PART ONE

Fundamental Techniques

In his essay *Sprech und Gesangsschule (Neue Vokalpratiken)* (School for Speech and Singing (New Vocal Practices)) (1972), the German composer and theologian Dieter Schnebel provides a thorough and innovative investigation into the production, usage and latent possibilities of new vocal techniques. Schnebel comments on the inherent limitations of conventional vocal performance practices and notations and their relative inaccuracy and unsuitability when applied to extended vocal techniques. He goes on to outline the main physical mechanisms and processes involved in the production of vowels and consonants, suggesting new ways in which to expand the range of attainable sounds through the combination of various mouth (vowel) positions with different noise (consonant) sounds. He discusses the importance of breath control in the shaping of these sounds, as well as the actual use of various modes of breath as compositional material, stating that 'Audible and inaudible breath processes would, so to speak, form the basis of the artistic processes... Breath, crossing the line of audibility, becomes art itself, built from shaped fricatives'[1]

However, such techniques are seen not only as processes with which to create new sounds but also as a means of liberating both the voice and the consciousness of the performer. Schnebel writes: 'Such operations lead, so to speak, into the sound production itself. Taken further, they can themselves attain form, without the overlapping unities like words or even sentences being adjusted. That is, the articulation process becomes the object of the composition...This demands a conscious knowledge of what is happening with the articulation'[2] With reference to his composition *Maulwerke* he adds: '...the content is expressed no longer mediated through the vehicle of a still rudimentary text, but rather directly. Speech as content arises through articulation' (456).[3] The necessity of understanding the physiology of vocal production is repeatedly emphasised. Since music utilising such techniques no longer has the benefit of a text with which to stir the imagination into a realisation of the articulation process, bodily control 'will become no longer intuitive but rather will be applied consciously and methodically, with the systematic result of the most varied derivations of speech or song'[4]

The *Index to a Recorded Lexicon of Extended Vocal Techniques,* published by the Center for Music Experiment and Related Research, University of California in 1974, lists seventy-four techniques classified under three main headings:

Monophonic	Reinforced harmonics, Whistle stop, Ululation Fry, Shake, Flutters, Voiced whistle;
Multiphonic	Chant, Glottal overpressure, Forced blown (inhaled and exhaled), Multiphonic buzzes and squeaks;
Miscellaneous	Buzzes, Squeaks, Clicks, Belches, Death rattle, Car crash, Buccal speech (+ finger), Bat sound, Glottal speech, Tongue squish, Tongue-teeth slap, Cricket sound, Water drops.

Unfortunately, the lexicon does not explain the methods of production for the majority of the sounds listed. While many can be discerned from the accompanying taped examples several sounds remain tantalisingly elusive to the listener. Nor does it classify the techniques in a comprehensively scientific manner, but tends to use aurally descriptive titles rather than those based on the internal morphology of the sounds (this highlights the problem of detailing ostensibly scientific and physical phenomena in a manner both understandable and useful to musicians).

The American composer Richard Jennings published *A Catalogue of Extended Vocal Techniques* in 1977 detailing sixty-seven techniques, mostly based on those listed in the CME lexicon but organised according to their basic sonic source. They are:

WAMP	Straight or Nasalised, Reinforced harmonics, Ululation, Cross-register Yodel, Uvulation.
Fry	
Breath	Coloured Breath, Snore, Choking, Buccal, Belches, Breath through nasal passages, Buzzes and squeaks, Whistles, Finger controlling coloured breath.
Body	Clicks, Smacks, Tongue flop, Tongue slap (bat wings), Taps;
Combinations	WAMP-fry, WAMP-body, WAMP-snore, Fry-body.

Trevor Wishart's *Book of Lost Voices* (1979) is a comprehensive examination into all aspects of extended vocal techniques resulting from four years research into the subject. Written specifically with the performer in mind (it accompanies his *Anticredos* (1980), for six amplified voices) it also offers practical information on how to produce the sounds. Of his research Wishart (1980) has written:

> *As a composer, my motivation for researching extended vocal techniques was a desire to be able to transform sounds of totally different timbre and pitch content into one another in a continuous process. Conventional musical instruments, constructed on the assumption that timbre should be held (relatively) constant, are unsuited to this aim. The human voice, however, is particularly suitable as an 'instrument' both because of its amazing flexibility and variety of sound generation ands its direct accessibility (compared, for example, with computer synthesis). As a composer, therefore, I was concerned to explore and categorise the range of sounds the voice can produce, regardless of any traditional assumptions about what is and what is not 'musical'. Pursuing my research has led me to re-examine traditional assumptions about the 'parameters' or 'internal architecture' of sounds. In particular I have had to revise notions of pitch and of the limits of a sound Gestalt (e.g. coherent sound object with unstable components).*

The Book of Lost Voices lists 169 items arranged under the following headings:

> Reinforced Harmonics; Mongolian Harmonics; Subharmonics; Break Multiphonics; The Larynx; Whistling; Lip and Cheek Vibration; Tongue Vibration; Noise Sources; Subaudio Vibrations; 'Grit' Sources; Complex Sounds; Intermodulation; Multiplexing; Articulation; Multiple Articulation; Manual Aids; Flutters; Clicks; Sounds Related to the Consonants K, T, D, P, B; Short Sounds; Combinations; Manually Assisted Short Sounds; Plosive Noise Bands.

In addition the notes to *Anticredos* provide notational equivalents to almost all the sounds listed along with numerical reference numbers. In 1985 Wishart considerably expanded his findings in 'The Human Repertoire', chapter 12 from *On Sonic Art* (1985, 1996). Along with 170 recorded examples he reclassified the sounds under the following broader headings dispensing with titles such as 'constipation multiphonics' and 'monkey' in favour of a less suggestive but clearer nomenclature.

> Oscillators and other sources; Filters; Noise; Double and treble production; Air-stream and other effects; Water effects; Transformations; Inhaled sounds; Pulses; Transitionals and percussives and Multiplexes and complex articulations.

What follows is a detailed explanation of certain fundamental techniques and their production. This list is based on an amalgamation of the techniques surveyed in the previous section but also includes several personal observations drawn from my own practical experience in the field of extended vocal techniques.

VOICED SOUNDS

Harmonic Filtering - The reinforcement or suppression of individual harmonics or bands of harmonics over any given fundamental tone is made possible by filtering the sound either at its point of resonance or its point of emission. Wishart suggests four methods: varying the size of the oral cavity, varying the position of the tongue's arch, greater or lesser rounding of the lips and greater or lesser nasalisation (Wishart 1985). To these methods we can also add manual filtering, whereby the hands are placed in a cup or funnel position in front of the mouth. Filtering using the first two methods is most easily demonstrated in the formation of different vowel sounds on a stable sung pitch (see Notation). This kind of filtering, combined with the controlled articulation of voiced or unvoiced friction noises (consonants), forms the basis of speech. Lip rounding is by far the easiest form of filtering and can be used to isolate and reinforce individual harmonic overtones quite clearly (generally speaking, increasingly higher harmonics are reinforced as the mouth opening is made larger). Both these forms of filtering become markedly more pronounced with increased nasalisation (allowing the sound to resonate through the nasal cavity) which tends to suppress the fundamental tone. Indeed, it is possible to suppress the fundamental tone to such an extent that it becomes significantly weaker than the resultant harmonic(s). Such a style of vocal production has reached a high level of development in some folk musics of Mongolia. An analysis of a sonogram of a Mongolian singer's voice reveals that, over a very weak drone at a frequency of 165Hz (E3), a melodic line is produced centred around two bands of harmonics. 'The first one occurs around harmonics 7, 8, 9, 10, 12 and 13, giving D6, E6, F#6, G#6, B6 and C7; the second zone occurs around harmonic number 20 or G7 with a frequency of 3,220 Hz. Other analyses have shown that the singer emitted 48 harmonics'.[5]

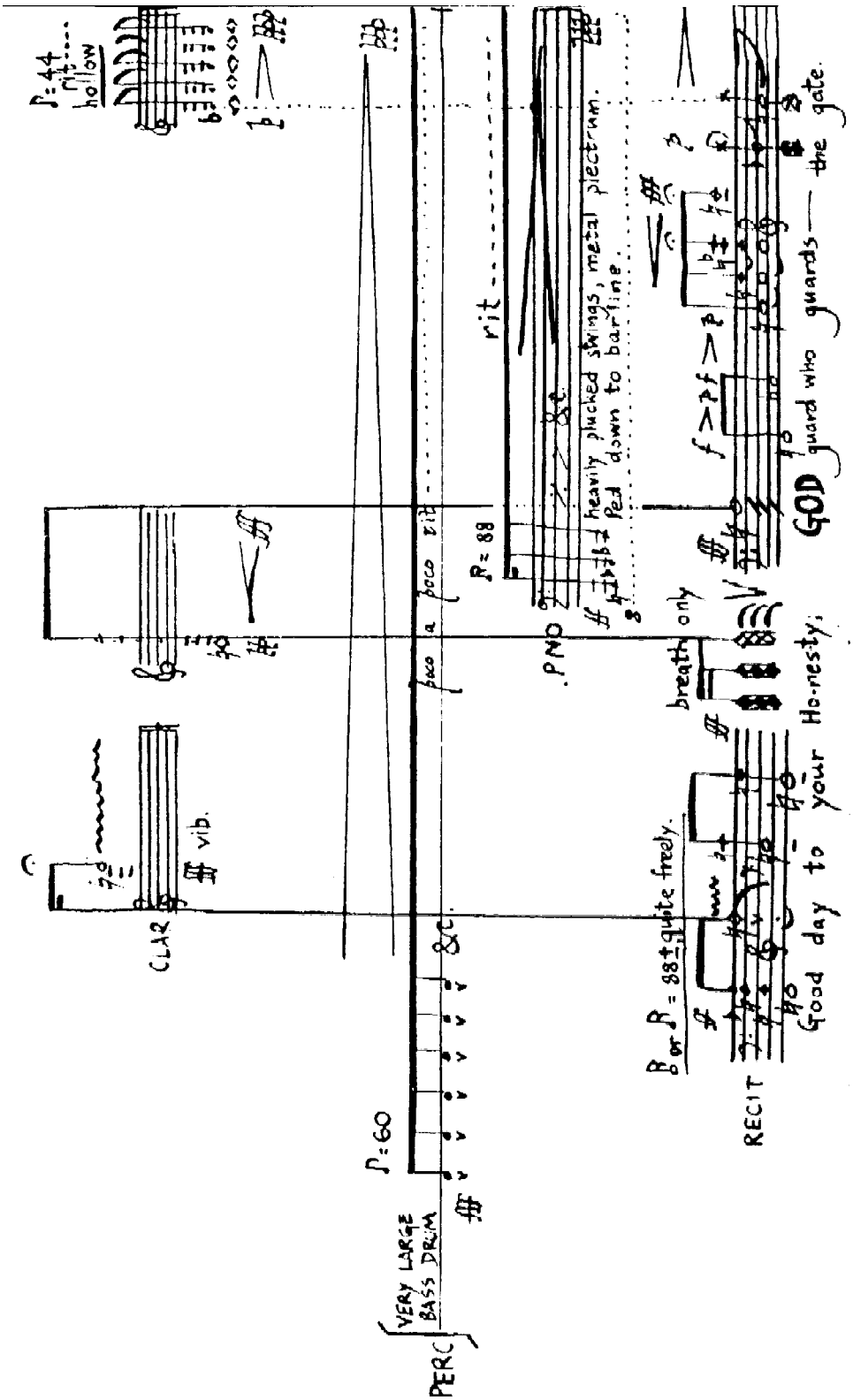
Filtering can be applied, with varying degrees of efficacy, to almost any orally produced sound, whether ingressive (inhaled) or egressive (exhaled). Double filtering may be produced by applying one or more of the above mentioned methods with external (manual) filtering.

Iterated Production (vocal fry) - Vocal fry, glottal scrape or subaudio production (TW) are the rather inappropriate names given to the effect that occurs when the rate of oscillation of the vocal chords during glottal production is slowed to a point where individual attacks or 'clicks' can be discerned. Jennings describes vocal fry as 'the action of gently phonating the vocal chords at a slower pace (often matching the pulsation rate of the breath source) than that used in WAMP' (Jennings, 1977). This can be achieved either egressively or ingressively (ingressive production tends to produces a much more striking sound capable of a great variation of timbre), usually over a range of about two octaves. Interestingly, there appears to be little difference in the attainable range of vocal fry in the male or female voice (ca 560 Hz).[6] In normal modes of production, lip, cheek and tongue vibrations occur within a similar frequency range as that of vocal fry. However, by manipulation with the hands or fingers, cheek vibrations may be further slowed down to the region of 45 Hz.

Octave Doubling - The lower register of the male or female voice may be increased by a range of up to two octaves by a particular method of vocal octave doubling (subharmonics (TW), chanting (CME)). This mode of production is similar to that practised in certain Tibetan tantric rituals[7] and is also to be found in the folk music of some Middle European countries. It is produced by first singing a quiet, relaxed glottal tone and then further relaxing the vocal chords until the note appears to split and drop an octave lower.[8] The tone produced has a rough, rasping quality and sounds simultaneously with the original tone although at a greater dynamic level. 'With practice, a note one octave and a fifth, or even two octaves below the glottal note can be produced' (Wishart, 1979). Production is made significantly easier by increased nasalisation. Once the tone is established it is possible to alter the pitch whilst retaining this mode of production. It is also possible, although more difficult, to produce this effect ingressively. The physical explanation of this method of vocal

production seems unclear. Wishart writes: 'the quality of these tones, and the apparent production of a 'subharmonic series' (physically impossible as such) suggest that some process of intermodulation between the original pitch and another, much lower, pitch is occurring' (Wishart, 1979). However, the explanation offered by the authors of the *Index*, i.e. that 'the physical sensation together with the perceived quality indicates that the production may involve the combination of light WAMP with vocal fry' (*Index*, 1974), seems more likely especially when one considers that the frequency range of both octave doubling and vocal fry is similar for both the male and female voice.[9]

Supraglottal and Subglottal Production - It is possible to produce continuous vocal sounds that appear to emanate from a source either above or below the glottis. Such sounds require considerable air pressure from the lungs and as such are difficult to sustain. Those sounds emanating from above the glottis, which we might call supraglottal, are produced by an extreme tightening of the throat muscles during WAMP production. This has the effect of restricting the movement of the vocal chords thus allowing the generation of a wide variety of extraneous friction noises of a generally complex and unstable nature. A certain amount of control is possible by mouth filtering. Various strained glottal tones tend to occur during production and it becomes possible to identify two or more distinct pitch areas and even (with practice) 'tune' these in the creation of genuine biphonic or multiphonic production. Ingressive supraglottal production can generate clearly defined pitches or pitch complexes well above the normally attainable range. Intermodulation with inhaled glottal tones can also occur giving ever more complex structures. However, this technique can be particularly stressful and potentially damaging. Subglottal production arises when large quantities of air are forcibly expelled through an open, relaxed throat. A definite, very low, pitch centre may be established although not altered. Again, intermodulation with glottal tones is possible.



Example 1 - a detail from the score of Peter Maxwell Davies' *Eight Songs for a Mad King*, a composition that makes extensive use of a great variety of supraglottal or 'strained' methods of production. The vocal part is written throughout in a mixture of conventional and semi-graphic notation, leaving many aspects of performance to the discretion of the singer.

UNVOICED SOUNDS

This category includes those sounds, both lunged and unlunged, whose primary sonic source is that other than the vocal chords. In addition, it will be shown that a great number of these sounds may be intermodulated with WAMP producing entirely new sounds.

Air-Stream Effects - The filtering of high-pressure air streams (i.e. forced breath) through the oral cavity produces coloured noise-type sounds. Filtering is achieved by adopting those mouth shapes that correspond to noise based consonants, namely, [ɸ], [f], [θ], [s], [ʃ], [ç], [x], [χ], [ħ], [h] and their various combinations (see Notation and Appendix I). Further filtering and pitch variation may be achieved through alteration of the shape of the oral cavity by adopting various vowel mouth positions. Voiced noise sounds (noise intermodulated with WAMP) can produce entirely new effects: [s], [ʃ] and [f] become [z], [ʒ] and [v] respectively. Whistling may also be regarded as a form of filtered noise. 'Pitch is varied by altering the size of the oral cavity and two registers may be distinguished according to whether the tongue arch is to the back or to the front of the mouth cavity. A third and higher register can be produced by filling the cheeks with air and pushing the tongue forward' (Wishart 1985). Unlunged whistling is made possible by rapidly drawing air in and out of the mouth through pursed lips by use of movement of the tongue.

Lip, Tongue and Cheek Vibrations - The lips, tongue or cheeks can all be used as sound generators when made to vibrate. The lips may be vibrated in the normal position, loosely relaxed (pouted) or tightly constricted (inverted). Pitch control is made possible by manually stretching or relaxing the lips. Alternatively, the top lip can be made to vibrate on its own by restricting the movement of the bottom lip with the top teeth. Ingressive lip vibrations produce complex sounds when lunged or click trains with a frequency as low as 2Hz when unlunged. Either the tip or middle of the tongue may be vibrated against the soft palate or the rear of the tongue against the uvula. A wide variation of pitch is possible by altering the position and pressure of the tongue against the soft palate. Pitch variation of cheek vibrations is controlled by altering the tension of the cheeks with the hand. Simultaneous production of lip and

tongue, lip and cheek, tongue and cheek and lip, tongue and cheek vibrations are all possible. In addition, these techniques can act as amplitude modulators when combined with WAMP or air stream effects.

Short Sounds (lunged and unlunged) - Short sounds can be defined as those sounds that cannot be produced as a continuum. These are nearly always based around the articulation, either lunged (voiced or unvoiced) or unlunged, of a consonant. Such sounds may be plosive or non-plosive. The following productions are possible:

voiced lunged	plosive combined with a very short vowel sound and some noise
voiced half-lunged	as above but with tight throat restricting air flow
unvoiced lunged	plosive combined with air stream noise
unvoiced half-lunged	as above but with tight throat restricting air flow
unlunged	unvoiced and non-plosive, produced by articulation of consonant without friction noise

Lunged or half-lunged sounds may be combined with different kinds of air stream ([f], [s], [ʃ], [x], etc.). The proportion of consonant to voice or consonant to air stream can be varied thus increasing the available range of sounds. Indeed, the very act of restricting the air flow in half-lunged and unlunged production or the use of strong plosive articulation can produce quite different sounds.

Lunged or half-lunged sounds may further be 'stopped' by cutting off the air stream with the epiglottis, tongue ([g], [k] or [t] positions) or the lips ([p] or [b] positions). A particularly interesting example in this category is the tongue click. This is used in certain native languages of southern Africa and is represented by the symbols [ɻ], [ɩ] or [ƪ] in the International Phonetic Alphabet. Suction created between the surface of the tongue and the roof of the mouth is audibly released with a sharp downward motion of the tongue. The resultant sound contains little or no friction noise and is perceived as having a definite pitch. The pitch can be altered by different placings of the tongue on the roof of the mouth.

AMPLITUDE MODULATION

Ululation - Sung tones may be ululated, that is, perceived to rapidly pulse, on either a stable pitch or a glissando. This is achieved by repeatedly stopping and starting the flow of breath with the throat muscles (the basis to this technique is similar to that of the early 17th century vocal *trillo*). When used across the natural break of the voice a rapid shake between two pitches is produced.

Manual Articulation - Manual pulsation of the diaphragm provides a controlled method of repeated articulation. This usually produces a similar effect to ululation, although when applied to a high glottal tone the sound can be broken into staccato like repetitions. In addition Wishart lists 'shake-head flutter, drum-glottis flutter, shake-body flutter, drum-cheeks flutter, strum-lips flutter, strum-nose flutter and hand-cup flutter' (Wishart 1985). Since these are all external articulations it follows that many may be applied to other, non-glottal vibrations.

Lip, Tongue and Cheek Vibration - If these vibrations are combined with glottal production and the intensity of the vibration is made sufficiently weak it then acts as an amplitude modulator to that glottal production. This is particularly effective when uvular vibration is combined with a quiet, low WAMP tone. Similar effects can be perceived when applied to air-stream noises.

MULTIPLE PRODUCTION

Multiphonics may be defined as the production of two or more stable sounds or pitches generated from the same sonic source. Harmonic reinforcement of WAMP, octave doubling and certain supraglottal tones all fall into this category. Harmonic filtering of a subharmonic tone can be used to produce a three note chord. Various types of ingressive and egressive multiphonic lip vibrations may be produced by placing a finger across the middle of the lips thus allowing the two edges to vibrate independently. Manual articulation can also be used to allow the two cheeks to vibrate independently. In addition, the tongue can be vibrated at the tip and at the rear simultaneously.

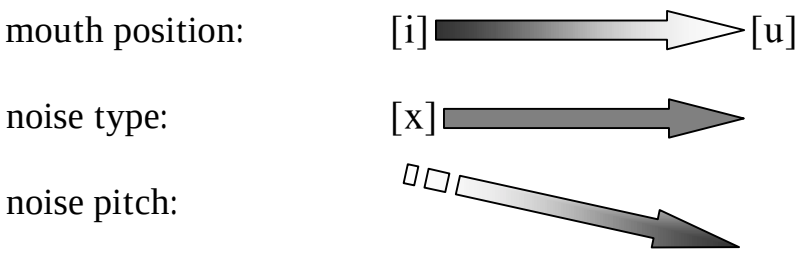
Intermodulation occurs when two or more sounds produced from separate sonic generators interact in the creation of a totally new sound. Most examples of multiple production fall into this category. They include WAMP with subglottal or supraglottal production, WAMP with whistle (when tuned to different pitches), WAMP with lip, tongue or cheek vibrations, WAMP with noise, simultaneous vibrations of lip and tongue, lip and cheek, tongue and cheek and lip, tongue and cheek, noise with lip, tongue, or cheek vibrations, noise with whistle, etc. A specific type of intermodulation occurs when two continuous pitched sounds are 'mistuned'. If a whistle tone and a glottal tone are tuned to a unison or octave they are heard as separate sounds. However, if either tone is moved away from that unison or octave, intermodulation occurs. This tends to occur automatically in the production of glottal and subglottal tones or glottal and pitched supraglottal tones. Since lip, tongue and cheek vibrations all possess pitch content it should be possible to obtain a similar effect when they are combined (individually) with glottal tones. However, this may only be achieved if the frequency of the lip, tongue or cheek vibrations, which is normally very low, is increased to a level approaching that of the glottal tones. A striking example of this kind of intermodulation occurs between glottal tones and tight, inverted lip vibrations.

Distinct Simultaneous Production - This category consists of sounds that are produced by separate sonic generators but which do not intermodulate. Only a very few examples of this type of production may be observed WAMP with tongue clicks, WAMP with unlunged [p], WAMP with ingressive unlunged lip vibration, WAMP with whistle (when tuned to the same pitches).

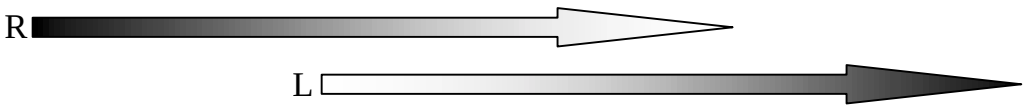
TRANSFORMATIONS

Two methods of transformation may be discerned: 1) transformation between two different spectra of a single sound. With continuous glottal tones this is achieved by alteration of the size of the mouth opening and tongue position. Transformation can occur between adjacent spectra, e.g., adopting the mouth position for [a] and arching the tongue to form [e] or between nonadjacent spectra by passing through an

intermediate stage, e.g., adopting the mouth position for [a] and gradually reducing the size of the mouth aperture, passing through [o] to reach the [u] position. Transformation of continuous noise based sounds is effected by altering the point of articulation. In this manner it is possible to trace a continuous line from the bilabial [ɸ] to the glottal [h], e.g. [ɸ]-[f]-[θ]-[s]-[ʃ]-[ç]-[x]-[χ]-[ħ]-[h]. As mentioned above, noise-type sounds may also be transformed through the alteration of mouth shape and tongue position, e.g.:

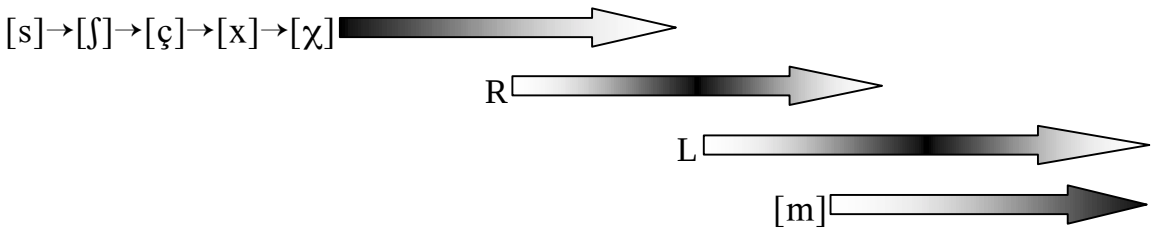


2) Transformation between sounds generated from different oscillators or between noise sounds and oscillated sounds. This is achieved by an overlapping of the two sounds as shown in the diagram (where R = tongue vibration, L = lip vibration):



Intermodulation may or may not take place at the point where the two sounds overlap.

By combining the above methods complex chains of transformations may be established. The following example represents a transition through various noise-type sounds to tongue vibration, lip vibration, and finally, nasalised glottal tone.



NOTATION

As yet there appears to be no standardised notational system amongst composers for extended vocal techniques (nor for many other contemporary music performance techniques). Conventional systems of notation are insufficiently precise and, moreover, inherently exclude the very information that is of most importance in the production of extended vocal techniques. However, the International Phonetic Alphabet (see appendix I) provides a useful and fairly comprehensive system for the codification of mouth and tongue positions (as with the formation of vowels) and points of articulation (as with the formation of consonants). As mentioned earlier, varied vowel formation over a continuous glottal tone provides an effective method of harmonic filtering, whilst the extension of consonants into a continuum yields a wide variety of noise based sounds. This range of available sounds is further increased if we include sounds that are formed in the transformation from one mouth position to another (e.g. diphthongs and tripthongs). These types of transitional sounds need not be limited to vowel-vowel combinations but can also include consonant-consonant combinations. In addition, elements of both can be utilised, e.g., vowel-consonant and consonant-vowel. Superpositions of two vowel, two consonant or vowel and consonant mouth positions still further extend the range of aural possibilities. Indeed, it should be remembered that the phonetic alphabet provides only abstract points of reference along the total oral auditory continuum in much the same manner that the tones in our Western pitch system are but predetermined divisions of a total auditory sound band.

Notes to Part One

[1] 'Hörbare und unhörbare Atemvorgänge müssen dann gewissermassen die Basis der künstlichen Vorgänge bilden... das Atmen selbst könnte, die Hörbarkeitsschwelle überschreitend, selbst Kunst werden und sich in ausgeformten Frikativen fortsetzen' (Schnebel 1972: 455).

[2] 'Solche Eingriffe führen sozusagen in die Lautproduktion hinein. Weitergehend ist diese selbst zu gestalten, ohne dass gleich übergreifende Einheiten wie Wörter oder gar Sätze visiert werden. Das heisst den Artikulationsprozess zum Gegenstand der Komposition machen... Das verlangt nach dem Bewusstsein dessen, was bei der Artikulation geschieht' (Schnebel 1972: 449).

[3] '...äussert sich Inhalt überhaupt nichr mehr vermittelt, durchs Vehikel eines noch so rudimentären Textes, sondern direkt. Sprache wie Inhalt entstehen in der Artikulation' (Schnebel 1972: 456).

[4] 'Diese Einbildung wird dann freilich nicht mehr intuitiv, sondern bewusst und methodisch angewandt, wobei systematisch die verschiedensten Ableitungen von Sprechen oder Gesang resultieren' (Schnebel 1972: 4467).

[5] Notes to UNESCO record *The Music of Mongolia*.

[6] 'The physiology of vocal fry is still in doubt. However, it seems that pitch is regulated by the mass of the [vocal] folds. There seems to be little influence of air flow and air pressure in frequency regulation although there is some speculation that these parameters may be important for the control of vocal intensity' (Colton and Hollien, 1973: 132).

[7] Kauffmann's authoritative *Tibetan Buddhist Chant* (1975) disapointingly provides rather vague, impressionistic descriptions of the vocal production used by the Tibetan monks. He writes: 'In some songs an unusual voice production can be noticed that occasionally is indicated by the word *rtsing* ('rough')... It is a deep, throaty sound that cannot be forgotten if once heard performed correctly.'(2) 'Voice timbres are called by various names... A frequently used term... is *stag mahi ngar skad*, 'the roar of the tigress', which means a rough, low sound... *Mgrin dbyangs* ('throat singing') also means rough and throaty sounds...' (15).

[8] In the notes to her record *Voice is the Original Instrument* (Wizard Records, New York, RVW 2266, 1976) the American singer and composer Joan La Barbara writes 'The lower tones or 'undertones' are produced by singing the original pitch, mentally directing the sound to a deep resonance area, relaxing the throat completely, pulling back slightly on the sound and breath pressure and allowing the sound to drop.'

[9] A combination of both these explanations also seems possible. If a glottal tone of 220 Hz (a') were to intermodulate with vocal fry at a frequency of 55 Hz (AA), i.e. two octaves lower, the resultant sound might yield a combination of WAMP and vocal fry at a frequency of 110 Hz (A), i.e., one octave lower than the original glottal tone.