



**HAL**  
open science

## Phonological use of the larynx: a tutorial

Jacqueline Vaissière

► **To cite this version:**

Jacqueline Vaissière. Phonological use of the larynx: a tutorial. *Larynx 97*, 1994, Marseille, France. pp.115-126. halshs-00703584

**HAL Id: halshs-00703584**

**<https://shs.hal.science/halshs-00703584>**

Submitted on 3 Jun 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## PHONOLOGICAL USE OF THE LARYNX

J. Vaissière

UPRESA-CNRS 1027, Institut de Phonétique, Paris, France

### RÉSUMÉ

Cette communication concerne le rôle du larynx dans l'acte de communication. Toutes les langues du monde utilisent des configurations caractéristiques du larynx, aux niveaux segmental, lexical, et supralexical. Nous présentons d'abord l'utilisation des différents types de phonation pour distinguer entre les consonnes et les voyelles dans les langues du monde, et également du larynx comme lieu d'articulation des glottales, et la production des éjectives et des implosives. Nous abordons ensuite l'utilisation de patrons caractéristiques de la fréquence du fondamental au niveau lexical pour contraster entre les mots dans les langues à tons et à accent mélodique. La participation du larynx pour transmettre des informations linguistiques et paralinguistiques au niveau de la phrase est traitée en dernière partie.

### INTRODUCTION

The larynx is one of the main articulators in speech production, with the tongue root, the tongue body, the tongue blade, the velum and the lips. The goal of this tutorial is to review succinctly the different uses of the larynx for communicative purposes. The first part concerns the segmental use of the larynx: (i) the *phonation types* for differentiating consonants and vowels, (ii) the larynx as a place of articulation for the *glottal consonants*, and (iii) the initiation of airstream in *glottalic consonants*. The second part concerns the nonsegmental (prosodic) use: the participation of the larynx for contrasting higher linguistic units such as words, phrases and sentences (*tone, accent and intonation*). The third part tackles the problem of the

larynx used as a carrier of *paralinguistic information*.

### THE PRIMARY FUNCTION OF THE LARYNX IS PROTECTIVE

As stated by Sapir, 1923, physiologically, "*speech is an overlaid function, or to be more precise, a group of*

*overlaid functions. It gets what service it can out of organs and functions, nervous and muscular, that come into being and are maintained for very different ends than its own*" [1]. The primary function of the larynx is *protective*. It prevents the entrance of foreign material into the lungs. It also permits the building up of intrathoracic pressure.

### PHONATION IS NOT PARTICULAR TO HUMAN SPEECH

The secondary function of the larynx is to provide phonation. During the production of speech, the vocal folds (VFs) are mainly in adducted position. They vibrate, alternatively opening and closing the glottis (the space between the VFs) for a very short period of time. The principles underlying vibration are now rather well understood. The excessive subglottal airpressure forces the closed VFs to go apart and the elastic recoil and the Bernoulli effect suck them again together without any muscle action. The vibrations divide the continuous stream of expired air coming from the lungs into puffs of air that will excite the supraglottic cavities. The quasiperiodic modulation of the respiratory airstream by the vibrations of the VFs provides the primary source of energy for the production of voiced sounds and this process is called phonation.

Most of the speech sounds (vowels and most of the consonants) are voiced. The degree of voicing however greatly varies

from one language to the next. A study by Catford has shown that French is voiced 78% of the time, while Chinese (Canton) only 41% (English: 72%, Russian: 61%, and Alkhaz: 56%) [2].

Phonation is not properly linguistic. The acoustic energy in nonspeech production of amphibians, primates, and human infants is generated in the same manner. The interruption of phonation (mainly during the consonants) will be used to create contrasts between consonants (as seen later).

### EVOLUTION HAS FAVORED FINER VOCAL ADJUSTMENTS

A morphofunctional consequence of the various neuro-cranial rearrangements caused by an increase in extreme prefrontal cortex size during hominization [3] has been a descent of the larynx in the neck. The descensus of the larynx in the human neck had rendered larynx and tongue movements more independent than otherwise. The intrinsic fundamental frequency (FO) of the vowels shows however a tongue-larynx interaction, supposedly via the hyoid bone [4]. A lower position of the larynx has made possible finer adjustments at the level of the larynx, particularly of the tension of the vocal folds (VFs) and the spacing between them. In amphibians, the tension of the VFs cannot be varied and the vocal repertoire is monotonous.

For what concerns the spacing, an interesting fact can be noticed. The posterior cricoarytenoid (PCA) is the only abductor muscle and it is found only in man [5]: it will be particularly active for the production of voiceless and aspirated sounds.

### THE POSSIBLE CONTROLS OF THE LARYNX

Let us resume succinctly the controls which are used linguistically. (consult [6] on the mechanisms underlying the control fundamental frequency). Figure 1 summarises the general trends.

First, the *subglottic pressure* and the pressure across the glottis may be varied.

Heightened subglottic pressure generally assists in making focus and prominence.

Second, *the length and the mass of the VFs* can be directly controlled. Two muscles are mainly involved: the cricothyroid (CT) and the vocalis (VOC). The contraction of CT stretches the VFs. Lengthening results in an increase in the longitudinal tension of the VFs, which increases the rate of vibrations of the VFs (and therefore of the perceived pitch). Pitch variations is mainly used for *tone, accent and intonation*. The contraction of the VOC assists CT for pitch raising and may favor voicing at high frequency rate of vibration by slackening the cover of the folds [7], [8].

Third, the *glottis (airspace between the VFs)* can be finely adjusted. The arytenoids cartilages may be pressed together or abducted, and there may be a medial compression of the vocal fold tissue. Slightly abducted glottis or slightly adducted glottis allows vibration of the VFs, in breathy voice, murmur, or slack voice, and in creaky voice, laryngealized voice, and stiff voice, respectively. The abduction of the glottis leads to air leakage and to the presence of noise. Strongly abducted or adducted VFs hinder voicing. Combined spacing between the VFs and tension are used in *phonation types* for segments (breathy, creaky vowels and consonants). Voice quality over larger units also carries also *affective message*, as seen later.

Fourth, the *whole larynx can be raised or lowered*. First, the raising of the larynx decreases the volume of the supraglottic cavities, increases the VFs tension by stretching them, and increases the coupling stiffness in the vertical dimension of the larynx. A lowering movement has the reverse effects. The larynx movements are essential for the production of glottalic consonants. Second, they are used to assist F0 raising and lowering, through their effect on VFs length. Third, a lowered larynx may assist voicing by a 'bunching' of the surface issue (slackening) the VFs and participate to the phonetic implementation of the [+slack] feature for voiced consonants [9], [10].

An increased effort on a vowel decreases the Open Coefficient (OQ: the ratio of the open glottis state to the total period duration) and increases the abruptness of closing, leading to reduction in the spectral tilt.

All the controls can be used for carrying an affective message (exasperation), to express the attitude of the speaker toward the listener (drawing his attention) , and toward what he is saying (doubt, conviction) in particular the range of pitch variations and voice quality.

Subgl. pressure		Focus, Prominence
Spreading of VFs tension of VFs	Phonation types	
Stretching of VFs		Tone, accent, intonation
Vertical mvt of larynx	ejectives implosives	
all		speaker affect

Figure 1: Summary of the main trends

### WHY IS THE LINGUISTIC FUNCTION OF THE LARYNX DIFFICULT TO MODEL?

Much is known on the functional aspects of the larynx, but not enough to make the task of writing a tutorial on the larynx an easy task. The reasons are the following.

First, the glottis size, the tension of the VFs, the larynx height, flow resistance at the glottis and subglottal air pressure appear *interdependent*.. Changes in rate of vibration of the VFs lead not only in change in FO, but also in change in the voice source [11]. The stiffening of the VFs to achieve voicelessness for a consonant can lead to higher rate of vibration of the VFs at the onset of the following vowel. This interdependency contributes to interesting sound changes. The voicing contrast on a consonant may be "transphonologized" on the following vowel [12], [13]. The voicing contrast of onset consonants (/p/ versus /b/, for example, may be lost and replaced by a tonal contrast on the following vowel (high versus low). In this respect, laryngeal constraints have shaped up to a certain extent the phonological systems of the languages.

Second, the same muscles (e.g. CT) are involved in both the realisation of the distinctive features of the segments (voice/voiceless), and higher level factors (FO patterns, for example): the observed data conflate too *many conditioning factors* to be directly interpreted. Local gestures for phonetic contrasts and more global gestures may conflict with each other or reinforce each other. Alignments between word pitch pattern and the segments are not invariant: peak of FO tends to occur earlier during a vowel after voiceless a consonant, and later after a sonorant. The interference leads to interspeaker and interstyle variability. The speaker may choose to compensate or not by adopting more complex articulatory strategies, or, on the contrary or even by suppressing one of two antagonistic gestures. The range of corrective gestures depends on the rate of speech, and on the style (hyper- and hypo-articulation).

Third, there is a *lack of physiological data* on larynx behavior in speech. Such data are difficult to obtain, because the methods of investigation are invasive. Most of the available data come from the University in Tokyo, and from the Haskins Laboratories. They are very useful, but limited. Recovering physiological information on the source uniquely from the signal is difficult. The acoustic consequences of a single laryngeal gesture is much more complex to be interpreted than the consequence of a gesture in the supraglottic cavity.

### ON THE AMBIGUITY OF THE TERMS

The terms "voice quality" , "type of phonation" and "laryngeal settings", which all refer to the same adjectives such as modal, creaky, breathy are ambiguous [14]. Voice quality is determined not only by laryngeal, but also by articulatory settings (such as labialisation, pharyngealisation, etc. [15], [16]). Laver [15], [17]) classifies the different types of settings into linguistic, paralinguistic and extralinguistic settings. Laver remarks that quality seems to differ more in the time-scale involved than in their phonetic form, but it is not exactly true: an utterance may be

the domain of application of both linguistic and paralinguistic settings.

"*Neutral setting*" is often taken as a reference. It is characterised by the following facts: only the true VFs are vibrating, there is no audible noise, and a moderate tension.

First, **linguistic settings** concerns segmental and suprasegmental aspects of speech. A particular pitch contour (rise, fall) may be used linguistically to contrast tones in a tone language, or to contrast a question with a response. Breathiness may span over a short segment and be used as a distinctive feature for the consonants and the vowels in a language, to contrast with modal voicing (an unmarked phonation type [16]). The terms "Phonation types" include not only different types of "phonation" (voice); they include also the abducted VFs in voiceless sounds (called *nil phonation* by Catford) [2].

Second, **paralinguistic settings** span over a long stretch of speech (such as a sentence). For example, breathiness may signal confidentiality. A clear separation between linguistic and paralinguistic functions of the larynx is often difficult to draw. The expression of modalities (assertion versus question) by FO movements is clearly considered as linguistic (and categorical), while the interpretation of the expression of incredulity as paralinguistic (and scalar). Note that a system, like Tobi, largely used for describing FO contours in English, does not differentiate clearly between linguistic and paralinguistic functions. The expression of the speaker's attitude is clearly a continuum ("in a little bit upset voice"), such as the possible pitch range.

Third, **extra linguistic setting** is a permanent feature of voice. It is a marker of identity of the speaker and of its physical state. For example, the breathy quality of the voice may be pathological. It may also be a characteristic of a group of speakers. Tense setting characterises the way of speaking of many North Germans speakers [2], while the American English sound to speak in a amusingly relaxed way to French ears.

While the same terms are used (modal, breathy, creaky, etc.) for the different

types of settings, it is not proved that the different kinds of voice quality are completely equivalent from a physiological point of view.

The part of the variations in laryngeal configurations which have audible consequences on the acoustic properties of the source signal (spectral tilt, formant bandwidth, presence of turbulent noise, etc.) are eligible for linguistic purpose. The capability of the larynx to assume a number of different configurations, and therefore modify the spectral properties of the source and to introduce dynamic variations is in fact used by all the known languages.

## PART I: SEGMENTAL ASPECTS

Our presentation below is not exhaustive. The reader is referred to articles and books by Catford, Stevens, Ladefoged and Maddieson for the use of laryngeal settings in the sounds of the world's languages [2], [10], [18], [19].

### A) Phonation types

#### I) Consonants

The main contrasts involve the presence versus absence of voicing and of aspiration.

#### - Voiceless stops [p], [t], [k].

*Obstruents* (stops and fricatives) tend to be naturally devoiced. The building up of supraglottal pressure due to the supraglottic constriction results in a decrease or cessation of transglottal pressure. No vibration is possible without transglottal pressure. As expected, voiceless obstruents are more common than voiced obstruents in languages [20]; most of the languages with only one series of stops are reported to have voiceless stops. The optimal state for occlusive consonants are generally considered by the phonologists to be *voiceless, unaspirated, unglottalized*, such as French [p], [t] and [k] [20]; [21]. In contrast, *sonorants* consonants are generally voiced. There are exceptions:

Hindi, for example, contrast between breathy-plain and voiced laterals. The reader is again referred to the book written by Ladefoged and Maddieson for more information.

Most of the languages use difference in laryngeal settings to differentiate among stops and fricatives. They generally have a two-way contrast for the stops. The phonetic implementation of the two way contrast differs from one language to the next. French contrasts between voiced and unvoiced stops, while Chinese contrast between voiceless unaspirated and voiceless aspirated stops. Some languages have a three-way contrast. Thai contrasts between voiced unaspirated, voiceless unaspirated and voiceless aspirated. Hindi has a four way contrast: voiced unaspirated, voiced aspirated, voiceless unaspirated, and voiceless aspirated (see figure 2).

#### **- Voiced, unaspirated [b,d,g]:**

A two-way contrast may typically consist in opposing voiceless to *voiced*. The main articulatory difference is the presence versus absence of VFs vibration during the consonant production. According to the model developed by Halle and Stevens [22], the difference in voicing is essentially a function of VFs tension (see also Stevens, [23]). In voiceless consonants, the stiffness of the VFs hinders vibration, while in voiced consonants, their slackness renders vibration possible, even at low transglottal pressure. The terminology [stiff] and [slack] is not unanimously accepted (some prefer the more traditional feature [ $\pm$ voice]). A difference in stiffening of the VFs is difficult to measure and there is no indication that the glottal constriction is looser for /b/ than for /p/. But the commonly observed higher fundamental frequency (FO) in the vowel after voiceless consonants may be interpreted as an progressive assimilation of an increased stiffness. According to Stevens [10], stiffness of the VFs may be indirectly achieved by raising the larynx, and slackness by lowering it.

In accordance to the observation that voiceless is considered as the unmarked state for obstruents by phonologists, voicing is

indeed difficult to maintain in voiced stops. Vibration of the VFs is maintained for a limited number of pulses through a passive extension of the vocal tract volume, due to heightened supraglottal air pressure. Active maintenance of voicing can be done in a number of ways. First, voicing can be maintained by expansion of the pharyngeal walls and supraglottal cavities [24]. Active extension of the supraglottic volume is easiest for labials (by puffing the cheeks). It is the most difficult for back stops. Accordingly, /g/ is less frequent in languages than /b/ and /d/ [20]. Second, it can be sustained by an active slackening and thickening of the VFs. The slackening gesture may be absent in the phonetic realisation of /b,d,g/ in English, which often lack a voice bar during the closure [21]. In Stevens' description, the tension cannot be neutral for stops; it has to be either slack or stiff [22] [23].

#### **-voiceless aspirated [p<sup>h</sup>,t<sup>h</sup>,k<sup>h</sup>]**

The presence of aspiration versus its absence is often contrastively used in languages. The aspirated stops are generally preaspirated stops, but there may be postaspirated stops. Aspirated means postaspirated.

The articulatory manoeuvres to achieve contrast based on aspiration have received a number of interpretation. Halle and Stevens (1971) interpret it essentially in terms of spreading of the glottis at the time of the release or after: [+spread] for aspirated consonants and [-spread] for unaspirated stops [22]. Lisker and Abramson relate the contrast mainly to a different timing of laryngeal activity relative to the suprasegmental constriction (Voice Onset Time or VOT). VOT is longer in the case of aspirated consonants [25]. The contrast can be interpreted also in terms of greater abduction for aspirated stops. PCA is more active in aspirated voiceless stops than in unaspirated stops, and it is not just therefore a question of timing. The dominant manoeuver does not need to be the same for all languages.

Acoustically, aspiration is characterised by the presence of turbulence noise at the time of

the release and after, with non-harmonic varying components.

	ejective	implosives	breathy	voiced	voiceless	aspirate
French				b	p	
English				b		p <sup>h</sup>
Chinese					p	p <sup>h</sup>
Korean	p*				p	p <sup>h</sup>
Thai				b	p	p <sup>h</sup>
Hindi			b <sup>H</sup>	b	p	p <sup>h</sup>

Figure 2: Source: Ladefoged and Maddieson, 1996.

**- breathy (aspirated) voiced [b<sup>H</sup>, d<sup>H</sup>, g<sup>H</sup>] AND slack consonants**

Voiced aspirated (also called breathy or murmured) stops are less common than voiceless aspirated stops. The VFs are vibrating during the closure of the stop, and the arytenoids are spread after the release. The VFs are supposed to be slack (by lowering the larynx). A more or less great deal of air is allowed to pass through the slightly open glottis.

The figure below illustrates the glottal aperture aspirated and breathy stops, based on fiberoptic data.

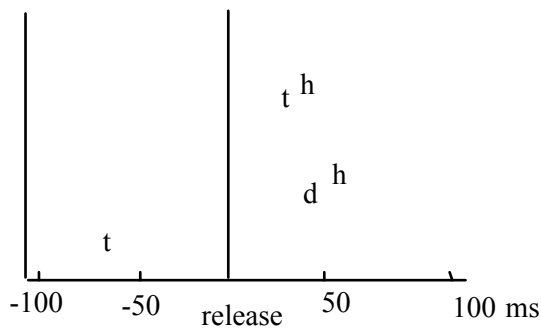


Figure :3: relative glottal aperture in Hindi consonants (after Kagaka and Hirose (1975) [26])

For slack consonants, the larynx is lowered (F1 lowers), the VFs are vibrating more loosely than in the modal voice, there is a slightly increased glottal aperture, and a moderate increases in air flow, FO is lower in the following vowel. There is reduced energy in the upper frequency range of the spectrum. There is a continuum between slack and breathe voice.

Breathy voiced and slack consonants do not to contrast [19].

**- creaky and stiff (laryngealized) consonants**

In creaky consonants, the arytenoids are closer than in modal voice [+constricted][- spread], and the tension is high [+stiff]. The VFs do not vibrate as a whole, or the parts of the VFs vibrate out of phase. Hausa contrasts between voiced, voiceless, “glottalized” consonants (In the bilabial and alveolar stops, the glottalization affects the following vowel as well). (for more information [19]).

In stiff consonants, like in Thai, there is a slight degree of laryngealisation. No languages contrast modal voice with no more than one degree of laryngealized voice.

**2) Vowels**

	breathy	modal	creaky
glottal configuration			
volume velocity			
spectral tilt			

acoustic characteristics	bandwidth increase noise high frequencies decrease glottal formant		bandwidth decrease no noise high frequencies increase
--------------------------	---	--	---

Figure 4: phonation types of vowels. Adapted from Stevens, 1977 [10]. Glottal configurations, state of the arytenoids, volume velocity, spectrum of glottal pulse, schematised spectrograms, and acoustic characteristics in three phonation types for vowels: *breathy, modal and creaky*

Most of the languages do not use any differences between glottal states to distinguish among vowels. A number of languages use two types of phonation (modal versus creaky or breathy). A few rare languages, like Mazatec, contrast between modal, breathy and creaky vowels [19].

**- creaky and slack vowels**

The glottis is constricted. The glottal pulse is sharper as compared to the modal vowels, leading a very low intensity of the fundamental. There is more energy in higher frequencies, and the bandwidth of the formants are reduced.

**- breathy and stiff vowels**

The arytenoids are spread. Voice source pulses have a larger open coefficient and a nonabrupt glottal closure. There is a relative decrease of the amplitude of the harmonics in the middle and upper part of the spectrum and a consequent increase in very low frequencies: there is a higher amplitude level of the fundamental compared to the second partial [27]. There is a greater random component. Due to open glottis and subsequent loss due to the subglottal system, there is widening of the bandwidth, and the first formant, in particular, is less well defined (see figure below).

**- voiceless vowels**

In some languages (North America, see [19]), voiceless vowels seem to contrast with modal vowels. Devoicing of generally is not contrastive; it is mostly due to assimilation with the surrounding voiceless context and concerns mainly close vowels, where the heightened supraglottal pressure results in a delay or a suppression of voicing. This process is regular in Japanese, and occasional in French.

The next figure illustrates the features used by Stevens for describing the different types of phonation in vowels. The dimensions are the same than for the consonants. The different degrees of tension of the VFs (stiff, neutral, slack) contrast the vowels with high mid and low tones, respectively.

**B) The glottal consonants [h ], [H ][? ]**

The larynx serves as the primary place of articulation for the stops [h ], [H ][? ].

[h ], [H ] are voiceless and voiced, respectively.

The glottal stop [? ] (coup de glotte in French) requires a very constricted glottis, and there is no voicing contrast possible. It is a speech sound in many languages (Arabic). In German, it is a regular characteristic of the beginning of the stressed initial vowels. In English, it is a device for marking morpheme boundaries in pair of words such as "an ice man" and "a nice man". It may also be a substitute for [t].

[Let us note a very interesting sound change: [H ] (or [h ] and [? ] in coda position may disappear and be replaced by a falling tone on the preceding syllable and a raising tone, respectively [12].

	stiff VFs	neutral VFs stiffness	slack VFs
spread arytenoids	voiceless vowels ëœ	breathy vowels A5	low-pitched breathy vowels
neutral glottal position	high pitched vowel A'	vowel in the mid-range A	low-pitched vowel A'
constricted glottis	glottalized vowels	creaky vowels Aε	low-pitched creaky vowels?

Fig 5 Phonetic categories associated with different ways of producing vocal-fold stiffness-slackness and with various types of abduction-adduction of the glottis for the vowel (Stevens, 1977).



**C) The glottalic consonants**

glottalic pressure			
ejective			
	closure	pressure rise	egressive air
glottalic suction			
implosive			
	closure	pressure down	ingressive air

Figure 6 glottalic consonants

The larynx can be used as a kind of piston to create initiation of airstream (see figure 6).

The consonants in languages can be classified into three types, depending on the initiation of airstream. Most of the consonants in the world languages are *pulmonic*, and produced with the help of the respiratory system. *Velaric* consonants are produced by motion of the tongue (clicks). *Glottalic* consonants involve the motion of the (closed) larynx, which is used as a kind of a piston.

For the production of glottalic consonants, there are two constrictions and the air is trapped in between. The arytenoids are constricted and there is a occlusion in the oral cavity. The larynx is moving up (for ejectives) or down (for implosives), compressing or rarefying the air trapped in the supraglottic cavities, respectively. The direction of the airstream is egressive for the ejectives and ingressive for the implosives.

Igbo, for example, contrasts between voiced, voiceless implosives, ejectives and affricates. Dagara has both voiced and unvoiced implosives. 18% of the world languages have ejective, often velar [20].

Since the glottalized consonants never contrast with ejectives (seen later), the difference between glottalized and ejectives may be considered as a matter of phonetic implementation.

Figure 7 illustrates the laryngeal features used by Halle and Stevens [22] to describe phonologically the different types of phonation used to contrast the consonants. The two dimensions are the tension of the VFs (stiff versus slack) and the aperture of the glottis. The very few features seem sufficient to describe the contrasts used by the languages. For example, no language contrasts between voiced breathy and slack consonants, or between creaky and stiff stops: the difference seems rather a matter of phonetic implementation (for a somehow different view, see [28]).

	Stiff VFs (raised larynx)	Slack VFs (lowered larynx)
spread arytenoids	p <sup>h</sup> (aspirated)	b <sup>h</sup> (breathy)
neutral glottal position	p	b
constricted glottis	p? (ejective)	ɓ (implosive)

Figure 7 Adapter from Stevens, 1977

**PART II: SUPRASEGMENTAL ASPECTS**

Vowels in languages are generally voiced. The ear is particularly sensitive to FO variations in the most stable part of the vowel [30]. All known languages seem to make use of pitch variations significantly, taking advantage of the ability to produce fine variations of FO, which are audible.

Mainly lexically determined pitch pattern	Mainly postlexically determined pitch patterns
---	--

Syllable-based	word-based	bounded to a special (stressed) syllables in the word	bounded to word boundaries
<b>TONE LANGUAGES</b>	<b>PITCH-ACCENT LANGUAGES</b>	<b>STRESS LANGUAGES</b>	<b>BOUNDARY LANGUAGES</b>
<i>Mandarin Chinese</i>	<i>Swedish Japanese</i>	<i>English</i>  expressive stress falls on stressed syllable	<i>French</i>  several syllables are accentuable

Fig 8: a possible typology (inspired in part by Garde, 1968 [33])

The pitch contour in the utterances in the languages of the world is the result of a combination of morphological, lexical, syntactic, pragmatic and/or paralinguistic factors. At each level, pitch may have a distinctive, culminative, demarcative function [31]. The way the languages combine functions and levels is language-dependent [32]. For example, at the word level, pitch has mainly a distinctive function in tone languages (like Chinese) a culminative function in stress languages (like English), and a demarcative function in boundary languages (like French), respectively. The extensive use of pitch at one level restraint its use at higher level.

Distinctive pitch pattern are used for levels higher than the word in stress languages and boundary languages (for example to contrast between interrogative and question, but much less in tone languages).

There is no standard typology for classifying the prosodic systems of the languages (see however the classification by Pike, [33]). A language may have features to would classify it in two categories. The following figure illustrates a possible classification (relatively conventional, except for what concerns boundary languages).

The exact pitch patterns found in the words can be MAINLY determined by lexical factors (tone- and pitch accent languages), or by higher factors (stress- and boundaries languages).

**A) Syllable-based pitch pattern: the tone languages**

The majority of the world languages are syllable-based tone languages. Pitch is used to distinguish *morphemes*. The pitch contour of the syllables is mainly determined by the lexicon. Pike distinguished between register, and contour tone languages and mixed languages. In Yoruba, a typical register tone language, the relative height of the syllabic pitches is important. In Mandarin Chinese, a contour tone, the shape and trajectory are the most important [33].

**B) word-based pitch pattern: the pitch accent languages**

Some languages, such as Swedish [34], Norwegian, Slovenian, Japanese have word-based tones. The pitch patterns of the words is lexically determined. Higher level features, such as syntax, focus, etc... may influence the FO range of variations and other parameters, the general FO pattern of the word (more rising or more falling, or higher or lower), but the changes in the pitch pattern will be not dramatic.

**C) stress and boundaries languages**

The particular pitch pattern that the word will receive in a sentence, is not a characteristic of the word. In stress languages, a syllable in the word is marked, and the realisation of the pitch patterns (or pitch-accents) are bounded relatively to the position of that stressed syllable (one for every lexical word). The marked syllable is the only syllable which can be made more prominent (except in case of contrastive stress). In boundary languages, the peripheral syllables in the word are singled out. The pitch movements are bounded to the left and right boundary of the word in French. The

position of emphatic stress allows to draw a clear distinction between stress and boundary languages. When emphasised, the same (marked) stressed syllable is made more prominent in stress languages. In a boundary languages, emphatic stress falls in word beginning (while boundary tones are bounded to the final syllable).

Word stress and word boundaries are realised by a combination of factors, not necessarily FO. The physiological correlates of stress are very complex and involves an higher voice excitation strength (calculated as the derivative of the glottal flow closing). Final lengthening mark the right boundary.

The number of pitch patterns that the whole word may receive is rather large; ([35] for English). In neutral, isolated sentences, the choice of pitch patterns for the successive words are in close relation with the syntactic structure [36]. Other factors play a greater role in other contexts. Rather simple tests allow to separate languages in different type(see for example [37] for ten languages, etc...

#### **D) Intonation**

Pitch has an intonational function when it is used in domain larger than the word.

There seem to be a number of similarities between languages in the use of FO at the supralexical levels. Two basic patterns can be recognised. The rising means non finality, openness, arousal of interest, intentness, non definiteness, unfinishedness, continuation, questions. The falling pattern involves semantic finality [38], [39].

#### **E) Positional effect on the larynx**

##### **- The suprasegmental features [±tense]**

Supraglottic and laryngeal configurations for phonemes are different, depending in their position (initial, internal or final) in larger units. I have proposed to add systematically the positional feature [±tense] to the phonological specification of each phoneme for interpreting their allophonic realisation [40]. What are the consequences

of the feature [±tense] on the larynx behavior in speech?

- **the initial position** is characterised by a general tensening of all articulators. The stiffening of the adductory muscles laryngeal muscle results in less voicing or a delay of voicing, for voiced consonants. It results in glottalization or in glottal stop in words beginning by a vowel (for German [41], for English: [42], [43]). It leads to a larger glottal width for voiceless consonants (like in English) and then to aspiration. Diachronically and synchronically initial phonemes are resistant to change and less likely to disappear. Great care should be taken when taking data: the synchronisation between the activation of all articulators during the speech ready gesture is random to a certain extent [44].

- **In medial, unstressed position**, there is a tendency to diminish the amplitude or even suppress some of the gestures. The omission of the glottal opening gesture for the voiceless stops leads to voicing; /t/ is often transformed into a glottal stop ("butter" in American English). Synchronically and diachronically (sounds changes), intervocalic voiceless consonants are prone to be voiced.

- **Final position** is marked by relaxation of all articulators, a slow down of the movements (related to final) lengthening, a decrease of the degree of stricture and even disappearance. At the laryngeal level, it is often marked early devoicing and breathiness. In a quite large number of languages (like German), the voicing distinction for stops and fricatives is neutralised in word final position and all final consonants are voiceless. Final position in sentences may also be marked by creakiness, due to irregular vibrations of the vocal fold (probably concomitant to a very low transglottal pressure).

### **PART III: PARALINGUISTICS**

Finely coded, attitudinal tones of voice are used by adult speakers to shed nuances on the spoken material. Controlled

tones of voice inform on expressiveness, incredulousness, sarcasm, annoyance, bored resignation, apology, supplication, etc. The realisation of the tones of voice includes manipulation of mean FO (and intensity), FO range, voice quality, etc [45]. For example, whispery voice signals confidentiality, breathy voice intimacy, harsh voice dissatisfaction. The outward expression of attitudes directly through tones of voice may be more or less standard, depending on the cultures, and on the degree of formality (attitudes may be expressed by lexical and grammatical means). The study of the culturally specific contribution of the setting of the tone of voice is an area still largely unexplored (with the exception of Fonagy [46] and Laver [17]).

There may be a link between the emotional aspects of speech, where low FO is bounded to dominance and high FO to submission and the linguistic use of high FO, through the frequency code [47], [18], [38]. It is not easy to draw the separation between strictly linguistic and paralinguistic aspects, since the two aspects are intimately joined in every day conversations. The paralinguistic information is very important in human relations since the tune of voice is often as important as the literal sense of the speech form (as in the speech of the human mothers addressed to the new-born).

## CONCLUSION

We may conclude that the larynx play a very important role in speech. Its most important function is to provide phonation. Phonation does not require a sophisticated laryngeal mechanisms (voice is possible without larynx). The possibility of a fine control of the tension of the VFs, the aperture of the glottis, the movements of the larynx, have been largely exploited by the languages, for distinguishing phonemes, words and sentences, etc... The degree of involvement of pitch at the lexical level differ greatly from one language to the next (such as in tone and stress languages). There are cross-languages similarities in the use of the laryngeal

configurations at supralexical levels. The linguistic functions of the larynx are very important, but no less important are the paralinguistic functions in every day exchanges. Human listeners are trained since they are born to pay attention to nuances of voices. Human mothers exaggerate their melodies in their speech addressed to their new-born, to draw their attention and to pacify them. The "tunes" of voice convey information and they are used prelinguistically by the infants, before they understand speech. Much more work is needed to elaborate a model of the interaction between segmental, linguistic, para- and extra-linguistic factors on pitch and voice quality.

## REFERENCES

- [1] Sapir, A., (1921), *Language: an introduction to the study of speech*, New York, Harcourt, Brace and World.
- [2] Catford, C. (1977), *Fundamental problems in Phonetics*", Edinburgh University Press.
- [3] Camperio Ciani, A., and B. Chiarelli. , (1988), "A systematic relationship between brain size increase and larynx transformation during hominization." in *Language Origin: a multidisciplinary approach*. Wind et al (Eds.), Dordrecht, Boston, London: Kluwer Academic Publishers, 51-66.
- [4] Honda, K. and Fujimura, O., (1991), "Intrinsic vowel FO and phrase-final lowering; phonological Vs biological explanations", in Gauffin, J., and Hammarberg; B., (1991), *Vocal Fold Physiology: acoustic, perpetual and physiological aspects of Voice Mechanisms*, Singular Publishing Group, San Diego, 149-158.
- [5] Hirose, H., (1976), "Posterior cricoarytenoid as a speech muscle", *Ann. Otolaryngology, Rhinology and Laryngology*, 85, 334-343.
- [6] Titze, I.R., (1991), "Mechanisms underlying the control of fundamental frequency" in Gauffin and Hammarberg (Eds), 129-148.

- [7] Hirano, M., (1981), *Clinical Examination of Voice*, Vienna: Springer-Verlag.
- [8] Fujimura, O., (1981), "Body-cover Theory of the vocal fold and its phonetic implication", in Stevens, K.N., and Hirano, M. (Eds), (1981), *Vocal Fold Physiology*, Tokyo University Press, Tokyo.
- [9] Ohala, J., (1973), "The physiology of tone" in Hyman, J.L. (Ed.), *Consonant Type and Tone*, Sth. Calif. Occ; Papers, 1, 1-14.
- [10] Stevens, K.N., (1977), "Physics of laryngeal behavior and larynx modes" , *Phonetica*, 34, 264-279.
- [11] Pierrehumbert, J., (1989), "A preliminary study of the consequences of intonation for the voice source", KTH, Speech Transmission Laboratory, Quarterly Progress Status Report, 4, 23-36.
- [12] Hagège, C., and A. Haudricourt. , (1978), *La Phonologie Panchronique*. PUF, Paris.
- [13] Hombert, J-M, Ohala, J., and Ewan, G., (1979), "Phonetic explanations for the development of tones", *Language*, 55, 1, 37-58.
- [14] Ananthapadmanabha, T., (1995), "Acoustic factors determining perceived voice quality", in Fujimura and Hirano (Eds), 113-126.
- [15] Laver, J., (1980), *The Phonetic Description of Voice Quality*, Cambridge University Press.
- [16] Ladefoged, P., (1988), "Discussion of Phonetics: a note on some terms for phonation types", in Fujimura, O. (Ed), (1988), *Vocal physiology: Voice Production, Mechanisms and Functions*, Raven Press, Ltd., New York, 373-375.
- [17] Laver, J., (1991) , *The gift of speech*, Edinburgh University Press.
- [18] Ladefoged, P., (1983), "The linguistic use of different phonation types", in Bless, D., and Abbs, J., (Eds), *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, San Diego: College-Hill Press., 351-360.
- [19] Ladefoged. P. et Maddieson, I., (1996), *The Sounds of the World's Languages*, Blackwell Publishers.
- [20] Maddieson, I., (1984), *Patterns of Sounds*, Cambridge: Cambridge University Press.
- [21] Kenstowicz, M., (1994), *Phonology in Generative Grammar*, Blackwell).
- [22] Halle, M., and Stevens, K., (1971), "A note on laryngeal features", *Quarterly Progress Report* 101, 198-212, Cambridge, mass.: Research Laboratory of Electronics, MIT.
- [23] Stevens, K.N., (1991), "Vocal-fold vibration for obstruent consonants", in Gauffin and Hammarberg (Eds), 29-36.
- [24] Westbury, J.R., (1983), "Enlargement of the supraglottic cavity and its relation to stop consonant devoicing", *JASA*, 73:1322-1336.
- [25] Lisker , L. and Abramson, A.S., (1964), "A cross-language study of voicing in initial stops: acoustical measurements", *Word* 20, 384-422.
- [26] Kagaya, R., and Hirose, H., "Fiberoptic, electromyographic and acoustic analyses of Hindi stop consonants", *Ann. Bull. Res. Inst. Logopedics and Phoniatics*, 9, 27-46.
- [27] Bickley, C., (1982), "Acoustic analysis and perception of breathy vowels", *MIT Working Papers*, 1, 71-81.
- [28] Ladefoged, P., Maddieson, I. and Jackson, M. (1988), "Investigating phonation types in different languages", in Fujimura, O. (Ed), 297-318.
- [29] Ladefoged, P., (1973), "The features of the larynx", *J. Phon.*, 1, 73-83.
- [30] House, D., (1990), *Tonal Perception in Speech*, Trav. Inst. Ling. Lund, Lund University Press.
- [31] Trubetskoy, (1957), *Principes de Phonologie*, Paris.
- [32] Vaissière, J., (1983), "Language independent prosodic features", in Cutler, A. and Ladd, D.R. (Eds), *Prosody: Models and Measurements*, Berlin: Springer-Verlag., 53-66.
- [33] Pike, K., (1948), *Tone languages*, university of Michigan Press, Ann Arbor, Michigan.
- [34] Gärding, E., (1983), "A generative model of intonation", in Cutler and Ladd (Eds.), 11-21.

- [35] Beckman, M. and Pierrehumbert, J., (1986), "Intonational structure in Japanese and English", *Phonology Yearbook*, 3, 15-70.
- [36] Lehiste, I., (1980), "Phonetic manifestation of syntactic structure in English", *Ann. Bull., RILP, University of Tokyo*, 14, 1-27.
- [37] Vaissière, J., (in press), "Langues, prosodie et syntaxe", *Revue ATALA*.
- [38] Delattre, P., (1961), "La leçon d'intonation de Simone de Beauvoir: Etude d'intonation déclarative comparée", *The French Review*, 35, 59-67.
- [39] Vaissière, J., (1995), "Phonetic explanations for cross-linguistic similarities." *Phonetica*, 52, 123-130.
- [40] Vaissière, J., (1986), "Variance and invariance at the word level", in *Invariance and Variability in Speech Process*, Perkell, J. S., & Klatt, D. H. (Eds.), Lawrence Erlbaum Associates, Publishers, 534-539.
- [41] Kohler, K., (1980), "Prosodic boundary signals in German", *Phonetica*, 38, 89-134.
- [42] Umeda, N., (1978), "Occurrence of glottal stops in fluent speech", *JASA*, 64 (1), 88-94.
- [43] Pierrehumbert, J., (1995), "Prosodic effects in glottal allophone, in Fujimura, O. and Hirano, M., *Vocal Fold Physiology: Voice Quality Control*, Singular Publishing Company, San Diego., 39-60.
- [44] Vaissière, J., (1986), "Comment on Abbs' paper", in *Invariance and Variability in Speech Process*, Perkell, J. S., & Klatt, D. H. (Eds.), Lawrence Erlbaum Associates, Publishers, 220-22.
- [45] Scherer, K., and P. Ekman, (1984), *Approaches to Emotion*. Lawrence Erlbaum Association, 1984.
- [46] Fonagy, I., (1983), *La vive voix: essais de psycho-acoustique*, Payot, Paris.
- [47] Ohala, J., (1984), "An ethological perspective on common cross-language utilisation of Fo of voice." *Phonetica* 41 (1984): 1-16.[46].