EKPA 3

A FEATURE-BASED PHONOLOGICAL FRAMEWORK AND *FONETIK* ALPHABET

PHONETIC KEY

These symbols describe observable qualities of phones as opposed to the supposed feature specifications. The correspondence between the feature specifications of a segment and the physical realization is many-to-many. Therefore it is useful to have one way to describe superficial characteristics and another for featural analyses. The notations for representing the feature specifications of a segment are given separately. This section explains the use of the superficial phonetic symbols.

Note: These symbols are only meant to be the phonetic key of this document and are *not intended to be integral to EKPA*.

In text, a single segment is bounded by ‹ ›, and a sequence of segments or a sub-segmental component is bounded by « ». Where clarification is necessary, segments in a sequence as well as the surrounding symbols « » are spaced, e.g., « æ p l » or «æpl» for *apple*.

Complex phones can be represented by combining existing symbols of similar qualities, e.g., ‹dl› for prestopped lateral liquid.

Where necessary, language-specific adjustments are made. For example, for some languages the symbol ‹p› may be used for an aspirated voiceless plosive and ‹b› for an unaspirated voiceless plosive.

Phones for which appropriate symbols cannot be found are described with a hyphenated one-word description suffixed to a symbol of a reasonably similar sound, e.g., ‹l-velar› or ‹w̥-lateral› for a velar lateral liquid. The hyphenated label may be substituted with a subscript label: ‹l_{velar}›, ‹$\text{w̥}_{\text{lateral}}$›. The same is done to distinguish multiple phones that fall into the definition of the same basic symbol, e.g., ‹t_{apical}› and ‹$\text{t}_{\text{laminal}}$›.

<u>Optional affixes</u>:

Labialized or labiovelarized	_w	Dorsal high-fronting	_y
Rhotacized	_r	Simultaneous _w and _y	_ẅ
Velarized	_w̥	Pharyngealized	_ħ
Aspirated	_h	Breathy	_ɦ
Glottalized	_ʔ	Long or tense	_:
Voiceless	h_	Preglottalized	ʔ_

<u>Simple consonants</u>:

		Labial		Coronal: Anterior		Posterior		Dorsal: Central		Back	
Stop:	Oral	b	p	d	t	ḍ	ṭ	g	k	ġ	q
	Nasal	m		n		ñ		ŋ		ŋ̇	
Fricative:	Mellow	β	φ	ð	θ			ɣ	x		
	Strident	v	f	z	s	ž	š			ɣ̇	x̣
	Lateral			lʒ	ł						
Affricate:	Mellow	bβ	pφ	dð	tθ			gɣ	kx		
	Strident	bv	pf	dz	c	ǰ	č			ġɣ̇	qx̣
	Lateral			dlʒ	tł						
Liquid:	Trill			r̃		ṛ̃				ʀ	
	Flap			ř		ṛ̌					
	Lateral			l		l̩					

The symbol on the right is voiceless, and on the left is voiced.

<u>Vowels</u>:

	Front		Central		Back	
High	i	ü	ɨ	ʉ	ï	u
Mid	e	ö	ə	ɵ	ë	o
Low	æ	ɔ̈	a	ə	æ̈	ɔ

The symbol on the right is rounded, and the on the left is unrounded.

<u>Vowel-like glides or extra-short vowels</u>:

	Front		Back or rounded
High	ĭ	ə̆	ŭ
Mid	ĕ		ŏ
Low	æ̆	ă	ɔ̆

‹ə̆› and ‹ă› are central with no strict height specifications. ‹ə̆› is higher than ‹ă›.

<u>Other glides</u>:

Labial or labiovelar	w	High-front dorsal	y
Rhotic	r	Simultaneous «w» and «y»	ẅ
Velar	w̥	Pharyngeal	ħ
Voiceless	h	Breathy	ɦ
Glottal stop	ʔ		

‹ŭ› and ‹w› as well as ‹ĭ› and ‹y› are used differently for convenience where two kinds of y-like or w-like glides are present.

INTRODUCTION TO EKPA

OVERVIEW

THE SEGMENTAL COMPONENT

The segmental component of the EKPA framework is fully feature-based. Each segment is represented as a complex of twelve binary features. Every feature is strictly binary. *n* features contain exactly *n* bits of information. Any use of features in which a single feature conveys ternary or greater information is prohibited. The only valid values for a feature are positive (+) and negative (–), and features are never privative. “Unspecified” is never a value that can be referenced by the grammar. In EKPA, if a segment is underspecified or unspecified for a feature, it can only mean that “it doesn’t matter” because the feature value can be positive or negative and the grammar generates the correct surface forms in either case or that the value of the feature remains implicit because it is irrelevant to the topic.

The feature system is loosely based on the traditional Jacobsonian system but with radical modifications. The chart on the right side of this page is a summary of the feature system. The phonetic key is provided separately but is similar to the Americanist phonetic notation.

The correspondence between the feature specifications and the phonetic realizations is loosely explained in this chapter. For a more precise description of the correspondence, in a separate chapter, formal illustrations are attempted.

“L” stands for “Loud” and it is the only feature that separates oral sonorants from oral obstruents. For each subclass, “B” (“Bright”) further divides segments into two groups. The feature ±Bright was called “Strident” in EKPA 2, but as the use of the feature is different from the traditional usage, to avoid confusion, the new name “Bright” is used in EKPA 3. For obstruents, +Bright is similar to conventional strident. For sonorants, +Bright is for liquids and flaps and –Bright for vowels and glides.

The features ±Constrictive and ±Dynamic are responsible for further classifications. +Constrictive is for central narrowing of the oral cavity, and +Dynamic is for extra-short, transitional, or complex articulations.

The features ±Peripheral and ±Overt are used for basic place specifications of consonants and vowels. These features are very similar to the traditional features “grave” and “compact.”

±Nasal separates nasals and nasalized segments from non-nasals. ±Tense distinguishes quantitative properties of segments.

±Expanded and ±Raised are responsible for finer tonality properties. These features are similar to traditional “flat” and “sharp.”

The segmental model of EKPA

L	–	–	–	–		+	+	+	+
P	–	–	+	+		–	–	+	+
O	–	+	–	+		–	+	–	+
B D C									
Obstruents:					Vocoids:				
– – –	d	ḍ	b	g		e	æ	ë	æ̈
– – +	ð		β	ɣ		i	er	ï	ër
– + –						ĕ	æ̆		
– + +	dð		bβ	gɣ			r		ẉ
					Sonorant consonants:				
+ – –	lʒ					l	ḷ		
+ – +	z	ž	v	ɣ̇		r̃	ṛ̃		ʀ
+ + –	dlʒ					ř	ṛ̌		
+ + +	dz	ǰ	bv	ġɣ̇					

	W	+	+	–	–
Laryngeal settings:	G	–	+	–	+
	Vowel	he	eɦ	e	eʔ
	Glide	h	ɦ	ĕ	ʔ
	Nasal	hn	nɦ	n	nʔ
	Plosive	th	dɦ	d	t

±Nasal and ±Tense:
+N is for nasal segments. ±T is responsible for quantity specifications, i.e., +T is longer or “heavier” than –T.

±Expanded and ±Raised:
For consonants, +E corresponds to o-like modifications and +R i-like modifications. For vowels, +E corresponds to a lower F2 and +R a lower F1.

In a textual or verbal form of communication, the following opposing terms may be used in place of the prefixed notation (in parentheses are the terms used as feature labels when prefixed and capitalized): modal (glottal), calm (whispered), lax (tense), oral (nasal), quiet (loud), dim (bright), static (dynamic), loose (constrictive), medial (peripheral), covert (overt), shrunk (expanded), lowered (raised).

The features are grouped into three quartets: the manner features: ±L, ±B, ±D, ±C; the place features: ±P, ±O, ±E, ±R, and the exterior features: ±W, ±G, ±T, ±N. Within each group, the features are preferably ordered in this order. This is a recommended convention for readability, and ordering or grouping them differently does not change the meaning.

THE PROSODIC COMPONENT

EKPA assumes the following prosodic hierarchy. The model is loosely based on Nespor & Vogel (1986) and Kodama (2008) but is less layered, and with markedness distinction for each unit. It is also notable that mora is incorporated into the hierarchy. Markedness is for representing surface contrasts associated with a given prosodic unit. A mora can be contrastively high or low, for example. The hierarchy is intended to complement the surface representation of EKPA as the segmental feature system only accounts for segmental properties and ignores pitch and stress. The present prosodic model is compatible with the N & V style phonology, where segmental rules are stated in terms of strict-layered limits of segmental sequences, but it should be noted that the primary purpose of adding the prosodic component is to account for surface contrasts, not simplifying segmental rules.

An utterance (U) is composed of a non-zero whole number of prosodic phrases (φ), each of which is composed of a non-zero whole number of prosodic words (ω), each of which is composed of a non-zero whole number of syllables (σ), each of which is composed of a non-zero whole number of moras (μ). Each prosodic unit is either unmarked or marked. A marked unit is indicated by placing an asterisk (*) on the right. Prosodic units, boundaries, and markedness are associated with phonetic properties, such as pitch, but the actual phonetic content depends on the language and the context. At this time, no strict limits are imposed on what phonetic qualities are associated to these prosodic specifications. Any contrastive phonetic properties can be associated with markedness, but phonetic contents that can be expressed by the segmental features should be encoded in segments. (This should not discourage an analysis where segmental rules are stated in terms of prosodic units, e.g., vowel harmony within the limits of a prosodic word or initial devoicing of a certain prosodic unit such as φ or ω.) The division of roles between the two components should help clarify what changes are necessary for the future versions of EKPA.

The prosodic model of EKPA

<u>The hierarchy of prosodic units</u>:

U > φ > ω > σ > μ

<u>A sample analysis of Tokyo Japanese</u>:

	“imakara	sotchi	ittemo	ii?”
	« i m a k a ṛ̌ a	s ɵ č: i	i t: e m o	i: »
U:	3			
φ:	1	2		1
ω:	2	3	3	1
σ:	1 1 1 1	1 2	1 2 1	2*
μ:	0* 0 0 0	0* 0* 0*	0* 0* 0* 0	0* 0

	“Ashita	omatsuri	iku	watashi.”
	« a š t a	o m a c ɨ ṛ̌ i	i k u	w a t a š i »
U:	2			
φ:	1	3		
ω:	3	4	2	3*
σ:	1 2	1 1 1 1	1 1	1 1 1
μ:	0* 0* 0*	0* 0* 0* 0*	0* 0*	0* 0* 0*

The units above the level of syllables are aligned, but syllables, moras, and segments are not aligned. The number indicates how many sub-units it contains. For example, the number 2 in the syllable tier means that the syllable contains two moras. Moras are shown as zeros because they are terminals in the hierarchy. Here it is assumed that a marked mora is associated with the non-falling tone and an unmarked mora with the falling tone, and a marked prosodic word is associated with the boundary fall (a pitch drop over the boundary of two non-falling moras. See Iida 2024b). As a tentative hypothesis, I suppose that the predicative intonation (Iida 2024b) is associated with syllable markedness.

Segments are not part of the prosodic hierarchy because incorporating them into the hierarchy would break the strict layer hypothesis. In Tokyo Japanese, a short onset (–T) is non-moraic while a long vowel (+T) is bimoraic, which means that segments cannot be lower nor higher than moras in the hierarchy.

At the level of syllable and mora, segmental features provide information for associating (linking) the segmental and prosodic components. The grammar knows the moraic weight and the syllabicity of a segment by referencing the segmental features. The prosodic hierarchy does not reduce segmental specifications of syllabicity or moraicity.

Word- and phrase-level prosodic boundaries can be stored in the lexicon. If a lexical item is specified with a prosodic boundary, the boundary is a word- or phrase-level boundary.

Under this model, each mora bears 1 bit of prosodic information. If a syllable is bimoraic, the syllable contains 3 bits. If a word has two moras and one syllable, it contains 4 bits. This is probably enough for Tokyo Japanese but no other languages have received a serious assessment. It is of interest whether in any language a one-mora, one-syllable word distinguishes more than eight tones, for example.

It should also be tested whether five levels are too few or too many. N & V assumes seven distinct types of prosodic units, sans mora, but their framework is focused on describing phonological rules while EKPA is more about representing surface contrasts.

PREFIX NOTATION

Features are represented by either the full label (e.g., ±Tense) or the capital letter (e.g., ±T). When multiple features are mentioned for the same segment, the features can be bracketed (e.g., [+T –W]), in which case the brackets indicate the limit of the segment—if multiple features are shown in the same brackets, they belong to the same segment. The prefix (e.g., "+") is not separated from the feature label, but features are separated by space unless they form a *chunk*.

A *chunk* is formed when one or more features of the same segment are concatenated under one prefix. [+T +W] can be alternatively represented as +TW. These two notations are equivalent. Here, the prefix "+" indicates the value of multiple features. This prefix reads, "All of the following features are positively specified." By the same token, the prefix "–" means, "all of the following features are negatively specified."

Prefixes can take up to two chunks to form a *pair*. The head prefix of a pair is directly attached to whichever component is written on the left side, graphically concatenating two prefixes (e.g., ++TW –PO). The binary value of a feature or a feature complex is interpreted in the usual way of binary logic. ++TW –PO is equivalent to [+TW –PO] *and* to ––TW +PO. (Negative times negative is positive, and negative times positive is negative.) A component of a pair can be a pair, but the use of trailing prefixes should be kept minimal.

The following prefixes are used:

	Meaning:	Unicode name:
+	All of the following are positive	Plus sign
–	All of the following are negative	En dash
=	All of the following are equal	Equal sign
≠	Not all of the following are equal	Not equal to
?	At least one of them is positive	Question mark
¿	At least one of them is negative	Inverted question mark
±	No specific values	Plus-minus sign
$	(See text)	Dollar sign

The dollar sign prefix is used with uppercase and lowercase letters representing features, denoting that the uppercase features are positive and the lowercase features are negative. ++TW –PO is written with this prefix as $TWpo. This notation proved to be especially convenient (mostly thanks to @pyapi379, who heavily used this notation since the early days of EKPA 2) and is used several times in this document. The plus-minus sign prefix is used to mention features without mentioning the values. The other EKPA 2 prefixes (!, ¡) are discontinued due to lack of proven use.

SEGMENTAL SYMBOLS

Segmental shorthands (symbols that stand for a segment and its feature specifications) will continue to be used, but due to underuse, the system of symbols is drastically reduced.

Consonants ($ldertg):							
			P	+	–	–	+
			O	–	–	+	+
B	C	W	N				
–	–	+	–	P	T	Ṭ	K
–	–	–	–	B	D	Ḍ	G
–	–	–	+	M	N	Ṇ	Ŋ
–	+	+	–		Þ	Þ̣	X
–	+	–	–		Đ	Đ̣	Ɣ
+	+	+	–	F	S	Ṣ	
+	+	–	–	V	Z	Ẓ	

Vowels ($Lbdrwgnt):					
	P	–	–	+	+
	E	–	+	–	+
C	O				
+	–	I	Ị	Ụ	U
–	–	E	Ẹ	Ọ	O
–	+	Æ	Æ̣	Ɔ̣	Ɔ

High glides ($LbDCPOwgnt):				
R	+	+	–	–
E	–	+	–	+
	Y	Ỵ	Ẉ	W

Loose glides ($LbDcpoernt):				
W	+	+	–	–
G	–	+	–	+
	H	Ḥ	ʔ̣	ʔ

Unspecified segment: ß (see text)

ß is a special shorthand that is unspecified for all features. This is used to represent a segment without mentioning features other than that is shown by featural diacritics (see below). The other shorthands are defined for all features. Unknown or irrelevant feature specifications can be ignored, but context should be sufficiently provided as to what features are at issue.

Optionally, lower case letters may be used. A lower case letter indicates that it is relatively superficial than the corresponding upper case letter.

A featural diacritic overwrites a feature specification of the shorthand it is attached to. Any number of diacritics can be added to any of the segmental shorthands although adding more than two diacritics to a segment may be better avoided as it will make the notation cumbersome.

+Loud	ß̇	+Bright	ß̌	+Dynamic	ß̆
+Constrictive	ß̋	+Peripheral	ß̀	+Overt	ß̈
+Expanded	ß̂	+Raised	ß́	+Whispered	ß̊
+Glottal	ß̉	+Nasal	ß̃	+Tense	ß̄

THE SEGMENTAL FEATURES

±LOUD

The feature ±Loud divides segments into two major classes. It has to do with the degree of obstruction in the oral cavity: –L for a significant constriction made for the primary articulation in the oral cavity and +L otherwise. Except for nasals, roughly, +L segments are sonorants and –L segments are obstruents.

±BRIGHT

The ±Loud and ±Bright features form some familiar natural classes. +LB are liquids and flaps, and +–L +B are strident obstruents such as ‹s, c, ɬ, tɬ›. A +Bright segment is "noisier" than the –Bright counterpart. For obstruents, +Bright is the conventional strident plus lateral obstruents. For sonorants, laterals and flaps are considered +Bright, and vowels and glides are –Bright. In the following, the segment on the right is +Bright and on the left is –Bright.

		B	–	+
L	D	C		
–	–	–	d	ɮ
–	–	+	ð	z
–	+	–		dɮ
–	+	+	dð	dz
+	–	–	e	l
+	–	+	i	r̃
+	+	–	ĕ	ř
+	+	+	ĭ	$\tilde{\text{r}}_{\text{fricative}}$

Brightness has implications on fine distinctions of articulatory places. The bright anterior coronal fricative ‹z› cannot be interdental, while the dim counterpart ‹ð› is often interdental. Similarly, the bright labial fricative ‹v› is usually labiodental while the dim counterpart ‹β› is usually bilabial, and the bright dorsal fricative ‹ẏ› is usually uvular while the dim dorsal fricative ‹ɣ› is typically velar. The exact identity of the dim counterpart of ‹ž› is unclear although by definition it is a less noisy posterior coronal fricative. SPE (p. 329) assumes the German *ichi-laut*, commonly written with IPA [ç], is the non-strident counterpart of ‹š›, but some (e.g., Hall 2009, p. 220) think it is dorsal. Here, the dim counterpart of ‹ž› is written as ‹ð$_{\text{palatal}}$›.

	Labial	Dental	Palatal	Dorsal
B				
–	β	ð	ð$_{\text{palatal}}$	ɣ
+	v	z	ž	ẏ

Labiodental plosives and interdental plosives are physically possible, but as far as I am aware, they are not known to be contrastive, i.e., there is no evidence that ‹p$_{\text{bilabial}}$› and ‹p$_{\text{labiodental}}$› are *categorically* different.

In some languages including Ubykh (Beguš 2020, p. 9) labialization (±E) is contrastive for velars *and* uvulars. To accommodate the four-way contrast, the bright (then “strident”) dorsal is ambiguously lateral or uvular.

	Labial	Dental	Palatal	Dorsal	
B			E	–	+
–	b	d	ḍ	g	gw ġ
+		lʒ	lʒ$_{\text{palatal}}$	lʒ$_{\text{velar}}$ ġ	lʒw$_{\text{velar}}$ ġw ġ

±DYNAMIC

±Dynamic is related to shortening, transitional articulatory gestures, and complex articulations.

+Dynamic represents affricates (+) as opposed to fricatives (–), glides and extra-short vowels (+) as opposed to plain vowels (–), fricative trills (+) as opposed to regular trills (–), flaps (+) as opposed to lateral approximants (–), doubly articulated plosives, including clicks, (+) as opposed to plosives with a non-occlusive secondary constriction (–), obstruent nasals (+) as opposed to plain nasals (–), and implosives and electives (+) as opposed to modally voiced or tenuis plosives (–).

		D	–	+
L	B	C		
–	–	–	d	
–	–	+	ð	dð
–	+	–	lʒ	dlʒ
–	+	+	z	dz
+	–	–	e	ĕ
+	–	+	i	ĭ
+	+	–	l	ř
+	+	+	r̃	r̃$_{\text{fricative}}$

The feature ±Dynamic is related to diverse types of oppositions, but in a minimal pair, the +D segment is more consonant-like than the –D segment. Affricates are more consonant-like than fricatives, glides are more consonant-like than vowels, extra-short vowels are more consonant-like than regular vowels, and fricative trills are more consonant-like than regular trills. It may be less clear whether flaps are more consonant-like than laterals, but ‹ř› is probably less likely to be a syllable nucleus than ‹l›.

±CONSTRICTIVE

+C segments are produced with a central constriction in the oral cavity. If the segment lacks a significant central constriction, it is –C. Flaps are –C because they do not have a maintained period of central constriction before or after the occlusion. Central fricatives and central affricates are +C as opposed to non-affricate stops and laterals (–), high or rhotic vocoids are +C as opposed to nonhigh nonrhotic vocoids (–), and trills are +C as opposed to lateral approximants (–).

In EKPA 2, both rhotic taps and rhotic approximants were represented as $LsCD (then ±S equivalents to now ±B). In EKPA 3, taps and flaps are represented as $LBcD while rhotic approximants remain $LbCD. As a result, the distinction between lateral flaps and central flaps is lost. Whether this is a problem or an improvement will hopefully be seen.

		C	−	+
L	B	D		
−	−	−	d	ð
−	−	+		dð
−	+	−	lʒ	z
−	+	+	dlʒ	dz
+	−	−	e	i
+	−	+	ĕ	ĭ
+	+	−	l	r̃
+	+	+	ř	r̃fricative

The loose glides ($LbDc) include pharyngeal consonants since for these consonants the oral cavity is free of obstruction ($Lbc). The oppositions between different types of pharyngeals are expressed in the same way as for vowels. The pharyngeal obstruction may be greater when the tongue is low (+O), posterior (+P), or RTR (−R). +E can also mean a posterior displacement of the tongue. Agul is known for having multiple pharyngeals. This does not mean that the articulatory correlates of these features must be identical in both contexts. It is possible that the lowest and most posterior ($POEr) loose glide ($LbDc) is a pharyngeal stop ‹ħstop› and the same setting ($POEr) for a loose vowel ($Lbdc) is a regular low back vowel ‹ɔ› in the same language.

In Japanese, the dental plosive ‹t› at the end of a verb stem becomes an affricate ‹č› or ‹c› (+C) when followed by a high vowel ‹i› or ‹ɨ› (+C). In English, when a dental plosive ‹t, d› is followed by the palatal glide ‹y› (+DC) over a word boundary, the plosive is optionally realized as an affricate ‹č, ǰ› (+DC), e.g., *don't you* is often pronounced *don*‹č› *you*. Dentals stops ‹t, d› is optionally affricated when followed by ‹r› (+DC) in words such as *dragon* «drægə̆n»~«ǰrægə̆n» or *tree* «tri:»~«čri:».

±PERIPHERAL

±Peripheral is associated with labials and dorsals (+) as opposed to dentals and palatals (–) and with back vocoids (+) as opposed to front vocoids (–).

±OVERT

±Overt is associated with posterior consonants (palatals and dorsals) (+) as opposed to anterior consonants (labials and dentals) (–) and with low and rhotic vocoids (+) as opposed to mid and high vocoids (–).

P	+	−	−	+
O	−	−	+	+
	Labial	Dental	Palatal	Dorsal
	v	z	ž	ɣ̇

±EXPANDED

+E indicates an o-like co-articulation. This is typically done by labialization or tongue body retraction. (Both lower the second formant.)

±RAISED

+R may indicate an i-like co-articulation, tongue root advancement (ATR), or laminal articulation. −R may indicate a low vowel-like co-articulation (pharyngealization/RTR) or apical articulation. A +R vowel is slightly higher than the −R counterpart. Languages make diverse choices for implementing this feature. When −R leads to a salient phonetic consequence, +R may be "plain" and vice versa.

					P	−	−	+	+
					E	−	+	−	+
						Front			Back
C	O	R							
Non-rhotic									
+	−	+	High	Higher		i			u
+	−	−		Lower					
−	−	+	Mid	Higher		e			o
−	−	−		Lower					
−	+	+	Low	Higher		æ			ɔ
−	+	−		Lower					
Rhotic									
+	+	+		Higher		ir			ur
+	+	−		Lower		ær			ɔr

The vowel space is broadly divided in eight by ±Peripheral (back (+) or front (–)), ±Overt (low (+) or nonlow (–)), and ±Constrictive (high or rhotic (+) or otherwise (–)). The backness axis is subdivided by ±Expanded, which as a result yields four levels of vowel backness.

±TENSE

±Tense is responsible for moraicity and duration. A +T segment has an equal to or greater moraic weight than the –T counterpart, *ceteris paribus*. If two segments are minimally contrastive in duration, the longer one is +T. In Japanese, short vowels are monomoraic (–T) and long vowels are bimoraic (+T).

The syllable structure of Tokyo Japanese (my analysis):

T	Consonant	Consonant	Onglide	Nucleus	Offglide
–	1 mora	0 moras	0 moras	1 mora	1 mora
+	2 moras	1 mora	(unattested)	2 moras	(unattested)

The rhyme (the nucleus and offglide; the structural standing of the onglide is unclear) is restricted to be bimoraic or shorter. As five mora syllables are unattested, the syllable may be restricted to four moras or shorter.

An example of a four mora syllable:
«šp:aĭ» "failure"

	š	p:	a	ĭ
T	–	+	–	–
# moras	1	1	1	1

Tense segments are often articulated with a greater effort than the corresponding lax segments. Korean tense consonants involve greater muscular activities than lax counterparts. Korean is an example of a language with a ±W contrast of +T consonants.

Korean plosives (based on Kim 1965, p. 356.):

The lax segments ‹d› and ‹t› are generally considered to be positional allophones.

	T	–	+
W			
–		d	d:
+		t	t:

The lax ‹t› is aspirated but more weakly than ‹t:›. ‹d:› is voiceless unaspirated. The tense series ‹d:, t:› has a greater duration of closure than the lax ‹t›.

In the literature, the term "tense" is often used in contrast to "lax" to refer to an elevated tongue position or the corresponding quality of a vowel, especially for the Germanic languages. EKPA's tense-lax opposition primarily concerns the quantity, not the quality, of the intrinsic content of a segment, e.g., a longer duration of articulation, an additional mora, greater muscular activities, etc. It is a "degree" feature rather than a "kind" feature, as in Kim (1965). The opposition between the higher and the lower ‹i› is represented by ±R, although in some languages ±R seems concomitant with ±T. In General American, the raised high vowels, here denoted as ‹i:, u:›, are generally taken to be longer or heavier (bimoraic) than the lowered counterparts ‹i, u› (e.g., Kawagoe 2014, pp. 51-54).

A partial vowel inventory of General American:

				P	–	+		
C	O	R	T					
+	–	+	+		i:	u:	"he"	"who"
+	–	–	–		i	u	"hit"	"hook"

±NASAL

Nasal or nasalized segments are +Nasal and non-nasal segments are –Nasal.

±WHISPERED

+Whispered indicates that the segment is pronounced with a greater airflow rate across the glottis. For voiceless plosives, +Whispered is for aspirated segments, and –Whispered is for unaspirated segments.

±GLOTTAL

For +Whispered sounds, +Glottal means "more voicing." For –Whispered sounds, +Glottal means "less voicing." In either context, modal voicing is the most "voiced" and non-voicing is the least "voiced" category, and non-modal voicing, e.g., breathy and creaky, is in-between.

In the later versions of EKPA 2, the features ±WG were interpreted to work together like a single four-valued feature. In EKPA 3, the definition of each feature is revised with more accessible phonetic cues. Voicing can be heard by the ears and detected on the computer. The air stream rate can be felt by placing a hand in front of the mouth. The previous framework assumed that the tenuis plosives are produced with a narrow glottis, but in reality it is not necessarily narrow. It may be wide or narrow for Hindi (Dixit 1989). It is still important to think of contexts when interpreting features. A +W segment has a faster air stream than a –W segment *ceteris paribus*.

THE THEORETICAL FRAMEWORK

PHONOLOGY AND PHONETICS

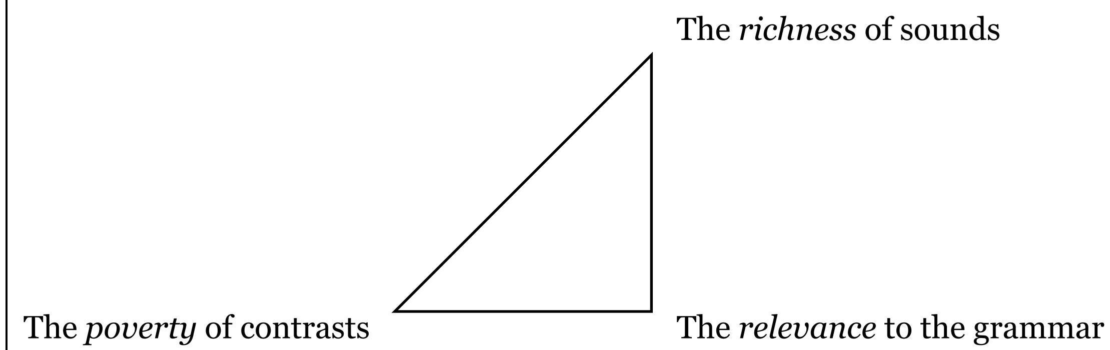
THE TRILEMMA OF PHONOLOGY

It is generally uncontroversial to assume that the speech signal of an utterance can be approximated as a sequence of discrete linguistic units, conventionally called segments. In phonology, these segments are usually characterized by classificatory devices such as distinctive features. Traditionally, distinctive features are defined in terms of acoustic properties or articulatory movements associated with speech sounds. Using measurable qualities as categorical labels is an effective strategy to fill gaps between the grammar and the sounds, for as long as the categories are discrete the grammar can operate on them essentially in the same way as it does on the other grammatical categories such as syntactic categories.

The question is as to how such phonetic categories can be chosen from speech signals so complex and non-discrete by nature. I broadly follow the SPE approach (Chomsky & Halle 1968, pp. 295-298). I consider it necessary to choose the properties that the human linguistic capacity is inherently sensitive to. Coding every detectable phonetic detail into the segments is not a sensible method of description. Many kinds of physical quantities are irrelevant to the grammar, and it is also possible that some grammatically important categories are never realized as part of physical sounds of the given language. The other end of the extremes—defining arbitrary abstract features for the grammar, as illustrated as a strawman in SPE, is not acceptable either, since that would be simply neglecting phonetics. An acceptable solution would be to find distinctive phonetic qualities that are inevitably relevant to the grammar.

The obtained list of features must account for all minimal contrasts present in any spoken language that can be naturally acquired as the first language. A competing requirement for the feature list is that it has to be economical. Specifically, a distinctive feature theory is expected to account for the *poverty of contrasts*—the contrastive segmental inventory of a language seems to never exceed 256 or 2^8. That is, as much as classifying and approximating the speech signals and associating utterances with grammar are competing needs, explaining the diversity of segments and predicting the lack of rich contrasts are two ends that must be reconciled.

The trilemma of phonology



Richness vs. poverty: The number of contrastive segments of a particular language has never been attested to exceed 256 or 2^8, which is far smaller than the total number of distinctive segments of the human language by any reasonable estimate.
Richness vs. relevance: Most of the physical properties of the phones are irrelevant to alternations and wellformedness.
Poverty vs. relevance: Phonological alternations are often opaque. Specifications for the features that cannot be observed from the speech signals are necessary to account for the distributions of allomorphs.

THE PHONOLOGY-PHONETICS INTERFACE

EKPA addresses the poverty of contrasts by reducing the size of the segmental feature space to 2^{12}.

However, abstract features detached from physics cannot be said to be a phonetic approximation unless some function that maps a featural representation onto a more specific phonetic representation is provided. In EKPA 2, the correspondence between the features and the phonetic categories was only informally explained. To provide a more concrete and formal explanation, EKPA 3 uses low-level phonetic features (or “the phonetic features” for short). They are not meant to be a complete list of what the phonetic representation should be composed of. Currently, the phonetic features are mere expository devices that are only meant to help illustrate the phonology-phonetics correspondences.

Phonology:

- The manner features
  - ±Loud ±Bright ±Dynamic ±Constrictive
- The place features
  - ±Peripheral ±Overt ±Raised ±Expanded
- The exterior features
  - ±Whispered ±Glottal ±Nasal ±Tense

Phonetics:

Major class features	
	loud, bright, constrictive
Duration and syllabicity features	
	quick
Manner of articulation features	
	occlusive, fricative
Active articulator features	
Jaw:	open
Blade of the tongue:	laminal
Dorsal space:	high, raised, front, retracted
Place of articulation features	
	coronal, posterior
Airstream features	
Nasal cavity:	nasal
Larynx:	voiced, whispered

The phonological representations are described with the phonological features only, in the same sense as they were in EKPA 2. In the output of the phonological component of the grammar, the phonological features define the surface representation. This surface representation is converted to a series of segments defined in turn with the phonetic features. The correspondence between the EKPA features and the physical sounds is explained by stating how this conversion is made.

ABSTRACTION

If two types of articulations produce similar sounds, any one of the two can substitute the other, i.e., when we know two articulation types have similar auditory consequences and the contrast between them does not seem necessary, we can consider one as the other.

The *r* sound in American English is pronounced in multiple ways. Some speakers pronounce the sound as a retroflex, and some speakers make it with a raised ("bunched") tongue body. The auditory effects of these articulations are similar and usually considered interchangeable (Pycha p.c.). Due to the lack of a comfortable way of representing the "bunched" version of this sound in the framework of EKPA, we assume that the sound is a retroflex for the purpose of formal representation. For the consonant sound, it is a regular rhotic retroflex approximant $LbDCpOEr, and for the vowel, it is a regular rhotic vowel $LbdCO.

Implosive stops and creaky voiced stops are not known to be contrastive (Ladefoged & Jonson 2014, p. 184). In EKPA 3, a creaky voiced stop is usually assumed to be a variation of an implosive stop, as in EKPA 2. The reverse treatment is permissible. If in some language $lbDcwgnE needs to be interpreted as a modal-voiced labial-velar stop, a contrastive labialized implosive velar stop can be represented as a labialized static creaky voiced stop: $lbdcwGnE.

PHONETIC FEATURES

MAJOR CLASS FEATURES

The major class features, {loud, constrictive, bright}, characterize the primary articulation of the segment. It is therefore important to distinguish between the primary and non-primary articulations. A segment can have multiple articulatory targets, and one of them is considered primary. The others are secondary. Although one segment may have multiple secondary articulations, oftentimes it is only necessary to discuss one secondary articulation. A secondary articulation can greatly alter the auditory impression of the segment. It is possible that the primary articulation is sonorant-like while the secondary articulation is obstruent-like.

These features directly reflect the phonological features of the same names: ±LBC (+ to 1, – to 0).

The feature {loud} distinguishes segments with a significant obstruction for the primary articulation (0) from those without (1). The degree of obstruction sufficient to be {0 loud} is just the degree at which spontaneous voicing is impossible provided that the other factors remain neutral.

Spontaneous voicing is an assumed mode of voicing proposed in SPE as part of the definition of sonorant. They supposed that the type of voicing characteristic to sonorants was only possible if the supraglottal pressure was nearly equal to the ambient pressure. While the post-SPE literature (e.g., Ladefoged 1971, pp. 109-110) points to weak physiological bases of this notion, it is operationally not difficult to define. When one pronounces ‹ɨ› with a modal voicing and while maintaining the pulmonic airstream and the supraglottal muscular tension gradually strengthens the constriction made at the soft palate, at some point the voicing suddenly and automatically stops. What follows is ‹x› unless the speaker makes adjustments. When the constriction becomes sufficiently intense, voicing stops because, presumably, due to the pulmonic air the supraglottal pressure builds up behind the constriction to the point where the difference between the supraglottal and the subglottal pressures is insufficient for spontaneous voicing. Just this level of obstruction of the primary articulation is sufficient for a segment to be {0 loud}.

The obstruction must be maintained long enough to make this effect for the segment to be {0 loud}. Even if the obstruction is very strong, if the obstruction is maintained only momentarily, the segment is {1 loud}. This is why flaps form a complete closure and are still {1 loud}.

Unlike the conventional notion of sonorant, as long as a sufficient degree of obstruction is made for the primary articulation both in the shape and in the duration, the segment is {0 loud} even if the nasal cavity is open.

The feature {constrictive} distinguishes a primary articulation made with an open but centrally narrowed oral tract. Simple stops and simple laterals are {0 constrictive} and central fricatives, central affricates, and high vocoids are {1 constrictive}. ‹e› and ‹a› are {0 constrictive}.

The question is then how narrow is narrow enough for a segment to be {1 constrictive}. This feature intends to include segments with a centrally concentrated airflow, but the requirement is much lower than for fricatives so as to include high vocoids such as ‹i, u, ĭ, ŭ› because fricatives and vocoids are already separated by the feature {loud}. At the same time, this feature is responsible for excluding mid and low vowels and non-approximant glides. Hence, it is reasonable to see the acoustic effect of the airflow of high vowels to draw the line between the high and mid vowels. If one pronounces ‹hæ› or ‹ha›, and while maintaining the airflow gradually raising the tongue body toward the position of the corresponding high vowel at a point the perceptual noise level sharply increases. An approximated degree of central constriction just sufficient for {1 constrictive} is obtained this way. If the central constriction for the primary articulation is weaker or is a complete closure, it is {0 constrictive}.

Trills are {1 constrictive} because when voiceless they produce a noise of a similar quality, but flaps are {0 constrictive} due to the absence of this quality. (The articulatory bases are discussed later.)

> Some linguists take virtually any kind of voiceless noise as a sufficient reason to regard a given non-syllabic continuant as a fricative. Akamatsu (1992, p. 39) claims, citing O'Connor (1973, p. 61) making essentially the same point, that voiceless approximants are by definition either nonexistent or silent. Presumably for the same reason the voiceless glide ‹h› is traditionally called the "glottal fricative" despite the fact that a strong constriction is made nowhere for this phone. This means that in the conventional usage "fricative" does not always imply {0 loud, 1 constrictive}. It is therefore necessary to take caution when consulting an existing description of a language that marks a segment as a voiceless fricative.

The feature {bright} distinguishes segments with a greater level of noise (1) from those without (0). For loud segments, relative noisiness is achieved by interfering the central airflow, e.g., ‹w_{flap}, ř, l›. Strident obstruents such as ‹c› are {1 bright} and mellow obstruents such as ‹tθ› are {0 bright}. Lateral fricatives such as ‹lʒ› are {1 bright} and simple plosives such as ‹d› are {0 bright}.

DURATION AND SYLLABICITY FEATURES

A vowel is {1 quick} iff it is either non-syllabic or extra-short. A loud segment with a strong obstruction for the primary articulation is {1 quick} if the prolongation of the obstruction makes the spontaneous voicing impossible. Glides, including consonant-like glides ($LbDC), are {1 quick}. Flaps are {1 quick} because prolongation of the closure would turn it into a plosive. Trills are {0 quick} because prolongation of it would be prolongation of trilling, not closure.

MANNER OF ARTICULATION FEATURES

The manner of articulation features {occlusive, fricative} describe both primary and non-primary articulations.

{occlusive} is a multi-valued feature. The coefficient represents the number of occlusive places for the segment. Occlusions can be made in the oral cavity, the pharynx, or the larynx. If the nasal cavity is closed, it counts towards the number of occlusions. Only complete closures count towards the coefficient of this feature. A simple oral stop is {2 occlusive} because a closure is made in the oral cavity and the nasal cavity is shut, but a simple lateral liquid is {1 occlusive} because no complete closures are made in the oral cavity and the nasal cavity is closed. Conversely, if a complete closure is formed within the oral cavity while the nasal cavity is open, it is also {1 occlusive}. Successive closures (trilling) at the same place for the same segment are counted as one. A nasal vowel is {0 occlusive} because no closures are made in the oral cavity and the nasal cavity is open.

A nasal obstruent, a.k.a. a post- or pre-stopped nasal, for which the nasal cavity remains open only for a partial period, is {2 occlusive} because one closure is made in the oral cavity and the nasal cavity is at least momentarily closed. A doubly articulated oral stop, whether it is a click or not, is {3 occlusive} because two closures are made in the oral cavity and the nasal cavity is closed. If the segment accompnies a glottal stop as its component, the laryngeal occlusion counts towards the coefficient of the feature.

A segment is considered {1 fricative} if it has a strong enough spraglottal constriction to prevent spontaneous voicing, provided that this constriction is maintained for a sufficient duration. If a segment lacks a sufficiently long fricative-like component, it is {0 fricative}.

Unlike {occlusive}, {fricative} is binary because it does not seem necessary to distinguish segments with two fricative components from those with only one.

ACTIVE ARTICULATOR FEATURES

The active articulator features describe the articulatory movements of individual articulators.

A segment is {1 open} if the jaw is significantly dropped for the segment. Otherwise, it is {0 open}. As a rough approximation of how open is open enough for {1 open}, ‹v› is difficult to produce if the jaw is {1 open}. Conversely, ‹$w_{\text{labiodental}}$› is difficult to produce with {0 open}. In describing minimal contrasts, discriminating jaw positions is rarely or perhaps never necessary, but {open} is referenced when we define {high}.

The feature {laminal} specifies the tongue tip position. A segment is {0 laminal} if it is apical, {1 laminal} if laminal.

The tongue body positions are closely related to the first and second formant frequencies. The higher the tongue body, the lower the first formant, and the more back the tongue body, the lower the second formant. However, other organs also affect acoustic properties. Protruding the lips lowers the second formant. Lowering the tongue root raises the first formant. The acoustic properties are a result of complex combinations of gestures of independently controllable articulatory organs. For the purpose of classifying phones, it is easier to construct an abstract two-dimensional space of idealized height and backness. It is akin to simulating a language only with simple tongue body movements where the real language utilizes complex combinations of several other gestures. We assume the following virtual dorsal space with seven levels of height and four levels of backness.

Backness index:		0	1	2	3
Height index:	6				
	5	i			u
	4				
	3	e			o
	2				
	1	æ			ɔ
	0				

We use the features {high, raised, front, retracted} to describe idealized tongue body positions in the way that the values of the features correspond to these indexes. The height is determined by {high, raised}, and the backness is determined by {front, retracted}. Not using the height index and backness index directly as phonetic features is for keeping the phonology-phonetics interface less painful. Decomposing them into combinations of multiple categories is motivated by phonology.

{high} distinguishes three degrees of tongue body height. A segment is {2 high} iff the tongue body is sufficiently high to make a {1 constrictive} primary articulation, even if the segment is not dorsal. A segment is {0 high} iff the tongue body is so low that it is required to be {1 open}. High vowels are {2 high}, mid vowels are {1 high}, and low vowels are {0 high}. In addition, rhotic vowels are assumed to be {1 high}.

{raised} defines finer degrees of height. {2 raised} is a raised position, {0 raised} is a lowered position, and {1 raised} is between the two.

The correspondence between the height index and the features {high} and {raised} as well as {constrictive} and {open} is summarized below:

constrictive	open	high		raised			H. index
				2			6
1	0/1	2	High	1			5
				0	2		4
0	0/1	1	Mid		1		3
					0	2	2
0	1	0	Low			1	1
						0	0

{1 high} partially overlaps in H. index with {0 high} and {2 high}. This is intentional. {raised} appears as subcategories of {high} in some languages, but some West African languages with ATR harmony have *cross-hight pairs*, in which the higher subcategory of a major category overlaps in height with the lower subcategory of the adjacent major category. In Foodo, the non-ATR high vowels are as low as the ATR mid vowels without neutralizing, distinguished by F2 (p. 113). Describing a language like this is easier with a feature system that allows convergence.

<u>Foodo vowel system (A sample analysis based on Anderson 2008, pp. 104-113)</u>:

		B. index:					
		H. index:	0	1	2	3	Phonetic features:
High	ATR	6	i			u	{2 high, 2 raised}
	Non-ATR	4	i			u	{2 high, 0 raised}
Mid	ATR	4		ə	ɵ		{1 high, 2 raised}
	Non-ATR	2		ə	ɵ		{1 high, 0 raised}
Low		1		a			{0 high, 1 raised}

Similar contrasts are often found in Germanic languages. Unlike West African ATR languages, the articulatory gestures responsible for the fine contrasts of vowel height seem to be made primarily on the tongue body, at least for English (Ladefoged & Maddieson 1996). In Swedish, {raised} is concomitant with the vowel length. The corresponding properties are sometimes interpreted with the notions of tense and lax (Löfstedt 2010). (These cannot be equated with our notions of tense and lax.) In a vowel pair, the tense vowel is longer and higher than the lax vowel. Although less obvious than Foodo, the Swedish vowel system shows what looks like a height convergence between phonological categories. A tense high vowel ‹ʉ:› and a lax high vowel ‹u› appear to be very close in height. This is potentially another instance of height convergence which can be captured in the present phonetic framework.

<u>Swedish vowel system (A sample analysis based on Engstrand 1999, pp. 140-141)</u>:

		B. index:					
		H. index:	0	1	2	3	Phonetic features:
High	Tense	6	i:	ü:		u:	{2 high, 2 raised}
		5			ʉ:		{2 high, 1 raised}
	Lax	5				u	{2 high, 1 raised}
		4	i	ü			{2 high, 0 raised}
Mid	Tense	3	e:	ö:		o:	{1 high, 1 raised}
	Lax	2	e	ö	ɵ	o	{1 high, 0 raised}
Low	Tense	1	æ:		æ̈:		{0 high, 1 raised}
	Lax	0			a		{0 high, 0 raised}

In strictly articulatory terms, the vowel ‹ʉ:› is front (Ladefoged & Maddieson 1996, p. 295).

The features {front} and {retracted} are both binary. They correspond to the backness index as illustrated by the following:

B. index:	0	1	2	3
front	1	1	0	0
retracted	0	1	0	1

Four levels of backness seem to be optimal, that is, sufficient and necessary. While it is common in the literature to suppose three levels of backness and two levels of roundness, it is unlikely that in a language all of these categories are minimally contrasted. The maximum number of contrasts along the horizontal axis known for a language is four (Rice 1995).

Dorsal co-articulations of consonants are described similarly. What is conventionally called “palatalization” of a consonant is raising and fronting of the tongue body, hereafter referred to as “high-fronting,” during the course of the primary articulation. This can be described by the same dorsal space features. Other dorsal or labial co-articulations are treated similarly.

The following chart exemplifies dorsal co-articulations of a coronal stop, indicated by the symbol ‹t› along with co-articulation affixes, «_y, _w, _ħ». Since consonants put certain limits on the dorsal position, the height features {high, raised} are abstracted as ad hoc Higher, Plain, and Lower. The default position of ‹t› is tentatively considered as Plain and B. index: 0.

B. index:	0	1	2	3
Higher	ty			
Plain	t			tw
Lower	tħ			

> We assume that for central approximants with {1 laminal, 2 high, 1 front}, the coronal and dorsal places of the primary articulation are neutralized. This mitigates the problem of the blurry borderline between coronals and dorsals while making high-front glides ‹y, ẅ› more accessible. A coronal articulation is often found for front vocoids, e.g., Czech ‹y›, which is articulated with the blade in addition to the tongue body (Keating & Lahiri 1993, p. 81). A coronal central approximant, e.g., IPA [ɹ], is often taken for granted to be apical. It seems that high-front approximants are typically simultaneously coronal and laminal coronal approximants are seldom necessary. (Presumably, the latter would sound like a dorsal approximant.) Consequently, high-front glides are available in three different places: anterior coronal (−PO), posterior coronal (+−P +O), and dorsal (+PO). These glides may be distinguished in a language by the other phonetic features including {raised} or by properties we have not yet formalized, but it is not necessary that, say, the “coronal” ‹y› is more coronal than the “non-coronal” ‹y›. (Our phonetic features are illustrative devices while the phonological features are theoretical constructs; contrasts are available if and only if phonological features specifications are different.)

The same set of features describes the primary articulation if the consonant is dorsal. For the primary articulation, the following diagram partially explains the correspondences. The default position of the dorsal consonant is assumed to be Plain and B. index: 2. Pharyngeal cannot be the primary place of a consonant because the pharynx is not in the oral cavity.

B. index:	0	1	2	3
Higher	ky			
Plain			k	kw
Lower				

Three levels of abstract heights are maintained for dorsal consonants although in reality the tongue body height is essentially fixed. Virtual heights are still useful for representing tongue root positions and potentially other factors.

Loose glides, ‹h, ɦ, ʔ›, etc., are pronounced at a specific position, but unless pharyngealized, the tongue position rarely matters. It is usually assumed that the tongue position is similar to that of the adjacent segments, usually vowels.

PLACE OF ARTICULATION FEATURES

The place of the primary articulation is specified by {coronal} and {posterior}. The features {coronal, posterior} are both binary. A segment is {1 coronal} iff the primary articulation targets the coronal place. A segment is {1 posterior} iff the primary articulation targets the posterior place.

Place:	labial	dental	palatal	dorsal
coronal	0	1	1	0
posterior	0	0	1	1

The division between coronal and dorsal is often problematic. ‹š› as in *shy* in English is uncontrovertially coronal and ‹k› as in *kite* is clearly dorsal. ‹ky› as in *Keating* is more front than ‹k› as in *kite* but is still dorsal. Front vocoids such as ‹y› as in *yacht* are usually considered dorsal but are often found to have a concomitant coronal articulation, which seems to mean that front vocoids are near the borderline, at least in some languages. It is expected that fine adjustments are necessary for particular languages. It is also expected that extremely fronted dorsals and extremely raised posterior coronals partially overlap. It would be useful to see if a borderline segment patterns as dorsal or coronal in the given language. For reference, the following is an excerpt from Keating (1991):

“[…] Alveolar stops and fricatives can be produced with the tongue tip down behind the lower teeth and a part of the tongue further back forming the constriction at the alveolar ridge. The phonological notion “coronal” surely depends on such articulations being made with the blade of the tongue, yet they are formed more than 1 cm behind the tip. […] In my own case this suggests a blade length on the order of 15-20 mm. […] [T]he maximum estimate for blade length is that part of the tongue in front of the part used to produce velar, that is, some 3 or 4 cm. […] [W]e will consider the blade of the tongue to be, conservatively, the movable part extending from 1 to 2 cm behind the tip, and we will consider the tip to include a small rim around the edge of the tongue.” (Keating 1991, p. 31.)

AIRSTREAM FEATURES

A segment is {1 nasal} iff the nasal tract is open for at least a partial period of the targeted articulations.

A segment is {2 voiced} if it is pronounced with a modal voice, {1 voiced} if it is pronounced with a non-modal voice, and {0 voiced} if it is voiceless.

A segment is {1 whispered} iff it is pronounced with a higher rate of pulmonic airstream across the glottis, *ceteris paribus*. A breathy voice has a higher rate of airstream than a creaky voice. An aspirated stop has a higher rate of airstream than an unaspirated stop. A glottal stop is {0 whispered} because for this type of segment the airstream is blocked at the glottis.

Sonorants:

whispered		0	1
voiced	0	glottal stop	voiceless
	1	creaky	breathy
	2	modal	modal

Stop obstruents:

whispered		0	1
voiced	0	tenuis	aspirated
	1	creaky	breathy
	2	modal	modal

It is not recommended to use ±W for contrastive pitch unless evidence suggests the given pitch feature is strictly segmental in the given language.

PHONES

FLAPS AND TRILLS

Flaps and trills have a closure component. Flaps are often seen as short plosive. They are sonorant despite having a phase of complete blockage of the air stream. Unlike flaps, trills can have multiple stop phases per a single articulation.

In articulating ‹r̃›, the blade is positioned to form a central constriction, like ‹r›, during the open phase. Because this segment is sonorant (it is sonorant-like in acoustics and is traditionally considered so), I assume that the degree of narrowing during this open phase is similar to that of an approximant. The articulator is appropriately positioned so that the air stream causes vibration. The air stream reduces the pressure behind the constriction and causes the articulator to raise, after which the pressure builds up again until it breaks the closure (SPE, p. 318). Taken together, the “base” or preparatory position of a trill is approximant-like (hence it is $LC and {1 loud, 1 constrictive}), and it differs from approximants in that the active articulator and the air stream mechanism are configured so that the air stram causes trilling.

Some may argue that if the open phase of a trill is approximant-like, this segment has a secondary articulation—having a stop component and an approximant component, just as a velarized dental stop has a stop component similar to ‹d› and an approximant component similar to ‹w̥›. However, due to the lack of a more basic counterpart, that is, a trill with a non-constrictive open phase, a simple trilling is considered atomic. (Presumably, a trill without a preparatory constriction is either very difficult or impossible.)

SPE proposes a potential distinction between a regular flap and a “tongue flap,” known as the *flap D,* of American English. According to SPE (p. 318), the regular flaps are similar to trills but made with a lower subglottal pressure while the tongue flap may be made with active muscular movements of the tongue, like the regular plosives. We do not distinguish these two types of flaps. We assume that flaps differ from plosives (and laterals) in that they are shorter, and from trills in that they lack a preparatory constriction. Flaps are {1 loud, 0 constrictive, 1 quick}.

The Czech fricative trill ‹$\tilde{r}_{fricative}$› is interpreted to have open phases with a more radical (fricative-like) constriction. This segment is still considered +L and {1 loud}. When the open phase of a trill is fricative, the fricative component is considered as a secondary articulation. Non-primary articulations can be like an obstruent without making the segment {0 loud} because major class features only concern primary articulations. This prevents some phonological sonorants from being classified as –L.

	loud	bright	constrictive	quick	fricative
d	0	0	0	0	0
l	1	1	0	0	0
r	1	0	1	1	0
ř	1	1	0	1	0
r̃	1	1	1	0	0
$\tilde{r}_{fricative}$	1	1	1	0	1

PHONOLOGY-PHONETICS INTERFACE

FIRST APPROXIMATION

For the following pairs of features, the correspondences are rather straightforward.

{loud} and ±Loud
{1 loud} iff [+Loud], {0 loud} iff [–Loud].
{bright} and ±Bright
{1 bright} iff [+Bright], {0 bright} iff [–Bright].
{constrictive} and ±Constrictive
{1 constrictive} iff [+Constrictive], {0 constrictive} iff [–Constrictive].
{whispered} and ±Whispered
{1 whispered} iff [+Whispered], {0 whispered} iff [–Whispered].
{nasal} and ±Nasal
{1 nasal} iff [+Nasal], {0 nasal} iff [–Nasal].

These correspondences are stated in the following format.

ΔL(loud) = 1
ΔB(bright) = 1
ΔC(constrictive) = 1
ΔW(whispered) = 1
ΔN(nasal) = 1

ΔL(loud) = 1 reads "{loud} for +L is one point greater than {loud} for –L with all other phonological features equally specified." Since {loud} is binary, this can only mean that it is {1 loud} if +L and {0 loud} if –L.

These conditions must be sufficed simultaneously in the surface form, i.e., they are inviolable constraints.

We will be adding constraints until we obtain some reasonable approximation of an easily imaginable (partial) setting.

PLACE SPECIFICATIONS

For consonants, {1 coronal} iff –P, and {1 posterior} iff +O. For vocoids (vowels and dorsal glides), a segment is {0 coronal, 1 posterior} regardless of the values of ±PO. Rhotic vowels are {1 high}. For non-rhotic vowels, {0 high} iff +O and {0 front} iff +P. Hence, the conditional statements must distinguish between some basic categories like vocoids and non-vocoids. If a segment is –L then it is necessarily a consonant. A +LB segment is roughly a sonorant consonant. A $Lbd segment is necessarily a vowel. $LbDc is a non-high dorsal glide. $LbDC is a consonant-like glide, for which ±PO will map onto {coronal, posterior}.

contoid:
ΔP(coronal) = –1
ΔO(posterior) = 1
vocoid:
ΔP(front) = –1
ΔE(retracted) = 1
$C:
$O(high) = 1
$o(high) = 2
$c:
$O(high) = 0
$O(high) = 1
where,
contoid = {$l, $LB, $LbDC}
vocoid = {$Lbd, $Lbc}

We use the symbols **= { }** to group multiple classes together and to assign a group name. The classes to be grouped are placed in the curled brackets, separated by commas, and the group name is placed before the equal sign.

Indentation (shown by tab spaces) indicates nesting.

> In EKPA 2, $LbDC was taken to be an ambiguous group. It was allowed to interpret a segment of this group like a consonant or like a vowel for the place specifications. In EKPA 3, if a segment is $LbDC, ±PO obligatorily maps onto {coronal, posterior}. This is a design decision made in light of the fact that parallelism between the vowels and the glides is rarely necessary—it is fairly common for a language to have more than ten contrastive vowels, but it is unlikely for a language to have more than ten contrastive dorsal glides. In sum, a segment is {1 coronal} if $lp, $LBp, or $LbDCp, {1 posterior} if $lO, $LBO, or $LbDCO, {1 front} if $Lbdp or $LbDcp, and {0 high} if $LbdO or $LbcO.

In addition to ±PO, the features ±ER have different consequences depending on the type of the segment. In the vowel context, ±E corresponds to {retracted}, and ±R to {raised}. For the non-dorsals, the correspondences are more complex. +R may imply, depending on the language, high-fronting. This means that the feature ±R must condition the specifications of {front, retracted} in addition to {high, raised}.

> Keating & Lahiri (1993, p. 99, paragraph 4) suggest that the backness of a velar constant is unspecified in surface forms in languages like Czech, where fronting is non-contrastive, while in Russian it is specified because backness is contrastive. This is an example of a ternary system built with ostensibly binary features. Under K & L's proposal, if in a given language the backness of a velar stop is unspecified, it receives automatic fronting before a front vocoid, while in a language where, like Russian, the backness of a velar stop is specified, contextual fronting is canceled. The problem of this system is that it distinguishes a front dorsal ([– back]), a back dorsal ([+ back]), and a dorsal with backness unspecified. If the grammar (or whatever system that "reads" the features) is sensitive to the "unspecified" status while it is also sensitive to the pluses and minuses of the features, it is a non-binary system. A ternary system is not necessarily bad—base ten is not inherently better than base sixteen. But a pseudo-binary system like this is deceptive—it appears as binary when it is not. Speaking of practicality, a ternary system like this, where one of the possible values is deemed "unspecified," is less economical than a binary system that is capable of expressing the same number of phones. The aforementioned K & L system is only capable of expressing two levels of backness (1 bit) on the surface despite that the feature carries a ternary digit (about 1.6 bits of information). A strictly binary feature would carry 1 bit of information and represent two levels of backness (1 bit). A binarist solution to the Russian problem would be to say that Czech has a fronting rule while Russian does not. In EKPA, a ternary interpretation of a binary feature is never allowed.

Here, we assume that a velar consonant is by default at the velarized place while a coronal consonant is by default at a different place. The treatment is essentially the same as SPE. In the following, it is conveniently assumed that the default tongue position of a coronal consonant is {1 front, 0 retracted, 1 high} and that of a velar consonant is {0 front, 0 retracted, 2 high}. With these settings, high-fronting of a coronal consonant is just raising, i.e., shifting the tongue position to {2 high}, and that of a velar consonant is just fronting, i.e., shifting the place to {1 front}. By the same token, a labial consonant is assumed to be {0 front, 0 retracted, 1 high}, from which, both raising and fronting are available. In any of the three cases, {raised} may or may not be involved.

Note that it is not intended to claim that some language has this exact setting. It intends to show that if in a given language a certain class of consonants is immune to certain types of co-articulation, the immunity can be explained by supposing the default positions, and that at least some theoretical dorsal position with which contrastive co-articulations are implemented must be available.

front:	1	1	0	0
retracted:	0	1	0	1
high:				
2			k	
1	t		p	
0				

±E relates to horizontal settings, but not vertical. Hence, if implementing velarization strictly requires raising, in addition to backing, it needs to be given by +R.

Strident dentals like ‹s, z› are conventionally assumed to be pronounced with a high tongue body (e.g., Keating 1988, p. 276). If these consonants allow both high-backing and high-fronting, the default position may be in the middle of the two: {1 front, 1 retracted, 2 high}. But a simpler solution would be to assume that strident dentals are by default {1 front, 0 retracted, 2 high}, and the high-fronted and the plain form are distinguished by {raised}, that is, the categorical difference between ‹s› and ‹sy› is assumed to be not that the latter is further fronted, but that the latter is further raised—although in reality ‹sy› is more front than ‹s› (Keating 1988, p. 276). For particular languages, phonetic studies may falsify one or both of these hypotheses.

front:	1	1	0	0
retracted:	0	1	0	1
raised:				

high:

2	2	sy		
	1	s		sw
	0	sħ		

For simplicity, let us just say that a coronal consonant is {1 front}, a non-coronal consonant is {0 front}, +E is {1 retracted}, and –E is {0 retracted}.

```
contoid:
    ΔP(coronal) = –1
    ΔP(front) = –1
    ΔE(retracted) = 1
where,
    contoid = {$l, $LB, $LbDC}
```

VOCOIDS

For vocoids, $Co is {2 high}, $cO is {0 high}, +E is {1 retracted}, and –E is {0 retracted}. The setting for +R varies, but +R must be associated with an equal to or greater value of {raised} than –R. A rhotic vowel is +CO and {0 laminal}. {height} of a rhotic vowel may vary, but for convenience we can suppose that it is always {1 high}.

```
vocoid:
    ΔP(front) = –1
    ΔE(retracted) = 1
    ΔR(raised) ≥ 0
    ΔCO(laminal) = –1
    $C:
        $O(high) = 1
        $o(high) = 2
    $c:
        $O(high) = 0
        $o(high) = 1
where,
    vocoid = {$Lbd, $Lbc}
```

The symbol ≥ in place of = indicates that, for example, the difference between the value of {raised} for +R and that of the same phonetic feature for –R is equal to or greater than the number given on the right.

ΔCO(laminal) = –1 reads "If $CO, {laminal} is one point smaller than otherwise." This ensures that rhotic vowels ($CO) are {0 laminal} and non-rhotic vowels (¿CO) are {1 laminal}.

LARYNGEAL SETTINGS

The correspondence between ±G and {voiced} depends on the value of ±W. If +W, +G implies a greater value of {voiced}. If –W, +G corresponds to a smaller value of {voiced}. The syllabic context may also affect the correspondence. It is unlikely that $wG is a glottal stop for a syllabic vowel, but such a situation is fairly common for a glide. The following is an example of a relatively simple situation.

sonorant				
W	+	+	–	–
G	–	+	–	+
	h	ɦ	ə	ʔ
voiced	0	1	2	0
obstruent stop				
W	+	+	–	–
G	–	+	–	+
	th	dɦ	d	t
voiced	0	1	2	0

Based on this,

```
$W:
    ΔG(voiced) ≥ 1
$w:
    ΔG(voiced) ≤ –1
```

SUMMARY OF THE RULES

<u>A partial representation of a sample phonetic grammar</u>:

ΔR(raised) ≥ 0
ΔL(loud) = 1
ΔB(bright) = 1
ΔC(constrictive) = 1
ΔW(whispered) = 1
ΔN(nasal) = 1
ΔE(retracted) ≥ 1
$W:
    ΔG(voiced) ≥ 1
$w:
    ΔG(voiced) ≤ 1
contoid:
    ΔP(coronal) = −1
    ΔP(front) = −1
    ΔO(posterior) = 1
vocoid:
    ΔP(front) = −1
    ΔR(raised) ≥ 0
    ΔR(laminal) ≥ 0
    $C:
        $O(high) = 1
        $O(laminal) = 0
        $o(high) = 2
    $c:
        $O(high) = 0
        $O(high) = 1
where,
    contoid = {$l, $LB, $LbDC}
    vocoid = {$Lbd, $Lbc}

Language-specific adjustments are necessary. This model only partially approximates one of several possible types of idealized languages. This model does not account for all phonological features.

ILLUSTRATIONS

The illustrations chapter is planned to be filled in later. See EKPA 2 for earlier attempts.

CONCLUSION

REFERENCES

ABBREVIATIONS

EKPA 2		Iida (2024a) and its earlier versions since 2.0 (also available at the same GitHub repository.)
IPA	1999	*Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.
SPE		Chomsky & Halle (1968).

REFERENCES

Some online sources are hyperlinked in the body.

Akamatsu, T.	1992	A critique of the IPA chart (revised to 1951, 1979, and 1989).
Anderson, C. G.	2008	The acoustic correlates of the harmony in seven- and nine- vowel African languages: A phonetic inquiry into phonological structure. *Dissertation*. The University of Texas at Arlington.
Beguš, G.	2020	Segmental phonetics and phonology. [Polinsky 2020]
Chomsky, N. & Halle, M.	1968	*The Sound Pattern of English*. Harper and Row, Publishers.
Dixit, R. P.	1989	Glottal gestures in Hindi plosives. *Journal of Phonetics*. 17,213-237
Engstrand, O.	1999	Swedish. [IPA, 141-142]
Hall, K. C.	2009	A Probabilistic Model of Phonological Relationships from Contrast to Allophony. *Dissertation*. The Ohio State University.
Iida, J.	2024a	*EKPA: A feature-based fonetık alphabet*. Unpublished version 2.7. https://github.com/NihongoTopics/EKPA/
	2024b	*Proposing a new romanization of Japanese*. Unpublished version 5.1. https://github.com/NihongoTopics/Kyouro/
Kawagoe, I. (川越いづえ)	2014	音節とモーラ. [Sugahara, M. 30-55].
Keating, P.	1988	Underspecification in phonetics. *Phonology*.
	1991	Coronal place of articulation. *Phonetics and Phonology,* Volume 2, 29-48.
Keating, P. & Lahiri, A.	1993	Fronted Velars, Palatalized Velars, and Palatals. *Phonetica*. 50. 73-101.
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.
Kodama, N.	2008	曲線声調と日本語韻律構造. ありあけ　熊本大学言語学論集 7 1-40.

Ladefoged, P.	1971	*Preliminaries to linguistic phonetics*. University of Chicago Press.
Ladefoged, P., & Jonson, K.	2014	*A Course in Phonetics, 7th ed*. Cengage Learning.
Ladefoged, P. & Maddieson, I.	1996	*The Sounds of the World's Languages*. Wiley-Blackwell.
Löfstedt, I.	2010	Phonetic Effects in Swedish Phonology: Allomorphy and Paradigms. *Dissertation*. University of California Los Angeles.
Nespor, M. & Vogel, I.	1986	*Prosodic Phonology*. [Nespor & Vogel 2007].
	2007	*Prosodic Phonolog: with a new forward*. De Gruyter Mouton.
O'Connor, J. D.	1973	*Phonetics*. Penguin Books Ltd.
Polinsky, M. (Ed.)	2020	*The Oxford Handbook of Languages of the Caucasus (Oxford Handbooks)*. Oxford University Press.
Pycha, A.	p.c.	A lecture of phonology at the University of Wisconsin-Milwaukee in 2020.
Rice, K.	1995	On vowel place features. *Toronto Working Papers in Linguistics* 114(1), 73-116
Sugahara, M. (菅原真理子) (Ed.)	2014	音韻論. 朝倉書店.

ABOUT THIS DOCUMENT

This is a major revision of EKPA 2, written by Junichi Iida, a.k.a. @awesomenewways (It's not Lida).

To cite this document:

Junichi Iida. 2024. EKPA3. Unpublished version 3.0. Available at https://github.com/NihongoTopics/EKPA/.