposé complet et pratique de la grammaire d'une langue quelconque? *Actes du premier Congrès international de linguistes à La Haye*. 33–36.
Jassem, Wiktor. 1965. The formants of fricative consonants. *Language and Speech* 8(1). 1–16.
Keating, Patricia A. 1988. *A survey of phonological features*. Bloomington: Indiana University Linguistics Club.

Keating, Patricia, A. 1991. Coronal places of articulation. In Carole Paradis & Jean-François Prunet (eds.), *Phonetics and Phonology 2. The special status of coronals: internal and external evidence*, 29–48. San Diego: Academic Press.
Kewley-Port, Diane & Charles S. Watson. 1994. Formant-frequency discrimination for isolated English vowels. *Journal of the Acoustical Society of America* 95(1). 485–496.
Kewley-Port, Diane & Yijian Zheng. 1999. Vowel formant discrimination: Towards more ordinary listening conditions. *Journal of the Acoustical Society of America* 106(5). 2945–2958.
Kinsler, Lawrence E. & Austin R. Frey. 1962. *Fundamentals of Acoustics*. New York: John Wiley & Sons.
Ladefoged, Peter. 1957. Use of palatography. *Journal of Speech and Hearing Disorders* 22(5). 764–774.
Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell Publishers.
Lee, Sungbok, Alexandros Potamianos & Shrikanth Narayanan. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America* 105(3). 1455–1468.
Lehiste, Ilse. 1964. *Acoustical characteristics of selected English consonants*. Bloomington: Indiana University.
Lindau, Mona. 1985. The story of /r/. In Victoria Fromkin (ed.), *Phonetic linguistics*, 157–168. Orlando: Academic Press.
Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
Ménard, Lucie, Jean-Luc Schwartz, Louis-Jean Boë, Sonia Kandel & Nathalie Vallée. 2002. Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. *Journal of the Acoustical Society of America* 111(4). 1892–1905.
Mielke, Jeff. 2011. Distinctive features. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology I. General issues and segmental phonology*, 391–415. Malden: Wiley-Blackwell.
Miller, James D. 1989. Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America* 85(5). 2114–2134.
Moore, Brian C. J. & Kengo Ohgushi. 1993. Audibility of partials in inharmonic complex tones. *Journal of the Acoustical Society of America* 93(1). 452–461.
Mrayati, M. & R. Carré. 1976. Relations entre la forme du conduit vocal et les caractéristiques acoustiques des voyelles françaises. *Phonetica* 33(4). 285–306.
Mrayati, M. & B. Guérin. 1976. Étude des caractéristiques acoustiques des voyelles orales françaises par simulation du conduit vocal avec pertes. *Revue d'Acoustique* 36. 18–32.
Narayanan, Shrikanth S. 1995. *Fricative consonants: An articulatory, acoustic, and systems study*. Ph.D. dissertation, University of California, Los Angeles.
Ohde, Ralph N., Katarina L. Haley & Christine W. Barnes. 2006. Perception of the [m]-[n] distinction in consonant-vowel (CV) and vowel-consonant (VC) syllables produced by child and adult talkers. *Journal of the Acoustical Society of America* 119(3). 1697–1711.
Parker, Stephen G. 2002. *Quantifying the sonority hierarchy*. Ph.D. dissertation, University of Massachusetts.
Pennington, Mark. 2005. *The phonetics and phonology of glottal manner features*. Ph.D. dissertation, Indiana University.
Peterson, Gordon E. & Harold L. Barney. 1952. Control methods used in a study of vowels. *Journal of the Acoustical Society of America* 24(2). 175–184.
Pruthi, Tarun. 2007. *Analysis, vocal-tract modeling and automatic detection of vowel nasalization*. Ph.D. dissertation, University of Maryland, College Park.
Reddy, D. R. 1967. Phoneme grouping for speech recognition. *Journal of the Acoustical Society of America* 41(5). 1295–1300.

Repp, Bruno H. 1984. Categorical perception: Issues, methods, and findings. In Norman J. Lass (ed.), *Speech and language: Advances in basic research and practice* (vol. 10), 243–335. New York: Academic Press.
Rosenblum, Lawrence D. 2008. Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science* 17(6). 405-409.
Sakthivel, Subbiah. 1977. *A grammar of the Toda language*. Annamalainagar: Annamalai University.
Shalev, Michael, Peter Ladefoged & Peri Bhaskararao. 1994. Phonetics of Toda. *PILC Journal of Dravidic Studies* 4(1). 19–56.
Shannon, Robert V., Angela Jensvold, Monica Padilla, Mark E. Robert & Xiaosong Wang. 1999. Consonant recordings for speech testing. *Journal of the Acoustical Society of America* 106(6). L71–L74.
Shupljakov, V., G. Fant & A. de Serpa-Leitão. 1968. Acoustical features of hard and soft Russian consonants in connected speech: a spectrographic study. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *Royal Institute of Technology, Stockholm* 9(4). 1–6.
Sreenivasan, K. R., A. Prabhu & R. Narasimha. 1983. Zero-crossings in turbulent signals. *Journal of Fluid Mechanics* 137. 251–272.
Stevens, Kenneth N. 1985. Evidence for the role of acoustic boundaries in the perception of speech sounds. In Victoria Fromkin (ed.), *Phonetic linguistics*, 243–255. Orlando: Academic Press.
Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge: MIT Press.
Stevens, Kenneth N. & Sheila E. Blumstein. 1981. The search for invariant acoustic correlates of phonetic features. In Peter D. Eimas & Joanne L. Miller (eds.), *Perspectives on the study of speech*, 1–38. Hillsdale, New Jersey: Lawrence Erlbaum.
Story, Brad H. 2006. Technique for "tuning" vocal tract area functions based on acoustic sensitivity functions. *Journal of the Acoustical Society of America* 119(2). 715–718.
Story, Brad H., Ingo R. Titze & Eric A. Hoffman. 1996. Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America* 100(1). 537–554.
Story, Brad H., Ingo R. Titze & Eric A. Hoffman. 1998. Vocal tract area functions for an adult female speaker based on volumetric imaging. *Journal of the Acoustical Society of America* 104(1). 471–487.
Sweet, Henry. 1877. *A handbook of phonetics*. Oxford: Clarendon Press.
Tabain, Marija & Andrew Butcher. 1999. Stop consonants in Yanyuwa and Yindjibarndi: locus equation data. *Journal of Phonetics* 27. 333–157.
Takemoto, Hironori, Kiyoshi Honda, Shinobu Masaki, Yasuhiro Shimada & Ichiro Fujimoto. 2006. Measurement of temporal changes in vocal tract area function from 3D cine-MRI data. *Journal of the Acoustical Society of America* 119(2). 1037–1049.
Toda, Martine, Shinji Maeda & Kiyoshi Honda. 2010. Formant-cavity affiliation in sibilant fricatives. In Susanne Fuchs, Martine Toda & Marzena Żygis (eds.), *Turbulent sounds: an interdisciplinary guide*, 343–374. Berlin: Walter de Gruyter.
Traunmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America* 88(1). 97–100.
Walsh Dickey, Laura. 1997. *The phonology of liquids*. Ph.D. dissertation, University of Massachusetts.
Weitzman, Raymond S. 1992. Vowel categorization and the critical band. *Language and Speech* 35(1–2). 115–125.
Welmers, William E. 1973. *African language structures*. Berkeley: University of California Press.
Yang, Chang-Sheng & Hideki Kasuya. 1994. Accurate measurement of vocal tract shapes from magnetic resonance images of child, female and male subjects. *Third International Conference on Spoken Language Processing (ICSLP-1994)*. 623–626.
Zhang, Zhaoyan & Carol Y. Espy-Wilson. 2004. A vocal-tract model of American English /l/. *Journal of the Acoustical Society of America* 115(3). 1274–1280.
Zhang, Shanjie & Jianming Jin. 1996. *Computation of special functions*. New York: Wiley Interscience.

Zhou, Xinhui, Carol Y. Espy-Wilson, Suzanne Boyce, Mark Tiede, Christy Holland & Ann Choe. 2008. A magnetic resonance imaging-based articulatory and acoustic study of “retroflex” and “bunched” American English /r/. *Journal of the Acoustical Society of America* 123(6). 4466–4481.