FACTORS AND THEIR FUNCTIONS IN THAI TONAL ANALYSIS

†P.J. Bee

An attempt to state possible discrete factors in the analysis of basic tones in Thai is bound to make certain initial assumptions. For instance, it will be taken for granted that the Thai tonal system resembles that of Chinese and, in a most general sense, that of Burmese, or Vietnamese. It will be taken for granted that the problem of tone in the 'tonal' languages of South East Asia is, at heart, a common one and that disparities such as actually occur in present-day speech are attributable to processes of change not so much in the nature of the basic tones themselves as in their environment within the syllable. A brief explanation of the influences due to syllable environment must be given, of course, but the main point in this paper is to focus attention on the basic, tonal, phonematic units that persist even after the various levels of conditioning due to surrounding 'extra-tonal' factors have been allowed for.

Levels of influence in tonal analysis[1]

(i) At the deepest level of influence there is a dichotomy imposed by syllable-finals. Syllables closing in an unexploded occlusive (-k, -t, -p and -ʔ) constitute a totally different system from that of continuant-final syllables.[2] In the latter case, there are three members of the system that commute in a prosody of tone. This holds true for old Chinese and old Thai. In the former case, stop-finals -- since there is a system of one -- it could be argued that this is not a tone at all since it is non-contrastive. Issue will not be taken here, though, on a terminological point. Suffice it to say that three tones always were (and still are, in a deferred, analytical sense) clear and contrastive in the continuant-final system: these will be the central topic of further enquiries - our 'basic' tones.

(ii) One stage up from the base, we can discern the influence of syllable quantity. In Thai, but not in Chinese, the stop-final syllables usually have their tonal inventory doubled (i.e. pitch now becomes contrastive in a new system of two) by distinguishing short and long quantities by complementary pitch phenomena. In the continuant-final system of three, quantity is phonemic but is not accompanied by special pitch exponents at all. The neatest solution, therefore, is to acknowledge the primacy of syllable length by assigning the distinguishing feature to the phoneme 'quantity' and to interpret the stop-syllables' pitch exponents as secondary features. Thus, although pitch in stop-final syllables behaves in a 'tonal' manner, it is not the obvious, prime factor in phonemic distinction. Note that we cannot say in these cases that pitch features (the two 'tones') are allophones of quantity since the phone of long vowel or short vowel is not *replaced.* We can, however, argue the merits of quantity as being the stable and most easily accountable distinguishing feature of an apparently composite phoneme. This becomes clear when comparisons between dialects are undertaken.

Even though the whole problem of analysing stop-final syllables is, by this argument, not central to the issue of basic tones, the natural and analytically justifiable preference for *discrete* factors (is the syllable short or long?) is symptomatic of linguists' aversion to pitch phenomena, which vary so startlingly from dialect to dialect and which take on the characteristics of a continuum divided into ranges rather than on/off, yes-or-no, type of assessment. Whilst it has become fashionable to reduce all processes to ordered binary oppositions in emulation of the computer, there is no gainsaying the simple truth, that it is easier to work with a limited number of 'hard' factors than to characterize things, e.g. tone classes, by a bundle of pitch-contour graphs representative of dialect and idiolect samples, many of which show only the vaguest resemblances in shapes and levels. More will be said on this point when the general theme of these speculations is set forth.

(iii) The stage of development which is uppermost, because it is demonstrably the most recent in time for both Thai and Chinese, is the displacement of levels of pitch in all tonal phenomena, basic and secondary 'stopped' tones alike. The effect is rather akin to the moving of a kaleidoscope out of focus: the image becomes a multiple one, often with some overlapping which produces what appear to be areas in common. If the images were interpreted as sets, one might be tempted to think that the common areas would yield some common factors, since they were shared between two or more sets. The overlapping images would, in other words, be regarded as Venn diagrams.* A moment's thought, however, will make it clear that the all-important factor is how much the kaleidoscope is moved: this is the actively controlled 'input', whilst the degree of displacement of image is its direct consequence. For tonal displacement, therefore, we might expect to find that the reason

---

* Venn diagrams, named for the British logician John Venn (1834-1923) who devised them in 1880, are a means of depicting mathematical sets and subsets, a graphical aid in the study of Boolean algebra. (Ed.)

that caused the shift in levels would be more useful in accounting for the actual tonal picture as it is today than a set-diagram style of analysis (superimposition of the pitch-contours of one 'tone' as found in various dialects, for instance). It is commonly held that the reason for the displacement of levels is to be found in a reduction in the inventory of features (consonant types) in syllable-initial prosody. In Thai and Chinese it is the loss of voice from initials (in Chinese usually agreed to have been originally voiced aspiration viz. stop + / ɦ / or / ɦ / alone), the conflation being either with unvoiced aspirates or, as in many dialects of North Thailand and the Shan States, with unvoiced non-aspirates. Of course, other features of initial prosody such as the pre-glottalization of voiced initials, may also have played a part, but this merely adds one more dimension to what is a clear, underlying general statement, namely, that initial phoneme depletion is reflected in increase of tone-level phonemes.

By great good fortune the rules of Thai orthography for 'spelling' these new levels caused by displacement find their traditional expression in the categorizing of all syllable initials as High, Mid and Low class consonants. It is also interesting to speculate that the choice of terms to describe these classes must originally have had some mnemonic value, High class consonants presumably yielding a 'high' band in pitch,[3] Mid a 'mid' and Low a 'low' band, but both time (successive developments within a given dialect tending away from neat correspondences for rules laid down at an earlier stage) and geography (the whereabouts of the dialect upon which these early correspondences were based[4]) have rendered the names of these consonant classes quite arbitrary today. The procedure, however, of spelling tones by a combination of consonant class and tone class-mark is certainly not arbitrary. It reflects the stages of which we have been speaking: the tone class-marks represent our concept of basic tone, and the consonant classes represent the later displacement phenomenon due to loss of distinctive features in the initials, the burden of maintaining these distinctions now being a tonal one.

Theorizing about tones

The following sequence of diagrams shows the steps or strata referred to above in a purely statistical way, indicating the *possibilities* of tonal items as one passes from stage to stage. The fulfilment of these possibilities in actual tone phenomena, however, is not represented nor even attempted. No language has a one hundred per cent tally between possibilities, probabilities, and actualities. Conflation, attrition, assimilation by analogy, and so forth, are typical forces at work to perturb the calculation of predictable end-product items (i.e. the probabilities are not equally weighted).

Nevertheless, for inclusive, cross-dialect purposes and for the upholding of an argument that all these stages may be set aside from the problems of *basic* tones, the straightforward mathematical count will be allowed to stand.

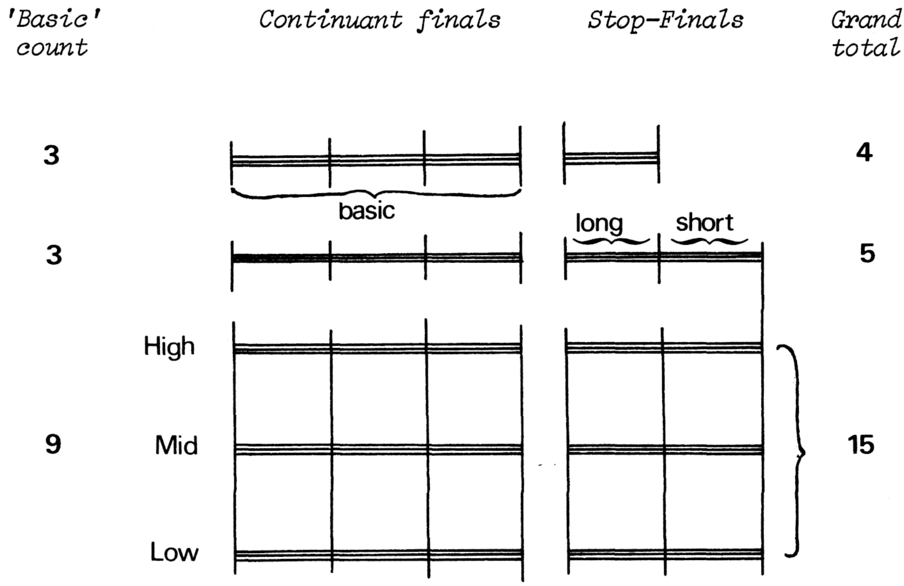


Diagram I : *Continuant and stop finals in Thai*

Having come up, as it were, by reasoning through successive stages, can we try to go down -- in reverse -- using the same arguments that tones tend to be derived from other, say crassly segmental features, and break down the *basic* tone barrier? In other words, can we extrapolate backwards and postulate a 'pre-basic' stage represented thus:

All continuant finals All stop finals

where further factors can be presumed to have existed, non-tonal in character (i.e. where such pitch phenomena as may have existed were not phonemic), whose disappearance produced a threefold classification of syllables of a certain kind, the distinctive features now being transferred to and incorporated in three derived tones? Note that all our effort and ingenuity must be concentrated on the continuant-final set, recalling that it has already been admitted that a stop-final set can be accounted for without resorting to tone distinction at all on this basic level.

The similarity between our new hypothesis challenging the primal nature of the three basic tones and the accepted fact that initial-phoneme prosodies have influenced dispersal of pitch-levels encourages speculation which probably springs from the mind's fascination with symmetry. If we can reconstruct initial phonemes along the lines of

k-	kh-	g-	(or kɦ-)
t-	th-	d-	(or tɦ-)
p-	ph-	b-	(or pɦ-)
ʔ-	h-	ɦ-	(or p -)

can we not balance these by assuming some similar inventory for archaic finals, like:

-χ	-θ (or -s)	-f	-h
-γ	-ð (or -z)	-v	-ɦ

in addition to the

-k	-t	-p	-ʔ

that we do, in fact, find to occur. This temptation is made all the stronger by the fact that the nasal consonants are subject to both initial and final placement.

Yet most scholars reject this. It seems too facile and, furthermore, no evidence of these elusive finals has been found in any reflexes of any dialect. More promising is the line of argument that presumes morphological residues (suffixes or possibly infixes) to be embodied in the basic tones. If this were the case, however, we might expect to find indication of *morphological* distribution in the three basic tones -- one tone predominantly one word class, say, or one tone predominantly characterizing one set or bundle of grammatical categories -- for instance, a 'tone-set' of all plural nouns, all passive verbs and all locative and instrumental function-words against another tone-set of singulars, actives and function-words of motion and intention. Whilst it is true that word-families in Chinese do often show what are apparently grammatical or morphological relationships within the coverage of one referential sememe, and whilst the same kind of relationship might be demonstrated in minimally contrastive tone pairs in Thai such as, for example:

Trans-cription	*Initial consonant class*	*Tone class mark*	*Pitch contour**	*Word class*	*Meaning*
/baaŋ/	Mid	zero (ø)	mid-level	numeral-term (partitive)	'some'
/bâaŋ/	Mid	2	falling	pronoun (partitive)	'some'
/muan/	Low	zero (ø)	mid-level	noun-classifier	'enumerator for roll-shaped objects, e.g. cigarettes'
/múan/	Low	2	high	verb	'to roll something, to curl'
/phɛ̌ɛn/	High	zero (ø)	rising	noun	'flat, sheet-like layer, e.g. map, plan'
/phɛ̀ɛn/	High	1	low-level	noun-classifier	'enumerator for flat things e.g. sheets of paper, gramo-phone records'

Fig.1 : *Thai contrastive tone-pairs by referential sememes*

* Pitch contours given for the Bangkok dialect which might well be termed Siamese R.P. nowadays.

The number of examples is small, the set of morphological or grammatical properties subsumed under one tone is hard to imagine, and there is evidence that morpho-syntactical relationships are reflected in many ways other than those which might be supposed to lead to an embodied tone differentiation. Initial features such as prefixation or clustering as against simple initials clearly betray morphological family connections. For example, the trio:

/râap/ /kràap/ /pràap/

probably belong together, the first meaning 'flat, level, as of a plain', the second meaning 'to make oneself level, i.e. to prostrate oneself' and the third meaning 'to raze or level down', with its extended meaning 'to subjugate, subdue, wipe out'. What tone differences there are stem from the prefixing of the first cluster-element and not from any latent or embodied quality. To a Thai scholar they are quite regular and predictable. Final alternations also offer food for thought. The two words:

/hàak/ and /hàaŋ/

seem bent on preserving the same pitch contour and to show a relationship peculiarly by homorganic final alternation between stop and nasal. They both mean 'to be apart, separate' and are both verbs. Collocational distribution is the criterion for the use of one rather than the other. It is even possible to find something akin to ablaut -- in Thai it appears as quality -- in a pair like:

/râap/ and /rîap/

wherein the 'levelness' of the first seems to apply to large tracts ('no hills and hollows') whereas the second serves on a more delicate, aesthetic scale to mean 'without roughness or ragged bits', hence 'smooth, neat and tidy'.[5] The evidence for all this morphological activity on the segmental phoneme-strings is at least as sure and as plentiful as it is for tonal pairings. If tone has failed to swallow *some* morphology, then we may fairly ask whether we ought to believe that it has swallowed *any*?

A great number of linguists, therefore, are content to settle for the assertion that three basic tones existed. Obviously they must have been mutually contrastive, but the features that made them so have to be left unspecified. Thai orthography shows the three as unmarked (here termed zero, or Tone ø), Tone 1 and Tone 2, which is the reading code.

Distributions and comparisons

Turning back to earlier remarks about the difficulty of finding a 'common denominator' for the Thai basic tones, the hopelessness of deciphering any common feature by superimposing pitch-contour graphs showing the ways a certain basic tone is realized in different dialects has been admitted. In any array of such graphs it will be normal to find any name we might choose to give to a 'basic' contour made inappropriate by contradictions from among the various samples. In many dialects, the general pitch-blend may be high but there will usually be some dialects having the same basic tone low in pitch. Many samples of a tone may be rising in contour but some will be level and, perhaps, falling. Some, in another basic tone, will arch up, whereas other dialects will make it dip down or stay level. Whilst comparative procedures on every other level -- regular segmental phoneme correspondences and semantic identity -- guarantee that we are dealing with cognate words, the tonal correspondences are regular only in a statistical way. They might more properly be called frequencies of co-incidence rather than linguistic correspondences, for, as yet, no convincing argument has been brought forward to explain how a distribution such as the following might plausibly be found where no common linguistic factor can be abstracted under the vertical column entries other than the capacity for being distinctive when read across as sets listed in horizontal rows:

	Tone ∅	*Tone 1*	*Tone 2*
Dialect A	high	mid	low
Dialect B	mid	low	high
Dialect C	low	high	mid

Fig.2 : *Tone distribution by dialect*

Such changes could occur on any three contours -- rising, falling, arching -- but this conclusion demanded by comparative methods peculiar to linguists is dismally unsatisfactory to me. It is as if we were to agree that the phonemes of *quinque, pente* and *fünf* had to be corresponding because all three words meant 'five', but at the same time insist that further enquiry about the nature of the correspondences was inadmissable. Tonal comparison is still waiting for its *Lautgesetze* ('Sound laws, phonetic laws' (Ed.)): we are still relying on distributional tabulation in which the filling in of the entries seems to be more important than pondering the implications of our headings and listings. If they are not entirely arbitrary or heuristic, they must point to a line of attack -- a possibility for abstraction and reconstruction.

Input and in-built conditioning

Tone displacement has earlier been compared with the unfocusing of a kaleidoscope (p.12 above). The suggestion was made that the 'controlled input' (the distance the image was moved out of focus) would give a discrete code-type number, much handier to work with than a series of descriptions of the visual impressions received at each interval of viewing whilst the unfocusing was going on (overlap, darkening, lightening, neutralization and so forth). Can this comparison be applied to basic tones as well as derived -- that is, displaced -- tones? Clearly, it will only be viable if some simple factor (or factors, discrete and few in number) can be found that will be on a par with the initial-prosody feature (loss of voicing) which yielded such a complex stratification of pitch bands and contours in the different dialects. What possible input can there be, that makes a tone what it is, other than the pitch function of the vocal cords? And, if this is the case, what possible output can be monitored other than the pitch behaviour of the voice? This will arrive at the impasse we are at present trying to avoid, namely, the comparison of representational data obtained from pitch-measuring and pitch-delineating machines. We are back with our sheaf of pitch-contour graphs that never throw up a common factor.

However, recalling that the loss of voicing brought about tone-displacement, might we not rephrase the matter to our advantage by abandoning the concepts of phonemic or prosodic analysis and putting the emphasis on speech production as a machine-like operation? Could we not, for the sake of illustration, and some incidental amusement too, imagine a speaker in possession of his speech-producing apparatus rather like a driver owning a car? Let us make believe that a Thai speaker (of some 600 years ago!) takes his machine to the speech-mechanic with the complaint:

> My machine seems to be getting a bit out of control nowadays. All I have done is comply with the new law banning voiced plosion but this seems to be affecting the whole syllable. Moreover, other syllables seem to show signs of becoming infected, as it were, and, all in all, I cannot rely on it any more. I'm turning it in for an overhaul.

In the report of his findings the mechanic explains the fault and suggests some possible repair-jobs:

> The logistics for carrying out the articulation for starting syllables with a voiced plosion did not take into account the possible withdrawal of voiced plosion from the availability list; even with its legitimate goal removed, the programme went on as before, continuing to put in for forces and supplies appropriate to its old objective. Thus, the evenly fed air-pressure suitable for voicing throughout the initial stop became either (i) too great for unvoiced non-aspirates or (ii) too little for unvoiced aspirates. In both cases the *time* this air-pressure

was being extended was less (no air-movement, i.e. wastage, until the release of the consonant) than that scheduled for the original voiced plosive but, in the case of aspiration, the sudden air-loss after release could not be made up from the amount of pressure originally indented for. Moreover, the original lenis closure was too weak to withstand the relative increase in pre-release pressure. Without the accompanying voiced-setting of the vocal cords, as used to be the case in voiced plosion, there was no cushioning effect. The access of pressure directly to the bilabial occlusion demanded something more to hold it in than lenis tension. To compensate by ordering a change to fortis tended to confirm the new situation as similar to *normal* unvoiced aspiration. To compensate by having a simultaneous glottal release (keeping the pre-release pressure sub-glottal) confirmed the new articulation as a non-aspirate and allowed lenis tension to be maintained -- the glottis having borne the brunt of the work. The original, unamended pressure supply, however, still produced either a loud (overbreathed) vowel after non-aspiration or a soft (underbreathed) vowel after aspiration-release. As for possible adjustments, one alternative is to keep the original indent for breath-pressure and reduce overbreathing by lower pitch-resonators (cushiony absorption) and assist under-breathing by higher pitch-resonators (highly strung and sensitive). The other alternative is to scrap the old pressure order and re-indent for *normal* breath-pressures appropriate to the new situation but mark the 'old' syllables, say, as high instead of loud and low instead of soft.

Accompanying the report we may imagine three drawings, little more than histograms really, that put these findings diagrammatically.

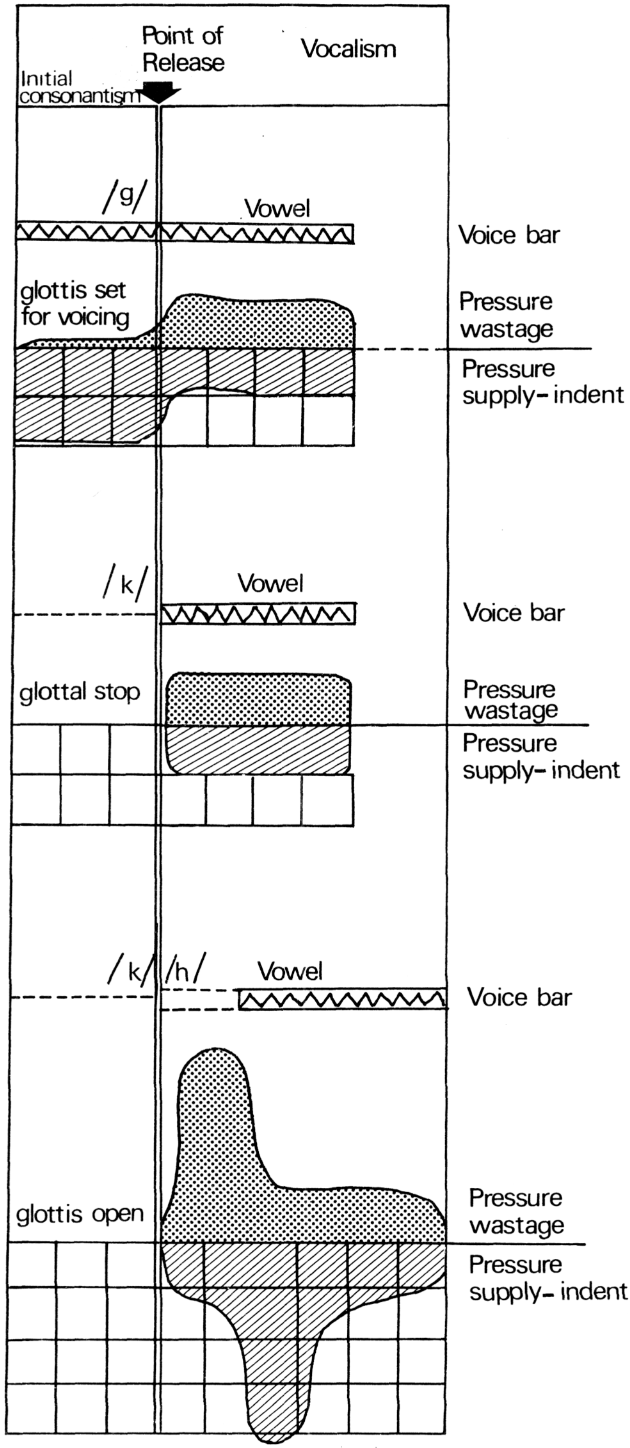


Diagram 2 : *Articulatory histograms in Thai*

So far this interlude, though made to seem like a visit to a garage or a consultation with a time-and-motion study expert, has at least got to grips with tonal displacement (acknowledged to be a derived phenomenon) and has uncovered a possible determining factor -- the amount and flow-type of breath-pressure 'put in for' -- that offered an explanation for the 'repair' reactions thrown up by different dialects facing the same problem. The rest of this disquisition will be devoted to an extended analogy, along similar lines, in an attempt to find simple determining factors for basic tones. Before going on with this, however, it is important to note that these factors are not directly susceptible to the recording instruments normally at the service of linguists. They are far from being phonemes and only on the fringe of acceptability to main-line prosodic analysis. They cannot be heard, logged on a graph or abstracted from the usual bundle of qualities or characteristics familiar to phoneticians. They are emphatically input-orientated concepts. They are not even the elements of a code, for if *ouput* is an encoded manifestation of message-units in specific detail, then there is an intermediate level at which an analysis of the components of the units might take place (distinctive features, say) which none the less remains an aspect of output. There is below all this an input level which concerns itself in a pragmatic way with rendering the code amenable to the medium of communication. It asks just what the medium can do and then puts in instructions to do this or do that. In a case like this, ironically enough, the medium *is* more important than the message.

It is as if one group agreed to analyse the Morse code no further than the listing of all available minimal units (the alphabet: here taken to be the phoneme inventory) without perceiving that another group was busy analysing the units, in turn, in terms of sound, silence, duration and sequence (here representing the factors analysable by prosodic analysis). For both groups the datum is the output -- in the strictly mechanical sense of the word, the emission of sounds. There is room for another group, however, who choose to deal with the obvious. Given the final objective -- some units of code to be transmitted that are known to us -- and given that all of the rules from the previous 'prosodic' abstractions are held over as conditioning environments for a performative skill, then the final, actual input instructions are reduced to two: now press the key; now release the key. With languages, tone languages no less than others, for sure, we are not dealing with the obvious. Nor are we dealing with such a clear-cut system as a code. Nevertheless, that hand on the Morse-key does remind us that we do speak with something other than our head: there is that necessary adjunct, the rest of our body.

The limits of 'Invention'

To repeat: from here on the intention is to offer an analogy between a machine and a speech production process, which might explain how many factors and what kind of factors might be held to be common to a basic tone and how a small number of input instructions might be listed, combinations of which might produce a set of basic tones. The gist of the argument will be that the lung-air pressure-controls are of overriding importance. Of these, the pair of intercostal muscles (or rather the instructions under which they are made to operate) will be suggested as primary controls, whilst the glottis will be held to function as a secondary control.

The machine, of course, will be an ideal one in the sense that it is devised to serve a specific purpose of illustration only. I should like to call it a naïve analogical model; naïve in the sense that it is plausible as a machine only as long as an engineer does not try to construct it. In other words, it needs to be consistent within its own terms but not efficient -- or even possible -- in practical terms. It will be analogical for reasons already stated. All that needs to be added on this score is the *caveat*: no analogy is perfect. It will be a model in the sense that it might predict patterns and processes if allowed to run on in exhaustive experimentation with its possibilities. Here, though, it must be admitted that it falls far short of the built-in inexorability of a mathematical model. This machine functions well only if oiled by charity!

The model: normal functioning

Imagine a tall, narrow pressure-chamber with a small outlet at the neck. Inside it is a piston with the rod passing through a pressure-tight seal in the base of the chamber. The piston-thrust comes from a mainspring beneath the pressure-cylinder. Effort is required to wind the spring (contraction). Some sort of switch is required to bring power to provide for the contraction of the spring at the proper time. Similarly, a trip mechanism will be needed to cut off the power and release the spring at the proper time. Therefore, on completing the up-stroke, the piston-head touches the shoulders of the chamber and this contact must be imagined as the signal calling for work to be done on contracting the spring and drawing the piston down. At the lowest point of the down-stroke, the contact between the chamber-base and the underside of the piston-head must be thought of as the instruction for work to cease, thus allowing the spring to unwind (release) sending the piston back up the chamber again. Notice that the spring is not programmed for stopping or restraining itself of its own accord. There is no way, for instance, of stopping or changing direction whilst the piston is in a mid-chamber position.

Such a simple regime will not work, however. One obvious necessity is to allow the piston to be drawn down without generating pressure underneath itself. Without some provision for such a contingency, the piston-head would never reach the base of the chamber at all. A valve in the piston-head must be supplied such that sensitivity to pressure from below causes it to open in order to facilitate the downstroke. Pressure from above has the opposite effect, ensuring that the piston-head fills to the maximum the chamber-space and fits flush with the walls. This is like saying that the piston-head has two states, expanded (which we now call **P**) and contracted (**p**) and these two states are *normally* coincident with the upstroke and the downstroke respectively.

Short reflection will bring us to the conclusion that such a pressure-pump still would not work, however. Having overcome the difficulty of pressure generated by the downstroke, we now face the difficulty of a vacuum behind the piston-head at the beginning of the upstroke. Assuming a pressure-tight fit for both the piston-head and the hole in the base through which the piston-rod passes, it will be seen that an equilibrium between the mainspring's effort and the braking effect of the vacuum in the piston-head's wake must cause the arrest of the upstroke after a very short time. Precisely at this point, just where the piston-head is stopping, another valve must be imagined in the wall of the chamber, tensed so as to be opened only by maximum negative pressure. Upon its opening, the rarefaction is relived, the piston proceeds upwards, the valve closes once more until further rarefaction retards the piston, opens the valve again, and so on until the piston has risen the whole height of the chamber. This crude, mechanical counterpart of a feedback control represents the means whereby the power of the mainspring is moderated. It prevents, as it were, our pump 'expiring' with a bang.

Clearly, the positioning of the valve in the wall -- its distance up from the cylinder base -- is just as important as the degree of its sensitivity to negative pressure. Both in fact are interrelated because both depend ultimately on the force of the upthrust generated by the mainspring. The positioning and the degree of sensitivity are empirically determined, we might say, always remembering, however, that we are dealing with an ideal machine. We can also say, be it noted, that we can restrict our symbols to the **P**, **p**, a pair **W**, **w** (standing for Wall-valve) with the two states valve-shut and valve-open respectively, and a position **c** (which will be held as 'crucial' for the syllable both in quantity and in tone) at an empirically determined distance up from the chamber-base. We have no need to refer to the operation or the force of the mainspring at all, as long as we are dealing with normal performance.

The downstroke is simply **pW**. The upstroke is **PW**, except for **Pw** at **c**, then a further series of **Pw** intermissions within the general **PW** situation until the state **pW** is again reached.

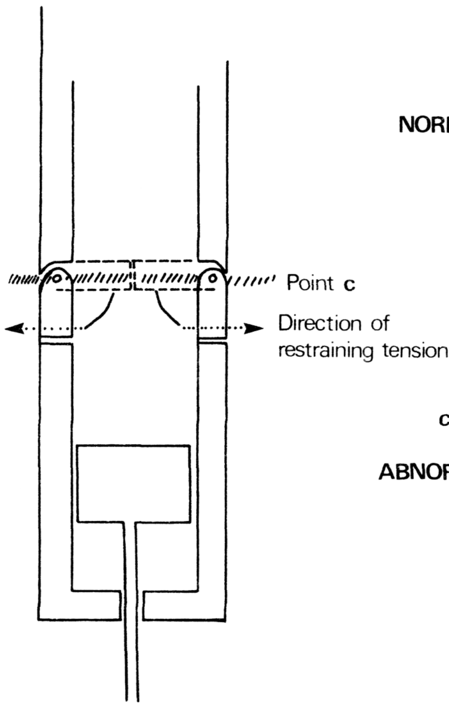
Thus far the model should have created the impression that its working is virtually automatic. The supply of power to the mainspring should be taken for granted as should also the right degree of tension or sensitivity in the valves. In other words, no 'deliberate' instructions need be given to carry out any stage of the piston's operation. In electronic terms, everything so far has been built into the circuit: there is as yet no need for a control panel nor for a choice of settings on the part of an 'operator'.

The analogy we have made is of a machine to represent the normal breathing cycle, indulging our tendency to the naïve in order to make the point that, so far as speech production is concerned, the breathing apparatus and its functioning must be considered a datum: *we postulate this to begin with. Whether inspiration and expiration are, in fact, totally conditioned reflexes or innate systole*, or not, *is for the physiologist to* say. All we assert here is that, for our purposes, the process might as well be taken to be automatic and self-perpetuating. So far we acknowledge no input.

Abnormal function

We ignore two possibilities. First, that the piston-head valve might fail to open on the downstroke. This would allow of no inspiration of air into the lungs for the simple reason that the piston would not overcome the pressure built on behind itself (non-function). Secondly, we ignore the possibility that the mainspring might fail to respond automatically to its contract/release cycle of instructions (laboured and self-conscious breathing). Neither do we contemplate deliberate interference with these automatic reactions. (Our machine is ideal: breakdown is impossible. Only deliberate contravention of the rules is imaginable.) We are left with the chance that the two valves might 'disobey' their control-settings: they might be ordered to 'refuse' to follow the promptings of their in-built sensitivity. As we have seen, the failure of the vacuum-relieving valve in the chamber-wall would lead to the arrest of the piston. Diagrams 3 and 4 will make it clear how crude our valve is imagined to be. They will also make clear, it is hoped, that no matter whether the abnormality is in the overruling of the valve's sensitivity, i.e. obstinate, complete closure leading to lack of feedback -- just as if there were no valve fitted in the chamber-wall at all -- or whether the flanges of the valve simply swing loose and uncontrolled, i.e. a temporary cancellation of all response and restraint, the result must be envisaged as the same: the arrest of the piston.

In the latter case, of course, the flanges merely swing shut in the wake of the piston's passing and are held there by the rarefaction occasioned by the piston's attempt to continue upward. Thus, the overall result of abnormality in the valve at **c** will be to terminate the upstroke prematurely: to prevent the succession of 'feedback' **Pw** intermissions characteristic of normal performance once **c** has been passed. This, therefore, is the malfunction that will be imagined as a deliberate device to attain syllabicity: one input and one event only at **c**.



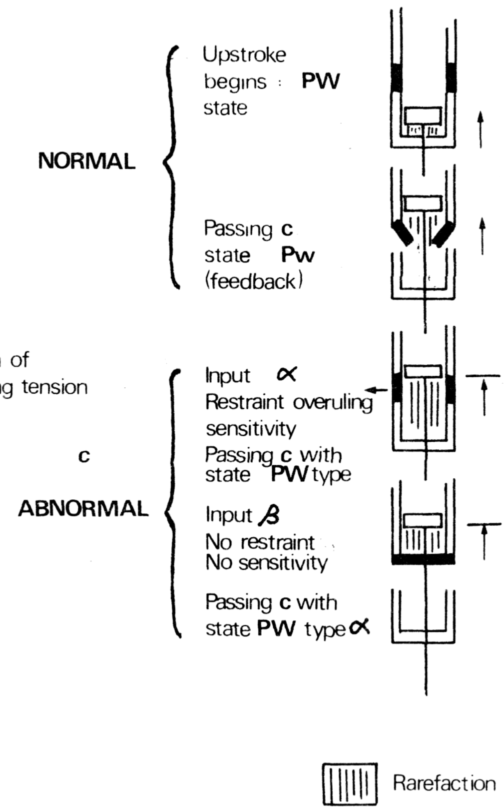


Diagram 3 : *Speech valving*

Diagram 4 : *Speech valving*

Turning to the alternative source of abnormality, a refusal of the piston-head valve to fill the chamber-space on the upstroke would lead to multiple disaster. There would be a proportional drop in the pressure of air pumped out of the chamber: the piston would 'leak'. Worse still, the vacuum brake effect would be absent. The piston would hurl itself upward with the full, unrestrained force of the released mainspring. Two safeguards may be postulated to prevent this occurrence. First: suppose there is a rapport between the two valves. Suppose that the piston-head is conditioned to expand fully on the upward passing of point **c** no matter what the pressure situation be above and below the piston-head. This would reinforce the expectation of normal conditions on the upstroke. It would be a sort of double-check. A cricketing enthusiast, for instance, will always value a good stumper, particularly when the bowling is mediocre. Our machine, likewise, leaves little chance for 'byes', even with faulty 'delivery'. Second: if even this safeguard is deliberately flouted, we can postulate -- rather like a long-stop in cricket -- one last resort: the emergency measure of peremptory closure of the neck-outlet, thus forcing up the internal chamber-pressure and steadying the perverse piston-head in its upward career. (If you can put a finger over the hole of a bicycle pump it will act out the effect of this emergency stop.) This procedure will later be interpreted as a stop-final syllable of a certain kind. Once again, though, the passing of point **c** will be the triggering off of this emergency reaction. The built-in programme would read:

> 'Any lack of evidence of an abrupt reversal of pressure at **c** (positive to negative) means the piston has bolted and the message to stop up the air-outlet will be sent immediately.'

The rules at **c** are ordered ones, then, and can be expressed thus:

Rule 1. On passing **c**, if state **P**, no instructions (= normal);

Rule 2. On passing **c**, if state **p**, first demand **p** → **P**;

Rule 3. On passing **c**, if Rule 2 disobeyed notwithstanding, then demand ʔ (emergency glottal occlusion).

The analogy here is obvious: the two valves and the rapport between them represent the intercostal pair of muscles and the feedback relationship we suppose to exist between them. The emergency 'long-stop', of course, is the glottis. Again we rely on our admission of naïveté to excuse the obvious shortcomings of the analogy in other respects; a cylinder with a piston and a side-wall valve makes a very odd looking lung!

This is really beside the point, however. What the machine is doing is to represent the counterpoise and interplay of two valves which determine pressure supply as being analogous to the work done by the intercostal muscles. We further arrange for these two valves to operate under abnormal conditions and perverse instructions in order to generate a deliberate syllable rather than an 'unconscious' systolic process. Our model emphasizes here that the input-points for discrete, deliberate control are the intercostals. The analogy implies that evidence of 'syllable pulse' at the diaphragm or elsewhere is secondary, 'referred' or, more bluntly, incidental. No detailed physiological substantiation of these claims is adduced because the analogy aims only at the *minimum* number of factors of primary importance in respiration that are in any way viable for syllable generation. Physiology may later tell us that there are more. What it tells us now (Greene 1964) is that there cannot in all common sense be less. Three emerge as irreducibles: the two intercostals and, in certain circumstances only, the glottis.

Syllable types

If we work out the possible combinations of our three deliberate-input chances, **P** or **p**, **W**-type **α** or **W**-type **β**, and the presence or absence of **?**, we must be careful to observe the limitations of our model. **P** for instance, will never co-occur with **?**. It is best done by diagrams showing the states of the valves (and the glottis) and the input-instructions to be applied at the key points in the piston's career. The one symbol not found in our alternatives, namely **w**, can serve, when written in square brackets, to mean an intervening open state of the wall-valve at **c** causing a momentary relief of rarefaction. This, for syllable production, will only occur in course of proceeding from **PW** after **c** (Abnormal type **α**) to **PW** after **c** (Abnormal type **β**) or vice versa, i.e. the flapping over from one shut position to another. The symbol cannot now occur as a simple **w** in its own right: **w** occurs *normally* in this way (i.e. not *abnormally*) during the upstroke and hence is out of consideration. Syllables are abnormal and will not tolerate a **w**-setting which is to preserve automatic systole.

Several observations might be helpful. We can, for instance, put in instructions only with respect to the actual start of the upstroke and to the passing of point **c**. We cannot, for instance, expand the piston-head during its course (**p** → **P**) without bothering to enquire what triggered off the sending of such an instruction at that time. By the same token, we cannot fly in the face of reason (even with our ideal machine!) and, say, open the wall-valve right at the commencement of the upstroke. Thus, if, in the diagrams, we read an entry 'Input Abnormality **β**' viz. no restraint or sensitivity at all at **c**, this does not mean that the valve *opens*: it means that it goes where circumstances (in this case, pressure) dictate that it should go. It should also be remembered that, whilst generating a syllable is a result of abnormal functioning, yet there is within the abnormality a degree of regularity. All syllables have to begin with a **W** state. They all have to end with a **PW** state. These appear to be normal. There will always be, however, some underlying, 'latent' instruction (α or β in our diagrams) compelling the wall-valve to act abnormally as soon as circumstances allow (i.e. when the piston passes point **c**). This co-existence of normality and abnormality is not surprising: though we want our machine to work oddly, we also need to have it work at all.

The diagrams

Only the Greek letters α, β and Π are actual *inputs*.

P means the state when the piston-head fills the chamber space.

p means the state when the piston-head is contracted and therefore 'leaks'.

α means the input order: 'Wall-valve performs abnormally in type α', i.e. 'Now shut!'

β means the input order: 'Wall-valve performs abnormally in type β', i.e. 'Be flaccid - do nothing'.

Π means the input-order: 'Contract the piston-head'.

2 means Rule 2 for **p** $\rightarrow$ **P** is applied here (a built-in emergency circuit).

3 means Rule 2 has failed, therefore demand **?** (an alternative circuit).

I^1 means the first stage or opportunity at which input can be fed in.

I^2 means the second stage where input can be fed in.

i means the initial run of the piston up to point **c**.

e means an episode to allow any operations directly occasioned by the passing of **c** to come into effect.

t means the tail-end of the syllable; in particular, the upshot of I^2 input at **c**.

	I^1	i	C e	I^2	t	
Piston	β	P	P			Syllable (i)
Wall valve		W	[w] → W type β			
Glottis		O	O			
Piston	β	P	P		P	Syllable (ii)
Wall valve		W	[w] → W type β	α	[w] → W type α	
Glottis		O	O		O	
Piston	α	P				Syllable (iii)
Wall valve		W				
Glottis		O				
Piston	π	p	p	2	P	Syllable (iv)
Wall valve	α	W	W		W	
Glottis		O	O		O	
Piston	π	p	p - - - - -	2	- - - - - - ▶ P	Syllable (v)
Wall valve	β	W	[w] → W type β		W^{β}	
Glottis		O			O	
Piston	π	p	p - - - - -	2	- - - - - ▶ P	Syllable (vi)
Wall valve	β	W	[w] → W type β	α	[w] → W type α	
Glottis		O	O			
Piston	π	p	[see relevant	3	[p] → P?	Syllable (vii)
Wall valve	α or	W	type]		W either type	
Glottis	β	O	O		?	

Diagram 5 : *Syllable samples*

From the entries seen above it could be agreed that a handier description of each syllable-type would be in terms of our chosen symbols rather than a mere enumeration (i), (ii), (iii), etc., so here are their suggested 'names'. Reference to the diagrams will, it is hoped, again provide the commonsense argument for choosing the naming symbols:

(i)	will be called a	**Pe** syllable;
(ii)	will be called a	**Pt** syllable;
(iii)	will be called a	**Pi** syllable
(iv)	will be called a	**p**α syllable;
(v) & (vi)	will be called a	**p**β syllable;
(vii)	will be called a	**p**ˀ syllable.

Already it looks as if quantity might feature as a by-product of input-combinations. **Pi** looks very short; **Pt** and all three **p**-syllables are long. A closer look -- and a stricter argument -- would lead us to expect three lengths (as demonstrated by **Pe**) and not the usual two, but this argument can be countered by stressing the nature of point **c** as a crux: quantity would be a before-**c** / after **c** dichotomy and not a question of linear (chronological) measurement. On our diagrams' showing, therefore, **Pi** is short and all other syllables are long. It must be pleaded that one of the virtues of our machine is that it satisfactorily (at least, analogically) explains why quantity is universally a two-way split wherever it is phonemic. There are no 'medium-lengths' or super-longs or super-shorts phonemically because each syllable has but one crux. Quantity -- in the behaviour of the piston with respect to point **c** -- becomes a discrete event rather than a relative scale of performance. The relative aspect, as Daniel Jones (1967: 124-34) saw in his *chronemes*, will always be in evidence but it will always emanate from the *fact* of quantity. Objectively recorded length seemed to Jones to be an accidental attribute of quantity but not an inherent definitive property of it. Phonemic 'length' and phonetic 'length' seemed somewhat at odds. It is not too bold to claim that our machine resolves this paradox: 'length', after all, turns out to be a question of whether something happened or not and is irrespective of whether the speaker took a long time over it or not.

The diagram as it stands, however, gives a slightly misleading impression. We must not mistake the length of a list for the length of time or distance that a piston moves in a chamber. While events are of prime importance, duration, as we see, remains a factor to be considered.

Graphs

We cannot let the diagram of input-combinations be the sole source of our impressions about likely syllables that are generated by our machine. We can, with reason and imagination, re-draw the tabulated information as graphs which will show some linear manifestation of the effects of the chosen input in a certain syllable. Time is an obvious 'dimension' for such a graph but, as we are dealing with a moving piston, so are distance and speed. Since our machine is an analogy of syllable production, we must also not lose sight of the continuum of pressure. After all, our model looks like a pump and works like a lung! There is going to be an outflow of air at the chamber's 'glottis', the regulation of which is bound to be central to the problem of syllable utterance and, it will be argued, central to the characteristics of syllable tone too.

Perhaps the most straightforward and informative graph would be one of speed of piston as against location of piston. **P** will be the state in which the piston is slowing (it is 'striving'); **p** will be the state in which the piston is hurtling along unchecked; and [w] offers the opportunity for the decelerated piston to win back some momentum during the brief relief from the braking action of rarefaction.

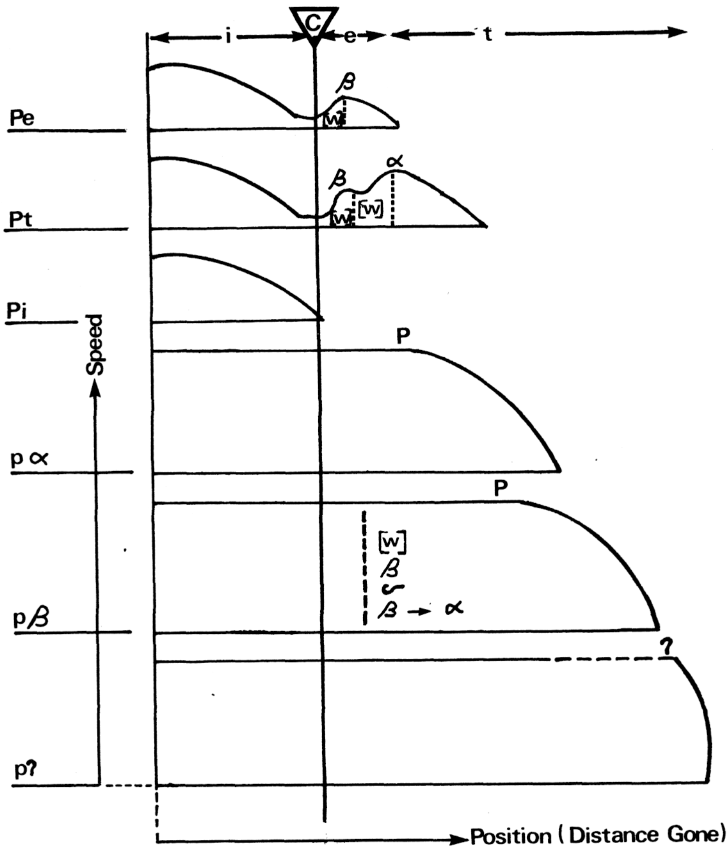


Diagram 6 : *Graphing of syllable samples*[7]

Side by side with the speed graphs we can place an imaginative estimate of corresponding pressure-readings for air being pumped out of the chamber. Note that a high speed (state **p**) gives a lower pressure reading than deceleration (state **P**) for the beginning of each peak. Apart from this, however, the pressure curves correspond generally in outline to the speed curves.

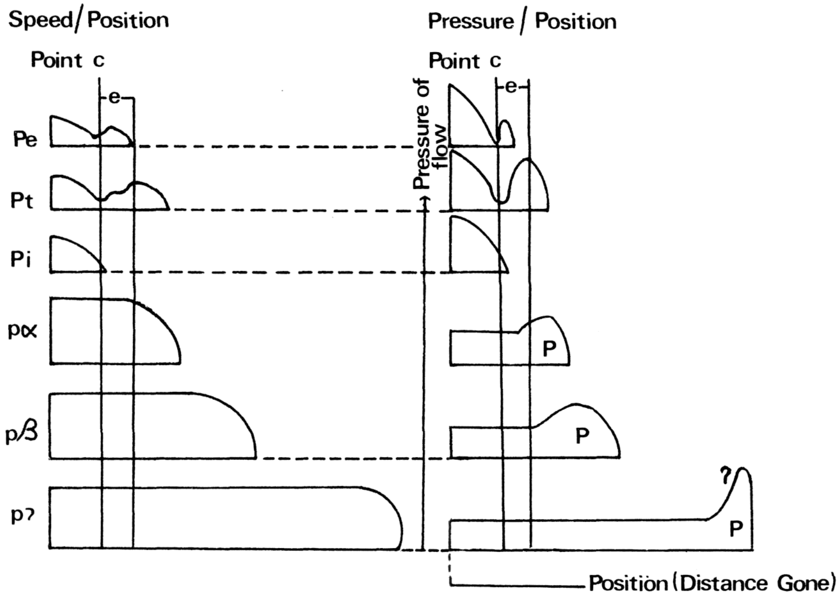


Diagram 7 : *Comparison of speed and pressure curves*

Modulations

The striking aspect of the graphs (Diagrams 6-9) is that they show such great disparity in overall lengths. The piston, in other words, travels only a short distance in some but a very long distance in others. Now, whilst the shorter distances will tend to take the longer times and the long distances in **p**-state will be covered extremely rapidly, thus tending towards isochronism*, some adjustments must be envisaged as necessary to achieve isochronism without such a wastage of space. Uniform syllable length will then be true of time - a convenience for the speaker - and also of distance - a convenience for chamber-capacity.

A modulation, a 'tempering' of the speed and pressure rates will take place so that the speaker may concatenate syllables with some degree of fixed but natural rhythm. The tail must be shortened, i.e. slowed down or, if felt convenient, the slow speeds in **P**-state round about the crux may be accelerated a little. In this fashion, not only will isochronism be the syllable norm demanded by the speaker but there will be a driving motive to get the crux (point **c**) placed in the middle of all syllables (except, of course, **Pi**) rather than to one extremity or the other as shown in the graphs in an unmodulated state.

A comment must be inserted here to answer the query as to whether isochronism *is* in fact a norm for the syllable. Firstly, it must be re-emphasized that syllable quantity is not at issue here. For short syllables specifically, we can either resort to **Pi** syllables, or have the point **c** in *any* type of syllable represent a trigger which fires a message off to stop the syllable there and then by glottal or oral occlusion or partial occlusion (consonantism). It is assumed, though, that there will be a tail-juncture (silence or held-over consonantism) to even up the rhythm of short-stop-finals ready for the down-beat which ushers in a new syllable-start. Secondly, it must also be pointed out that isochronism does not rule out overall speed variations due to idiosyncratic style, emotional state or syntactic intonation-patterns. It merely means that, all other things being equal (viz. the things just mentioned), the natural tendency of the human body's functions is towards isochronism (walking, breathing, sleeping/waking metabolic rates, the heart-beat, etc.). Between two peaks in a repetitive wave-trace there will be a striving towards maintaining a steady wave length although item-distribution within the wave length is not, of course, rigorously scalar. A cycle is symmetrical in its repetition; it is not a repetition of symmetrical units. So, if we do not take this for granted, then the deliberate or conditioned variations aimed at in stylistic effect will be rendered unsubtle or even null and void -- a mere accretion of spasmodic rhythms because not offset by the expectation of a 'line' of homogeneous 'feet'.

---

* Isochronism is the character or property of oscillating or of taking place in equal spaces of time, therefore equal in duration or in intervals of occurrence. (Ed.)

Time discrepancies, then, can not have been the missing factors that were embodied in our basic tones. To reconstruct hypothetical phonemes of infinitesimal duration-differences will not do. Not only would they be beyond the limits of natural performance and perception but they would also go against the best interests of a speaker who seeks to use a comfortable and convenient reference-frame in which to assess the effects of *rallantando* or *accelerando*.

Modulation in the model

Just before this brief discussion of isochronism, the likelihood of a need for pressure adjustments to slow down piston-speed and on occasion, perhaps, to increase it, had been suggested. It now remains to supply some sort of sensitive machinery for doing this. To resort to further adjustments of the two valves in our model will get us nowhere. The whole point of *abnormal* functioning was to have minimal discrete instructions (α, β, Π) in direct opposition to the *normal* homeostatic balance between states **W** and **w** in the valve at **c**. We must at all costs preserve the peremptory 'open-or-shut' case for the valves' operations because they are being put forward as the prime factors, the real input-conditioning for our basic tones. The only other site of pressure control we can appeal to is at the outlet -- the glottis.

Once already the outlet has played a role in the abnormal functioning of our model -- in the **Pt** syllable whereby an abrupt closure of the outlet on emergency instructions from the valve at **c** (Rule 3) stopped the runaway piston. If the outlet could be modulated in size and shape so as to influence the extrusion-rate of air pumped from the chamber, then we could once again 'cushion' the piston's thrust, slow it down, or, by widening the outlet we could help the piston on its way. The outlet would thus have to be sensitive to pressure and -- yet another tendency to homeostasis -- would be programmed to operate in favour of isopiestic* restraint. The attempt would be made -- glottal tension and configuration is the last resort in this matter -- to control the outflow of air so as to achieve fairly uniform pressure during the syllable for the sake of even voicing.

Glottal modulation

We have now arrived at a state of affairs not unlike

---

* Isopiestic means 'denoting equal pressure'. (Ed.)

the tone-displacement situation in which, it will be remembered, 'the withdrawal of voiced plosion' from the consonant-input list left an inappropriate breath-pressure and flow. The solution suggested there was to re-interpret loudness and softness as high or low pitches. This was brought about by a choice of two adjustments. Here the case is not one of inappropriate pressure but one of intra-syllabic fluctuations in pressure (depending on what syllable type is chosen) which must be evened out. The delicate sensitivity and the speed of response of the vocal cords make them ideally suited to be the agents for modulating pressure-fluctuations of each syllable-type. There will, presumably, be many solutions for each type. Some will reflect the pressure-variation directly as a pitch contour whilst adjusting the voice *quality* to counterbalance the lapses from steady pressure-supply. There is a suggestion here of the 'clear' (*qīng* 清) and 'muddy' (*zhuó* 濁) terminology in Chinese, though their use in that language is clearly applicable only to derived tones (post-tonal-displacement) and not to basic tones. Could it be that the 'inner' and 'outer' divisions of traditional lexicography reflect a Chinese typology conditioned by **P**-state as against **p**-state syllables? Be that as it may, an imaginative portrayal of such modulations might be as follows:

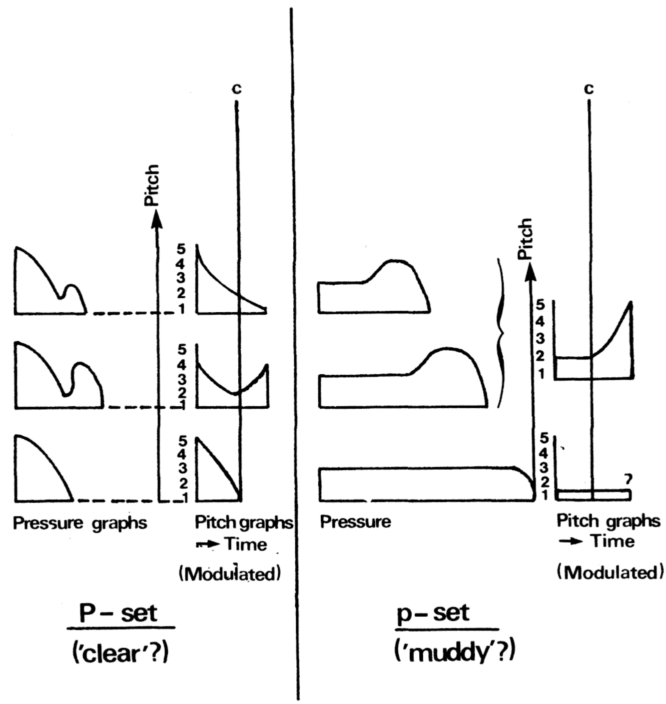


Diagram 8 : *Glottal modulation*

Some will use the vocal cords to counterbalance the pressure fluctuations by *pitch* alone and, to some extent, produce a mirror-image of the pressure-curve thus:

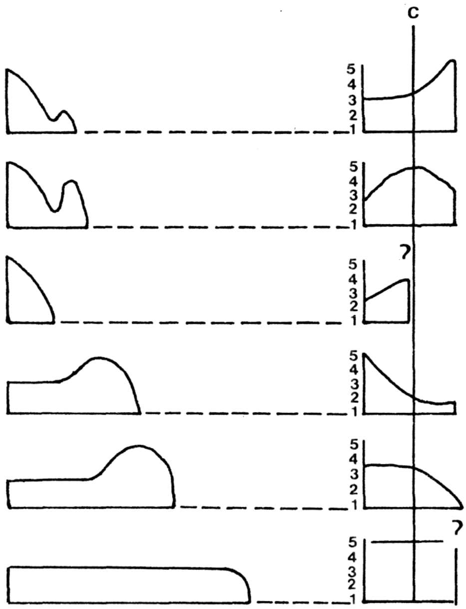


Diagram 9 : *Comparison of pitch and pressure curves*

Without detailed physiological knowledge, abstracted on a multi-dimensional diagram or graph, it would be rash to press further with an account of how the human body's glottal equipment might be reduced to terms amenable to our model. Indeed, it will be recalled that pitch and pitch-contours were acknowledged as being subtle, non-discrete and difficult to account for. This very admission was one of the main motives for imagining a model to investigate the possibility of less complex sub-glottal factors in tonal analysis. We have succeeded in describing a model with three such factors (input instructions) and, not surprisingly, it has in the long run *led us back to the subtleties of glottal activity.*

Conclusion

In an attempt to escape from the impasse of unanalysable recorded data (pitch phenomena) to account for basic tones, we looked for a few key factors. Granted that our machine analogy is informative, then these will be found in the relationship between the intercostal muscles as they undergo adapted programming of the work of expiration. We might expect to find in tonal languages of the South East Asian type not *one* favourite arrangement for this intercostal relationship habitually relied upon to produce any syllable but a *choice* of input instructions (α, β and Π in these diagrams) yielding a selection of syllable-types. Inherent anomalies of duration and pressure were imagined which had to be modulated into a pressure supply that the vocal cords could use to maintain a normal vocalism and syllable-quality. This modulation was done at the sacrifice of freedom of pitch. In these languages, then, pitch became contrastive between syllables of different types: it became an exponent of basic tone. The particular contour and pitch-band adopted to render the syllable 'comfortable' to the speaker's expectations of isochronism and isopiestic breath supply varied from language to language and dialect to dialect, but it is *implied that the sub-glottal input did not. These intercostal* syllabification schedules emerged as the common, stable basic *tone factors. Their functioning* involved the glottis and, consequently, pitch phenomena.

Corollary

Can these speculations point to new departures in practice? Certainly, they re-inforce a demand for more rigorous physiological investigation of the human body during the activity of speaking. To prove whether the analogy presented in this paper was right or wrong or merely irrelevant, the nerve-impulses sent to the intercostal pair of muscles would have to be monitored and interpreted alongside a record of lung pressure and emergent (oral) pressure and glottal activity. Both voice quality and glottal occlusion must be seen as likely concomitants, if not definitive characteristics, of certain tones. Some analyses should be attempted which treat syllable-final, quantity, and voice quality (creak, reediness, final 'crumble', final 'strangulation', etc.) as at least the co-equals of pitch phenomena and rightly to be subsumed under the single rubric 'tone'. Field workers must therefore listen again -- or, perhaps, record again with higher fidelity -- and be on the look-out for such features. They must also be prepared to dissect their recorded pitch-contours into **i** and **t** sections and to note the presumed occurrence of **e** at **c**, especially if the contour shows a curve or a hump. By doing this, they might find more stability between dialect samples of a tone in one section, say, **i**, than the other.

Most of all, though, it is hoped that model-imagining and combinatorial tabular systems will encourage attempts to see larger 'wholes' in human utterance. If we are happy now to have grappled with the syllable as a whole because of our preoccupation with tone, perhaps we shall later be less daunted by the challenge of rigging up a model in the mind to deal with concatenation of syllables in a phonologically and physiologically integrated way.

NOTES

This paper was originally presented by Peter Bee several years ago to a staff seminar in the Department of Phonetics and Linguistics of the School of Oriental and African Studies of London University. It was rediscovered after his death in May 1982. On my request to include it in, and edit it for, this collection, Peter's family kindly gave me permission to prepare it for publication. Now, it can reach a wider public. (Ed.)

1. I am greatly indebted to J. Marvin Brown (1965) in which tabular presentation of Tone-contours is particularly compendious and clear.

2. Included in this term are zero-final syllables, the decisive factor being whether the syllable closes *abruptly -- by an occlusion of the air-flow, a hiatus in* the pressure gradient -- or not.

3. The word 'band' here is used to avoid ambiguity: 'level' might be construed as a 'flat, even contour' whereas 'band' allows for contours to be of any shape so long as they are within the width defined.

4. Before the stabilization of effective, centralized statehood under Ayutthaya, the dialect groups are hard to recognize and even harder to locate with any precision. It would be rash to suppose that the present inhabitants of Sukhothai speak anything like the same dialect as that, say, of King Ramkamhaeng in the thirteenth century.

5. Disyllabism (syllabic infixation) has been omitted from consideration since it is most likely to be a trait borrowed from Mon or Khmer morphology.

6. Syllables (v) and (vi) are so similar as to be classified as minor variants of the general input-family Π β, while syllable (vii) is so tabulated to avoid uninformative repetition. The point is that with **p** throughout, it does

not matter what the wall-valve is doing. The final **P?** in this syllable (column **t**) means that the state **P** is consequent upon the closure of the glottis. The piston-head is *forced* to expand by overhead pressure.

7. **pα** and **pβ** are similar. Whichever pattern of input instructions is chosen, the difference is simply one of *when* the curve begins to go down. The **pβ** descent is obviously slightly later (or 'further') than **pα** because of the additional [w] intermission, but there is no resemblance to the curve and distance of **p?** in the least.

REFERENCES

Brown, J.M. 1965. *From ancient Thai to modern dialects.* Bangkok: Soc. Sci. Assoc. Press of Thailand.

Greene, Margaret C.L. 1964. *The voice and its disorders.* London: Pitman Medical Publ.,2nd ed. (1st ed. 1957).

Jones, D. 1967. *The phoneme: its nature and use.* Cambridge: Heffer, 3rd ed. (1st ed. 1950).