Laryngeal Complexity in Otomanguean Vowels
Author(s): Daniel Silverman
Source: *Phonology*, Vol. 14, No. 2 (1997), pp. 235-261
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/4420102
Accessed: 29/08/2010 02:00

Phonology **14** (1997) 235–261. Printed in the United Kingdom

*Laryngeal complexity in Otomanguean vowels**

Daniel Silverman
University of Illinois at Urbana-Champaign

1 Introduction

Linguistic sound systems necessarily possess contrastive values that are sufficiently distinct from one another that their individual characters may be learned by the listener. In this way, any given value in any given system fulfils its functional role of rendering forms distinct which differ in meaning. Articulatory, aerodynamic, acoustic and auditory constraints serve to mediate between such sound–meaning correspondences in non-trivial ways. Indeed, if it can be shown that the sound patterns of language are in part explainable by these physical systems, then students of linguistic sound systems would do well to study in detail the phonetic base. Consider an example case. Laryngeal gestures and supralaryngeal gestures are by and large articulatorily independent of each other. Thus, for example, a voiceless aspirated stop consists of an oral occlusion, cued by silence, as well as an articulatorily independent laryngeal abduction, cued by broadband noise. Were the phonetic realisation of these two gestures strictly simultaneous, the cues signalling the laryngeal abduction would not be perceived as such by the listener (*[t̥]). A listener can tell that there is no voicing, but cannot recover more specific information regarding the state of the glottis during oral closure. Stated simply, the full closure here reduces the acoustic output to zero. With zero acoustic energy, no source information other than silence is transmitted to the listener. However, upon staggering the two gestures, the otherwise obscured information is rendered salient. Maximal laryngeal abduction is normally realised at or around the interval of release of the oral occlusion ([tʰ]), for example in English (Yoshioka *et al.* 1986), Danish (Fukui & Hirose 1983), Hindi (Dixit 1989) and Korean (Kim *et al.* 1992). As the maximally abducted larynx is phonetically realised across the transition from the stop into the following more sonorous gesture, sufficient acoustic energy is present to transmit source information other than silence to the listener. There are well-established aerodynamic, articulatory, acoustic and auditory reasons why specifically postaspirated stops are cross-linguistically preferred to their preaspirated counterparts ([ʰt]), and, relatedly, why the presence of preaspirates in a language implies the presence of postaspirates (see especially Kingston 1985, 1990, Bladon 1986). Thus, implicational relations within segment inventories may here (and, potentially, elsewhere) be accounted for, at least to some extent, by reference to the physical systems which mediate between the speaker and the listener (see also, for example, Liljencrants & Lindblom 1972, Ohala 1975, 1983, 1992, Löfqvist 1980, Bladon 1986, Lindblom 1983, 1984, 1986, 1989, 1990a, b, Lindblom *et al.* 1984, Flemming 1995, Jun 1995, Steriade 1995, Silverman 1996a, b, 1997a, b).

In this paper I explore another case of sound patterning which lends itself, again, at least in part, to a physical explanation. In contrast to stops, vowels, with their maximum acoustic energy, are able to accommodate simultaneously cues of oral origin (place cues) and cues of laryngeal origin (laryngeal spreading: [h] or [V̤]; laryngeal constriction: [ʔ] or [V̰]), or, of course, laryngeal configurations which affect rate of vocal fold vibration: tone. As airflow may persist relatively unimpeded for the duration of a vowel, sufficient energy is present to allow the full simultaneity of oral and laryngeal gestures, without risk of obscuring the cues of either component. Indeed, the existence of breathiness or creakiness accompanying any and all vowel qualities confirms this (for example, breathiness in Oriya (Dhall 1966) and Gujarati (Fischer-Jørgensen 1970), and creakiness in Sedang (Smith 1968, 1979)). Here, instrumental studies often show that breathiness or creakiness persists for the duration of the vowel.

The story becomes rather more complex when considering vowels that possess both contrastive phonation and contrastive tone. Such vowels, which I term 'laryngeally complex', are attested throughout the Otomanguean language group of southern Mexico. As I argue here, tone is most perceptually salient when occurring with plain or 'modal' phonation, and less salient when occurring with non-modal phonation (breathiness or creakiness). Largely for this reason, in laryngeally complex vowels, tone and non-modal phonation may be sequenced with respect to one another, so that tone may be realised in modal voice. In Jalapa Mazatec, for example, contrastive laryngeals precede the toned vowel ([V̤͡V˥], [V̰͡V˥]). In Comaltepec Chinantec, laryngeals may either precede the toned vowel, as in Jalapa Mazatec, or follow the toned vowel ([Vh˥], [Vʔ˥]). Finally, Copala Trique possesses both of these contrastive timing patterns, and in addition, allows contrastive laryngeal 'interruption', in which aspiration or glottal closure intrudes upon the middle of the toned vowel ([VhV˥], [VʔV˥]. And again, as argued herein, there are aerodynamic, articulatory, acoustic and auditory reasons why certain timing relations among these contrastive gestures are preferable to others, and relatedly, why the presence of certain timing patterns in a system usually implies the presence of other timing patterns. It is, by hypothesis, in part for these reasons that the presence of postvocalic laryngeals in Comaltepec Chinantec and Copala Trique implies the presence of prevocalic laryngeals, and the presence of interrupted forms in Trique implies the presence of both prevocalic and postvocalic laryngeals.[1]

Especially relevant here is that attested sub-optimal timing patterns are normally maximally distinct from preferred timing patterns: sub-optimal postvocalic laryngeals ([Vh˥], [Vʔ˥]) are maximally distinct from optimal prevocalic laryngeals ([hV˥], [ʔV˥]), and interrupted forms ([VhV˥], [VʔV˥]), being equidistant in timing from prevocalic and postvocalic laryngeals, are maximally distinct once again. As argued herein, sound systems typically maximise the perceptual distinctness among their contrastive values, mediated by natural biological constraints on the involved articulators and the auditory system. This approach to phonological investigations is not new. Rather, it is inspired by the computational models devised and employed by Liljencrants & Lindblom (1972), Lindblom (1983) and Lindblom *et al.* (1984), which predict vowel inventories by numerically modelling the maximal perceptual distinctness among their members. In the present investigation, which remains unquantified, perceptual distinctness is considered on a temporal scale rather than on a spectral scale. That is, I consider contrasts involving changes in phonation over time, whereas Lindblom and his associates primarily investigate contrasts involving static vowel and consonant qualities in isolation and combination.

In §2 I consider the three cases of laryngeal complexity in Otomanguean, providing data from Jalapa Mazatec (§2.1), Comaltepec Chinantec (§2.2) and Copala Trique (§2.3). In §3 I review the phonetic facts that might constrain timing patterns for laryngeally complex vowels, considering, in turn, sufficient acoustic distance (§3.1), sufficient articulatory compatibility (§3.2) and optimal auditory salience (§3.3).

In §4 I discuss some additional cases of laryngeal complexity which pattern differently from those in Otomanguean. Instead of sequencing the tonal and phonatory gestures, the Tibeto-Burman languages of Mpi and Tamang indeed realise their laryngeally complex vowels with simultaneous tone and non-modal phonation. Tellingly, the non-modal phonation, at least in Mpi, is much more lightly implemented than that in, for example, Jalapa Mazatec. In this way, full overlap of all contrastive cues is maintained, but perhaps at the expense of a genuinely robust realisation of either contrastive component. I argue that such superficial counterexamples to my claims regarding laryngeal complexity are actually expected to exist, if we take seriously Lindblom's (1983) hypothesis that 'sufficient articulatory compatibility' is an independent force acting on linguistic sound systems.

Finally, in §5, I consider the results of the previous sections, moving towards a quantitative model of timing contrasts in laryngeally complex vowels.

2 The data

2.1 Jalapa Mazatec

Jalapa Mazatec (Pike & Pike 1947, Kirk 1996, 1969, Bull 1983, Kirk *et al.* 1991, 1993, Steriade 1992, 1994, Silverman 1994, Silverman *et al.* 1995) is spoken by 10–15,000 people in the states of Oaxaca and Veracruz, Mexico, and is a member of the Popolocan branch of the Otomanguean language family. Jalapa Mazatec contains the segment inventory shown in (1) (from Silverman *et al.* 1995).

(1) *Jalapa Mazatec segment inventory*

p	t	ts	tʃ	k		i			u
pʰ	tʰ	tsʰ	tʃʰ	kʰ			æ	o	
ᵐb	ⁿd	ⁿdz	ⁿdʒ	ᵑg				a	
	s		ʃ						
m	n		ɲ	ŋ					
	l								
w		j							
h,ʔ									

(The labial obstruents and the lateral are limited to loanwords.)

Tone, nasality and length contrasts greatly expand the vowel inventory. Jalapa Mazatec possesses three level tones (˩, ˧, ˥), as well as a series of tonal contours (Kirk 1966).

Laryngeals may stand alone prevocalically, as exemplifed in (2):

(2) *Prevocalic laryngeals*

ʔa˧	'why'	hæ˧	'finish'
ʔã˧	'I'	ha˥	'men'

The sonorants of Mazatec are perhaps unique in that laryngeals may either precede or follow modal phonation. Some examples are shown in (3) (from Kirk 1966, Silverman *et al.* 1995):

(3) *Voiceless nasals and glides* / *Laryngealised nasals and glides*

m̥͡ma˧	'black'	ti˥ʔmæ˧	'he is sick'
n̥͡næ˧	'he falls'	ʔna˧	'shiny'
ɲ̥͡ɲa˧	'growth, bush'	ʔɲa˩	'chills'
j̥͡ju˧	'peace'	ʔja˩	'ants'
w̥͡wæ˧	'use up'	ʔwi˧	'drinks'

In addition, breathiness or creakiness may accompany Jalapa Mazatec vowels. In either case, non-modal phonation is realised primarily in the first portion of the vowel, actually beginning toward the end of any prevocalic sonorant. The second portion of the vowel usually possesses severely weakened breathiness or creakiness, verging on modal phonation. Some examples are in (4) (from Kirk 1966, Silverman *et al.* 1995):

(4) *Breathy vowels* / *Creaky vowels*

m͡m̤æ̤͡æː˦	'wants'	m͡m̰o̰͡o˧sæ̰͡æ˧	'eviction'
n͡n̤a̤͡a˦	'my tongue'	n͡n̰æ̰͡æ˧	'he says'
ɲ͡ɲ̤a̤͡a˧	'nine'	ɲ͡ɲ̰V̰͡V	(*unattested*)
j͡j̤æ̤͡æ˩	'boil'	w͡w̰a̰͡aj˧tsæ̃ːj˧	'he remembers'
w͡w̤o̤͡o˧	'hungry'	ti˥w͡w̰æ̰͡æ˧	'hits, gives birth to'

In Fig. 1 are prototypical wideband and narrowband spectrograms of two Jalapa Mazatec words, containing a plain vowel (left) and a breathy vowel (right) produced by one male and one female speaker, respectively. In the second form, note in particular the absence of a clear harmonic structure during the initial portion of vocalism, where breathy phonation resides. In §3.1 I argue that the wide bandwidths and low signal-to-noise ratio above the fundamental play a significant role in the timing patterns found in laryngeally complex vowels. In contrast, the lefthand form possesses a clear harmonic structure throughout, which is indicative of modal phonation.

In summary, non-modal phonation may be prevocalic in Jalapa Mazatec. In the context of a preceding voiceless stop, or in isolation, aspiration is voiceless; [tʰV], [hV] (see (1)–(2)). Elsewhere, aspiration or creakiness is implemented simultaneously with voicing, that is, as breathy, or creaky phonation, respectively, [V̤͡V, V̰͡V].

2.2 Comaltepec Chinantec

Comaltepec Chinantec is a member of the Chinantecan language group, spoken by approximately 1500 people in the village of Comaltepec, Oaxaca, Mexico (Grimes 1988). In (5) is the segment inventory of Comaltepec (Anderson 1989, Anderson *et al.* 1990, Pace 1990).[2]

(5) *Comaltepec Chinantec segment inventory*

p	t	tʃ		k	i	ɨ	u
ᵐb	ⁿd	ⁿdʒ		ᵑg	e	ʌ	o
f	s	ʃ	ʂ	x	æ		a
			ʐ				
m	n			ŋ			
	l						
		j		w			
h,ʔ							

([f] is a free variant of [w̥w] (/hw/). [s] and [ʃ] are allophonically related, as are [ʂ] and [ʐ].)

The tones exemplified in (6) are attested in morphologically simplex environments. (Additional tone patterns are present in morphologically and phonologically complex environments.)

(6) *Comaltepec Chinantec tones*

hi˩	'book'
ʔoː˧	'papaya'
l̥loʔ˥	'pretty'
ᵑgiŋʔ˩˧	'swing'
li˩˥	'tepejilote palm shoot'

Nuclei may possess postvocalic aspiration (Vh˥]). Syllables with such nuclei are traditionally considered to possess 'ballistic stress'. Ballistically

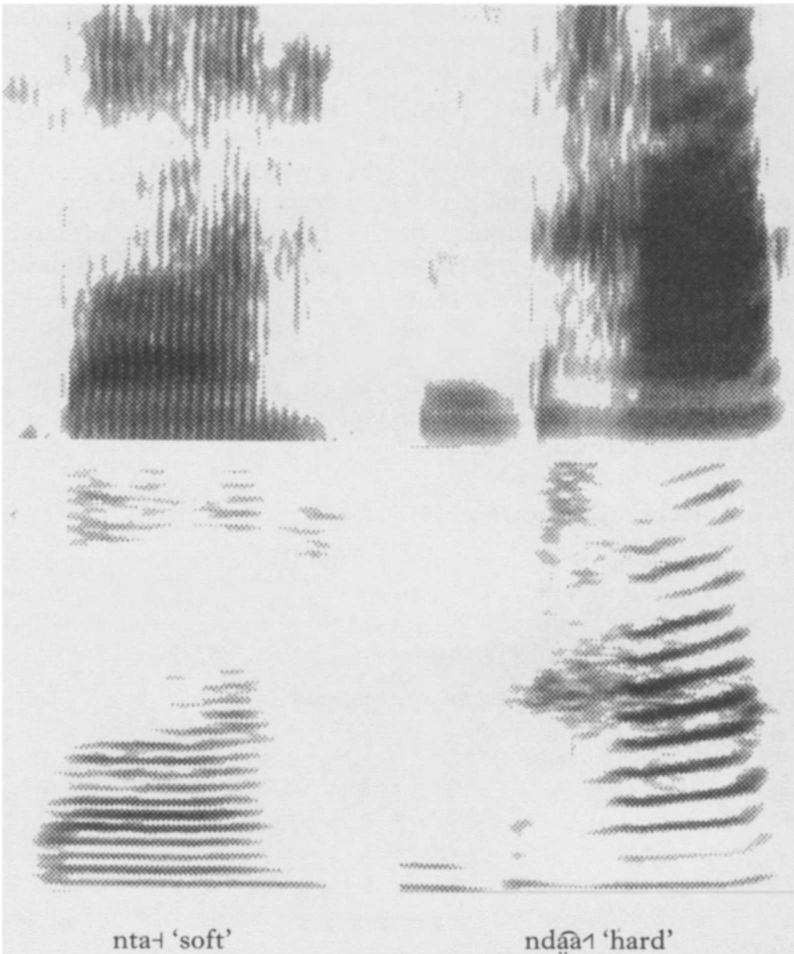


Figure 1
Spectrograms of plain vowel (*left*) and breathy vowel (*right*) in Jalapa Mazatec

stressed syllables are reportedly articulated more forcefully than 'controlled' (non-ballistic, or plain) syllables, affecting pitch, amplitude and phonation (Merrifield 1963, Rensch 1978). Some examples are presented in (7):

(7) *Plain vowels* (*'controlled'*) / *Vowels with postvocalic aspiration* (*'ballistic'*)

ta˩˥	'work'	lih˥	'flower'
heʔ˥˩	'frog'	huhʔ˥	'pineapple'
ᵑgwoː˩	'good (i)'	ᵑgjʌŋ̊˧	'hand'

In Fig. 2 are wideband and narrowband spectrograms for forms which near-minimally contrast for so-called ballisticity (from Silverman 1994, 1997a). The speaker is a forty-year-old native of Comaltepec. Wideband spectrograms indicate that the ballistic syllable (right) differs from the controlled syllable (left) in possessing significant postvocalic aperiodic noise, characteristic of aspiration. This is indicated by the faint markings towards this syllable's right edge, after the cessation of a defined formant structure. Note that this energy is aperiodic, indicated by the lack of vertical striations toward the right edge of the display. In contrast, the controlled syllable possesses periodic vibration for the duration of the vowel. This is indicated by the vertical striations which are present throughout.

Narrowband spectrograms clearly reveal several distinctions between the controlled syllable and the ballistic syllable. First, in comparison to the controlled syllable, the harmonic structure of the ballistic syllable is less well defined toward the right edge of the display. This loss of definition temporally correlates with the noise present in the wideband spectrogram. Often, the harmonic structure of controlled syllables indicates a gradual lowering of fundamental frequency. This gradual decrease in fundamental is not found in ballistic syllables, which in fact are often accompanied by slight rises in pitch as aspiration sets in. As postvocalic aspiration is aerodynamically and acoustically weak (Bladon 1986), extra respiratory muscular flexion at the internal intercostal muscles (Ladefoged 1958, 1962, 1968) in the context of postvocalic aspiration may serve to increase transglottal flow, which in turn increases acoustic energy. A byproduct of this increased transglottal flow is a moderate pitch increase on the latter portion of the vowel, around the onset of aspiration.

The pitch distinction between ballistic and controlled syllables is not limited to Comaltepec Chinantec. Robbins (1961: 246–247, 1968: 24), as well as Mugele (1982: 70), report that level tones in the Lalana and Quiotepec dialects possess a slight pitch rise in ballistic syllables (considered a distinctive tonal category in Robbins 1961), although a slight pitch *fall* is occasionally heard late in the syllable. Meanwhile, controlled counterparts involve a gradual pitch fall (e.g. Mugele 1982: 74), just as is reported for the Comaltepec dialect.

In summary, as in Jalapa Mazatec, Comaltepec Chinantec vowels possess both tone and phonation contrasts, that is, they are laryngeally complex. The non-modal phonatory gesture is timed either to precede tone (see (6)) or to follow tone (see (7)), so that all contrastive information is recoverable. In this postvocalic position, however, aspiration is potentially weakened. The specifically postvocalic realisation of aspiration in ballistic syllables is, by hypothesis, due to the presence of a contrasting timing configuration involving prevocalic aspiration. Due to its postvocalic location, subglottal pressure may be actively increased, thus enhancing the salience of the laryngeal abduction. Postvocalic laryngeals, of course, are maximally distinct in their timing from prevocalic laryngeals.

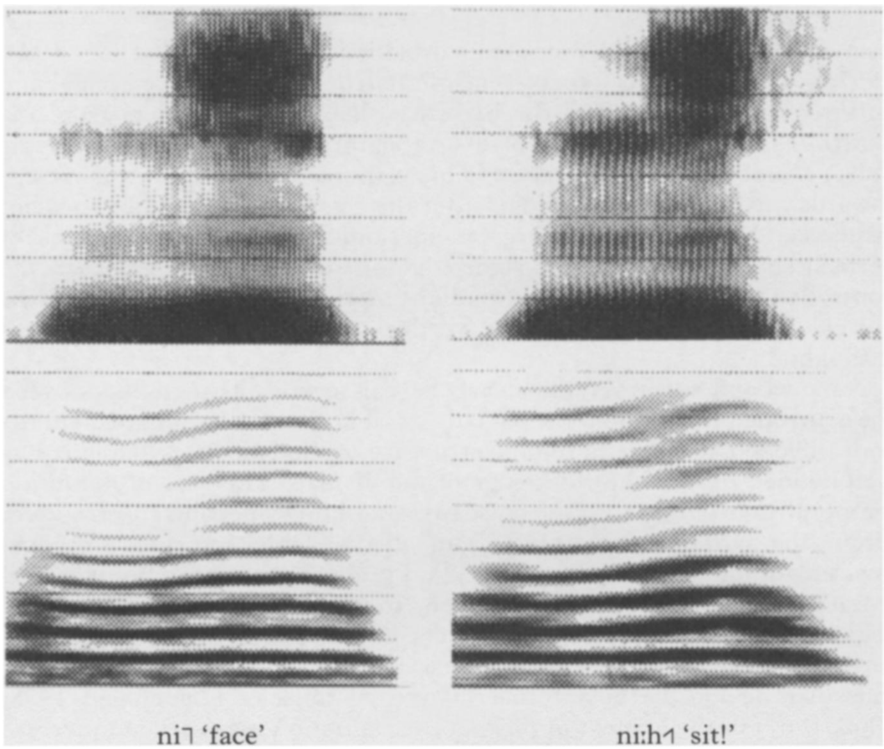


Figure 2
Spectrograms of 'controlled' (*left*) and 'ballistic' (*right*) syllables in Comaltepec Chinantec

2.3 Copala Trique

The Mixtecan language of Copala Trique implements a third realisation of laryngeally complex vowels. In addition to prevocalic and postvocalic laryngeals, Copala Trique also possesses laryngeally 'interrupted' vowels, in which the laryngeal gesture intrudes upon the central portion of the otherwise modal vowel; [VhV], [VʔV].

Copala Trique is spoken by approximately 8000 people in San Juan Copala, Oaxaca, Mexico (Grimes 1988). The word in Copala Trique normally consists of a bisyllabic root and subsyllabic inflectional material (consisting of tone, length, ablaut and/or consonantism) residing on the final syllable. Copala Trique possesses the segment inventory listed in (8) (from Hollenbach 1977; as in Jalapa Mazatec, labial obstruents occur only in loans).

(8) *San Juan Copala Trique segment inventory*

p	t				k	i ɨ u
b	d				g	e o
		tʃ	cç	tʂ		a
	s	ʃ		ʂ		
	z	ʒ				
	m	n		ɾ		
	l					
			j		w	
ʔ,h						

Copala Trique possesses eight contrastive tonal patterns, shown in (9):

(9) ˧˥ ˦˧ ˧˦ ˧ ˦˩ ˨ ˩ ˩˧

Copala Trique words are stress-final. Open final syllables are usually long, and there is a freer distribution of consonants in finals than in (unstressed) initial syllables. Examples of lengthened open finals are presented in (10), from the various dialects discussed by Longacre (1952, 1957) and Ruiz de Bravo Ahuja (1975).[3]

(10) ma˨ɾe:˧ 'red'
gu˧na:˦˧ 'to remain'
ɾa˧ʔa:˧ 'hand'
ɾi˧o:˧˦ 'trough, manger'

As in Jalapa Mazatec and Comaltepec Chinantec, laryngeals may be prevocalic, as exemplified in (11). Interestingly, in the native vocabulary, [h] has been lost in prevocalic position, evolving into a rising tone (Longacre 1957). This, of course, is a highly uncommon distribution. It is very rarely the case that a language allows postvocalic aspiration to the exclusion of prevocalic aspiration. This pattern, then, counterexemplifies my claims regarding the implicational hierarchy in the class of laryngeally complex vowels. However, prevocalic [h] is attested in Spanish loans:

(11) ʔɨ̃:˨	'nine'	liha˧˦	'sandpaper' (<Sp. *lija*)
ʔu˩ʔu:˩	'five'	hu˧lja:˧	'Julia'
ʔweʔe˩	'ice, frost'	gaw̥we:˧˦	'coffee'
ʔnɨh˩	'inside of'		

Only the laryngeals ([ʔ h]) may close syllables, and only final syllables may be closed. Examples of postvocalic laryngeals are presented in (12):

(12) waʔ˩	'the right'	ʒu˥kwah˥	'to be twisted'
jaʔ˦˧	'teeth'	jah˧	'ashes'
ni˩kaʔ˩	'short'	ɾah˧˦	'to grind'

Vowel interruption is exemplified in (13):

(13) ga˧tuʔu˧˦ 'incense-burner'
ɾi˧uhu˦˧ 'hollow reed'
na˧ʔaha˧˦ 'conversation'

While Longacre does not discuss the phonetic interaction of contour tones and interruption, I assume that each toneme manifests itself on one half of the interrupted vowel, thus /tuʔu˦˧/ = [tu˦ʔu˧]. This realisation would thus contrast with that of contour tones on non-interrupted forms, which are realised as normal tonal contours, that is, interpolating smoothly from one pitch target to the next.

Longacre (1952: 75, n. 2) presents six reasons to interpret interrupted vowels as laryngeal gestures timed to interrupt a single vocalic gesture, rather than as a bisyllabic sequence involving two distinct vowel gestures.

First, interrupted vowels are distinguished from true sequences of vowel–laryngeal–vowel in that interrupted forms do not undergo final lengthening. Thus [weʔe˧] 'house' is monosyllabic, while [we˧ʔeː˥] 'beautiful' is bisyllabic. Longacre hypothesises that since interrupted vowels manifest their vocalism on either side of the laryngeal, each half is a single mora in length. Therefore, length under stress is redundant.

Second, Longacre shows that interrupted forms lose their second vocalic component in phrasal contexts:

(14) jaha˧ *but* jah˧zi˧ŋaː˥ 'nasturtiums'
joʔo˧ *but* joʔ˧ga˧ciː˧˦ 'the past year'
nakihi˧˦ *but* nakih˧ɾu˧neː˧˦ 'bean-atole'

This elision is not reported for true [VʔV] sequences, although syllable-final [h] elides within the phrase.

Third, interrupted vowels often appear in otherwise canonical bisyllabic words, whereas true trisyllabic words are quite rare:

(15) na˨kihi˧ 'atole' ga˧uʔu˧˦ 'incense-burner'
gi˧ʔjaha˧˦ 'holy day, festival' ɾe˧kaʔa˧˦ 'stick'
na˥nɨhɨ˦˧ 'open' ɾe˧keʔe˧˦ 'splinter'
da˧kuhu˦˧ 'ascent'

Fourth, tonal sequences occurring on interrupted forms are limited to those which occur on single vowels.

Fifth, voiceless obstruents and 'fortis' nasal consonants may occur before interrupted sequences. Elsewhere, these consonants are limited to word-final syllables. If interrupted vowels are single nuclei, then a strong generalisation may be made regarding the distribution of voiceless and fortis consonants, that is, they are limited to final syllables.

Finally, interrupted vowels possess but a single vowel quality, whereas true [VʔV] sequences may possess two vowel qualities (reported in Longacre 1957, no examples given).

Based on this evidence, I concur with Longacre that interrupted syllables indeed consist of a single vowel specification interrupted by a non-modal phonatory gesture.

Let us now summarise the distribution of laryngeal gestures in Copala Trique final syllables. First, laryngeals may stand in onset position ([hV˥] (loans only), [ʔV˥]). Second, laryngeals may be postvocalic ([Vh˥], [Vʔ˥]). Finally, laryngeals may interrupt the vowel ([VhV˥], [Vʔ˥V]). Copala Trique is thus perhaps unique in allowing three distinct timing relations among phonatory gestures, tone and vowels.

2.4 Summary

(16) summarises the patterns presented in this section:

(16) *Summary of the patterns*

	prevocalic	*postvocalic*	*interrupted*
Jalapa Mazatec	hV˥, ʔV˥	—	—
Comaltepec Chinantec	hV˥, ʔV˥	Vh˥, Vʔ˥	—
Copala Trique	hV˥, ʔV˥	Vh˥, Vʔ˥	VhV˥, VʔV˥

Observe again that the presence of postvocalic laryngeals implies the presence of maximally distinct prevocalic laryngeals, and that interruption, again maximally distinct, implies the presence of the other two patterns.

3 Explaining the patterns

The observed timing patterns in Otomanguean laryngeally complex vowels represent only a minuscule fraction of those that are logically possible. Why should the attested patterns be so limited in their variability? What are the specific constraints operating here that so severely limit timing variation to the three distinct patterns attested, and moreover, how might these constraints relate to the implicational hierarchy among the observed timing patterns?

As discussed in the introductory section, sound patterns of language are often explainable in part by the aerodynamic, acoustic and articulatory properties of the speech mechanism, as well as properties of the hearing mechanism. In this section I introduce certain of these physical systems' influences on the observed patterns in an attempt to explain the observed limits on contrastive timing configurations, considering, in turn, sufficient acoustic distance (§3.1), sufficient articulatory compatibility (§3.2) and optimal auditory salience (§3.3).

3.1 Sufficient acoustic distance

It is well known that the listener does not attend primarily to the fundamental frequency during pitch perception, but instead attends to the harmonics which accompany the fundamental. Even when the fundamental frequency is masked, it may be recovered from the pulse period and the surviving harmonics. Regarding the harmonics, given an F0 between 100 and 400 Hz, the frequency range between 400 and 1000 Hz may be the most important for pitch perception, provided amplitude exceeds a minimum of 10 dB above threshold (Ritsma 1967). This region roughly corresponds to the third through the fifth harmonics, or, very approximately, the first formant region (Remez & Rubin 1984, 1993).

Now, the perception of pitch during modal phonation bears a correlative (though non-linear) relationship with the frequency at which the vocal folds vibrate. During modal phonation, as the frequency of vocal fold vibration increases, perceived pitch increases as well. As it turns out, however, a reliable and stable pitch percept which derives from glottal vibration may be disrupted during non-modal phonation. While creaky phonation may result in glottal wave quasi- or a-periodicity, breathy phonation may also disrupt the transmission of a periodic glottal vibration. In this subsection I consider certain of these interactions between pitch targets and non-modal phonation. Anticipating my conclusion, when a periodic glottal wave is either obscured or not present, the acoustic signal cannot encode a salient pitch value.

Acoustic analyses of breathy vowels indicate that the fundamental frequency is enhanced relative to the lower harmonics (Bickley 1982, Huffman 1987, Ladefoged *et al.* 1988, Cao & Maddieson 1992). While this enhanced fundamental might be argued to provide a salient pitch percept, recall that when analysing pitch the auditory system is less attuned to the fundamental frequency than to the higher harmonics and the pulse period. Recall now that harmonic structure possesses an overall weakening during breathy phonation, as well as an overall increase in noise, which in some cases has been shown to largely obscure the harmonic structure (Silverman 1994, 1996b, 1997b, Silverman *et al.* 1995). Moreover, Kirk *et al.* provide waveforms of the breathy portion of breathy vowels in Jalapa Mazatec: 'The breathy vowel is characterized by an onset of indiscernible pulses' (1993: 445). Given both the obscured harmonic structure and the indiscernible pulses that may accompany breathy phonation, pitch may not be reliably cued when implemented with breathy voice. Indeed, in Silverman (1996c) I provide experimental evidence that indicates listeners are less adept at discriminating pitch values during Jalapa Mazatec breathy phonation than during modal phonation.

Now, although it is only in an experimental setting, as opposed to a natural linguistic setting, that listeners might be called upon to determine just- and near-just-noticeable differences in pitch, it is not at all unexpected that languages should evolve to avoid less good contrasts in favour of better ones. For example, trained subjects are able to discriminate minor differences in voice onset time (VOT) that are never employed contrastively in language (see, for example, Strange 1972). However, languages typically employ VOT differences that are far less effortfully noticeable; positive VOT (aspirated), zero VOT (plain) and negative VOT (voiced). Thus, given that pitch differences seem to be better noticeable during modal phonation than during breathy phonation, languages which possess tonal contrasts and a strong breathy phonation contrast on a given vowel might avoid their simultaneous implementation, so that both values are more likely recoverable.

Creaky vowels also possess a potentially unanalysable harmonic structure. This is due to the aperiodic and/or unstable glottal vibration that results from vocal fold constriction (Kirk *et al.* 1993, Ladefoged & Maddieson 1996). For example, Kirk *et al.* (1993) provide waveforms indicating glottal pulse patterns in creaky *vs.* modal vowels in Jalapa Mazatec: 'creaky vowels have speech jitter (irregularly spaced pulses)' (1993: 445). Now note that Rosenberg (1965) finds that when a pulse period varies, or jitters, by more than 10%, an otherwise just-noticeable pitch difference within the 300–1000 Hz range is rendered indiscriminable. It might therefore be predicted that languages which possess tone on heavily creaky vowels – such as Jalapa Mazatec – might sequence their tonal and non-modal phonatory gestures, so that both contrastive tone and contrastive phonation are recoverable. This prediction, of course, is correct for the languages reported on thus far.

3.2 Sufficient articulatory compatibility

Lindblom (1983) investigates the principle of least effort in biological motor systems such as speech production, and its relevance to the study of linguistic sound patterns. Very briefly, a particular motor goal, or combination of motor goals in sequence or in simultaneity, seems to abide by a 'sufficient compatibility' condition: antagonistic articulatory postures tend to be avoided. That is, extreme displacements, either from a neutral position (quantified in Lindblom's model), or from a neighbouring target posture, are dispreferred.

While this avoidance of articulatory antagonism, again, may involve gestures in sequence or in simultaneity, Lindblom's illustrative examples – coarticulation, vowel reduction and aspects of syllable structure – all involve sequential articulatory targets. The present data, although presented as incompatible if implemented simultaneously (those needed to achieve the broadband noise which characterises breathiness or those needed to induce creak, and those needed to reach a particular pitch target) are nonetheless subject to this Lindblomian interpretation. Indeed, it is, by hypothesis, these gestures' very incompatibility that may be contributing to their *de facto* sequencing. In this section then, I turn my attention to these articulatory considerations in the context of Lindblom's 'sufficient compatibility' condition. I consider pitch production with modal phonation and pitch production with breathy phonation.

Ohala (1978) summarises the interacting muscular, articulatory and aerodynamic factors involved in pitch production. Briefly, pitch is controlled primarily by tensing (stretching) and laxing of the vocal folds via the cricothyroid muscle, in combination with the thyroarytenoid muscle. Provided that a steady transglottal airstream is maintained, tensing the vocal folds increases rate of vocal fold vibration (hence increasing pitch), while laxing the folds decreases rate of vibration (hence reducing pitch).

However, there are additional ways in which pitch may be more moderately influenced. First, subglottal pressure affects pitch: increases in subglottal pressure through increased respiratory muscle activity (the internal intercostal muscles) have been shown to have moderate effects on rate of vocal fold vibration. All else held constant, the higher the subglottal pressure, the higher the rate of vocal fold vibration; as transglottal flow is increased, vocal fold vibration increases as well. However, the subsequent increase in flow may lead to a *lowering* of subglottal pressure, perhaps lowering pitch as well. Additionally, larynx height correlates with pitch: raising the larynx is associated with pitch increases, while lowering the larynx is associated with pitch falls. However, lowering the larynx shrinks the subglottal cavity, thus increasing subglottal pressure, potentially *raising* pitch. Glottal aperture may also affect pitch, interacting in complex ways with airflow, subglottal pressure and also supraglottal stricture. All else being equal, a more open glottis may result in faster transglottal airflow, hence higher pitch. However, again, the consequent reduction in subglottal pressure may lead to a pitch fall.

Now consider the interaction of tone and non-modal phonation. While the primary articulatory correlates to pitch and phonation type are in theory independently manipulable, these potentially distinct laryngeal configurations may make conflicting demands on their respective accompanying mechanisms. Indeed, insufficient compatibility becomes especially problematic when considering that respiratory musculature and laryngeal musculature may be differently postured for a given phonation target in comparison to a given pitch target (see, for example, Hirano 1981, Titze 1994 and, for discussion of specific languages, Ladefoged 1962, Fischer-Jørgensen 1970 and Thongkum 1991).

Consider, for example, breathy voice and tone. Breathy voice is implemented primarily through abducting the vocal folds via the posterior cricoarytenoid muscles. Subglottal pressure increases may also enhance breathiness, as increased airflow increases acoustic energy (Fischer-Jørgensen 1970). Finally, there may be a correlation between breathy phonation and a moderate lowering of the larynx (Henderson 1952, Gregerson 1976, Hirano 1981, Thongkum 1991).

Now, Hombert (1978: 91) briefly discusses certain potential interactions between breathy phonation and pitch: 'although the rate of airflow is high upon release of breathy-voiced consonants [i.e. breathiness on a following vowel; DS], the vocal cords are not closely adducted, and thus the Bernoulli force should be weak'. Regarding the Bernoulli force, as air passes through the glottis its velocity is increased. This increased velocity reduces the pressure perpendicular to the flow of air, that is, against the vocal folds themselves. Consequently, the folds are rapidly drawn together until subglottal pressure increases sufficiently to blow them apart once again. During voicing, this cycle of events repeats in rapid succession. However, when the folds are spread farther apart, a greater force is required to achieve the same cyclic rate. Thus, all else held constant, as breathy phonation involves vocal fold spreading, breathy phonation may involve a lower rate of vocal fold vibration, hence producing a lower pitch than modal phonation.

Hombert also notes that several researchers report a less active response of the laryngeal adductor muscles (the lateral cricoarytenoids) during breathy phonation. These muscles serve to bring the vocal folds together, but are also involved in larynx lowering, which, as noted, is associated with pitch lowering. Indeed, it is possible that breathy and/or voiceless phonation has evolved (or is evolving) into a low-tone contrast in a number of languages, including Punjabi, Vietnamese, Middle Chinese (Hombert 1978 and references therein), Huave, Ojitlán Chinantec and Usila Chinantec (Silverman 1997b). Alternatively, if subglottal pressure is compensatorily raised in the context of breathy and/or voiceless phonation, a pitch *rise* may be the diachronic result. Possible cases along these lines include Jeh (Gradin 1996) and Quiotepec Chinantec (Robbins 1968). Thus pitch might be either decreased or increased during breathy phonation, depending upon the positioning of the involved mechanisms. In other words, pitch-based and phonation-based gestures may make conflicting demands on the involved musculature. Consequently, certain combinations of pitch and phonation are likely to be avoided in linguistic sound patterns.

Similar arguments may be made with respect to pitch targets and creaky phonation, the conclusions being the same: given the potential for insufficient articulatory compatibility here, it should not be viewed as surprising that languages tend to avoid laryngeal complexity. In those rare cases where laryngeal complexity is present, implementing the involved laryngeal gestures in sequence, rather than in simultaneity, may serve to mitigate some of the articulatory complications which might otherwise arise.

3.3 Optimal auditory salience

Upon investigating auditory nerve response to energy deriving from the acoustic source, it turns out that the exact same signal may be either more, or less recoverable, depending upon the acoustic context in which it is embedded: the auditory nerve is more sensitive to a particular acoustic event in certain contexts than it is in other contexts. Bladon (1986) proposes some of the major principles of auditory phonetics. For present purposes, his principles in (17) (1986: 5) are most relevant:

(17) a. *On/off response asymmetry*: spectral changes whose response in the auditory nerve is predominantly an onset of firing are much more perceptually salient than those producing an offset (Tyler *et al.* 1982).

b. *Short-term adaptation*: after a rapid onset of auditory nerve discharge at a particular frequency, there is a decay to a moderate level of discharge, even though the same speech sound is continuing to be produced (Delgutte 1982).

Thus, auditory nerve firing does not seem to respond exclusively to absolute levels of acoustic energy. Instead, the relevant fibres along the basilar membrane respond in part to local changes, especially increases in acoustic energy. Consequently, the same acoustic cue may evoke a greater response at the relevant fibres when in one acoustic context than when in another.

In (18) is a schematic of the acoustic energy and corresponding auditory nerve activity over time for [ha˥], at a particular characteristic frequency (CF). As the acoustic signal at the characteristic frequency suddenly increases in energy at the onset of modal voicing, there is a heightened response at the corresponding fibres along the basilar membrane.

(18) *Gross schematic of characteristics of* [ha˥] *sequences*

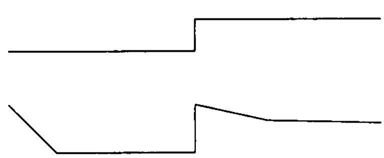
articulatory

supralaryngeal: vowel

laryngeal: abduction — approximation tensing

acoustic signal (amplitude)

auditory nerve response at CF



percept h — a˥

For [ah˥], auditory nerve response diminishes at the transition from voicing to voicelessness, as in (19).

(19) *Gross schematic of characteristics of* [ah˥] *sequences*

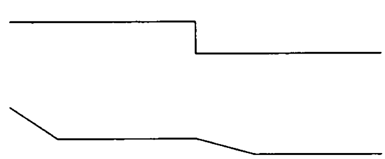
articulatory

supralaryngeal: vowel

laryngeal: approximation tensing — abduction

acoustic signal (amplitude)

auditory nerve response at CF



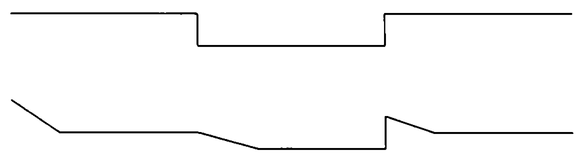
percept a˥ — h

For interrupted forms (schematised in (20)), energy diminishes in the transition from modal voicing to voicelessness (or breathiness). As stated, if instead [h] stands in isolation prevocalically, or especially if aspiration is realised at prevocalic stop release, this sudden onset of energy helps to better cue the noise associated with glottal spreading.

(20) *Gross schematic of characteristics of* [aha˥] *sequences*
articulatory
supralaryngeal: vowel
laryngeal: approximation tensing — abduction — approximation tensing

acoustic signal (amplitude)

auditory nerve response at CF



percept a — h — a˥

In summary, different timings of a given set of articulatory gestures may produce a stronger or weaker neurochemical response in the inner ear. In the case of laryngeally complex vowels (and potentially elsewhere as well) auditorily better timing patterns are more prevalent cross-linguistically than are auditorily worse timing patterns.

3.4 Summary

In summary, aerodynamic, acoustic and articulatory properties of the speech mechanism, as well as properties of the hearing mechanism, may account for aspects of the observed timing patterns in Otomanguean laryngeally complex vowels, in part explaining why only a tiny fraction of the possible timing patterns are in fact attested.

4 Superficial exceptions

If Lindblom's notion of sufficient articulatory compatibility is indeed an independent force acting to constrain the patterning of linguistic sound systems, then certain predictions might be made about alternative realisations of laryngeally complex vowels. Specifically, if sufficient articulatory compatibility is somehow achieved by reducing the demands on the involved mechanisms, then we might yet observe the fully simultaneous implementation of tone and non-modal phonation. Consider, then, Mpi and Tamang in this light.

4.1 Mpi

Mpi, a Tibeto-Burman language (Ladefoged & Maddieson 1996) possesses six contrastive tones, in addition to a phonation contrast involving laryngealisation. Any tonal pattern may occur with modal phonation or laryngealisation, and thus Mpi qualifies as a laryngeally complex language. In (21) are some examples:

(21)

Tone	*Modal*		*Laryngealised*	
low rising	si˩˧	'to be putrid'	sḭ˩˧	'to be dried up'
low level	si˩	'blood'	sḭ˩	'seven'
mid rising	si˧˥	'to roll (rope)'	sḭ˧˥	'to smoke'
mid level	si˧	(colour)	sḭ˧	(classifier)
high falling	si˥˧	'to die'	sḭ˥˧	(name)
high level	si˥	'four'	sḭ˥	(name)

Moreover, Mpi laryngealisation persists throughout the duration of the vowel, thus constituting an exception to the claim that non-modal phonation and tone are always sequenced in laryngeally complex vowels. How then might I account for this patterning? The answer lies in the degree to which Mpi laryngealised vowels are creaked. Ladefoged & Maddieson (1996) compare laryngealisation in Mpi to that in Jalapa Mazatec. They report that Mpi laryngealised vowels 'definitely have a less constricted glottis' (1996: 16). Recall that Jalapa Mazatec glottalised vowels are laryngeally complex as well, and indeed limit their non-modal phonation to the vowel's first portion.

In Fig. 3 are wideband and narrowband spectrograms of a pair of words which minimally contrast for creakiness, taken from the archives of the UCLA phonetics laboratory (originally recorded by Jimmy Harris in May 1976).

Observe in particular the fairly steady glottal pulse pattern of the creaked form (in the wideband spectrogram), as well as its by and large clear and steady harmonic structure (in the narrowband spectrogram). With their lesser degree of laryngeal constriction, Mpi vowels may simultaneously implement their tonal and phonatory features, without the risk of non-recoverability. That is, sufficient articulatory compatibility is achieved perhaps at the partial expense of acoustic discriminability.

4.2 Tamang

Tamang is a Tibeto-Burman language spoken by approximately 664,000 people in Nepal and Sikkim, India (Grimes 1988). It is traditionally characterised as possessing a register system of the Mon-Khmer variety, that is, involving pitch and voice quality distinctions. The four registers of Tamang consist of four pitch patterns and two phonation types. These are presented in (22), from Mazaudon (1973):

(22) *Tamang registers*

Clear	*Breathy*
High	
Mid-High	
	Mid-Low
	Low

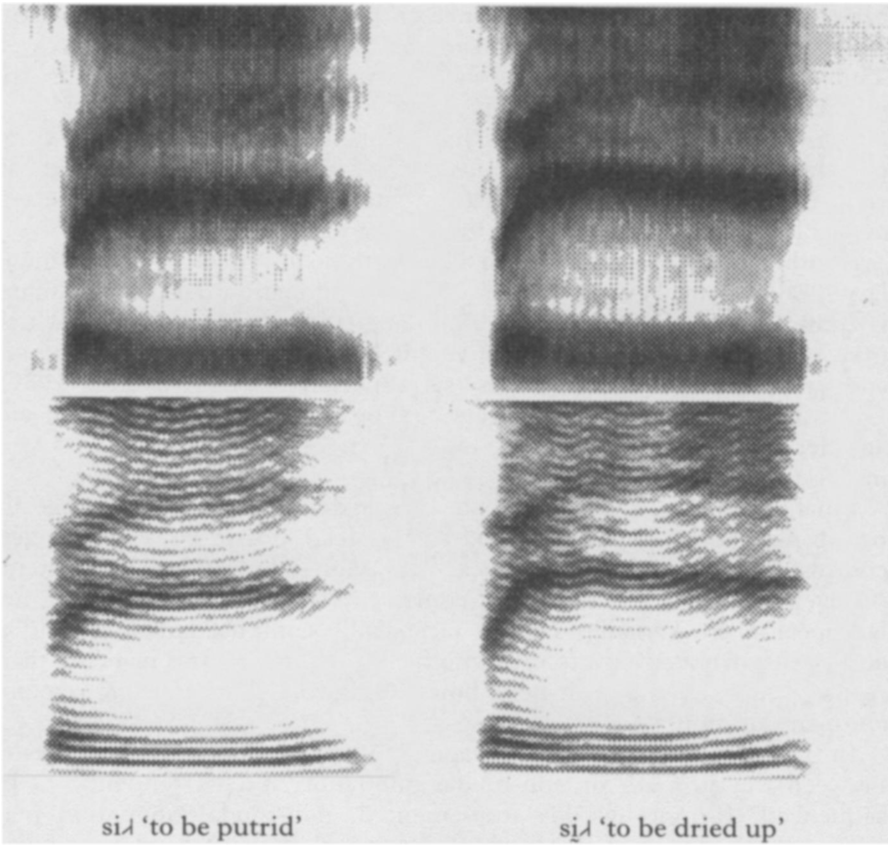


Figure 3
Wideband and narrowband spectrograms for Mpi forms minimally contrasting for creakiness

Regarding these four registers, Maddieson (1984: 132) notes that the system may be treated as one in which tone and phonation cross-classify. A reorganisation along these lines is presented in (23):

(23) *Tamang tone and phonation*

	V	V̤
H	high pitch, modal phonation	mid-low pitch, breathy phonation
L	mid-high pitch, clear phonation	low pitch, breathy phonation

As tone and phonation may be characterised as cross-classifying here, Tamang may be considered a laryngeally complex language. Yet researchers say nothing of a part-modal/part-non-modal realisation of the breathy registers. Given their silence on the subject, I assume that non-modal phonation probably persists for the duration of its associated vowel. Thus Tamang is probably an exception to my claims regarding the phonetic patterning of laryngeally complex vowels.

It is possible that Tamang is like Mpi, in that breathiness is comparatively light. But as no phonetic descriptions of Tamang register detail such information, and as no instrumental analyses of these vowels is available, I simply do not know the degree of breathiness here.

Another possibility centres on the relative simplicity of the Tamang laryngeal system. With only one tonal and phonation contrast, Tamang is quite distinct from Otomanguean languages such as Trique, which may possess up to five contrastive pitch levels, many pitch contours and up to three phonation types (see especially Longacre 1952, 1957). Given this simplicity, pitch targets may be sufficiently distant from one another to yet emerge distinct, even when breathiness is fully superimposed.

Finally, it is also conceivable that Tamang does not possess a laryngeally based phonation contrast at all, but instead possesses a pharyngeal contrast, as is sometimes the case in Mon-Khmer register systems (Gregerson 1976). If the relevant contrast in Tamang is pharyngeal, not laryngeal, then Tamang is not a laryngeally complex language, and is consequently not subject to the hypothesised constraints influencing their realisation. As it stands for now, however, the status of Tamang remains an open question.

In laryngeally complex languages, there is apparently a trade-off between the strength of non-modal phonation and its tendency to be sequenced. If more weakly implemented, non-modal phonation may persist throughout the duration of the vowel without rendering unrecoverable concomitant tone; sufficient articulatory compatibility is achieved. Mpi, for example, possesses tone and phonation contrasts which cross-classify. However, the relatively light implementation of non-modal phonation does not render contrastive pitch unrecoverable. Consequently, contrastive phonatory and tonal gestures here may be implemented simultaneously.

However, non-modal phonation may alternatively be strongly implemented, thus producing a greater acoustic distinction between modal and non-modal phonation; tonal contrasts are cued during modal phonation. This is the Otomanguean pattern. Here, gestural sequencing is observed, so that strongly implemented non-modal phonation and robust tonal contrasts are saliently cued to the listener.

5 Towards a quantitative model of timing contrasts in laryngeally complex vowels

A common criticism of so-called 'functional' explanations in phonology is their lack of formal rigour. But far from being an inherent shortcoming of an explanatory phonology, to the extent that such criticisms are valid, this neglect of formalism is perhaps best viewed as a mere consequence of intellectual history: so-called functionalists typically do not concern themselves with formal modelling, just as so-called formalists typically do not concern themselves with functional explanation. Herein, while a strict quantitative model of contrastive timing configurations is not formulated, the work of Liljencrants & Lindblom (1972) may be viewed as a blueprint from which a reliable model might ultimately be built. The perceptual space – in their case, the vowel space – is the product of particular articulatory – tongue, lip, jaw and larynx – postures, whose coordination is constrained by natural limitations on vocal tract configuration. From these modelled positions, a set of corresponding acoustic values may be derived. For vowels, this set includes the first three formants, which are most important in determining a given vowel quality. Formant values are plotted in a three-dimensional acoustic space (eventually collapsed into two), with each dimension corresponding to one of the three formant values. Now, with a sufficiently rich data set fed into a computer program written to maximise the linear distance between the points on the multi-dimensional scale (eventually converted from Hertz, an acoustic scale, to Mels, an auditory scale), maximally distinct vowel qualities may be calculated. As the number of points increases, naturally, linear distances between points are diminished. As Liljencrants & Lindblom report, 'the formant patterns which [the model] can generate are in close agreement with those of basic vowel qualities' (1972: 855), and sufficient variation in the search procedure exists to model the system-to-system variation found in actual inventories.

The present data set, of course, does not involve the shades and colours of acoustic/auditory quality which correlate with static formant values. Rather, we are here dealing with more grossly defined spectral characteristics which correlate with dynamic changes in phonation over time (see also Lindblom *et al.* 1984, who investigate CV transitions). Given this intrinsic dynamism, a proper model here requires overt reference to the temporal coordination – overlap and/or sequencing – of the articulatory gestures involved. Moreover, given this added dynamic dimension, we must additionally consider particular properties of the peripheral auditory system, such as those discussed in §3.3, for as shown, the exact same acoustic signal may trigger a more or less robust neural response, depending on the acoustic context in which it is embedded.

What is especially relevant in the Otomanguean data is that attested auditorily sub-optimal timing patterns are normally maximally distinct from preferred timing patterns, just as might be predicted in the Liljencrants & Lindblom model. Thus sub-optimal postvocalic laryngeals are maximally distinct from optimal prevocalic laryngeals, and, again, interrupted forms, being equidistant from prevocalic and postvocalic laryngeals, are maximally distinct once again. Now, although the Liljencrants & Lindblom model is not employed here, it should nonetheless be kept in mind as the overriding inspiration for the present approach to the contrasts under discussion. It should also be recalled, however, that the present data set differs in two important ways from that investigated by Lindblom *et al.* First, where the Liljencrants & Lindblom model has been applied to paradigmatic contrasts – primarily place of articulation in vowels, i.e. vowel quality – the present data contrast syntagmatically, in terms of timing contrasts among phonation types. Second, where the Liljencrants & Lindblom model has been applied to within-class contrasts that are primarily cued by relatively subtle changes in formant values, the present data involve between-class contrasts, that is, those involving gross spectral differences characteristic of differing phonation types.

In this section, then, I consider the various timing patterns observed in the systems investigated in light of the basic parameters which a computational model must be capable of quantifying.

First, largely for the auditory reasons presented in §3.3, prevocalic non-modal phonation followed by modal phonation is the optimal timing pattern to cue vocalism, voicelessness or glottal closure, and tone ([hV˥], [ʔV˥]): all these contrastive values are optimally recoverable from the speech signal if timed in this fashion. Therefore, if we take seriously the claim that the relative recoverability of contrastive cues has a direct and often overriding influence on linguistic sound patterns, we make the following prediction: if a language has laryngeally complex vowels, then it likely possesses this maximally recoverable timing pattern. This prediction, of course, is borne out by the languages investigated in §2. Jalapa Mazatec, Comaltepec Chinantec and Copala Trique, which all possess laryngeally complex vowels, all display this auditorily optimal timing pattern, with the notable exception that native prevocalic [h] has evolved into rising tone in Copala Trique.

Now consider a second prediction, directly inspired by the Liljencrants & Lindblom approach to contrast inventories: if a language possesses a timing contrast involving laryngeal complexity – that is, two distinctive timing patterns – then they are maximally distinct from one another, in order to ensure the contrast's salience. This prediction, again, is supported by the Otomanguean data. Two of the investigated languages, Comaltepec Chinantec and Copala Trique, possess (at least) two timing patterns within the laryngeally complex class. The second value, that is, the timing pattern that contrasts with the optimal timing pattern, is maximally distinct from the optimal pattern: where the optimal timing pattern involves prevocalic non-modality ([hV˥], [ʔV˥]), contrastive patterns in Comaltepec Chinantec and Copala Trique involve postvocalic non-modality ([Vh˥], [Vʔ˥]).

Let us now go one step further, and predict what timing pattern would contrast with the two maximally distinct patterns already considered. Well, if the optimal pattern involves prevocalic non-modality, and the maximally distinct pattern involves postvocalic non-modality, then an additional contrastive pattern should time non-modal phonation equidistant from these two extremes. That is, non-modal phonation should interrupt the vowel ([VhV˥], [VʔV˥]). And this is exactly what is found in Copala Trique.

Note that, quite naturally, as the number of distinctive patterns increases for a given system, the less perceptually distinct the contrasts become. We might thus predict a sharp fall-off rate in the number of contrastive patterns allowed, which would most likely correlate with a sharp fall-off rate in the number of languages which allow increasingly crowded systems along a given parameter, for the temporal resolution capabilities of the hearing mechanism quickly become taxed. As noted in the literature on universal implicational typology (for example, Maddieson 1984, and more formally, though obviously less exhaustively, in the Liljencrants & Lindblom model), this prediction, again, appears quite correct.

An adequate computational model must consequently be able to calculate and predict what sorts of timing patterns will be present in phonological systems. This requires then, not only standard notions of contrastiveness involving the raw spectral quality of the various combinations of sounds, but also, quantified values of temporal contrasts, and also, a measure of the consequent effect that these timing patterns have on the auditory nerve.

Finally, consider Mpi-like systems, in which the involved articulatory gestures are sufficiently lightly implemented that they may be implemented simultaneously. A promising method of quantifying the necessary changes in articulatory configuration which result in acoustic discriminability here might involve modelling stepwise changes in articulation which produce speech jitter, and experimentally investigating the discriminability limen, *à la* Rosenberg (1965).

Such a computational model would move beyond a characterisation of mere superficial patterning towards a more satisfying characterisation of underlying motivation for this aspect of linguistic sound structure.

6 Conclusion

In laryngeally complex languages, the sequencing of contrastive laryngeal configurations is often observed, so that all contrastive information is rendered recoverable by the listener. Sub-optimal patterns are typically maximally distinct in their timing from optimal patterns, and as more timing contrasts are present, temporal equidistance is by and large maintained, until the approaching point where resolution is jeopardised. Moreover, it is nearly always the case that the presence of a sub-optimal timing pattern implies the presence of an optimal timing pattern, where optimality correlates, at least in part, with the degree of auditory nerve response to the involved cues.

If weakly implemented however, non-modal phonation may persist throughout the duration of the vowel without rendering unrecoverable concomitant tone.

NOTES

* I extend my sincere thanks to two anonymous *Phonology* referees. This research was supported by NIH training grant T32 DC-00008.

[1] In some languages, both tonal contrasts and phonation contrasts are present, but the two do not cross-classify. Put another way, such languages do not possess tonal contrasts on vowels that bear non-modal phonation. However, a full array of tonal contrasts exists on modally phonated vowels. For example, White Hmong (Lyman 1974, Smalley 1976, Huffman 1987, Ratliff 1992) possesses five tones that may be realised on plain vowels. Vowels bearing non-modal phonation – creakiness or breathiness – never contrast in tone. These phonation contrasts are traditionally labelled 'creaky tone' and 'breathy tone', respectively. Corroborating the hypothesis that the so-called breathy tone is in fact a phonation contrast, not a tonal contrast, Ratliff (1992), in the context of a lengthy discussion of White Hmong fossilised tonal morphology, reports that the breathy tone bears different pitch patterns for male *vs.* female speakers. For male speakers, the breathy tone is implemented as a low, whispered pitch fall: [V̤˧˩]. For female speakers, the breathy tone is implemented as a high, whispered fall: [V̤˥˧]. 'My perception of this difference leads me to believe that the phonation contrast is the primary phonetic cue, fundamental frequency change ('contour') the secondary phonetic cue, and fundamental frequency itself ('pitch') only the tertiary phonetic cue of this tone' (Ratliff 1992: 12). That is, the relative pitch of the breathy tone is not crucial to the lexical contrast, as it varies with respect to other pitch patterns. Instead, the reliable and constant cue to the contrast is its breathy quality.

[2] Anderson *et al.* do not include the glides /j w/ in their consonant inventory, instead considering these non-syllabic /i u/, respectively.

[3] Longacre does not actually provide phonetic transcriptions which indicate length, but reports that 'non-phonemic stress and non-phonemic lengthening of unchecked vowels occur regularly on the final syllable' (1957: 15).

REFERENCES

Anderson, J. L. (1989). *Comaltepec Chinantec syntax*. Dallas: Summer Institute of Linguistics.

Anderson, J. L., I. H. Martinez & W. Pace (1990). Comaltepec Chinantec tone. In Merrifield & Rensch (1990). 3–20.

Bickley, C. (1982). Acoustic analysis and perception of breathy vowels. *MIT Speech Communication Working Papers* **1**. 73–83.

Bladon, A. (1986). Phonetics for hearers. In G. McGregor (ed.) *Language for hearers*. Oxford: Pergamon Press. 1–24.

Bull, B. E. (1983). Constraint-governed rule application: principles underlying the application of phonological rules in San Jerónimo Mazatec. *Linguistics* **21**. 431–486.

Cao, J. & I. Maddieson (1992). An exploration of phonation types in Wu dialects of Chinese. *JPh* **20**. 77–92.

Delgutte, B. (1982). Some correlates of phonetic distinctions at the level of the auditory nerve. In R. Carlson & B. Granström (eds.) *The representation of speech in the peripheral auditory system*. Amsterdam: Elsevier Biomedical. 131–150.

Dhall, G. B. (1966). *Aspiration in Oriya*. Bhubaneswar, Orissa, India: Utkal University.

Dixit, R. P. (1989). Glottal gestures in Hindi. *JPh* **17**. 213–237.

Fischer-Jørgensen, E. (1970). Phonetic analyses of breathy (murmured) vowels in Gujarati. *Indian Linguistics* **28**. 71–140.

Flemming, E. (1995). *Auditory representations in phonology*. PhD dissertation, UCLA.

Fromkin, V. (ed.) (1978). *Tone: a linguistic survey*. New York: Academic Press.

Fukui, N. & H. Hirose (1983). Laryngeal adjustments in Danish voiceless obstruent production. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* **17**. 61–71.

Gradin, D. (1966). Consonantal tone in Jeh phonemics. *Mon-Khmer Studies* **2**. 41–53.

Gregerson, K. J. (1976). Tongue-root and register in Mon-Khmer. In P. N. Jenner, L. C. Thompson & S. Starosta (eds.) *Austroasiatic studies*. Part 1. Hawaii: University Press of Hawaii. 323–369.

Grimes, B. F. (1988). *Ethnologue*. 11th edn. Dallas: Summer Institute of Linguistics.

Henderson, E. J. A. (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies* **14**. 149–174.

Hirano, M. (1981). *Clinical examination of voice*. New York: Springer.

Hollenbach, B. E. (1977). Phonetic vs. phonemic correspondence in two Trique dialects. In W. R. Merrifield (ed.) *Studies in Otomanguean phonology*. Dallas: Summer Institute of Linguistics. 35–67.

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In Fromkin (1978). 77–111.

Huffman, M. K. (1987). Measures of phonation types in Hmong. *JASA* **81**. 495–504.

Jun, J. (1995). *A constraint-based analysis of place assimilation typology*. PhD dissertation, UCLA.

Kim, D.-W., H. Hirose & S. Nimi (1992). A fiberoptic study of laryngeal gestures for Korean intervocalic stops. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* **26**. 13–21.

Kingston, J. (1985). *The phonetics and phonology of the timing of oral and glottal events*. PhD dissertation, UCB.

Kingston, J. (1990). Articulatory binding. In J. Kingston & M. E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 406–434.

Kirk, P. L. (1966). *Proto-Mazatec phonology*. PhD dissertation, University of Washington.

Kirk, P. L. (1969). The development of Jalapa Mazatec voiced aspirates. Paper presented at the 43rd Annual Meeting of the Linguistic Society of America, San Francisco.

Kirk, P. L., J. Ladefoged & P. Ladefoged (1991). Acoustic analysis of Jalapa Mazatec complex vowels. Paper presented at the 47th International Congress of Americanists, New Orleans.

Kirk, P. L., J. Ladefoged & P. Ladefoged (1993). Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec. In A. Mattina & T. Montler (eds.) *American Indian linguistics and ethnography in honor of Lawrence C. Thompson*. Ann Arbor: University of Michigan. 435–450.

Ladefoged, P. (1958). Syllables and stress. *Miscellanea Phonetica* **3**. 1–14.

Ladefoged, P. (1962). Sub-glottal activity during speech. In A. Sovijärvi & P. Aalto (eds.) *Proceedings of the 4th International Congress of Phonetic Sciences*. The Hague: Mouton. 73–91.

Ladefoged, P. (1968). Linguistic aspects of respiratory phenomena. *Annals of the New York Academy of Sciences* **155**. 41–151.

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

Ladefoged, P., I. Maddieson & M. Jackson (1988). Investigating phonation types in different languages. In O. Fujimura (ed.) *Voice production, mechanisms, and functions*. New York: Raven. 297–317.

Liljencrants, J. & B. Lindblom (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Lg* **48**. 839–862.

Lindblom, B. (1983). Economy of speech gestures. In MacNeilage (1983). 217–245.

Lindblom, B. (1984). Can the models of evolutionary biology be applied to phonetic problems? In M. P. R. van den Broeke & A. Cohen (eds.) *Proceedings of the 10th International Congress of Phonetic Sciences*. Dordrecht: Foris. 67–81.

Lindblom, B. (1986). Phonetic universal in vowel systems. In J. J. Ohala & J. J. Jaeger (eds.) *Experimental phonology*. Orlando, Fl.: Academic Press. 13–44.

Lindblom, B. (1989). Phonetic invariance and the adaptive nature of speech. In B. A. G. Elsendoorn & H. Bouma (eds.) *Working models of human perception*. New York: Academic Press. 139–173.

Lindblom, B. (1990a). Explaining phonetic variation: a sketch of the H & H theory. In W. J. Hardcastle & A. Marchal (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer. 403–439.

Lindblom, B. (1990b). Phonetic content in phonology. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)*. 101–118.

Lindblom, B., P. MacNeilage & M. Studdert-Kennedy (1984). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie & Ö. Dahl (eds.) *Explanations for language universals*. Berlin: de Gruyter. 181–203.

Löfqvist, A. (1980). Interarticulator programming in stop production. *JPh* **8**. 475–490.

Longacre, R. (1952). Five phonemic pitch levels in Trique. *Acta Linguistica* **7**. 62–82.

Longacre, R. (1957). *Proto-Mixtecan*. Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.

Lyman, T. (1974). *Dictionary of Hmong Njua*. The Hague: Mouton.

MacNeilage, P. F. (ed.) (1983). *The production of speech*. New York: Springer.

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

Mazaudon, M. (1973). *Phonologie tamang*. Paris: SELAF.

Merrifield, W. R. (1963). Palantla Chinantec syllable types. *Anthropological Linguistics* **5**. 1–16.

Merrifield, W. R. & C. R. Rensch (eds.) (1990). *Syllables, tone, and verb paradigms*. Dallas: Summer Institute of Linguistics.

Mugele, R. L. (1982). *Tone and ballistic syllable in Lalana Chinantec*. PhD dissertation, University of Texas, Austin.

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

Ohala, J. J. (1978). Production of tone. In Fromkin (1978). 5–39.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In MacNeilage (1983). 189–216.

Ohala, J. J. (1992). What's cognitive, what's not, in sound change. In G. Kellerman & D. Morrisey (eds.) *Diachrony within synchrony: language and cognition*. Frankfurt: Peter Lang. 309–355.

Pace, W. J. (1990). Comaltepec Chinantec verb inflection. In Merrifield & Rensch (1990). 21–62.

Pike, K. L. & E. V. Pike (1947). Immediate constituents of Mazateco syllables. *IJAL* **13**. 78–91.

Ratliff, M. (1992). *Meaningful tone: a study of tonal morphology in compounds, form classes, and expressive phrases in White Hmong*. Northern Illinois University: Center for Southeast Asian Studies.

Remez, R. E. & P. E. Rubin (1984). On the perception of intonation from sinusoidal sentences. *Perceptual Psychophysics* **48**. 313–325.

Remez, R. E. & P. E. Rubin (1993). On the intonation of sinusoidal sentences: contour pitch and height. *JASA* **94**. 1983–1988.

Rensch, C. R. (1978). Ballistic and controlled syllables in Otomanguean languages. In A. Bell & J. B. Hooper (eds.) *Syllables and segments*. Amsterdam: North-Holland. 85–92.

Ritsma, R. J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *JASA* **42**. 1191–1198.

Robbins, F. E. (1961). Quiotepec Chinantec syllable patterning. *IJAL* **27**. 237–250.
Robbins, F. E. (1968). *Quiotepec Chinantec grammar*. Mexico: Museo Nacional de Antropología.
Rosenberg, A. E. (1965). Pitch discrimination of jittered pulse trains. *JASA* **39**. 920–928.
Ruiz de Bravo Ahuja, G. (1975). *Trique–San Juan Copala, Oaxaca*. Mexico: Instituto de Investigación e Integración Social del Estado de Oaxaca.
Silverman, D. (1994). A case study in acoustic transparency: [spread glottis] and tone in Chinantec. *NELS* **24**. 559–572.
Silverman, D. (1996a). Phonology at the interface of morphology and phonetics: root-final laryngeals in Chong, Korean, and Sanskrit. *Journal of East Asian Linguistics* **5**. 301–322.
Silverman, D. (1996b). Voiceless nasals in auditory phonology. *BLS* **22**.
Silverman, D. (1996c). Pitch discrimination during breathy phonation. Poster paper presented at the 3rd Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu.
Silverman, D. (1997a). Tone sandhi in Comaltepec Chinantec. *Lg* **73**. 473–492.
Silverman, D. (1997b). *Phasing and recoverability*. New York: Garland.
Silverman, D., B. Blankenship, P. Kirk & P. Ladefoged (1995). Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* **37**. 70–88.
Smalley, W. A. (1976). The problems of consonants and tone: Hmong (Meo, Miao). In W. A. Smalley (ed.) *Phonemes and orthography: language planning in ten minority languages of Thailand*. Canberra: Department of Linguistics, Australian National University. 85–123.
Smith, K. D. (1968). Laryngealization and de-laryngealization in Sedang phonemics. *Linguistics* **38**. 52–69.
Smith, K. D. (1979). Sedang grammar. Canberra: Australian National University.
Steriade, D. (1992). Segments, contours and clusters. Ms, UCLA.
Steriade, D. (1994). Complex onsets as single segments: the Mazateco pattern. In J. Cole & C. Kisseberth (eds.) *Perspectives in phonology*. Stanford: CSLI. 203–292.
Steriade, D. (1995). Neutralization and the expression of contrast. Ms, UCLA.
Strange, W. (1972). *The effect of training on the perception of synthetic speech sounds: voice onset time*. PhD dissertation, University of Minnesota.
Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs: Prentice-Hall.
Thongkum, T. L. (1991). An instrumental study of Chong registers. In J. Davidson (ed.) *Essays on Mon-Khmer linguistics in honour of H. L. Shorto*. London: School of Oriental and African Studies. 141–160.
Tyler, R. S., Q. Summerfield, E. J. Wood & M. A. Fernandez (1982). Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. *JASA* **72**. 740–752.
Yoshioka, H., A. Löfqvist, H. Hirose & R. Collier (1986). How voiceless sound sequences are organized in terms of glottal opening gestures. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* **20**. 55–67.