3. Consonants

Michael Proctor

3.0 Abstract

Consonants are speech sounds produced with a closure or near complete constriction of the vocal tract. All languages systematically exploit place of articulation to differentiate consonants. Eight other phonetically independent parameters are used to create consonant contrast: airstream, constriction degree, laryngeal setting, nasality, laterality, length, articulator stiffness, and respiratory strength. Aspiration, affrication, pre-stopping, secondary articulations, and other properties of 'complex' consonants are best described as patterns of coordination in the underlying gestures.

3.1 Introduction

A fundamental distinction can be made between two types of speech sounds: *vowels*, produced with a more open vocal tract; and *consonants*, which are characterized by greater stricture at one or more places between the glottis and the lips. This chapter provides an overview of the phonetic properties of consonants in the world's languages and key issues in the phonetic description of consonants.

Although the vocal tract is capable of producing infinitely many consonant sounds, only a fraction of these are exploited in human languages, and the fundamental linguistic consideration is whether they are contrastive with other sounds (Trubetzkoy, 1969). The phonetic properties of consonants may therefore be described in terms of the different dimensions in which they contrast, and the types of contrasts exploited in each dimension. Because the categories of *'consonant'* and *'vowel'* are primarily phonological constructs, we first consider the phonological basis of the distinction.

3.1.1 Defining the Class of Consonants

The primary phonological distinction between consonants and vowels is phonotactic, based on the distribution of segments within the syllable. In most languages, consonants cannot form a syllable on their own, unlike vowels, which provide a sonorant nucleus that is the foundation of a syllable. All languages allow syllables of the form CV (Prince & Smolensky, 2008), and more complex syllable structures are formed mainly by adding more segments to onsets and codas (Gordon, 2016). Consonants can therefore be defined as segments that can occupy the margins of a syllable.

In most languages, consonants are also *restricted* to syllable margins, but this characterization is not inviolable. Tashlhiyt Berber allows any segment to occupy a nucleus, for example the fricatives in /sfq.qst/ *'irritate him'* (Ridouane, 2008). In other languages a subset of consonants may function as syllabic: the labial nasal forms part of an onset in Swahili /mbu.ni/ *'ostrich'* but a nucleus in the three syllable word /m̩.bu.ni/ *'coffee plant'* (Mohammed, 2001). Some consonants can also be variably realized as syllabic, depending on speech style. The final sonorants in English *'bottle'* [bɔt.l̩], *'button'* [bʌt.n̩], and *'chasm'* [kæz.m̩], for example, may form their own syllables in rapid speech. These phenomena notwithstanding, a class of consonants can be defined in most languages as the set of segments that do not form syllable nuclei (Chomsky & Halle, 1968).

Segments with identical phonetic properties may pattern differently across languages in this respect. Sonorants produced with a high tongue body may function as vowels when they occupy a syllable nucleus, and consonants when they appear at syllable margins. Spanish /j/ and /w/, for example, are primarily distinguished from /i/ and /u/ on this basis, and the approximants in *viejo* 'old' and *pues* 'then' have been described as non-syllabic allophones of high vowels: /bi̯e.xo/, /pu̯es/ (Hualde, 2005).

3.1.2 Phonetic Properties of Consonants

All consonants share the common property that they are produced with characteristic stricture in one or more regions of the vocal tract. Beyond that, they are a remarkably diverse collection of sounds. More than 600 different consonants have been described in the world's languages – more than three times the number of vowels (Ladefoged, 2005). Consonants are a less phonetically unified class than vowels, produced through a greater variety of articulations with a wider range of acoustic consequences.

At least nine phonetic parameters are needed to capture all the ways that consonants differ from each other in linguistically meaningful ways (Table 3.1). Although each parameter is phonetically continuous, only a discrete number of settings are exploited cross-linguistically to create phonemic contrasts, ranging from 17 (places of articulation) to two: nasality, laterality, length, and respiratory strength each distinguish at most two consonants, if all other parameters remain constant.

Table 3.1: **Phonetically Contrastive Properties of Consonants**.

PARAMETER	NUMBER OF CONTRASTS	EXAMPLES
Place of Articulation	17	ɸ-ɸ̼-f-θ̟-θ̪-θ̺-s̪-s̺- ʃ-s̠̪-s̠̺-ç-x-χ-ħ-ʜ-h
Laryngeal Setting	7	ʔ-p̰-b̬-b-b̥-b̤-p
Airstream Mechanism	4	p-p’-ɓ-ʘ
Constriction Degree	3	c-ç-j
Nasality	2	b-m
Laterality	2	j-ʎ
Articulator Stiffness & Damping	3	d-ɾ-r
Length	2	k-kː
Respiratory Strength	2	t-t*

No language systematically exploits all phonetic variables, and physiological, aerodynamic, and acoustic factors restrict the set of possible and felicitous combinations and settings of phonetic parameters (Ohala, 1983). Most languages utilize only a fraction of the possibilities afforded by the human vocal tract for encoding contrast in consonants. A hypothetical language that used all settings of every parameter in Table 1 would contrast 68,544 different segments, but no known language uses more than 122 consonant phonemes (!Xóõ, Southern Khoisan, Botswana; Maddieson, 2013a).

Consonants have traditionally been categorized according to four primary properties: place of articulation, manner of articulation, voicing state, and airstream (e.g. Nolan et al., 1999; Catford & Esling, 2006). In these taxonomies, several different articulatory properties are conflated under the single dimension of *manner*, including laterality and nasality (Abercrombie, 1967). The approach taken is this chapter is to survey the range of fundamental phonetic properties that differentiate consonants. Although they interact and affect each other to differing degrees, each parameter in Table 3.1 is a separate phonetic variable that can be controlled independently. For example, configuring the tongue to allow central or lateral airflow is a separate mechanism from articulation of constriction degree, so laterality may be treated as an independent phonetic parameter. This removes the need for a separate manner classification for lateral fricatives, which occupy a separate row of the IPA chart despite being indistinguishable in some cases (e.g. [ɬ-l̥]) from voiceless lateral approximants (Esling, 2010).

Likewise, nasalized fricatives cannot be accommodated within traditional taxonomies of manner of articulation, and have been argued to be phonetically impossible (Ohala, 1975). Yet because the velum can be raised or lowered independent of oral articulation, nasality can be treated as a separate dimension of consonant contrast, regardless of the aero-acoustic difficulties of nasalizing certain types of segment. Recent evidence suggests that voiceless segments like [f̃], [s̃] and [x̃] may occur in languages including Coatzospan Mixtec (Oto-Manguean, Mexico) and Scots Gaelic (Indo-European, UK), and may even play a role in nasal harmony systems (Shosted, 2006). These findings are neither surprising nor problematic to describe when consonants are analysed in terms of their independent phonetic parameters.

Some consonant contrasts are not exemplified in Table 3.1 because they result from differences in articulatory coordination rather than changes in a phonetic parameter. These include affricates, aspirated stops, pre- and post-stopped nasals, and multiply-articulated consonants. Some of these have also been treated as distinct manners of articulation, but many aspects of the behaviour of these consonants are best understood in terms of the timing and constituency of their underlying gestures (Browman & Goldstein, 1992). Under this approach, 'complex' consonants involve different combinations of gestures and patterns of temporal organization, but still use the same set of phonetic parameters that define other segments (Section 3.3.10).

All languages make use of *place of articulation* and at least one other phonetic dimension – either constriction degree or laryngeal setting – to encode consonant contrast. The primacy of place of articulation, constriction degree and laryngeal setting as parameters of phonological contrast is reflected in the design of the IPA chart, which primarily organizes consonants by these three dimensions. Ohala (1981), Goldstein & Fowler (2003), Mielke (2008), and Gordon (2016) address different factors which make some phonetic properties and combinations of properties more salient in the organization of consonant systems, and may account for typological preferences.

3.2 Historical Overview

There is a long history of description and classification of consonants in many languages and philosophical traditions. Developments in some early writing systems reflect an understanding of consonants as a distinct class of sounds which could be categorized by their articulatory properties. Egyptian hieroglyphs were used to represent consonantal sounds in the 4th millennium BCE (Coulmas, 1999). Brahmic scripts dating from at least the 3rd century BCE are organized by consonant place, manner, and voicing contrasts, reflecting the detailed systemic phonetic knowledge of ancient Indian grammarians, most famously expressed in the Shivasūtras of Pānini (Kiparsky, 1991). Coarticulatory processes affecting consonants at morpheme boundaries were systematically described in Sanskrit sandhi rules.

The term *consonant* originates in ancient Greek descriptions of sounds that differ from vowels in their phonotactic properties. In Platonic thought, the main division of sounds was between *phōnēenta*: vowels 'possessing voice'; and *aphōnia*: 'voiceless' sounds, denoting consonants. A later distinction was drawn between sounds that could be pronounced on their own, and those 'voiced in combination' with others: *sumphōna*, translated into Latin as *consonantes*. Aristotle, Dionysius Thrax and other grammarians of the 4th to 2nd centuries BCE also described shared manner, voicing and aspiration features of Greek consonants, defining classes including liquids, continuants, and semivowels (Allen, 1981, 116-120).

Early descriptions of consonants were largely language-specific, but Renaissance scholars began to identify commonalities across languages. Isaac Newton proposed a phonetic alphabet illustrated with reference to English and Hebrew, and described articulatory relationships between vowels and consonants differing only in stricture at the same place of articulation (Elliott, 1954). Wilkins (1668) outlined a comprehensive 'organic' alphabet and phonetic chart which classified consonants by airstream, voicing, place, manner, and perceptual-acoustic properties, distinguishing also between active and passive articulators.

With the development of comparative linguistics in the philological tradition, more universal descriptions of consonants emerged. Volney (1795), Lepsius (1863), Jespersen (1889), and others proposed phonetic alphabets designed to be able to describe consonants in any language. Transcription systems developed by Brucke (1856), Bell (1863), and Sweet (1877) were informed by new scholarship on speech physiology. The culmination of all these efforts was the publication in 1888 of the 1st Edition of the International Phonetic Alphabet, which has since defined the standard for transcription. The history of phonetic characterization of consonants is described in more detail in Tillmann (2006), Kemp (2006), Heselwood et al. (2013), and in Chapters 4 and 14 in this volume.

3.3 Critical Issues in Consonantal Phonetics

We have a good understanding of the phonetic properties of consonants in many of the world's languages, but debate continues about the number and nature of phonetic primitives, how they interact, how contrasts are encoded and perceived, and how the temporal properties of consonants are represented and controlled.

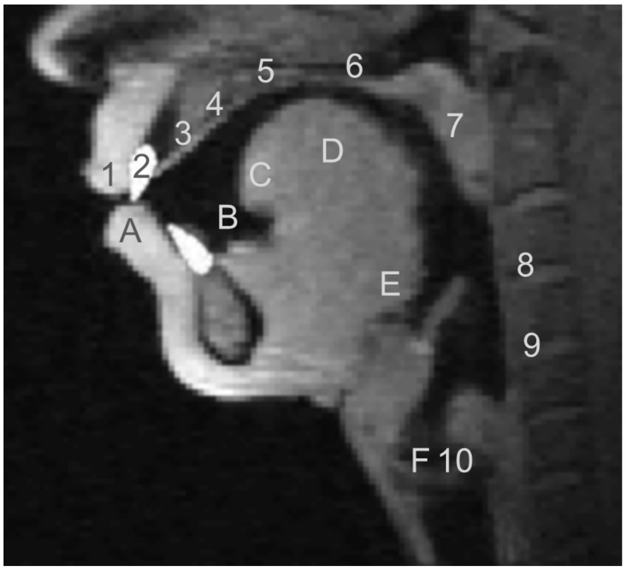
3.3.1 Place of Articulation

All consonants are produced by narrowing some part of the vocal tract to constrict airflow. The location of the primary constriction is one of the most fundamental determinants of the acoustic properties characterizing a segment (Fant, 1960; Stevens, 1998), and is a fundamental organizing principle in all phonetic descriptions of consonants.

A survey of phoneme inventories suggests that place of articulation is the only phonetic parameter systematically exploited by all languages to create consonant contrast. Rotokas (East Papuan, Bougainville) uses six consonants articulated at three places of articulation: /β-r-g/–/p-t-k/, but manner is not contrastive independent of voicing (Firchow & Firchow, 1969). Warlpiri (Pama-Nyungan, Australia) distinguishes stops, nasals, laterals, flaps and approximants produced at five different places of articulation, but voicing is not independently contrastive (Nash, 1980). Many other Australian languages also do not distinguish voicing independent of manner contrasts, nor does Seneca (Iroquoian, New York; Chafe, 1960).

In many smaller consonant inventories, place distinctions are primarily made with different sets of articulators. Hawaiian (Polynesian; 8 consonant phonemes) differentiates only labial, lingual and glottal places of articulation (Schutz, 1980), and Pirahã (Mura, Brazil; 8 consonants) uses only labial, coronal, dorsal and glottal place contrasts (Everett, 1986). Goldstein & Fowler (2003) observe that these organic contrasts provide a common phonological currency in a community of language users – phonological primitives intrinsically shared by speakers and listeners – which may account for typological preferences in small consonant systems. The most common arrangement of average-sized consonant inventories (18 to 24 consonant phonemes) divides the vocal tract into five primary places of articulation: labial, dental/alveolar, post-alveolar/palatal, velar, and glottal (Maddieson, 1980; Gordon, 2016).

Ladefoged & Maddieson (1996) identify 17 different places of articulation that are required to distinguish between all attested consonants. Six different *active* articulators are used to form constrictions at ten different *target* places in the vocal tract, illustrated on the midsagittal MR image of an adult male in Figure 3.1. IPA symbols are given for the voiceless fricative produced at each place of articulation.



A1	Bilabial	ɸ
A2	Labiodental	f
C1	Linguo-labial	s̼
C2	Interdental	θ̟
B2	Apical dental	θ̺
C2	Laminal dental	θ̻
B3	Apical alveolar	s̺
C3	Laminal alveolar	s̻
B4	Apical retroflex	s̠
C4	Palato-alveolar	ʃ
B5	Retroflex	ʂ
D5	Palatal	ç
D6	Velar	x
D7	Uvular	χ
E8	Pharyngeal	ħ
E9	Epiglottal	ʜ
F10	Glottal	h

Figure 3.1: **Contrastive Places of Articulation** (Ladefoged & Maddieson, 1996). *Active articulators*: A. lips; B. tongue tip; C. tongue blade; D. tongue body; E. tongue root; F. glottis. *Articulator targets*: 1. labial; 2. dental; 3. alveolar; 4. post-alveolar; 5. palatal; 6. velar; 7. uvular; 8. pharyngeal; 9. epiglottal; 10. glottal. Teeth do not image in MRI, and have been superimposed. (Image courtesy USC SPAN group).

Place contrasts are treated differently in Articulatory Phonology (Browman & Goldstein, 1989, 1992). In this framework, consonants may differ in the *constriction location* of their constituent gestures, which are defined in terms of *tract variables* (Saltzman & Munhall, 1989). Tract variables specify a target location and aperture, but these goals may be achieved in different ways through the coordinated activity of sets of articulators (Table 3.2). Sibilant fricatives, for example, vary in apicality/laminality and in the relative contributions of jaw and tongue articulation in different vowel contexts (Toda & Honda, 2003; Iskarous et al., 2011), but are all produced with a critical constriction at the same (speaker-specific) place in the alveo-palatal region. Goals of production of some consonants may therefore be better characterized in terms of constriction formation and release in different regions of the vocal tract, rather than articulator-specific actions.

Table 3.2: **Tract Variables and Associated Articulators** in Articulatory Phonology (Browman & Goldstein, 1992).

	TRACT VARIABLE	ARTICULATORS INVOLVED
LP	lip protrusion	upper & lower lips, jaw
LA	lip aperture	upper & lower lips, jaw
TTCL	tongue tip constriction location	tongue tip, tongue body, jaw
TTCD	tongue tip constriction degree	tongue tip, tongue body, jaw
TBCD	tongue body constriction location	tongue body, jaw
TBCL	tongue body constriction degree	tongue body, jaw
VEL	velic aperture	velum
GLO	glottal aperture	glottis

Obstruent consonants produced at the same place of articulation share acoustic properties influenced in large part by the size and geometry of the vocal tract anterior to the constriction. Place cues provide crucial insights into the relative invariance of acoustic and articulatory properties of consonants, and have been a focus of research into goals of production and the nature of phonetic primitives (Stevens & Blumstein, 1978; Remez et al., 1981; Sussman & Shore, 1996; Iskarous et al., 2010, etc.).

3.3.2 Airstream

All speech sounds involve displacement of air in some part of the vocal tract. This process can be initiated with three sets of speech organs, to produce airflow in two directions (Table 3.3), creating six hypothetical ways of generating sounds (Catford, 1939). Four of these airstream mechanisms are used to produce consonants. The typological gaps can be explained by physiological factors. Ingressive pulmonic speech is possible but tiring, and the larynx may asymmetrically configured for egressive voicing (Ohala, 1983). Egressive linguo-velaric sounds ('squeezed' clicks) would to be very difficult to generate with the transience necessary to excite the vocal tract.

Table 3.3: **Consonant Airstream Mechanisms**.

INITIATION	DIRECTION	
	INGRESSIVE	EGRESSIVE
Pulmonic	—	Pulmonic
Glottalic	Implosive	Ejective
Linguo-Velaric	Click	—

All languages use pulmonic consonants. Approximately one fifth of all languages use ejectives, and about 10% use implosives (Maddieson, 1992). Clicks are the least common type of consonant with respect to airstream initiation, found only in Khoisan and Bantu languages of Southern Africa, with a few exceptions (Maddieson, 1992).

In pulmonic consonants with an oral place of articulation, there is a clear division between the speech organs responsible for initiating and shaping the sound. These components are assumed to be independent in source-filter models of speech production (Chiba & Kajiyama, 1941; Fant, 1960; Stevens, 1998); however, the distinction between airstream and place becomes blurred in consonants where the same organs are involved in initiation and sound shaping, in particular, clicks and glottal/glottalized consonants. Even in non-glottalized pulmonic consonants, there are complex interactions between airstream and articulation which require further investigation (Hoole, 1998; Kim et al., 2005).

3.3.2.1 Glottalic Consonants

Glottalic consonants are produced by closing the vocal folds to prevent airflow between the trachea and the pharynx, and moving the glottis up or down to displace air in the supraglottal cavities. *Implosives* are initiated with a downward movement of the glottis, lowering the supraglottal air pressure and drawing air into the vocal tract. *Ejectives* are produced with the opposite movement: by rapidly raising the closed glottis, the air above it is compressed, and expelled when the velum is lowered or the oral constriction is released.

The distinction between true glottalically-initiated consonants and pulmonic consonants with glottal accompaniment appears to be more gradient than categorical (Ladefoged & Maddieson, 1996; McDonough & Wood, 2008). Consonants characterized as ‘glottalized’ may be variously realized as pulmonic with a glottal constriction, as purely glottalic, or with some combination of glottically- and pulmonically-initiated airflow (Aoki, 1970; Pinkerton, 1986; Maddieson, 2013b).

Ejectives

Key stages in the production of a bilabial ejective are illustrated in Figure 3.2. Rapid upward movement of the closed glottis can create greater compression of air behind an oral constriction than is typically achieved during pulmonic stop production, so ejectives are usually characterized by a stronger stop release burst (Ladefoged & Maddieson, 1996; Grawunder et al., 2010) and longer Voice Onset Time (VOT; Wagner & Baker-Smemoe, 2013); however, speakers and languages differ in the timing of the glottal and supraglottal gestures and the laryngeal setting during release (Lindau, 1984; Flemming et al., 2008; McDonough & Wood, 2008).

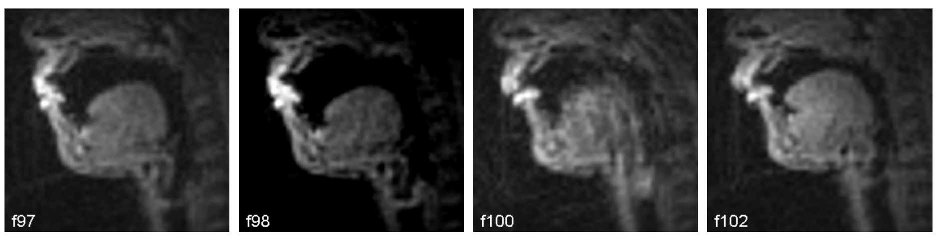


Figure 3.2: **Articulation of an affricated bilabial ejective [p͡fʼ]**. Frame 97: lowering of tongue and larynx, closure of velum; f98: fully-lowered larynx, glottalic closure; f100: rapid laryngeal raising and partial release of lips; f102: glottal abduction with larynx 21 mm higher than at start of initiation. (Proctor et al., 2013).

Ejectives often contrast with pulmonic consonants produced at the same places of articulation, e.g. Nez Perce (Sahaptian, Idaho): [tawn] *'make a stone tool'* vs. [t'awn] *'guess in gambling game'* (Aoki, 1970). Stops, affricates and fricatives can be fully or partially initiated with glottalic egressive airstream. The term *glottalic sonorant* is also used to describe contrasts such as Yurok (Algic, California): [kem] *'again'* vs. [kem'] *'food'* (Blevins, 2003), but Ladefoged & Maddieson (1996) argue that modal voicing is physiologically impossible during this initiation mechanism. If ejectives are intrinsically voiceless, these sonorants must be *glottalized* (produced with an accompanying laryngeal constriction) (Bird et al., 2008), rather than *glottalic* (generated through an ejective airstream mechanism). This analysis is consistent with the behaviour of these segments in some Algonquian and Wakashan languages, where glottalized sonorants can be realized as glottal stop-sonorant sequences (Esling et al. 2005) or re-syllabified as glottal stop-initial clusters in intervocalic environments (Blevins, 2003).

Implosives

Laryngeal lowering is observed during production of different types of stops (Ewan & Krones, 1974), and because multiple airstream mechanisms may be involved (Lindau, 1984), Ladefoged & Maddieson (1996, 82) propose a broad definition of implosives as "stops that are produced with a greater than usual amount of lowering of the larynx during the time that the oral closure for the stop is maintained".

Canonical implosives are initiated through complete closure of the glottis and downward action of the larynx (Catford, 1964). Voiceless glottalic ingressive stops of this type are found in Lendu (Central Sudanic, Congo): [ɓ̊áɓ̊á] *'attached to'* (Demolin, 1995), Isoko (Niger-Congo, Nigeria) and other languages. More commonly, implosives are produced with an incompletely constricted glottis, so that some trans-glottal airflow is possible. These *voiced implosives* may also involve some pulmonically-initiated airflow, resulting in modal or creaky voicing during some part of the stop closure interval. Voicing can actually increase during the stop interval (Lindau 1984), as the lowering of the larynx counteracts the oral pressure build up due to pulmonic airflow (which typically attenuates voicing in a plosive). Uduk (Nilo-Saharan, Sudan) contrasts modally-voiced implosives with voiced pulmonic stops: [ɓaʔ] *'back of neck'* vs. [baʔ] *'to be something'*; and Hausa (Chadic, Niger) prototypically contrasts creaky-voiced implosives with voiced pulmonic stops: [ɓ̰aːtáː] *'spoil'* vs. [ɓaːtáː] *'line'* (Ladefoged & Maddieson, 1996).

3.3.2.2 Clicks

Clicks are produced by creating a complete seal between the tongue and the roof of the mouth and lowering the middle of the tongue to create a low-pressure cavity into which air rapidly flows when part of the seal is released. The back of the tongue forms a seal against the uvular and velum, but clicks differ in the location of the front part of the seal and the way it is released. Clicks can be formed with a labial [ʘ], dental [ǀ], palatal [ǂ], retroflex [‼], or alveolar anterior constriction, which can be released centrally [ǃ] or laterally [ǁ]. Many languages contrast clicks produced at four different places of articulation, including Nama (Khoe, Namibia; Figure 3.3), and some Tuu languages use up to five place-contrastive click phonemes (Miller, 2011).

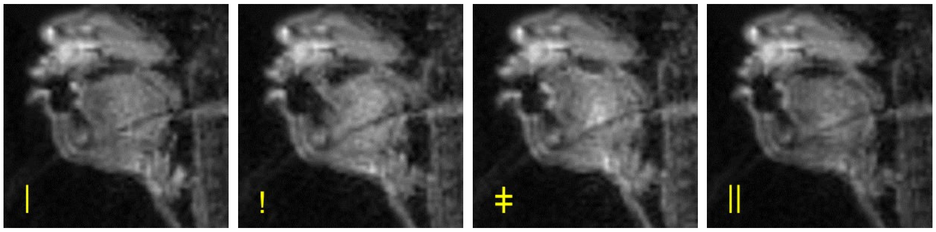


Figure 3.3: **Place contrasts in Nama (Khoekhoegowab) clicks.** Midsagittal lingual posture captured at acoustic onset of click; L-to-R: dental: **|om** /ŋ̊|ˀòm̀/ *'breathe'*, alveolar: **!om** /ŋ̊!ˀóm̀/ *'remove thorn'*, palatal: **ǂom** /ŋ̊ǂˀóm̀/ *'sew'*, lateral: **ǁom** /ŋ̊ǁˀòm̋/ *'sleep'*. (Proctor et al., 2016).

Clicks may be produced with a range of accompanying airstream mechanisms, and different nasal and laryngeal settings. Clicks in G|ui (KhoeKwadi, Botswana) can be entirely lingually-initiated (e.g. [!]) or produced with pulmonic ([!͡q]) and glottalic ([!͡ʼ]) airstream components (Nakagawa, 2006). Clicks are restricted to syllable onset position, so languages exploit the transition into the pulmonically-initiated syllable rime to encode further contrasts in voicing, aspiration, and affrication. Up to four contrastive settings of VOT are used with clicks in Ju|'hoansi (Ju, Namibia; Snyman, 1975). Miller (2011) identifies 12 different airstream and manner contrasts found in click inventories, but debate continues as to whether these are best analysed as clusters (Güldemann, 2001; Nakagawa, 2006), or segmental contrasts (Ladefoged & Traill, 1994; Miller, 2010).

3.3.3 Constriction Degree

By varying the aperture of a constriction, the rate and type of airflow and the acoustic response of the vocal tract can be manipulated (Fant, 1960). Along with laryngeal setting, *constriction degree* is the primary phonetic dimension used to create additional contrasts at the same place of articulation. The acoustic properties of a consonant depend on the configuration of the entire vocal tract, but the aperture of the narrowest part of the vocal tract is especially important in shaping the sound. Consonants produced with the same type of constriction at different places share many of the same acoustic properties (Stevens, 1998).

Although constriction degree is a continuous variable, languages systematically exploit only a limited range of strictures. Consonant gestures can be specified for *narrow* (approximant), *critical* (fricative), and *closed* (stop) targets (Browman & Goldstein, 1995). The number and range of divisions of the constriction degree continuum might be language-specific, and the exact apertures of gestural targets may be speaker-specific, selected "on the basis of quantal articulatory-acoustic relations (e.g. Stevens, 1989) and/or on the basis of adaptive dispersion principles (e.g. Diehl, 1989; Lindblom & Engstrand, 1989; Lindblom et al., 1983; Lindblom & Maddieson, 1988)" (Browman & Goldstein, 1992, 26). In the IPA (2005) consonant chart, three of eight manner distinctions are made on the basis of oral tract aperture: *plosive*, *fricative*, and *approximant*; all other manner distinctions involve changes in nasality, laterality, and articulatory stiffness.

3.3.3.1 Stops

Stops are produced with a complete constriction that prevents airflow through the oral tract. Stops have a privileged status as the only type of consonant used in all the world's languages (Ladefoged & Maddieson, 1996). Acoustically, stops are characterized by an interval of reduced or zero energy during oral closure, with a periodic or quasi-periodic component if phonation persists through or begins during the occlusion. An intense acoustic burst and/or a short interval of reduced energy frication may follow the closure phase, depending on the manner in which the oral constriction is released. The rapid formation and release of a complete vocal tract closure affords many different mechanisms for encoding contrasts, but languages exploit only the *closure* and *release* phases of stops phonologically: place-independent distinctions are not encoded in the *onset* phase, when the articulators are moving together (Henton et al. 1992).

The most frequently attested types of stop are *plosives* (pulmonic egressives), but stops are produced with a greater range of laryngeal settings and airstream mechanisms than consonants characterized by other degrees of constriction. Owerri Igbo (Niger-Congo, Nigeria) uses six stop phonemes at each place of articulation, illustrated in the minimal sextet [ípa] *'to carry'* – [ípʰa] *'to squeeze'* – [íɓ̥a] *'to gather'* – [íba] *'to get rich'* – [íɓʱa] *'to peel'* – [íɓa] *'to dance'* (Ladefoged et al., 1976). Most commonly, languages use a˚ two-way voicing and/or airstream contrast between stops produced at the same place: approximately half of all languages contrast a series of voiced vs. voiceless stops (Gordon, 2016).

3.3.3.2 Fricatives

Fricatives are consonants characterized by a turbulent airstream, produced with a critical constriction of the vocal tract. Turbulent airflow emerging from a narrow aperture creates a noise source with broadband energy whose spectral properties are partly determined by the geometry of the vocal tract downstream from the constriction, as well as the cross-sectional geometry of the constriction (Shadle, 1999). In *sibilant* fricatives, e.g. [s]–[ʃ]–[s̺], the turbulent jet emerging from the constriction strikes the teeth, creating one or more additional noise sources. Magnetic resonance images of American English voiceless sibilant fricatives (Figure 3.4) reveal large differences in tongue grooving, lip rounding, and the geometries of the constrictions and anterior cavities into which the turbulent jet emerges, all of which help shape the different sounds of [s] and [ʃ].

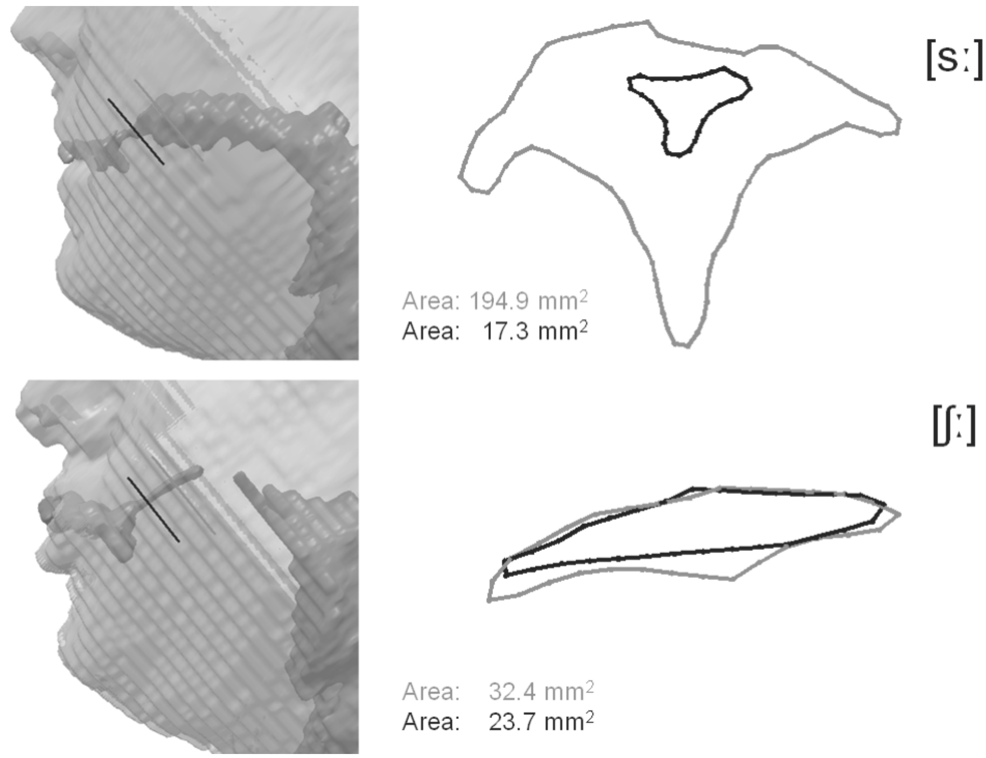


Figure 3.4: **Constrictions characterizing English sibilant fricatives.** Cross-sections through MRI volumes (normal to tract midline) reveal constriction geometries for [s] (top) and [ʃ] (bottom). Inner (dark) line: most constricted part of airway in alveolar region. Outer (light) line: cross-section of airway 10mm posterior to constriction (Proctor et al., 2010).

Voiced fricatives have an additional noise source at the glottis, which is acoustically coupled through the critical constriction. The interactions between constriction degree, airstream mechanism, rate of airflow, and laryngeal setting are complex and imperfectly understood, and remain an active area of research (Stevens, 1971; Shadle, 1990; McGowan & Howe, 2007; Pinho et al., 2012).

Turbulent airflow can be generated by constricting either central or lateral airstreams. Bura (Chadic, Nigeria) uses voiced [ɮ] and voiceless [ɬ] alveolar lateral fricatives, and a voiceless palatal lateral fricative [ʎ̥] (Ladefoged & Maddieson, 1996). Languages use up to eight contrastive places for central fricatives (Ubykh, Caucasian, Georgia; Vogt, 1963), but the most common number of phonemic fricatives is two. Voiceless fricatives are more commonly attested than voiced fricatives, and are preferred at all places of articulation other than labial and dental (Gordon, 2016).

3.3.3.3 Approximants

Approximants ('semi-vowels' or 'glides'), are produced with a narrow constriction of sufficiently wide aperture to allow non-turbulent airflow and spontaneous voicing. The resulting vocal tract configuration typically resonates in response to a modal glottal source, producing formant structures similar to those of vowels. In some cases, the distinction between approximant and vowel is purely phonological (Section 3.1.1); however, Maddieson & Emmorey (1985) show that in Amharic (Afro-Asiatic, Ethiopia) and other languages [j] and [w] are produced with significantly greater stricture than [i] and [u]. Similarly, the distinction between approximant and fricative is not always categorical, as the same degree of constriction may result in laminar or turbulent airflow, depending on aerodynamic factors and laryngeal state (Shadle, 1999).

Approximants can be produced with a central ([ʋ]-[ɹ]-[ɻ]-[j]-[ɰ]) or lateral airstream ([l]-[ɭ]-[ʎ]-[ʟ]), and with a single characteristic constriction ([ʋ], [ɰ]), or through the coordination of multiple constrictions. The American English rhotic approximant [ɹ] is typically produced with labial, coronal and pharyngeal approximations in onset positions (Alwan et al., 1997). More than 86% of languages use a palatal approximant, 81% use a lateral approximant, and about three quarters of known languages have a labial-velar approximant (Gordon, 2016).

3.3.4 Laryngeal Setting

Most commonly during speech, the vocal folds are configured in close approximation and at a tension that results in spontaneous vibration when air is expelled from the lungs. This process of *voicing*, or *phonation*, is the primary means used to acoustically excite the vocal tract, but a wide range of different configurations of the larynx are used during speech (Hanson et al., 2001; Gordon & Ladefoged, 2001). The vocal folds can be closed to prevent transglottal airflow, or opened wide so that egressive airflow from the pulmonary system passes relatively unimpeded – the prototypical setting for voiceless sounds. Both glottal aperture and vocal fold tension can be adjusted along a continuous range of settings between these extremes, and depending on the transglottal pressure maintained by the respiratory system, this will result in different modes of vibration and airstream turbulence (Stevens, 1977).

In addition to modal voice, several different laryngeal settings and associated types of phonation have been described (see also Chapter 2.4). Catford & Esling (2006) distinguish *breath*, *whisper*, *voice*, and *creak*, and combinations of *voice* with the other three modes. Esling & Harris (2005) propose three additional categories of *harsh* voice. Laver (1980) and Gobl (2010) describe six voice qualities: *modal*, *breathy*, *whispery*, *creaky*, *tense*, and *lax*. Ladefoged & Maddieson (1996) do not distinguish breathy from whispery, and propose that seven main types of laryngeal setting are used to create phonological consonant contrast across languages (Table 3.4).

Table 3.4: **Laryngeal Settings for Consonants.**

Glottal Aperture	Laryngeal Mode	Example
Closed	Glottalized/Stopped	Nuuchahnulth [ʔ]
Close	Creaky (Laryngealized) Voice	Fula [b̰]
Close	Stiff Voice	Thai [b̬]
Narrow	Modal Voice	Marathi [b]
Wide	Slack Voice	Javanese [b̥]
Wide	Breathy Voice (Murmur)	Igbo [b̤]
Open	Voiceless	Tongan [p]

Approximately half of the world's languages use a two-way laryngeal contrast in stops (Gordon, 2016). Another quarter of all languages use three or more laryngeal contrasts, but these are often implemented through differences in the coordination of laryngeal and supraglottal articulation, rather than distinct settings on the continuum represented in Table 3.4. Cross-linguistically, the third most common type of stop (after voiceless and voiced) is voiceless aspirated (Henton et al., 1992), exemplified in the English word *'team'* [t^{h}im] in many accents. Aspiration results when the glottis remains open after the release of a supraglottal constriction. When a stop is followed by a voiced segment, the length of the aspirated interval can be measured as voice onset time (VOT), a metric which can also be used to quantify voicing duration in partially voiced stops (Lisker & Abramson, 1964).

Stop phonation intervals and aspiration durations vary considerably across languages, and different regions of the voice onset time (VOT) continuum are exploited to create up to four phonemic stop contrasts. Eastern Armenian uses a three-way labial stop contrast implemented as [b] (mean VOT = -96 ms) / [p] (VOT = 3 ms) / [pʰ] (VOT = 78 ms), but the three contrastive labial stops of Korean all have positive mean VOTs (Lisker & Abramson, 1964). Stops contrasting in VOT may also contrast in laryngeal setting, as in some Indo-Aryan languages with four-way voicing distinctions. Bengali contrasts four orals stops at each place of articulation, e.g.: [pati] *'mat'* / [pʰati] *'I burst'* / [bati] *'bowl'* / [b̤ʱati] *'kiln'*. In each quartet of stops, the voiced aspirated plosive is realized with breathy voice (Khan et al., 2007).

Understanding the phonetics of laryngeal contrasts is complicated by the fact that these parameters also interact with suprasegmental factors including tone and intonation. Furthermore, voice quality is a key phonetic parameter for encoding indexical information (Laver, 1968; Podesva & Callier, 2015), and speakers manipulate laryngeal setting with emotion, mood and attitude (Gobl & Chasaide, 2003). Interactions between laryngeal setting, airstream, phonation type and supraglottal coordination are complex, and remain an especially active area of research (e.g. Titze & Story, 1997; Son et al., 2011; Moisik & Esling, 2014; Lancia et al., 2016).

3.3.5 Nasality

Oral consonants are produced with the velum raised to form a seal against the nasopharynx, closing off the nasal cavity (Figure 3.1). Most configurations of the oral vocal tract can be nasalized by lowering the velum without modifying other articulatory parameters, depending on the airstream mechanism and the location of the oral constriction with respect to the uvula. A complete oral constriction anterior to the velic opening will force all airflow through the nasal airway, producing a nasal stop (Figure 3.5).

Figure 3.5: **Articulation of a Nasal Stop**. Alveolar nasal stop [n] produced by a male speaker of General American English. Left-to-right: midsagittal slice; coronal slice through uvula; axial slice through top of tongue and mid-uvula. (Image courtesy USC SPAN group).

Opening the velo-pharyngeal port couples the nasal cavity to the oral cavity, which has several acoustic consequences. Formant frequencies are lowered, and the intensity ratio of F1 to higher formants is increased. The reconfiguration induces characteristic nasal anti-resonances (*N1 … Ni*) that attenuate frequency bands of the corresponding oral spectrum (Fant, 1960; Recasens, 1983; Stevens, 1998).

The vast majority of the world's languages (> 97%) make phonemic contrasts between oral and nasal consonants. Approximately 7% of attested nasal phonemes are either voiceless, laryngealized, or breathy-voiced, but no language uses non-modal nasals unless they also have a voiced nasal at the same place of articulation (Gordon, 2016). Marathi (Indo-Aryan, Maharashtra) contrasts voiced ([maːr] *'beat'*) and breathy voiced ([m̤aːr] *'a caste'*) nasals at three different places of articulation (Khan et al., 2007).

3.3.6 Lateralization

Most speech sounds are produced by directing airflow down a channel formed in the centre of the vocal tract. The sides of the tongue create a seal against the teeth and palate so that airflow primarily follows the groove of the tongue. In these *central* (or *medial*) consonants, the region of maximum aperture in the oral cavity is located around the midsagittal plane. In *lateral* consonants, the sides of the tongue do not form a complete oral seal, allowing parasagittal airflow through one or more side channels.

The contrast between central and lateral articulation is illustrated in Figure 3.6. MR images of an American English /l/ show the middle of the tongue to be raised and the sides lowered (top row). In contrast, a concave tongue grove and raised sides of the tongue can be seen in the central /z/ (bottom row).

As with central consonants, constriction degree can be manipulated to create additional lateral contrasts. Most languages use only a single lateral consonant, most commonly an approximant (Maddieson, 1980). Approximately 10% of languages use lateral fricatives (Maddieson, 2013a), articulated with more constricted side channels producing turbulent airflow and acoustic properties similar to central fricatives. As with central consonants, constriction degree is phonetically gradient, and the distinction between lateral approximants and lateral fricatives isn't always clear (Maddieson & Emmorey, 1984). In Welsh and other languages, /ɬ/ is realized as [l] in some environments (Ball, 1990).

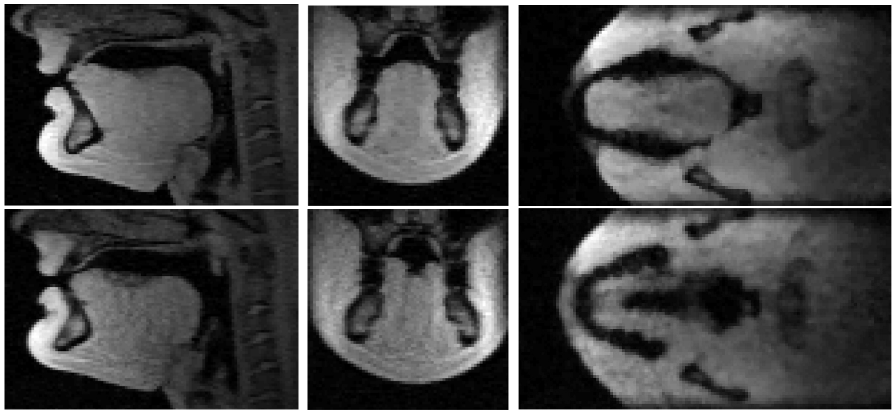


Figure 3.6: **Comparison of Lateral and Central Articulation**. Top row: alveolar lateral approximant [l]. Bottom row: voiced alveolar fricative [z]. Left-to-right: midsagittal slice; coronal slice through hard palate; axial slice through top of tongue. Adult male speaker of General American English. Both consonants elicited in [u_u] context. (Image courtesy USC SPAN group).

Lateral consonants can be produced with airstreams other than pulmonic. Lateral ejectives are found in Najavo, and many Khoisan and Bantu languages contrast alveolar and lateral clicks. Lateral clicks are prototypically released to one side, while laterals initiated with other airstream mechanisms are canonically bi-lateral. Additional lateral contrasts can be made by varying the timing and degree of lateralization in stop releases to produce lateral affricates.

Cross-linguistically, lateral consonants are produced at five different places of articulation: dental, alveolar, post-alveolar/retroflex, palatal, and velar (Ladefoged et al., 1977). No one language contrasts laterals at all five places of articulation, but Kaytetye (Arandic, NT Australia) uses four lateral approximant phonemes: [əl̪ə] *'nose'* – [ələ] *'hard-headed'* – [kəɭə] *'done quickly'* – [əʎə] *'shade'* (Turpin et al., 2012).

English laterals are canonically produced with coronal fronting and dorsal backing gestures that elongate the midsagittal profile of the tongue, creating lateral side channels in the process (Giles & Moll, 1975; Sproat & Fujimura, 1993). More data is required to understand how similar gestures are coordinated in laterals in other languages, and the extent to which lateralization is actively articulated by lowering the sides of the tongue. This issue is of particular interest in language with 'clear' laterals that don't involve the same degree of dorsal retraction as English coda [ɫ].

3.3.7 Articulatory Stiffness and Damping

Consonants sharing the same place can differ in the dynamic properties of the articulators that form the constriction. Some coronal consonants differ primarily in the nature of the contact between the tongue tip and the passive articulators: [d], [ɾ], and [r] are all voiced consonants produced with an apical closure against the alveolar ridge, but the number and duration of contacts varies. Stops, taps, flaps, and trills can be partly differentiated on the basis of the physical properties of the articulators that give rise to these differences under the same aerodynamic conditions, in particular *stiffness*.

An important distinction must be made between at least two different uses of the term *stiffness*. Anatomically, muscular stiffness refers to the strength of internal forces that resist deformation and promote restoration of a muscle to an original shape (Nichols & Huyghues-Despointes, 2009). In the speech production literature, stiffness is "a measurement of articulator movement that characterizes speed independent of its displacement" (Roon et al., 2007). Studies of patterns of articulation suggest that key components of speech motor control are achieved through specification of articulator stiffness (Kelso et al., 1983; Munhall et al., 1985). In the Task Dynamic framework of Articulatory Phonology (Saltzman & Munhall, 1989; Browman & Goldstein, 1989), stiffness and damping ratio are properties of gestures that determine their durational characteristics and behaviour upon reaching a target: whether they over- or under-shoot (Browman & Goldstein, 1990).

Different settings of stiffness and damping can account for patterns of articulatory motion in consonants that differ in velocity, duration and repetition of contact. Three sets of articulators – the lips, tongue tip, and uvula – are sufficiently elastic that they can be configured to oscillate spontaneously in an egressive airstream, producing a *trill*. The moving articulator in a trill needs to be placed in sufficient proximity to another less elastic surface for the Bernoulli Effect to operate (Barry, 1997), and many other factors are involved in sustaining tongue tip oscillation, including airstream velocity, inertia and damping of the tongue tip, and vocal tract wall compliance (McGowan, 1992; Solé, 2002). Labial /ʙ/ and alveolar /r/ trills are contrastive in Nias Selatan (Austronesian, Sumatra; Brown, 2001), and Siwi (Cariban, Venezuela) contrasts a uvular trill /ʀ/ with an alveolar tap /ɾ/ (Oquendo, 2004). The epiglottis may also be trilled in variant realizations of some pharyngeal fricatives (Moisik et al., 2010).

Taps and flaps may also differ from other consonants produced at the same place in gestural damping (Browman & Goldstein, 1992), or "vibrator tension" (Barry, 1997). Unlike trills, which are aerodynamically actuated, these segments are ballistic: articulator movement is initiated by muscle activity, rather than the airstream (Abercrombie, 1967; Barry, 1997). Unlike stops, the oral constriction created in taps and flaps does not create a large difference in air pressure, so these segments are not acoustically characterized by a strong release burst (Zue & Laferriere, 1979; De Jong, 1998).

Although the terms are often used interchangeably, and despite the fact that the IPA does not provide separate symbols to distinguish between them, taps and flaps are articulated through distinct mechanisms (Ladefoged, 1968). A *flap* involves tangential movement of the active articulator along a single trajectory (e.g. tongue tip brushing forward across the alveolar ridge), while a *tap* is produced with a relatively symmetrical articulatory trajectory towards and away from a target (e.g. tongue tip raising and lowering towards the alveolar ridge) (Ladefoged & Traill, 1994; Derrick & Gick, 2011). These differences in articulatory action are also reflected in asymmetries in formant trajectories into and out of a flap (Monnot & Freeman, 1972). Phonemic taps and flaps are attested at three places of articulation: alveolar /ɾ/, retroflex /ɽ/, and labio-dental /ⱱ/, found in a few languages including Mono (Banda, Congo, Olson, 2004).

3.3.8 Length

Some languages systematically manipulate duration at the segmental level to create additional contrasts. Consonant duration varies with suprasegmental factors such as prosodic context, and change in gestural coordination affects other timing factors, such as VOT, burst duration, and pre-stopping, but because some consonant phonemes differ only in their intrinsic duration while other factors remain consistent, length can be considered an independent phonetic parameter of consonants.

A variety of terms have been used to describe temporal contrasts in consonants, not all of which involve comparable phonetic timing differences: short/long, fortis/lenis, and singleton/geminate. Where these contrasts occur in word-medial positions, they may arise from a heterosyllabic sequence of similar consonants, but consonant length contrasts are also found at word boundaries. The Cypriot Greek words [ˈpefti] *'Thursday'* and [ˈpːefti] *'he falls'* differ only in the duration of the initial labial stop (Muller, 2002), and Tashlhiyt Berber contrasts stop length in word-final position: [fit] *'give it*-MASC*'* vs. [fitː] *'give it*-FEM*'* (Ridouane, 2008). Although some languages distinguish three phonemic vowel lengths (Remijsen & Gilley, 2008), consonant length contrasts appear to be maximally binary (Gordon, 2016).

Acoustically, stop length contrasts are primarily realized in the duration of the closure phase (Abramson, 1986; Ladefoged & Traill, 1994; Ridouane, 2010). Differences in duration between short and long consonants varies considerably. Madurese (Austronesian, Indonesia) word-medial voiced geminate stops are 1.7 times the duration of singleton equivalents, and 3.2 times as long in Toba Batak (Cohn et al., 1999). Consonant length contrasts may also be encoded in additional cues, such as differences in intensity, burst characteristics, and differential influences on surrounding segments (Smith, 1993; Kirchner, 2000). The relationship between consonant duration and other phonetic parameters is complicated, and is an ongoing topic of research (Hankamer et al., 1989; Ham, 2001; Idemaru & Guion, 2008, etc.).

A separate length parameter is not needed to describe consonant length contrasts if segmental duration is determined by the stiffness of the underlying gestures (Browman & Goldstein, 1992). Stiffness is inversely related to gestural duration, so intrinsically longer consonants may consist of gestures specified for lower stiffness than their short counterparts. Other types of geminates may be produced as adjacent overlapping gestures (Smith, 1993).

3.3.9 Respiratory Strength

Consonants in some languages may also be partially differentiated by the degree of respiratory strength employed in the airstream mechanism. Stops contrasting this dimension are typically described as *fortis* and *lenis*, although these terms are used inconsistently to describe many properties of stops including voicing, aspiration, length, and "increased tension in the vocal apparatus" (Loos et al., 2004).

Contrasts in degree of respiratory strength are rare (Ladefoged & Maddieson, 1996), and because they are realized through the complex interaction of multiple phonetic parameters, it is an open question whether this should be considered an independent phonetic property of consonants. Fortis-lenis distinctions are made for most consonants in Itunyoso Trique (Mixtecan, Oaxaca), in contrasts such as [kãh˩] *naked* vs. [kːãh˧] *sandal*. Fortis obstruents are realized with greater glottal aperture than lenis equivalents, but the primary phonetic correlate is duration, suggesting that the contrast involves a combination of length and glottal timing features, not dissimilar to stop aspiration contrasts in other languages (DiCanio, 2012).

Korean oral stops may be differentiated in part by respiratory differences. At each place of articulation, Korean contrasts aspirated vs. two types of unaspirated plosives, e.g. [t^hal] *'mask'* / [tal] *'moon'* / [t*al] *'daughter'*. The unaspirated stops are variously described as ‘plain’/‘tense’, ‘lenis’/‘fortis’, and ‘unaspirated’/‘stiff voice’ (Ladefoged & Maddieson, 1996). Fortis stops are characterized by “higher intraoral pressure before release, yet a lower oral flow after release, than corresponding lenis stops” Dart (1987, 138). However, coronal stops are also differentiated by intensity (Kim, 1965), supralaryngeal articulatory factors – most importantly constriction duration (Son et al., 2011) – and the f0 and voice quality of the following vowel (Cho et al., 2002). More recent studies also suggest a merger in VOT between fortis and lenis stops for some Korean speakers (Kang, 2014). The evidence from Korean and other languages (Butcher 2004) therefore suggests that respiratory strength is a phonetic parameter of consonant contrast which interacts with other properties of stops. More typological research is needed to understand how respiratory force may vary at a segmental level in other languages.

3.3.10 Complexity arising from Articulatory Coordination

Many other types of consonants differ from prototypical segments in ways typically described as ‘complex’: they may be accompanied by additional articulations, or their properties may change over time. Laver (1994) draws a distinction between *complex* and *simplex* segments based on the relative timing of articulators during the medial (steady-state) articulatory phases, so that pre-stopped nasals are classified complex segments, but affricates are simplex. Other taxonomies distinguish between *primary* and *secondary articulations*, and treat consonants as having fundamental articulatory structures which can be augmented though various *accompanying* and *enhancing* articulations.

Problematic distinctions between ‘simple’ and ‘complex’ segments are unnecessary when consonants are described in terms of the constituency and coordination of their underlying gestures (Browman & Goldstein, 1989, 2000). Some commonly attested complex consonants are listed in Table 3.5, with the gestural reconfigurations that differentiate them from canonical realizations, and illustrations of languages that use these consonants phonemically. Large consonant inventories systematically exploit additional coordinative structures of multiple gestures, but all may be described using some combination of parameters listed in Table 3.1.

Table 3.5: **Articulatory Characterisation of Complex Consonants**.

COMPLEXITY	MECHANISM	EXAMPLE
Aspiration	Glottal opening lags supraglottal gesture	Mandarin /t^{h}/
Pre-aspiration	Glottal opening leads supraglottal gesture	Icelandic /^{h}t/
Affrication	Critical constriction during stop release	German /t͡s/
Pre-nasalization	Velum lowered before oral closure	Guarani /^{n}t/
Pre-stopping	Oral closure precedes accompanying gesture	Kaytetye /^{t}n/
Nasal release	Velum lowered during release of oral gesture	Wolof [k^{ŋ}]
Lateral release	Lateralization lags lingual closure gesture	Chipewean /t^{l}/
Labialization	Synchronous labial gesture	Arrernte /t^{w}/
Palatalization	Synchronous TB palatal gesture	Irish /t^{j}/
Velarization	Synchronous TB velar gesture	Marshallese /t^{ɣ}/
Pharyngealization	Synchronous TB pharyngeal gesture	Arabic /t^{ʕ}/
Glottalization	Synchronous glottalic constriction gesture	Vietnamese /^{ʔ}d/

3.4 Current Research in Consonantal Phonetics

Research on consonants continues to focus on themes related to goals of production, timing, variation, and representation. Greater insights into these properties are emerging through new research methods in Laboratory Phonology and as more diverse types of data become available.

3.4.1 Goals of Production

Debate continues over the goals of consonant production, and whether they are primarily articulatory or acoustic in nature. Rhotic consonants have been a focus of much research on this topic; English rhotic approximants in particular have provided a case study in the nature and scope of segmental variability. Nieto-Castanon et al. (2005) argue that American English /ɹ/ is characterized by greater variability in the articulatory domain, but united by acoustic stability (Guenther et al., 1999), specifically a lowered third formant (Westbury et al., 1998; Espy-Wilson et al., 2000). On the other hand, there is evidence for articulatory commonalties. Despite considerable variation in the specific configuration of the tongue, and the tongue tip in particular (Delattre & Freeman, 1968; Hagiwara, 1994), the many rhotic variants appear to share three main goals of production: formation of coordinated constrictions at the lips, palate and pharynx (Gick et al., 2003; Gick & Campbell, 2003; Iskarous, 2006; Zhou et al., 2008). Mielke et al. (2016) show that rhotic articulation is often speaker-specific with idiosyncratic allophonic distributions that may factor in sound change.

Formant transitions – and what they reveal about the nature of coarticulation – are another set of phenomena at the centre of this debate. Because the formant patterns associated with coronals depend on vocalic context (Liberman et al., 1967), these and other acoustic properties of consonants have been described as variant cues to invariant articulatory primitives (Fowler, 1994, 2005). An alternative interpretation is that the invariance can be found at the acoustic point of convergence of families of formant transitions associated with consonants produced at the same place (Stevens & Blumstein, 1978). The interpretation of place cues, and *locus equations* in particular, continues to be an active area of research into the nature of phonetic primitives (Stevens, 2004; Iskarous et al., 2010; Lindblom & Sussman, 2012; Viswanathan et al., 2014; Chen et al., 2015, etc.).

Computational modelling offers new kinds of insights into goals of consonant production. DIVA (Tourville & Guenther, 2011; Guenther, 2016) facilitates neural network modelling of speech motor control and its acoustic consequences, including auditory feedback simulations. TADA (Nam et al., 2006) allows testing of hypotheses about gestural structure and inter-articulator coordination, and is facilitating new avenues of research into vocal tract control during consonant production (Lammert et al., 2013; Ramanarayanan et al., 2016).

3.4.2 Status of the Segment

Although the segment has been the most common unit of sound specification, a growing body of evidence suggests that many mechanisms of perception and production primarily operate at lexical and subsegmental levels, which raises questions about the status of consonants as atomic units.

Much current research on consonants deals with the properties of their constituent gestures, and intergestural timing (Byrd et al., 2009; Scobbie & Pouplier, 2010; Tilsen & Goldstein, 2012, etc.). Segments are modelled as stable coordinative relationships of gestures that recur in phonological systems (Browman & Goldstein, 1990). Although the status of consonants as independent phonological units is not precluded in this framework, phonological and phonetic phenomena are primarily described and explained at the sub-segmental level using gestures, the atoms of phonology.

Another body of work is exploring additional dimensions of phonetic encoding that present challenges for traditional segmental models of the consonant. Phonetic realizations of words include rich indexical information about the speaker, and this encoding is a source of a significant phonetic variability, much of which is not predictable (Foulkes & Docherty, 2006; Hay & Drager, 2007). Other sources of phonetic variability in consonants are word frequency and morphological structure (Pierrehumbert, 2001; Hay, 2004; Kuperman et al., 2007). To account for these phenomena, *exemplar models* (Johnson & Mullennix, 1997; Pierrehumbert, 2002, etc.) posit lexical representations built over cumulative sets of phonetically-rich word-level episodic traces. The relative roles of episodic and abstract representations in the lexicon, and the mechanisms by which segmental units might be derived from word-level exemplars in these models remain active areas of research (McQueen et al., 2006; Ernestus, 2014).

3.4.3 Temporal Organization of Consonants

A special focus of current research on consonant production is gestural timing, and how it is influenced by different structural and prosodic environments. The coronal and dorsal gestures of English laterals exhibit different timing patterns in onsets and codas (Sproat & Fujimura, 1993; Scobbie & Pouplier, 2010). Syllable position also influences gestural constituency and coordination in English rhotics (Zawadzki & Kuehn, 1980; Mielke et al., 2016), and in nasals (Krakow, 1999; Byrd et al., 2009). In complex onsets, consonant gestures overlap more with each other and with the tautosyllabic vowel than in coda clusters, where multiple consonants are realized more sequentially (Marin & Pouplier, 2010). Variations of this *C-centre* effect have been observed in English (Browman & Goldstein, 1988), Georgian (Chitoran, 2002), Tashlhiyt Berber (Goldstein et al., 2007), Romanian (Marin & Pouplier, 2014), Morrocan Arabic (Shaw et al., 2009), Italian (Hermes et al., 2008), and French (Kühnert et al., 2006), but Slovak coda consonants may be subject to similar principles of organization as in onset clusters (Pouplier & Beňuš, 2011). More research in a wider range of languages is required to better understand segmental and clustering properties of consonants.

3.5 Best Practices for Teaching and Learning

The phonetics of consonants can be illustrated in new ways with typological databases, images and models of the vocal tract, and annotated corpora. Consonant production and perception is best understood when these resources are used to enhance traditional pedagogies built around textbooks, practical exercises, and physical demonstrations of vocal tract acoustics. Many tools are freely available online, providing diverse and enriched new ways for students to learn about consonants.

3.5.1 Key References

Ladefoged (2005), Ladefoged & Johnson (2014), Johnson (2011), and Gick et al. (2012) all work well as foundational textbooks on consonant phonetics. References describing students' languages and the language of instruction should also be made available, including their own dialects if these differ from British and American English (e.g. Cox & Fletcher, 2017). Ladefoged & Maddieson (1996) and Gordon (2016) provide important cross-linguistic perspective on consonants in languages throughout the world.

3.5.2 Practical Exercises

The phonetic properties of consonants are best understood when students learn to manipulate them themselves. A useful exercise for learning about place and manner contrasts is to produce pulmonic *ingressive* fricatives while sensing the refrigerating effect of the airflow at the point of minimum constriction. Students can practise adjusting stricture and place while referring to the IPA chart, then reverse the direction of airflow to produce standard egressive consonants at the same places of articulation. Using this method, students can explore the range of constrictions associated with different consonants in their own vocal tract.

Ultrasound imaging is an especially engaging method of teaching articulatory phonetics. Video from a portable ultrasound machine projected onto a large screen in a classroom allows students to experiment with producing different sounds while monitoring their own articulations and those of their classmates in real time. Students typically respond well to invitations to demonstrate consonants in their own languages, which can initiate animated discussions about the phonetic differences between languages and sociophonetic variation amongst speakers and varieties. Online resources (Sections 3.5.5–3.5.6) can also be used to encourage experimentation with other accents and consonant sounds.

3.5.3 Physical Demonstrations

Many phonetic properties of consonants can be taught by physical demonstration, preferably involving students in experimentation. Principles of vocal tract acoustics can be dramatically illustrated through physical analogies in other bodies. Musical instruments, plastic pipes, rulers, water columns, and desks can all be manipulated to produce 'consonant' sounds. Ropes, guitar strings, sand patterns, water-filled bottles, and rubber bands can be used to demonstrate standing waves. Prisms can be used to separate light when teaching about frequency. Physical demonstrations of acoustic principles are most pedagogically effective when followed with production exercises involving consonants that exploit the relevant phonetic dimensions.

3.5.4 Phonological Databases

Cross-linguistic comparisons of consonants in the world's languages can be taught with reference to phonological databases such as PHOIBLE (http://phoible.org), P-base (Mielke, 2007), and WALS (http://wals.info), which visualizes typological data on configurable maps. These resources provide important perspective on language diversity beyond the Western European languages that dominate academic discourse.

3.5.5 Enhanced IPA Charts

Online IPA charts such as http://www.ipachart.com, and https://web.uvic.ca/ling/resources/ipa/charts/IPAlab/IPAlab.htm, provide hyperlinked audio recordings illustrating the consonant sounds represented by each IPA symbol. Other charts include MRI (http://sail.usc.edu/span/rtmri_ipa), ultra-sound, and schematic videos (http://www.seeingspeech.ac.uk) illustrating articulation of each consonant in different vowel contexts. Some excellent apps demonstrating consonant phonetics are available, including *Sounds of Speech* and *IPA Phonetics*.

3.5.6 Multimedia Resources

Video sharing sites host a remarkable range of videos useful for teaching phonetics, including lecture recordings, laryngeographic videos, demonstrations of phonation types and voice qualities, and recordings of rare languages. Configurable vocal tract models (e.g. http://smu-facweb.smu.ca/~s0949176/sammy/) allow students to experiment with articulatory parameters in an interactive sagittal section. https://dood.al/pinktrombone includes a synthesizer which provides real-time feedback on the acoustic consequences of articulatory reconfiguration. Sidney Wood's website (https://swphonetics.com) has annotated x-ray sequences and transcribed waveforms describing details of consonant production and coarticulation. The UCLA Phonetics Lab Archive (http://archive.phonetics.ucla.edu/) illustrates consonant contrasts with example language data.

3.6 Future Directions

Improved experimental technologies and the increasing availability of large-scale phonetic data are opening up new ways of investigating consonants. Important themes which can be addressed with these new research thrusts are individual speaker variability, multimodal aspects of consonant production and perception, and description of consonants in understudied languages.

3.6.1 New Methods in Instrumental Phonetics

New techniques for real-time MRI (e.g. Niebergall et al., 2013; Lingala et al., 2017) offer greater temporal and spatial resolution of the vocal tract, which is important in the study of clicks, trills, and other consonants involving rapid articulatory movement. Higher frame rate imaging is also important to examine details of connected speech processes and coarticulation (Fu et al., 2015). Improved imaging of the pharynx and velum will lead to a better understanding of pharyngealization (Israel et al., 2012), tongue root control in liquids and other consonants (Proctor & Walker, 2012; Proctor et al., 2013), and velic coordination in nasals (Carignan et al., 2015).

Dynamic three-dimensional (3D) imaging of the vocal tract is also becoming viable for speech studies (e.g. Zhu et al., 2013; Fu et al., 2017), and will eventually provide high frame rate volumetric images of the entire vocal tract. To date, most phonetic data have been restricted to the vocal tract midline: x-ray, XRMB, ultrasound, EMA, and MRI data are typically acquired or analysed exclusively in the midsagittal plane. Yet consonants differ in important ways in their parasagittal configurations, especially those with complex lingual geometries, such as liquids and fricatives (Fletcher & Newman, 1991; Alwan et al., 1997; Narayanan et al., 1999). Dynamic 3D MRI will allow detailed examination of consonant production in all parts of the vocal tract. These data will also provide volumetric geometries for more sophisticated 3D vocal tract models, to inform studies of articulatory-acoustic relationships in consonants.

3.6.2 Large-scale Phonetic Data

Large-scale speech corpora offer new opportunities for consonant research using 'big data' approaches. Specialized phonetically-annotated multimedia corpora (e.g. Cooke et al., 2006; Narayanan et al., 2014) provide companion acoustic and articulatory data of consonants produced in different phonological environments, by many more speakers than could typically be recruited for laboratory studies. Automatic analysis of large datasets using techniques such as Principal Components Analysis (PCA: Hueber et al., 2007) is shedding new light on patterns of coarticulation and individual speaker differences in consonant production (Mielke et al., 2016; Hoole & Pouplier, 2017).

A growing volume of general speech data has recently become generally accessible through the proliferation of online multimedia content, but remains largely underutilized as a phonetic research resource. Thousands of hours of acoustic speech recordings are available online in the form of audio books, many freely accessible in the public domain. These recordings constitute a distributed multi-speaker database which, combined with tools for forced-aligned automatic transcription (e.g. Katsamanis et al., 2011), are suitable for large-scale multilingual, multi-speaker acoustic studies of consonant realization and coarticulation in a variety of prosodic contexts, and for analysis of individual speaker variability.

3.6.3 Typologically Incomplete Data

Ladefoged & Maddieson (1996) remains the most comprehensive cross-linguistic survey of consonants, yet their analyses are based on data from 298 languages, which represents less than 5% of languages spoken today (Simons & Fennig, 2017). Phonological databases such as UPSID (Maddieson, 1992) and P-base (Mielke, 2007), contain data on up to 600 languages – still less than a tenth of the world's languages – and for many of those there is little or no phonetic data. Although it is likely that consonants in these languages will share many properties with those that have been studied, we cannot assume that this is always the case. Ubykh (North Caucasian, Turkey) appears to have the largest attested ratio of consonants to vowels (84:2), but there is still no consensus on the phonological structure of this language (Colarusso, 1992). The last Ubykh speaker died in 1992, so the opportunity for further fieldwork to shed more light on these issues has passed.

No definitive account of the type and scope of variation in consonant systems is possible while so many of the world's languages remain relatively unexamined, and the prospects of obtaining data from most of these languages are rapidly diminishing (Harrison, 2008). There is an urgent need for more phonetic fieldwork to document the properties of consonants that have not been examined, ideally with instrumental, as well as acoustic data. Portable ultrasound technology now provides a viable method for collection of articulatory data which would previously have been impossible during fieldwork in remote areas (Gick, 2002).

3.7 References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Abramson, A. S. (1986). The perception of word-initial consonant length: Pattani Malay. *JIPA*, *16*(1), 8–16.

Allen, W. S. (1981). The Greek contribution to the history of phonetics. In R. E. Asher, & E. J. Henderson (Eds.) *Towards a history of phonetics*, (pp. 115–122). Edinburgh: Edinburgh University Press.

Alwan, A., Narayanan, S., & Haker, K. (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. *JASA*, *101*(2), 1078–1089.

Aoki, H. (1970). A note on glottalized consonants. *Phonetica*, *21*(2), 65–74.

Ball, M. J. (1990). The lateral fricative: lateral or fricative. *Celtic Linguistics/Ieithyddiaeth Geltaidd. Readings in the Brythonic Languages: Festschrift for T. Arwyn Watkins. Amsterdam: John Benjamins*, (pp. 109–25).

Barry, W. J. (1997). Another r-tickle. *JIPA*, *27*(1-2), 3545.

Bell, A. M. (1863). *The Principles of Speech and Vocal Physiology, and Dictionary of Sounds, Etc*. London: Hamilton & Adams.

Bird, S., Caldecott, M., Campbell, F., Gick, B., & Shaw, P. A. (2008). Oral–laryngeal timing in glottalised resonants. *J. Phon*, *36*(3), 492–507.

Blevins, J. (2003). The phonology of Yurok glottalized sonorants: segmental fission under syllabification. *International Journal of American Linguistics*, *69*(4), 371–396.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2-4), 140–155.

Browman, C. P., & Goldstein, L. (1995). Dynamics and articulatory phonology. In T. van Gelder, & B. Port (Eds.) *Mind as motion: Explorations in the dynamics of cognition*, (pp. 175–193). Cambridge, MA: MIT Press.

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

Browman, C. P., & Goldstein, L. M. (1990). Gestural specification using dynamically-defined articulatory structures. *J. Phon*, *18*(3), 299–320.

Browman, C. P., & Goldstein, L. M. (1992). Articulatory phonology: an overview. *Phonetica*, *49*(3-4), 155–180.

Browman, C. P., & Goldstein, L. M. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée*, *5*, 25–34.

Brown, L. (2001). *A grammar of Nias Selatan*. Ph.D. thesis.

Brucke, E. (1856). *Grundzüge der Physiologie und Systematik der Sprachlaute für Linguisten und Traub-Stummenlehrer*. Vienna: Carl Gerold.

Butcher, A. (2004). 'Fortis/lenis' revisited one more time: the aerodynamics of some oral stop contrasts in three continents. *Clinical Linguistics & Phonetics, 18(6-8)*: 547–557.

Byrd, D., Tobin, S., Bresch, E., & Narayanan, S. (2009). Timing effects of syllable structure and stress on nasals: a real-time MRI examination. *J. Phon*, *37*, 97–110.

Carignan, C., Shosted, R. K., Fu, M., Liang, Z.-P., & Sutton, B. P. (2015). A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *J. Phon*, *50*, 34–51.

Catford, J., & Esling, J. (2006). Phonetics, articulatory. In K. Brown (Ed.) *Encyclopedia of Language & Linguistics (2nd Edition)*, (pp. 425–442). Oxford: Elsevier.

Catford, J. C. (1939). On the classification of stop consonants. *Le Maître Phonétique*, *65*, 2–5.

Catford, J. C. (1964). Phonation types: the classification of some laryngeal components of speech production. In D. Abercrombie (Ed.) *In Honour of Daniel Jones*, (pp. 26–37). London: Longmans.

Chafe, W. L. (1960). Seneca morphology I: introduction. *International Journal of American Linguistics*, *26*(1), 11–22.

Chen, W.-R., Chang, Y.-C., & Iskarous, K. (2015). Vowel coarticulation: Landmark statistics measure vowel aggression. *JASA*, *138*(2), 1221–1232.

Chiba, T., & Kajiyama, M. (1941). *The vowel – its nature and structure*. Tokyo: Kaseikan Pub. Co.

Chitoran, I. (2002). A perception-production study of Romanian diphthongs and glide-vowel sequences. *JIPA*, *32/2*, 203–222.

Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *J. Phon*, *30*(2), 193–228.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Cohn, A. C., Ham, W. H., & Podesva, R. J. (1999). The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. In In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. C. Bailey, eds., *Proc. 14th Intl. Congress of Phonetic Sciences* (Vol. 1, pp. 587–590). Berkeley, CA: ICPhS.

Colarusso, J. (1992). How many consonants does Ubykh have? *Caucasian Perspectives*, (pp. 145–156).

Cooke, M., Barker, J., Cunningham, S., & Shao, X. (2006). An audio-visual corpus for speech perception and automatic speech recognition. *JASA*, *120*(5), 2421–2424.

Coulmas, F. (Ed.) (1999). *The Blackwell Encyclopedia of Writing Systems*. Blackwell Reference Online. Blackwell.

Cox, F., & Fletcher, J. (2017). *Australian English pronunciation and transcription*. Cambridge: Cambridge University Press.

Dart, S. N. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. *JASA*, *81*(1), 138–147.

De Jong, K. (1998). Stress-related variation in the articulation of coda alveolar stops: Flapping revisited. *J. Phon*, *26*(3), 283–310.

Delattre, P., & Freeman, D. C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics*, *44*, 29–68.

Demolin, D. (1995). The phonetics and phonology of glottalized consonants in Lendu. In B. Connell, A. Arvaniti, J. Kingston, B. Connell, A. Arvaniti, & J. Kingston (Eds.) *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* , (pp. 368–385). Cambridge: Cambridge University Press.

Derrick, D., & Gick, B. (2011). Individual variation in English flaps and taps: A case of categorical phonetics. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *56*(3), 307319.

DiCanio, C. T. (2012). The Phonetics of Fortis and Lenis Consonants in Itunyoso Trique. *International Journal of American Linguistics*, *78*(2), 239–272.

Diehl, R. (1989). Remarks on Steven's quantal theory of speech. *J. Phon*, *17*(1-2), 71–78.

Elliott, R. W. V. (1954). Isaac Newton as Phonetician. *The Modern Language Review*, *49*(1), 5–12.

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, *142*, 27–41.

Esling, J. H. (2010). Phonetic notation. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.) *The Handbook of Phonetic Sciences*, (pp. 678–702). Oxford: Blackwell, 2nd ed.

Esling, J. H., Fraser, K. E., & Harris, J. G. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *J. Phon*, *33*(4), 383–410.

Esling, J. H., & Harris, J. G. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In W. J. Hardcastle, & J. M. Beck (Eds.) *A figure of speech: A festschrift for John Laver*, (pp. 347–383). Mahwah, NJ: Lawrence Erlbaum.

Espy-Wilson, C. Y., Boyce, S. E., Jackson, M., Narayanan, S., & Alwan, A. (2000). Acoustic modeling of American English /r/. *JASA*, *108*(1), 343–356.

Everett, D. L. (1986). Pirah˜a. In D. C. Derbyshire, & G. K. Pullum (Eds.) *Handbook of Amazonian languages*, vol. 1, (p. 200326). Berlin: Mouton de Gruyter.

Ewan, W. G., & Krones, R. (1974). Measuring larynx movement using the thyroumbrometer. *J. Phon*, *2*, 327–335.

Fant, G. (1960). *Acoustic theory of speech production, with calculations based on X-ray studies of Russian articulations*. s'Gravenhage: Mouton.

Firchow, I., & Firchow, J. (1969). An abbreviated phoneme inventory. *Anthropological Linguistics*, (pp. 271–276).

Flemming, E., Ladefoged, P., & Thomason, S. (2008). Phonetic structures of Montana Salish. *J. Phon*, *36*(3), 465–491.

Fletcher, S. G., & Newman, D. G. (1991). [s] and [ʃ] as a function of linguapalatal contact place and sibilant groove width. *JASA*, *89*(2), 850–858.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *J. Phon*, *34*(4), 409–438.

Fowler, C. A. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Attention, Perception, & Psychophysics*, *55*(6), 597–610.

Fowler, C. A. (2005). Parsing coarticulated speech in perception: effects of coarticulation resistance. *J. Phon*, *33*(2), 199 – 213.

Fu, M., Barlaz, M. S., Holtrop, J. L., Perry, J. L., Kuehn, D. P., Shosted, R. K., Liang, Z.-P., & Sutton, B. P. (2017). High-frame-rate full-vocaltract 3D dynamic speech imaging. *Magnetic Resonance in Medicine*, *77*(4), 1619–1629.

Fu, M., Zhao, B., Carignan, C., Shosted, R. K., Perry, J. L., Kuehn, D. P., Liang, Z.-P., & Sutton, B. P. (2015). High-resolution dynamic speech imaging with joint low-rank and sparsity constraints. *Magnetic Resonance in Medicine*, *73*(5), 1820–1832.

Gick, B. (2002). The use of ultrasound for linguistic phonetic fieldwork. *JIPA*, *32*(2), 113–121.

Gick, B., & Campbell, F. (2003). Intergestural timing in English /r/. In *Proc. 15th Intl. Congress of Phonetic Sciences*, (pp. 1911–1914). ICPhS.

Gick, B., Iskarous, K., Whalen, D. H., & Goldstein, L. M. (2003). Constraints on variations in the production of English /r/. In *6th International Seminar on Speech Production*, (pp. 73–78). Sydney: ISSP.

Gick, B., Wilson, I., & Derrick, D. (2012). *Articulatory phonetics*. Wiley.

Giles, S. B., & Moll, K. L. (1975). Cinefluorographic study of selected allophones of English /l/. *Phonetica*, *31*(3-4), 206–227.

Gobl, C. (2010). Voice source variation and its communicative functions. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.) *The handbook of phonetic sciences*, vol. 1, chap. 11, (pp. 378–423). Wiley, 2nd ed.

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, *40*, 189–212.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In *Proc. XVIth international congress of phonetic sciences*, (pp. 241–244). Citeseer.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N. O. Schiller, & A. S. Meyer (Eds.) *Phonetics and phonology in language comprehension and production: Differences and similarities*, (pp. 159–207). de Gruyter.

Gordon, M., & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *J. Phon*, *29*(4), 383–406.

Gordon, M. K. (2016). *Phonological typology*, vol. 1. Oxford University Press.

Grawunder, S., Simpson, A. P., & Khalilov, M. (2010). Phonetic characteristics of ejectives–samples from Caucasian languages. *Turbulent sounds. An interdisciplinary guide*, (21), 209–244.

Guenther, F. H. (2016). *Neural control of speech*. MIT Press.

Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., & Perkell, J. S. (1999). Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *JASA*, *105*(5), 2854–2865.

Güldemann, T. (2001). Phonological regularities of consonant systems across Khoisan lineages. *University of Leipzig Papers on Africa*, *16*, 1–50.

Hagiwara, R. (1994). Three types of American /r/. *UCLA Working Papers in Phonetics*, (pp. 55–62).

Ham, W. H. (2001). *Phonetic and Phonological Aspects of Geminate Timing*. New York: Routledge.

Hankamer, J., Lahiri, A., & Koreman, J. (1989). Perception of consonant length: Voiceless stops in Turkish and Bengali. *J. Phon*, *17*(4), 283–298.

Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *J. Phon*, *29*(4), 451–480.

Harrison, K. D. (2008). *When languages die: The extinction of the world's languages and the erosion of human knowledge*. Oxford University Press.

Hay, J. (2004). *Causes and consequences of word structure*. Routledge.

Hay, J., & Drager, K. (2007). Sociophonetics. *Annu. Rev. Anthropol.*, *36*, 89–103.

Henton, C., Ladefoged, P., & Maddieson, I. (1992). Stops in the worlds languages. *Phonetica*, *49*(2), 65–101.

Hermes, A., Grice, M., Mücke, D., & Niemann, H. (2008). Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. In R. Sock, S. Fuchs, & Y. Laprie (Eds.) *Proc. 8th Intl. Seminar on Speech Production*, (pp. 433–436). Strasbourg, France: INRIA.

Heselwood, B., Hassan, Z. M., & Jones, M. J. (2013). Historical overview of phonetics. In M. J. Jones, & R.-A. Knight (Eds.) *The Bloomsbury companion to phonetics*, (pp. 5–20). Bloomsbury Publishing.

Hoole, P. (1998). Do airstream mechanisms influence tongue movement paths? *Phonetica*, *55*(3), 131–146.

Hoole, P., & Pouplier, M. (2017). Öhman returns: New horizons in the collection and analysis of imaging data in speech production research. *Computer Speech & Language*, *45* (Supplement C), 253 – 277.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge: Cambridge University Press.

Hueber, T., Aversano, G., Cholle, G., Denby, B., Dreyfus, G., Oussar, Y., Roussel, P., & Stone, M. (2007). Eigentongue feature extraction for an ultrasound-based silent speech interface. In *IEEE Intl. Conf. on Acoustics, Speech and Signal Processing - ICASSP07*, (pp. I–1245 – I–1248).

Idemaru, K., & Guion, S. G. (2008). Acoustic covariants of length contrast in Japanese stops. *JIPA*, *38*(2), 167–186.

Iskarous, K. (2006). The Articulation of the Palatal Gesture in American English [r]. In *Proceedings of ISSP 2006. 7th International Seminar on Speech Production*. Ubatuba, Brazil: CEFALA.

Iskarous, K., Fowler, C. A., & Whalen, D. H. (2010). Locus equations are an acoustic expression of articulator synergy. *JASA*, *128*(4), 2021–2032.

Iskarous, K., Shadle, C. H., & Proctor, M. I. (2011). Articulatory-acoustic kinematics: The production of American English /s/. *JASA*, *129*(2), 944–954.

Israel, A., Proctor, M., Goldstein, L., Iskarous, K., & Narayanan, S. S. (2012). Emphatic segments and emphasis spread in Lebanese Arabic: a Real-time Magnetic Resonance Imaging Study. In *Proc. Int'l Conf. on Speech Communication and Technology*. Portland, OR.

Jespersen, O. (1889). *The articulations of speech sounds represented by means of analphabetic symbols*. NG Elwert.

Johnson, K. (2011). *Acoustic and Auditory Phonetics*. Acoustic and Auditory Phonetics. Wiley.

Johnson, K., & Mullennix, J. W. (1997). Speech perception without speaker normalization: an exemplar model. In K. Johnson, & J. W. Mullennix (Eds.) *Talker variability in speech processing*, (pp. 145–166). San Diego: Academic Press.

Kang, Y. (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *J. Phon*, *45*, 76–90.

Katsamanis, A., Black, M., Georgiou, P., Goldstein, L., & Narayanan, S. (2011). SailAlign: Robust long speech-text alignment. In *Proc. Workshop on New Tools* *and Methods for Very-Large Scale Phonetics Research*, (pp. 44–47). Philadelphia: University of Pennsylvania.

Kelso, J. S., Tuller, B., & Harris, K. S. (1983). A "dynamic pattern" perspective on the control and coordination of movement. In *The production of speech*, (pp. 137–173). Springer.

Kemp, J. A. (2006). Phonetic transcription: History. In K. Brown (Ed.) *Encyclopedia of Language & Linguistics (2nd Ed.)*, (pp. 396–410). Oxford: Elsevier.

Khan, S., Esposito, C. M., & Hurst, A. (2007). Breathy nasals and /Nh/ clusters in Bengali, Hindi, and Marathi. *Indian Linguistics*, *68*, 275.

Kim, C.-W. (1965). On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word*, *21*(3), 339–359.

Kim, H., Honda, K., & Maeda, S. (2005). Stroboscopic-cine MRI study of the phasing between the tongue and the larynx in the Korean three-way phonation contrast. *J. Phon*, *33*(1), 1–26.

Kiparsky, P. (1991). Economy and the Construction of the Sivasutras. In M. Deshpande, & S. Bhate (Eds.) *Paninian Studies. Festschrift for Prof. S.D. Joshi*, 37, (pp. 239–261). Ann Arbor, Michigan: Centers for South Asian Studies, University of Michigan.

Kirchner, R. (2000). Geminate inalterability and lenition. *Language*, (pp. 509–545).

Krakow, R. A. (1999). Physiological organization of syllables: a review. *J. Phon*, *27*(1), 23–54.

Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and C-center in selected French consonant clusters. In *7th Intl. Seminar on Speech Production (ISSP)*, (pp. pp. 327–334). Ubatuba, Brazil.

Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *JASA*, *121*(4), 2261–2271.

Ladefoged, P. (1968). *A Phonetic Study of West African Languages*. Cambridge: Cambridge University Press.

Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. Chicago: University of Chicago Press.

Ladefoged, P. (2005). *Vowels and consonants (2nd Ed.)*. Malden MA: Blackwell.

Ladefoged, P., Cochran, A., & Disner, S. (1977). Laterals and trills. *JIPA*, *7*(2), 46–54.

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics*. Nelson Education.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford; Cambridge, MA: Blackwell.

Ladefoged, P., & Traill, A. (1994). Clicks and their accompaniments. *J. Phon*, *22*, 3364.

Ladefoged, P., Williamson, K., & Elugbe, B. U. (1976). The Stops of Owerri Igbo. *Studies in African Linguistics*, *7*, 147–163.

Lammert, A., Goldstein, L., Narayanan, S., & Iskarous, K. (2013). Statistical methods for estimation of direct and differential kinematics of the vocal tract. *Speech Communication*, *55*(1), 147–161.

Lancia, L., Voigt, D., & Krasovitskiy, G. (2016). Characterization of laryngealization as irregular vocal fold vibration and interaction with prosodic prominence. *J. Phon*, *54*(Supplement C), 80 – 97.

Laver, J. (1980). The phonetic description of voice quality. *Cambridge Studies in Linguistics London*, *31*, 1–186.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Laver, J. D. (1968). Voice quality and indexical information. *British Journal of Disorders of Communication*, *3*(1), 43–54.

Lepsius, C. R. (1863). Standard alphabet. *London*, *275*, 277.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, *74*(6), 431.

Lindau, M. (1984). Phonetic differences in glottalic consonants. *J. Phon*, *12*, 147–155.

Lindblom, B., & Engstrand, O. (1989). In what sense is speech quantal? *J. Phon*, *17*(1-2), 107–121.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1983). Self-organizing processes and the explanation of phonological universals. *Linguistics*, *21*(1), 181–204.

Lindblom, B., & Maddieson, I. (1988). Phonetic universals in consonant systems. *Language, speech, and mind*, (pp. 62–78).

Lindblom, B., & Sussman, H. M. (2012). Dissecting coarticulation: How locus equations happen. *J. Phon*, *40*(1), 1–19.

Lingala, S. G., Zhu, Y., Kim, Y.-C., Toutios, A., Narayanan, S., & Nayak, K. S. (2017). A fast and flexible MRI system for the study of dynamic vocal tract shaping. *Magnetic Resonance in Medicine*, *77*(1), 112–125.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384–422.

Loos, E. E., Anderson, S., Day, D. H., Jordan, P. C., & Wingate, J. D. (2004). *Glossary of linguistic terms*, vol. 29. Dallas: SIL International.

Maddieson, I. (1980). Phonological Generalizations from the UCLA Phonological Segment Inventory Database (UPSID). *UCLA Working Papers in Phonetics*, *50*, 57–68.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, I. (1992). *UCLA phonological segment inventory database*. Los Angeles: UCLA.

Maddieson, I. (2013a). Consonant inventories. In M. S. Dryer, & M. Haspelmath (Eds.) *The World Atlas of Language Structures Online*, chap. 1. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Maddieson, I. (2013b). Glottalized consonants. In M. S. Dryer, & M. Haspelmath (Eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Maddieson, I., & Emmorey, K. (1984). Is there a valid distinction between voiceless lateral approximants and fricatives? *Phonetica*, *41*(4), 181–190.

Maddieson, I., & Emmorey, K. (1985). Relationship between semivowels and vowels: cross-linguistic investigations of acoustic difference and coarticulation. *Phonetica*, *42*(4), 163–174.

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*(3), 380–407.

Marin, S., & Pouplier, M. (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. *J. Phon*, *42*, 24 – 36.

McDonough, J., & Wood, V. (2008). The stop contrasts of the Athabaskan languages. *J. Phon*, *36*(3), 427–449.

McGowan, R. S. (1992). Tongue-tip trills and vocal-tract wall compliance. *JASA*, *91*(5), 2903–2910.

McGowan, R. S., & Howe, M. S. (2007). Compact green's functions extend the acoustic theory of speech production. *J. Phon*, *35*(2), 259–270.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30*(6), 1113–1126.

Mielke, J. (2007). P-base. URL: http://pbase.phon.chass.ncsu.edu/

Mielke, J. (2008). *The emergence of distinctive features*. Oxford; New York: Oxford University Press.

Mielke, J., Baker, A., & Archangeli, D. (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /r/. *Language*, *92*(1), 101–140.

Miller, A. (2010). Tongue body and tongue root shape differences in N|uu clicks correlate with phonotactic patterns. In S. Fuchs, M. Toda, & M. Zygis (Eds.) *Turbulent sounds: An interdisciplinary guide*, chap. 7, (pp. 245–280). Berlin/New York: de Gruyter.

Miller, A. L. (2011). The representation of clicks. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.) *The Blackwell Companion to Phonology*, vol. 1, chap. 18, (pp. 416–439). Malden: Wiley-Blackwell.

Mohammed, M. A. (2001). *Modern Swahili grammar*. East African Educational Publishers.

Moisik, S. R., & Esling, J. H. (2014). Modeling the biomechanical influence of epilaryngeal stricture on the vocal folds: A low-dimensional model of vocal–ventricular fold coupling. *Journal of Speech, Language, and Hearing Research*, *57*(2), S687–S704.

Moisik, S. R., Esling, J. H., & Crevier-Buchman, L. (2010). A high-speed laryngoscopic investigation of aryepiglottic trilling. *JASA*, *127*(3), 1548–1558.

Monnot, M., & Freeman, M. (1972). A Comparison of Spanish Single-tap /r/ with American /t/ and /d/ in Post-stress Intervocalic Position. In A. Valdman (Ed.) *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, (pp. 409–416). The Hague: Mouton.

Moran, S., McCloy, D., & Wright, R. (Eds.) (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: http://phoible.org/

Muller, J. S. (2002). On the theoretical implications of Cypriot Greek initial geminates. *Journal of Greek linguistics*, *3*(1), 115–137.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 457.

Nakagawa, H. (2006). *Aspects of the phonetic and phonological structure of the G|ui language*. Ph.D. thesis, University of the Witwatersrand.

Nam, H., Goldstein, L., Browman, C., Rubin, P., Proctor, M., & Saltzman, E. (2006). TADA (TAsk Dynamics Application) manual.

Narayanan, S., Byrd, D., & Kaun, A. (1999). Geometry, kinematics, and acoustics of Tamil liquid consonants. *JASA*, *106*(4), 1993–2007.

Narayanan, S., Toutios, A., Ramanarayanan, V., Lammert, A., Kim, J., Lee, S., Nayak, K., Kim, Y.-C., Zhu, Y., Goldstein, L., Byrd, D., Bresch, E., Ghosh, P., Katsamanis, A., & Proctor, M. (2014). Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research (TC). *JASA*, *136*(3), 1307–1311.

Nash, D. (1980). Topics in Warlpiri Grammar. New York: Garland Press.

Nichols, T. R., & Huyghues-Despointes, C. M. J. I. (2009). *Muscular Stiffness*, (pp. 2515–2519). Berlin, Heidelberg: Springer Berlin Heidelberg.

Niebergall, A., Zhang, S., Kunay, E., Keydana, G., Job, M., Uecker, M., & Frahm, J. (2013). Real-time MRI of speaking at a resolution of 33 ms: Undersampled radial FLASH with nonlinear inverse reconstruction. *Magnetic Resonance in Medicine*, *69*(2), 477–485.

Nieto-Castanon, A., Guenther, F. H., Perkell, J. S., & Curtin, H. D. (2005). A modeling investigation of articulatory variability and acoustic stability during American English /r/ production. *JASA*, *117*(5), 3196–3212.

Nolan, F., Handbook of the International Phonetic Association, et al. (1999). Introduction to the IPA. *JIPA*, *25*(1), 3–33.

Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.) *Nasalfest: Papers from a symposium on nasals and nasalization*, (pp. 289–316). Stanford: Language Universals Project.

Ohala, J. J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, & M. F. Miller (Eds.) *Papers from the Parasession on Language and*

Behavior, Chicago Linguistic Society, May 1-2, 1981, (pp. 178–203). Chicago: Chicago Linguistic Society, 1st ed.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.) *The Production of speech*, chap. 9, (pp. 189– 216). New York: Springer-Verlag.

Olson, K. (2004). Acoustic correlates of the labial flap. Boston. Note: LSA 78.

Oquendo, L. (2004). La vibrant uvular y la approximate labiodental en la lengua japreria como cultura fonológica. *Opción*, *20*(45), 60-74.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven, & N. Warner (Eds.) *Laboratory Phonology 7*, (pp. 101–139). Berlin: Mouton de Gruyter.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. L. Bybee, & P. Hopper (Eds.) *Frequency and the Emergence of Linguistic Structure*, (pp. 137–157). John Benjamins.

Pinho, C. M., Jesus, L. M., & Barney, A. (2012). Weak voicing in fricative production. *J. Phon*, *40*(5), 625–638.

Pinkerton, S. (1986). Quichean (Mayan) glottalized and nonglottalized stops: a phonetic study with implications for phonological universals. In J. Ohala, & J. Jaeger (Eds.) *Experimental Phonology*, (pp. 125–139). Orlando: Academic Press.

Podesva, R. J., & Callier, P. (2015). Voice quality and identity. *Annual Review of Applied Linguistics*, *35*, 173–194.

Pouplier, M., & Beňuš, Š. (2011). On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory phonology*, *2*(2), 243–273.

Prince, A., & Smolensky, P. (2008). *Syllable Structure Typology: The CV Theory*, (pp. 105–118). Blackwell.

Proctor, M., Bresch, E., Byrd, D., Nayak, K., & Narayanan, S. (2013). Paralinguistic mechanisms of production in human "beatboxing": a realtime magnetic resonance imaging study. *JASA*, *133*(2), 1043–1054.

Proctor, M., & Walker, R. (2012). Articulatory bases of English liquids. In S. Parker (Ed.) *The Sonority Controversy*, vol. 18 of *Studies in Generative Grammar*, (pp. 285–312). Berlin: De Gruyter.

Proctor, M., Zhu, Y., Lammert, A., Toutios, A., Sands, B., Hummel, U., & Narayanan, S. (2016). Lingual Consonant Production in Khoekhoe: a Real-time MRI

Study. In S. Shah, & M. Brenzinger (Eds.) *Khoisan Languages and Linguistics – Proc. 5th Intl. Symposium, July 1317, 2014, Riezlern/ Kleinwalsertal*, (pp. 337–366). Köln: Rüdiger Köppe Verlag.

Proctor, M. I., Shadle, C. H., & Iskarous, K. (2010). Pharyngeal articulation differences in voiced and voiceless fricatives. *JASA*, *127*(3), 1507–1518.

Ramanarayanan, V., Parrell, B., Goldstein, L., Nagarajan, S., & Houde, J. (2016). A new model of speech motor control based on task dynamics and state feedback. In *Proc. Int'l Conf. on Speech Communication and Technology*, (pp. 3564–3568).

Recasens, D. (1983). Place cues for nasal consonants with special reference to catalan. *JASA*, *73*(4), 1346–1353.

Remez, R. E., Rubin, P. E., Pisoni, D. B., Carrell, T. D., et al. (1981). Speech perception without traditional speech cues. *Science*, *212*(4497), 947–949.

Remijsen, B., & Gilley, L. (2008). Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *J. Phon*, *36*(2), 318–344.

Ridouane, R. (2008). Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology*, *25*(2), 321–359.

Ridouane, R. (2010). Geminates at the junction of phonetics and phonology. *Laboratory phonology*, *10*, 61–90.

Roon, K., Gafos, A. I., Hoole, P., & Zeroual, C. (2007). Influence of articulator and manner on stiffness. In *Proc. Int'l Congress on Phonetic Sciences*, (pp. 409–412).

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382.

Schutz, A. J. (1980). A reanalysis of the Hawaiian vowel system. *Oceanic Linguistics*, *20*, 1–43.

Scobbie, J. M., & Pouplier, M. (2010). The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in wordfinal English /l/. *J. Phon*, *38*(2), 240–259.

Shadle, C. (1999). The aerodynamics of speech. In W. J. Hardcastle, & J. Laver (Eds.) *The Handbook of Phonetic Sciences*. Blackwell Publishing.

Shadle, C. H. (1990). Articulatory-acoustic relationships in fricative consonants. In W. J. Hardcastle, & A. Marchal (Eds.) *Speech production and speech modelling*, vol. 55, (pp. 187–209). Dordrecht: Kluwer Academic Publisher.

Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2009). Temporal evidence for syllabic structure in Moroccan Arabic: Data and model. *Phonology*, *26*(1), 187–215.

Shosted, R. K. (2006). *The aeroacoustics of nasalized fricatives*. Ph.D. thesis, UC Berkeley.

Simons, G. F., & Fennig, C. D. (Eds.) (2017). *Ethnologue: Languages of the World, Twentieth edition*, vol. 2017. Dallas: SIL International, 20th ed. URL http://www.ethnologue.com

Smith, C. L. (1993). *The timing of vowel and consonant gestures*. Ph.D. thesis, Yale University.

Snyman, J. W. (1975). *Zhu/'hõasi fonologie & woordeboek*. Cape Town; Rotterdam: A.A. Balkema; University of Cape Town.

Solé, M.-J. (2002). Aerodynamic characteristics of trills and phonological patterning. *J. Phon*, *30*(4), 655–688.

Son, M.-J., Kim, S.-H., & Cho, T.-H. (2011). Supralaryngeal Articulatory Characteristics of Coronal Consonants in Korean. *Journal of the Korean society of speech sciences*, *3*(4), 33–43.

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *J. Phon*, *21*(3), 291–311.

Stevens, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants: Static considerations. *JASA*, *50*(4B), 1180–1192.

Stevens, K. N. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, *34*(4), 264–279.

Stevens, K. N. (1989). On the quantal nature of speech. *J. Phon*, *17*(1-2), 3–45.

Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge: The MIT Press.

Stevens, K. N. (2004). Invariance and variability in speech: Interpreting acoustic evidence. In J. Slifka, S. Manuel, & M. Mathies (Eds.) *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, (pp. B77–B85). MIT.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *JASA*, *64*(5), 1358–1368.

Sussman, H. M., & Shore, J. (1996). Locus equations as phonetic descriptors of consonantal place of articulation. *Attention, Perception, & Psychophysics*, *58*(6), 936–946.

Sweet, H. (1877). *A Handbook of Phonetics*, vol. 2. Oxford: Clarendon Press.

Tillmann, H. (2006). Experimental and instrumental phonetics: History. In K. Brown (Ed.) *Encyclopedia of Language & Linguistics (Second Edition)*, (pp. 374 – 389). Oxford: Elsevier.

Tilsen, S., & Goldstein, L. (2012). Articulatory gestures are individually selected in production. *J. Phon*, *40*(6), 764–779.

Titze, I. R., & Story, B. H. (1997). Acoustic interactions of the voice source with the lower vocal tract. *JASA*, *101*(4), 2234–2243.

Toda, M., & Honda, K. (2003). An MRI-based cross-linguistic study of sibilant fricatives. In *Proc. 6th Intl. Seminar on Speech Production*, (pp. 290–295).

Tourville, J. A., & Guenther, F. H. (2011). The diva model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981.

Trubetzkoy, N. S. (1969). *Principles of phonology*. Berkeley: University of California Press.

Turpin, M., Ross, A., & Staff, I. P. (2012). *Kaytetye to English Dictionary*. Alice Springs: IAD Press.

Viswanathan, N., Magnuson, J. S., & Fowler, C. A. (2014). Information for coarticulation: Static signal properties or formant dynamics? *J. Experimental Psychology: Human Perception and Performance*, *40*(3), 1228.

Vogt, H. (1963). *Dictionnaire de la langue oubykh: avec introduction phonologique, index français-oubykh, textes oubykhs*, vol. 52. Univ.-Forl.

Volney, c., Constantin-Fran (1795). *Simplification des langues orientales*. Paris: L'imprimerie de la République.

Wagner, K. O. C., & Baker-Smemoe, W. (2013). An investigation of the production of ejectives by native (L1) and second (L2) language speakers of Q'eqchi' Mayan. *J. Phon*, *41*(6), 453–467.

Westbury, J. R., Hashi, M., & Lindstrom, M. J. (1998). Differences among speakers in lingual articulation for American English /r/. *Speech Communication*, *26*(3), 203–226.

Wilkins, J. (1668). *An essay towards a real character, and a philosophical language*. London: Sa. Gellibrand.

Zawadzki, P. A., & Kuehn, D. P. (1980). A cinereadiographic study of static and dynamic aspects of American English /r/. *Phonetica*, *37*, 253–266.

Zhou, X., Espy-Wilson, C. Y., Boyce, S., Tiede, M., Holland, C., & Choe, A. (2008). A magnetic resonance imaging-based articulatory and acoustic study of 'retroflex' and 'bunched' American English /r/. *JASA*, *123*(6), 4466–4481.

Zhu, Y., Kim, Y.-C., Proctor, M. I., Narayanan, S., & Nayak, K. S. (2013). Dynamic 3D Visualization of Vocal Tract Shaping during Speech. *IEEE Trans. Medical Imaging*, *32*(5), 838–848.

Zue, V. W., & Laferriere, M. (1979). Acoustic study of medial /t, d/ in American English. *JASA*, *66*(4), 1039–1050.