# Acoustic and psychoacoustic aspects of vocal vibrato

Sundberg, J.

journal: STL-QPSR
volume: 35
number: 2-3
year: 1994
pages: 045-068

# Acoustic and psychoacoustic aspects of vocal vibrato

*Johan sundberg*

***Abstract***

*This article reviews research on vocal vibrato and related phenomena. As the vibrato corresponds to a modulation of the fundamental frequency, the primary acoustic attributes are the frequency and amplitude of this modulation as well as its waveform. The range of variation of these parameters in classical operatic singing are discussed as well as the boundaries to similar phenomena such as tremolo and trill. Also considered are some perceptual attributes, such as the effect of the vibrato to glue spectrum partials together to timbral* gestalts, *the effect on vowel intelligibility, and the relationship between the frequency modulation and the pitch perceived.*

## Introduction

Vibrato is not a well-defined phenomenon. In different music traditions and cultures different kinds of vibrato have developed, presumably with more or less differing acoustical and/or physiological characteristics. In Western operatic singing, a vibrato type is used which is characterised by an undulation of the fundamental frequency. A factor of great significane in the production of this type of vibrato seems to be pulsating contractions of the cricothyroid muscle. In popular singing and in singing in some non-Western cultures, another type of vibrato is frequently heard which seems driven by a pulsation of the subglottal pressure. In this paper I will summarise the state of knowledge regarding various acoustic and perceptual aspects of different types of vibrato. As the vibrato occurring in Western operatic singing has been much more successful than other kinds of vibrato in attracting researchers' attention, I will focus on this type of vibrato mainly. It develops quasi automatically as voice training successfully proceeds and can thus be observed in almost all professional Western opera singers (Björklund, 1961).

## Acoustic attributes

### 1. Theory

Insufficient knowledge of the acoustic theory of voice production has caused much of the research in the realm of vocal vibrato. The vibrato phenomenon can be understood only in the light of this theory, which I will briefly review (for a more exhaustive description, see Sundberg, 1987). The acoustic theory of voice production was established by Gunnar Fant and others during the 1950:s (Fant, 1960). It is illustrated schematically in Figure 1. It assumes that the characteristics of the transglottal airpulses are determined entirely by subglottal pressure and vocal fold adjustments and, thus, independent of the resonances of the vocal tract; this description seems to be reasonably realistic.

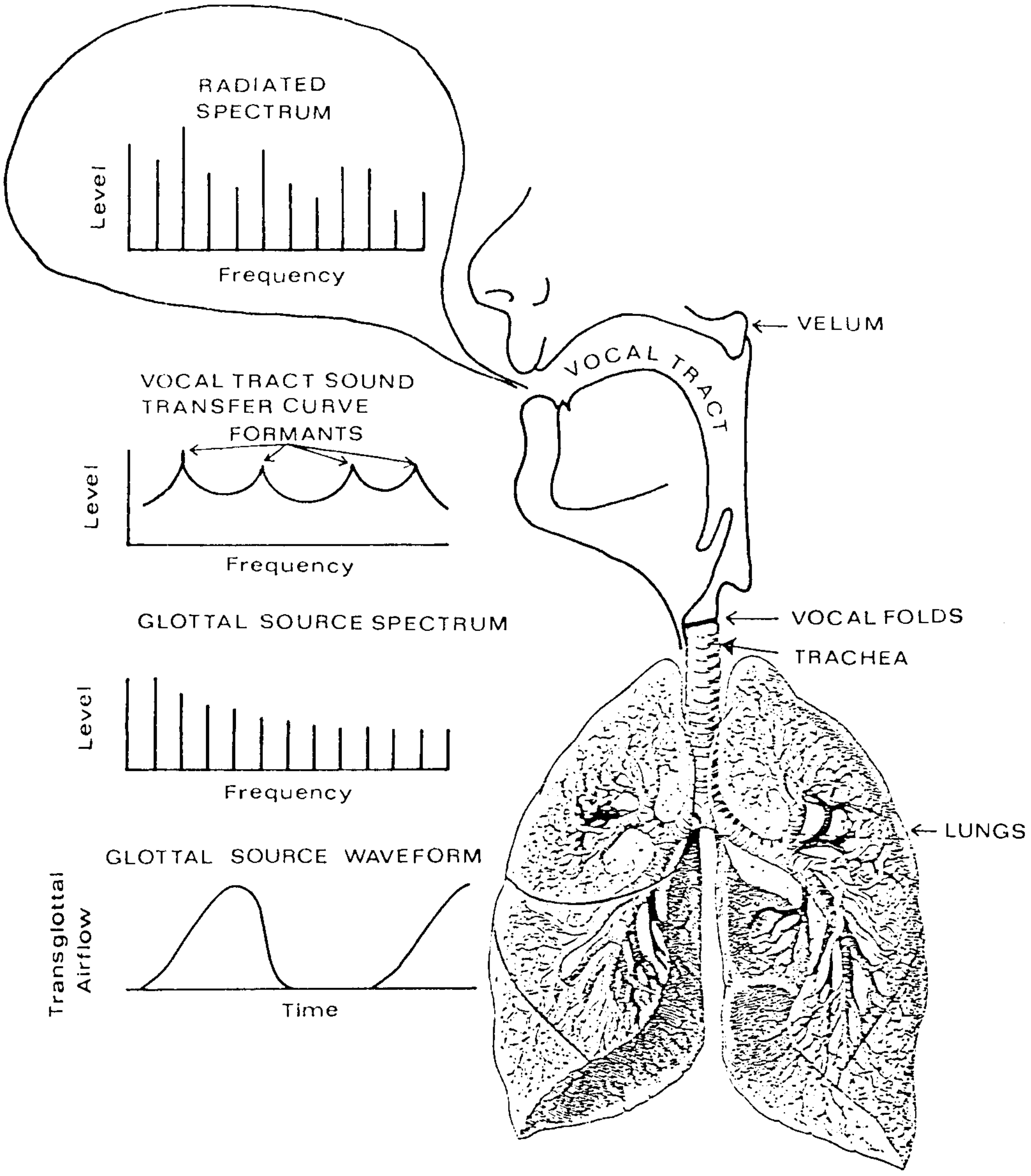


*Figure 1. Source - filter theory of voice production.*

Voice sounds are produced by a regularly pulsating airflow through the glottis. This train of airpulses is called the *voice source.* Its frequency corresponds to the vibration frequency of the vocal folds and determines the pitch. The pulses correspond to a sound constituted by a number of harmonic partials, the amplitudes of which decrease gradually with increasing frequency. The overall slope varies with vocal loudness and is about -6 dB of sound pressure per octave for intermediate loudnesses. The frequencies of the partials constitute a harmonic series, so that the frequency of partial number n equals n times the frequency of the fundamental. This means that the frequency of any partial can be calculated given the frequency of the fundamental. This is of particular relevance to vibrato tones, since they vary regularly in frequency.

When travelling from the glottis to the lip opening the partials of the voice source spectrum are fed to a *filter,* the vocal tract. As the vocal tract is a resonator, its frequency curve is characterised by peaks at the resonance or *formant* frequencies, and valleys in-between. Thus, partials lying close to a formant frequency become stronger than other partials. This means that the amplitude of a spectrum partial depends on two circumstances: first, its amplitude in the source spectrum, i.e. the number of the partial and how loudly the tone was produced, and second, how far the partial is from a vocal tract resonance frequency. This, again is of particular relevance for vibrato tones, since all their spectrum partials vary in frequency and, thus, continuously vary their distance to the closest formant frequency.

The frequency curve of the vocal tract, thus, depends on the formant frequencies. These frequencies are tuned by means of the vocal tract shape. For example, if the tract is narrow in the pharynx and wide in the mouth, the first formant appears at a frequency as high as, say, 900 Hz, and if the tube is wide in the pharynx and narrow in the mouth the second formant appears at a frequency as high as, say, 2500 Hz. This dependence of the formant frequencies on the vocal tract shape implies that the formant frequencies will be modulated if the vibrato involves also the vocal tract shape.

The characteristics of the voice source is determined by two major factors, subglottal pressure and the muscular adjustment of the glottis. Subglottal pressure is the tool for varying vocal loudness; when we want to increase vocal loudness, we increase subglottal pressure. The acoustic result is that the amplitudes of the source spectrum partials increase, more so for the higher partials than for the lower partials. Thus, when the partials near 500 Hz increase by 10 dB those near 3 kHz may increase by something like 15 dB (Gauffin & Sundberg, 1989).

The glottal adjustment is carried out by a set of intrinsic and extrinsic muscles. The cricothyroid muscle appears to play a simple and straight-forward role; it stretches the folds and so raises the frequency of phonation. The adductor muscles also play a very important role. When glottal adduction is changed, the mode of phonation changes along a phonatory dimension ranging from breathy in one extreme, over neutral, and to pressed phonation in the other extreme. Thus, breathy phonation is characterised by a low degree of adduction and pressed is produced with exaggerated adduction. Acoustically, a major correlate of these variations is the amplitude of the lowest spectrum component, the fundamental. In pressed phonation, the amplitude of the fundamental may be more than 10 dB weaker than in neutral phonation. To meet these different degrees of glottal adduction, subglottal pressure has to be accordingly tuned. When glottal adduction is exaggerated as in pressed phonation, the folds refuse to open the glottis and start vibrating unless subglottal pressure is high (Gauffin & Sundberg, 1989).

As mentioned, the vibrato physically corresponds to a periodic, rather sinusoidal, modulation of the phonation frequency (Schultz-Coulon & Battmer, 1981). Therefore, it can be described in terms of four parameters, see Figure 2:

1. the *rate*
2. the *extent,*
3. the *regularity,* and
4. the *waveform* of the undulations.

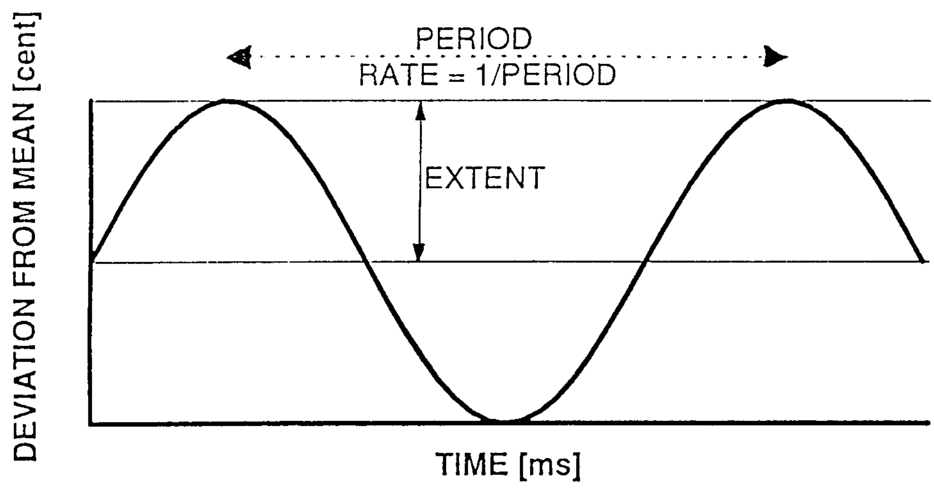


*Figure 2. Definition of vibrato period, rate, and extent.*

The vibrato *rate* specifies the number of undulations per second while the *extent* describes how far phonation frequency departs up and down from its average during a vibrato cycle. Often the extent is preceded by plus and minus signs so as to prevent confusion with an amplitude measure. The *regularity* shows how similar the frequency excursions are to one another. It is considered a sign of the singer's vocal skill: the more skilled the singer, the more regular the undulations. Moreover, in well-trained singers the regularity is generally not influenced appreciably even if the singer's auditory feedback is masked by noise (Schultz-Coulon, 1978). The *waveform* is generally more or less similar to a sine wave.

There is nothing to suggest that regularity and waveform are of lesser perceptual relevance than rate and extent. Still, only rate and extent have been analysed extensively in the past. Therefore, awaiting that regularity and waveform will be studied in future investigations, we will henceforth focus on rate and extent.

## 2. Definitions

The vibrato shows similarities with some other types of musical ornaments which occur in singing, namely, trill, tremolo and trillo. Let us first ask what characterises these different ornaments.

Using a computerised method of analyzing fundamental frequency, Schultz-Coulon and Battmer (1981) compared single examples of tremolo, trill and vibrato and observed certain differences. Later, Hakes, Shipp & Doherty (1987) compared more systematically fundamental frequency and amplitude characteristics of straight tones, vibrato tones, trills and trillos as performed by four internationally renowned solo singers. From these investigations the results shown in Figure 3 emerged.

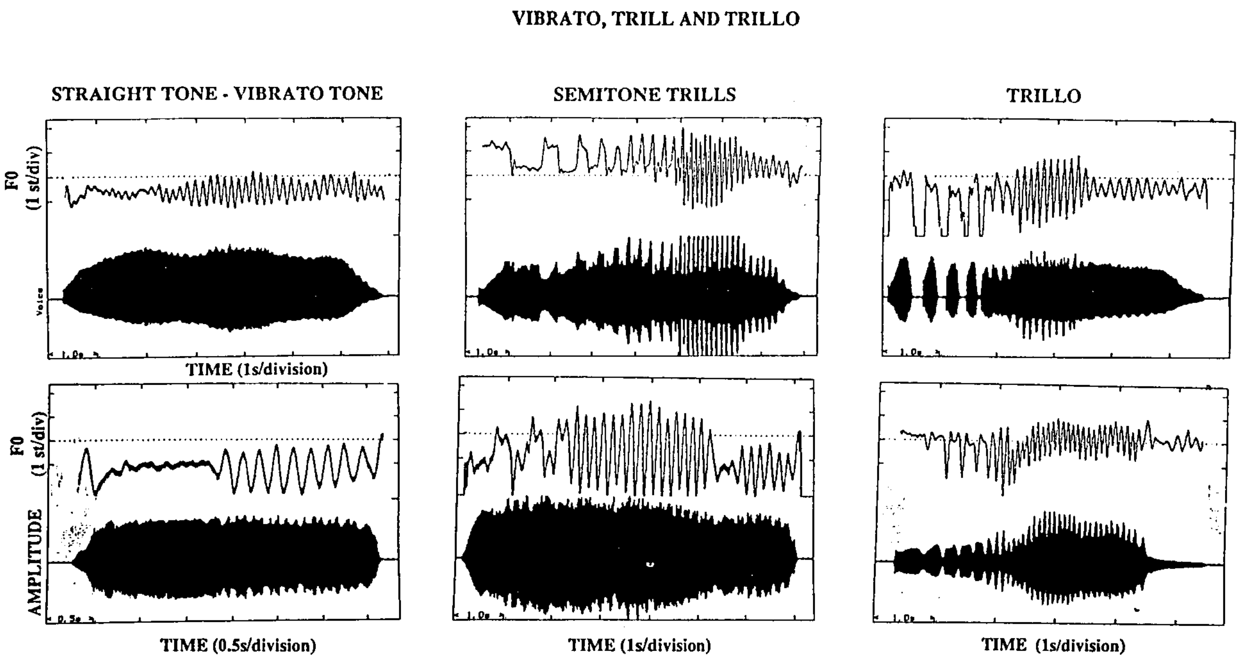


*Figure 3. Examples of fundamental frequency (F0) and amplitude patterns during straight tone, vibrato, trill and trillo according to Hakes & al. (1990) as produced by professional singers.*

The *vibrato* tones were characterised by a fundamental frequency undulation at a rate of 5 to 7 undulations per second and an extent of about ±1 semitone. The *trill* had a greater modulation width often amounting to several semitones. It consisted of a typical initiating pattern, where the singer presented the two target tones concerned, by extending their durations to about 0.5 s each, i.e. the alternation frequency was about 2 Hz. Gradually, the alternation speed was then increased to about 7 Hz, and during the final part the trill was similar to a vibrato tone with an exaggerated extent. These observations were later corroborated by Castellengo & Colas (1991). The *trillo* is used especially in early music. It is an ornament, in which the same pitch is repeated, interleaved with silent intervals. The repetition frequency is low at the beginning and gradually increases. These repetitions are accompanied by clear amplitude variations in the first part of the ornament. Likewise, fundamental frequency variations occur which are of wide extent in the first part and more vibrato-like toward the end.

Vibrato has been compared also to vocal tremor (Ramig & Shipp, 1987). The vibrato characteristics of nine opera singers were compared with the tremor characteristics of 6 patients of different diagnoses suffering from vocal tremor. Surprisingly, the results revealed only minor physical differences. The rate was 5.5 Hz for the singers and 6.8 Hz for the vocal tremor patients, and the regularity of the fundamental frequency variations appeared to be greater in singing. However, none of these differences reached statistical significance in their investigation. In any event, it is fair to conclude that there are similarities between vocal tremor and vibrato.

## 3. Frequency variation

### *a. Rate*

The rate of the frequency modulation is often considered to be constant within a singer. Thus, in an experiment which I did with Tom Shipp and Rolf Leanderson, we asked professional opera singers to change the rate of their vibrato, but none of them seemed able to do so (Shipp & al., 1980). However, singers do exist who are able to deliberately change their vibrato rate. Also, by means of special exercises singing teachers are mostly capable of correcting an inappropriate vibrato rate in a student.

The personal vibrato rate seems to depend on a number of factors, such as sex and age as well as the emotional involvement of the singer (Shipp et al., 1980). Bennett (1981) found that also the pitch affected the vibrato rate in the singers he studied, but this observation has not been corroborated in other investigations. Furthermore, the vibrato rate is not as constant within a singer as is generally assumed. In this volume, Eric Prame presents new data from phonogram recordings which, contrary to present assumptions and expectations, show that the vibrato rate actually varies during the course of the individual tones in a song, such that it accelerates during the last five vibrato cycles, just before the pitch change.

Disregarding such local variations of vibrato rate, female singers tend to have a slightly faster mean vibrato rate than male singers. Shipp & al. (1980) analysed the vibrato of ten professional opera singers, five males and five females. All subjects went through the same protocol singing sustained vowels in an anechoic chamber. The average vibrato rate found was 5.4 Hz for the male singers and 5.9 Hz for the female singers. No systematic dependence on pitch was observed. These values are similar to those reported in other laboratory investigations (Large & Iwata, 1971). On the other hand, they are clearly low as compared with results from investigations of sung performances. Analysing such phonogram recordings of seven famous singers, Winckel (1974) found vibrato rates ranging between 5.5 and 7.7 Hz, mean 6.9 Hz, and Seashore (1938), analysing recordings of 29 artists, arrived at an overall average of 6.6 Hz. Shipp & al. (1980) speculated that the emotional involvement of the singer may increase the vibrato rate.

I found some further support for this speculation by comparing two recordings of the same professional operatic baritone singer which I had stored in the lab archive. In one he sang sequences of sustained vowels and in the other he sang a song, a vocalise by Panofka. The mean vibrato rate for the sustained tones averaged to 5.4 Hz, SD 0.2 Hz, while that from the real performance of a song averaged to 6.2 Hz or about 15% higher, SD 0.2 Hz. However, the tones analysed from the sung performance were not long enough to avoid the just mentioned effect that the rate speeds up during the last five vibrato cycles. In any event, the hypothesis that emotional involvement influences vibrato rate seems worthwhile to analyse more thoroughly in the future. The results may shed some light on the physiological origin and the communicative function of the vibrato.

Damsté et al. (1982) studied the significance of age by comparing pairs of commercial recordings of the same seven singers. The first recording was made when the singers were in their third decade and the second in their sixth decade, toward the end of their careers. As shown in the left graph of Figure 4, most of the singers showed a decrease of the vibrato rate. The values range from -30% to +5%, and the mean over subjects was -11%. Only one singer, Tauber, showed an increase and one, Flagstadt, showed no change.

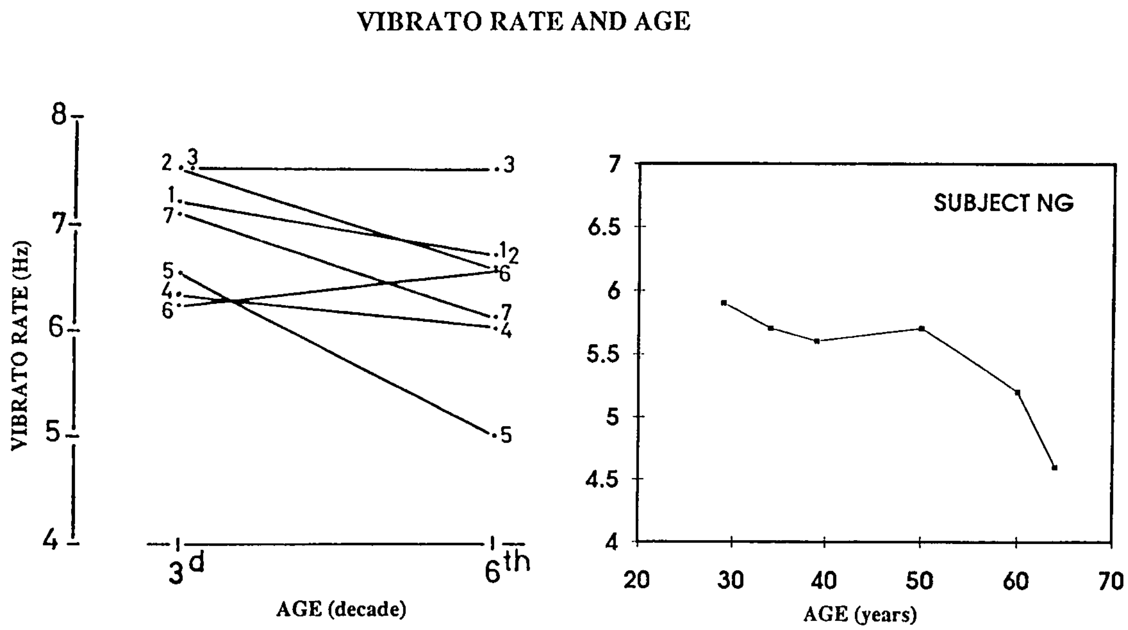


*Figure 4. Changes of vibrato rate with the singer's age. The left graph shows data by Damsté & al. (1982). The right graph shows data from a professional tenor's recordings of the same song over a period of 35 years.*

The right graph in Figure 4 illustrates this slowing of the vibrato with increasing age in greater detail for a particular, world famous tenor. It shows his vibrato rates measured on the same tone in the same aria as performed on six occasions over a period of 35 years; the first recording was made when he was 29 years old and the latest when he was 64 years old. The result shows a marked decrease above the age of 50. These data support the conclusion that vibrato rate tends to decrease with age. However, the effect may appear at different ages, most singers exhibiting the effect in their 60:s. Thus, as with most other signs of ageing, the period during which the slowing of vibrato rate appears, seems to vary greatly between individuals.

### *b. Extent*

As mentioned, the extent is generally narrower than ±1 semitones which corresponds to a frequency swing of ±6%, approximately. Choir singers show rather irregular vibratos with very small extents averaging to no more than 0.1 semitone. Wind and bowed instruments generally exhibit much smaller vibrato extents, typically less than ±0.5 semitones.

According to Winckel (1953) and others, the extent of the vibrato undulations varies with loudness of phonation. Schultz-Coulon & Battmer (1981) published data from a soprano who increased her vibrato extent from 1 to 1.5 semitones when she raised her vocal loudness by 15 dB, keeping the pitch constant, see Figure 5. Michel & Myers

(1991) tried to corroborate this dependence of extent on loudness by measurements on singing students who sang crescendos and decrescendos at low, medium and high pitches. They found a great intersubject variability and different effects at different pitches, but no quantitative data were reported on the sound level, the independent variable in their experiment. At high pitches, however, most subjects exhibited a similar effect; during a crescendo the vibrato extent increased slightly with vocal loudness, from 0.6 or 0.7 semitones in pianissimo to about 1.0 semitone in fortissimo, so the effect was quite small. However, the reverse effect was rarely observed in a decrescendo. Bennett (1981) studied the dependence of vibrato extent on pitch in four soprano singers; three of them remained at about 0.5 semitone throughout their ranges while one increased her vibrato extent dramatically with pitch up to ±1.8 semitones at her top pitch.

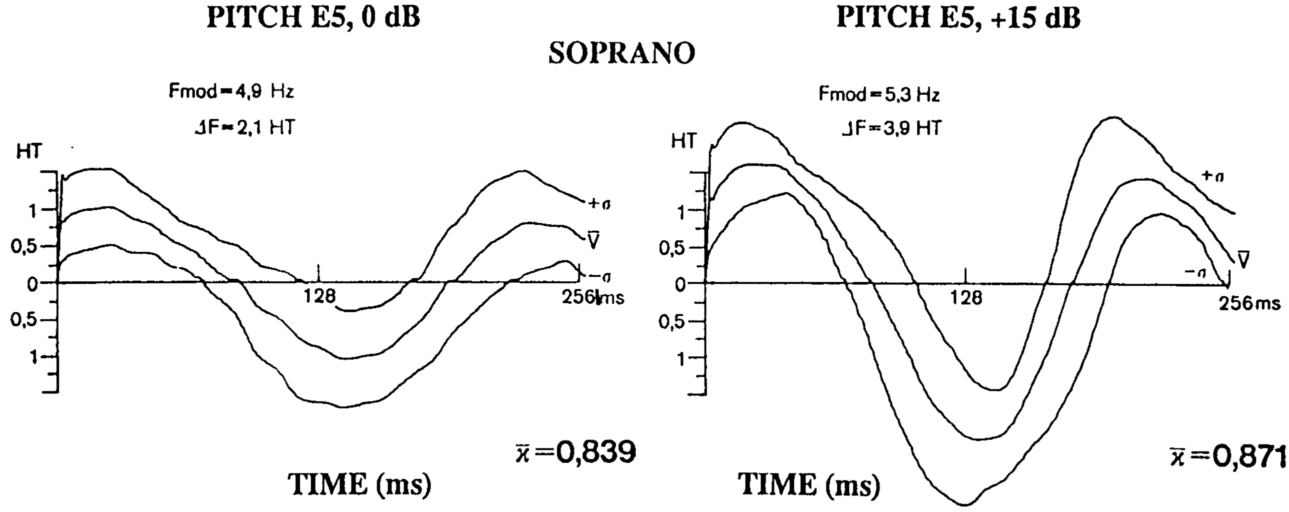


*Figure 5. Vibrato waveforms observed in a soprano singing the same pitch at two different degrees of vocal loudness. The middle curves represent the mean waveform, and the upper and lower curves show the standard deviation. The average rate and extent were 4.9 Hz and 2.1 semitones for the softer tone and 5.3 Hz and 3.9 semitones for the louder tone. Data from Schultz-Coulon & Battmer (1981). The $\bar{\chi}$ is a measure of the regularity of the vibrato.*

## 3. Waveform and regularity

The waveform of the vibrato undulations is generally more or less similar to a sine wave, as mentioned. However, considerable deviations from this waveform occur, as demonstrated by Schultz-Coulon (1976; 1978; Schultz-Coulon & Battmer, 1981). Examples of vaweforms observed for some singers are shown in Figure 6.

They also assessed the regularity of the vibrato of a few singers. This was realised by correlating each amplitude reading of a sampled waveform for a given vibrato period with those of the next vibrato period. The correlation coefficient was then used as a measure of the waveform regularity. For each portion of the vibrato waveform they obtained a mean and a standard deviation. Figure 6 also shows both this mean as the middle curve and the standard deviation as the adjacent upper and lower curves. It can be seen that the departures from the average waveform tend to vary with the phase of the vibrato waveform. In some cases particularly great deviations occur during the negative phase.

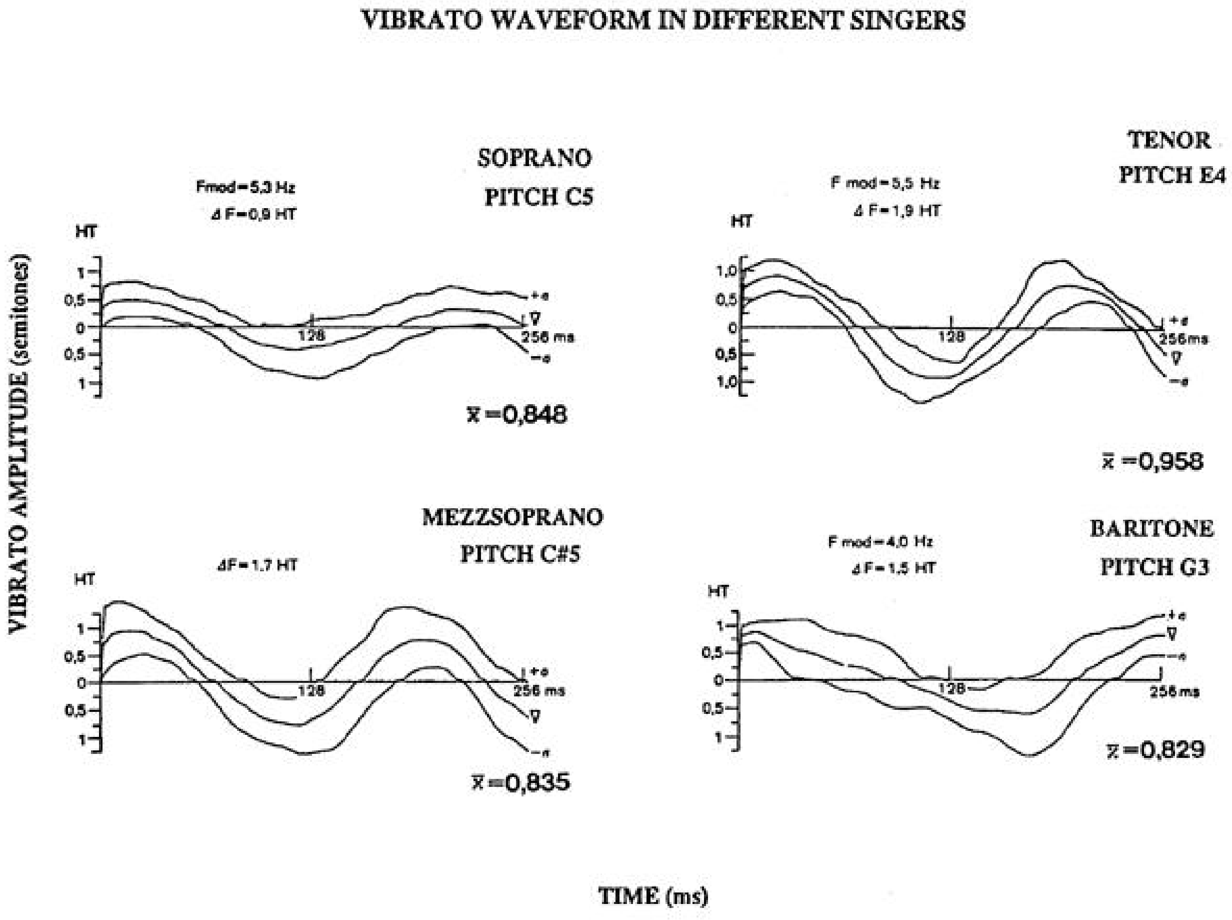


*Figure 6. Vibrato waveforms observed for different singers. In each panel, the middle curves represent the mean waveform, and the adjacent upper and lower curves show the standard deviation. $F_{mod}$ shows the average rate in Hz, ΔF the average extent in semitones, and $\overline{\chi}$ is a measure of the regularity of the vibrato. Data from Schultz-Coulon & Battmer (1981).*

## 4. Rate - Extent - Interaction

It has sometimes been assumed that there is an interaction between vibrato rate and extent, such that great extents often appear in combination with low rates and vice versa (Winckel, 1967, quoting Ramsdell). Not much empirical support has been reported for this assumption. Figure 7 shows an example from an investigation by Keidar & al. (1984). The data points pertain to 20 singers performing a high C (pitch C5, about 520 Hz) in an aria from *Il Trovatore* by Guiseppe Verdi. As can be seen in the graph there is no clear relation between the rate and extent although there is a slight trend for slow rates to appear in combination with wide extents.

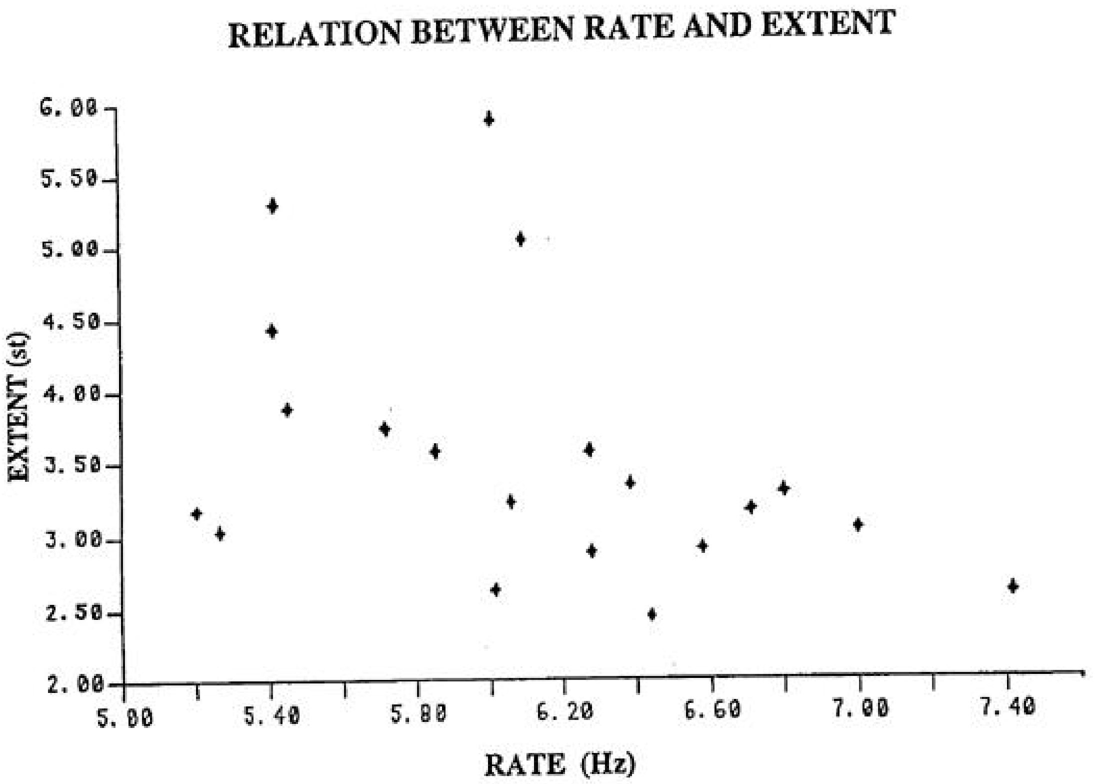


*Figure 7. Vibrato extent as function of vibrato rate observed at the pitch of C5 in 20 professional tenor singers' performances of the same tone in an aria from* Il Trovatore *by Guiseppe Verdi. Data from Keidar & al. (1984).*

## 5. Amplitude variation

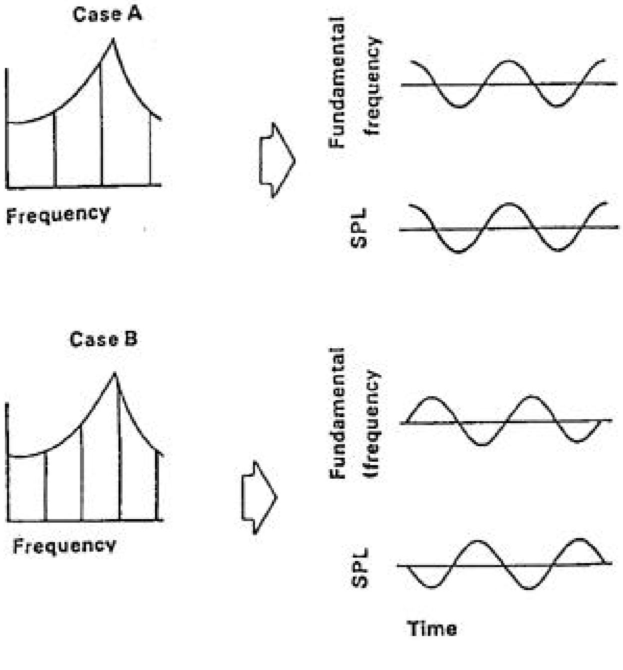
Vibrato tones generally vary not only in frequency, but also in amplitude, or, more precisely, the amplitudes of the various spectrum partials of a vibrato tone tend to vary. Although the perceptual significance of these variations seems limited, this amplitude variation has been analysed in a number of investigations. The phase relationship between the variations in frequency and overall amplitude has been an issue that has been repeatedly investigated. A fact which unfortunately has been realised only by few researchers is that this phenomenon cannot be correctly analysed without paying attention to the acoustic theory of voice production (Sundberg, 1984). Let us first review the different possible causes of the amplitude variation which this theory allows, noting that either one or any combination of them suffice to produce an amplitude vibrato.

Three different sources of amplitude modulation in vibrato tones can be identified. As was explained before, one cause is the *frequency variation* itself; as the frequency of a specific partial varies, its frequency distance to the closest formant and hence its amplitude varies. Another source may be the characteristics of the *voice source*, which can be brought to an amplitude variation due to a pulsation of the subglottal pressure or the glottal adjustment. A third source is *vocal tract shape*, which sometimes varies rhythmically, thus causing the formant frequencies to move slightly up and down in frequency.

Because the spectrum partials appear at integer multiples of the fundamental frequency, the vibrato modulation of the phonation frequency causes the frequencies of all the overtones to vary regularly and in synchrony with the frequency modulation of the fundamental. These frequency variations suffice to produce variations of the overall amplitude. The reason is that this overall amplitude often nearly equals the amplitude of the strongest spectrum partial, i. e. the partial closest to the first formant (Gramming & Sundberg, 1988).

The amplitude of every spectrum partial is determined by its distance from the formant frequencies; if a partial approaches a formant, its amplitude increases, other things being equal. Therefore, when the fundamental frequency increases in the vibrato cycle, the frequency of the strongest partial may either bring that partial closer to the first formant, so that its amplitude and hence the overall sound level increases, or it may bring the partial further away from the first formant and the opposite will happen (Sundberg, 1984; Horii, 1989; Imaizumi & al., 1993).

Both these cases are illustrated in Figure 8. In case A the fundamental frequency and the overall amplitude will undulate in phase (both increase and decrease in synchrony). In the opposite case B they will vary in opposite phase, so that an increase in one is associated with a decrease in the other. Which case occurs depends on whether the strongest partial is slightly lower or slightly higher than the first formant frequency.



RELATION BETWEEN FREQUENCY AND AMPLITUDE VARIATION

*Figure 8. Dependence of the amplitude of a spectrum partial on its frequency according to the acoustic theory of voice production. In case A the strongest spectrum partial is slightly lower than the first formant, so fundamental frequency and SPL vary in phase. In case B the strongest spectrum partial is slightly higher than the first formant, so fundamental frequency and SPL vary in counterphase. From Sundberg (1984).*

A third case may also occur: the frequency of the strongest partial may undulate symmetrically around the frequency of the first formant. Then the amplitude vibrato will actually be twice as fast as the frequency vibrato. In any case, the phase relationship between amplitude and frequency vibrato depends on the frequency relationship between the first formant and the strongest spectrum partial. The frequency variation of the vibrato is, thus, alone sufficient to induce variations of the overall amplitude.

Up to now, we have assumed that the voice source and the formant frequencies were constant, independent of the vibrato. This is not always the case, though. In certain types of vibrato, the voice source varies in amplitude. Such undulations may result from an undulation of the subglottal pressure. Particularly in popular singing such variations often seem to be the main physiological correlate of the vibrato. Small subglottal pressure variations may accompany also the Western operatic vibrato (Rothenberg & al., 1988). A modulation of the voice source may result also from an undulation of the adductive force in the glottis. Figure 9 shows an example taken from a professional singer, a counter tenor who was singing the same pitch at intermediate loudness in modal and in falsetto register. The figure shows the amplitude of the voice source as obtained by inverse filtering of the flow signal. The amplitude varies because of a varying glottal leakage which occurred in synchrony with the vibrato.

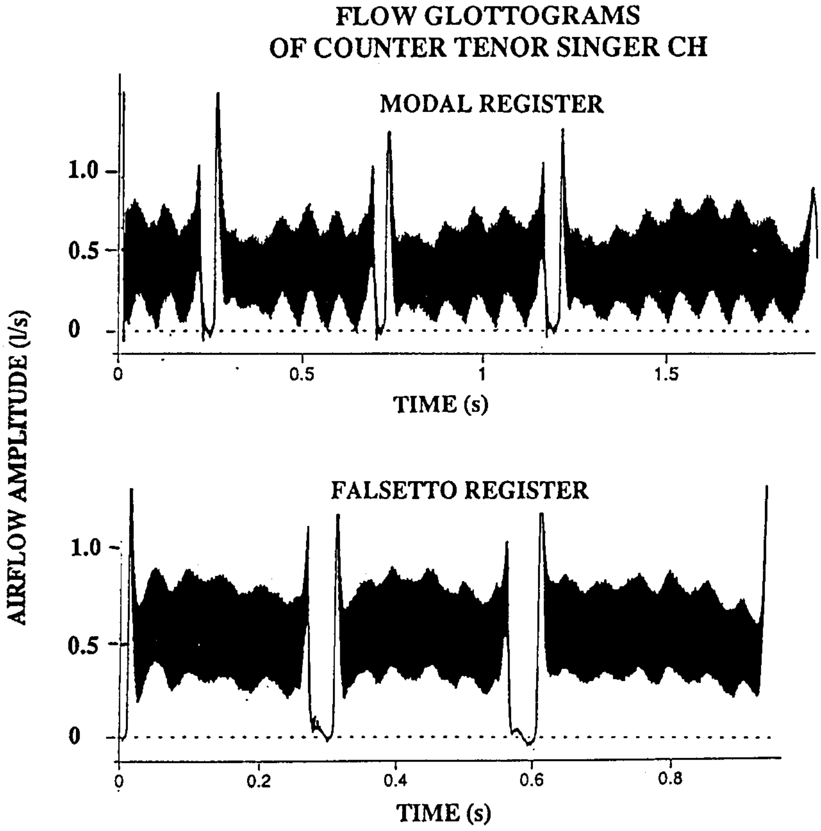


*Figure 9. Example of vibrato combined with an ad/abduction undulation in the glottis. The graphs show the amplitude of the transglottal airflow as revealed by inverse filtering of tones sung by a professional counter tenor singer in the modal and falsetto registers (upper and lower graphs). When glottal adduction is increased the glottal leakage decrease and vice versa.*

The cause of this phenomenon can safely be postulated: a pulsation of the glottal ad/abduction force. This case was also reported and documented by Rothenberg, Miller & Molitor (1988). The acoustic result of such an undulation of glottal adduction can be predicted. The glottal leakage consumes a good deal of the sound energy which travels along the vocal tract because it leaks down to the trachea which absorbs it. Therefore, the attenuation of the vocal tract resonances will increase when the glottal leakage is great, and decrease when it is reduced. In other words, the attenuation can be expected to vary with the vibrato. As the attenuation influences the amplitudes of the formant peaks in the vocal tract frequency curve, the amplitudes of the formant peaks will vary with the vibrato.

If such an adduction-based vibrato is combined with a non-leaking phonation, the situation will be different. Then, the amplitudes of the transglottal flow pulses will vary. This will induce a variation of the amplitudes of the lower source spectrum partials. This may result in an increased, decreased, or cancelled variation of the overall amplitude of the tone depending on whether the strongest partial is slightly below or slightly above the first formant and on how strong the variation of the source is.

In yet other types of vibrato, the formant frequencies vary in synchrony with the vibrato. The cause of this is that the vocal tract shape varies. For instance, the tongue shape or the sidewalls in the pharynx may move quite substantially, as has been shown in a fiberscope video film by Dr. Selkin, New York. In this film one of the singers sings with a slow and wide vibrato, which is associated with clear variations of the pharynx width. Such variations can be expected to cause particularly the first formant frequency to undulate.

Summarising, the main acoustic characteristic of the vibrato is an undulation of the fundamental frequency ranging from 5 to 8 Hz in rate and less than ±1 semitone in extent. These undulations cause the frequencies of the spectrum partials to vary in synchrony. As a result undulations occur in the overall amplitude of the tone. These undulations may be in phase or counterphase with the frequency modulation but seem to be of small perceptual relevance.

# Perceptual attributes

## 1. Rate and Extent

The vibrato adds a quite special quality to the tone. In the 30:s, Seashore found evidence for claiming that it adds "richness" to the timbre. On the other hand, it does not seem entirely clear what the quality of "richness" stands for. Perhaps it would be more appropriate to claim that the vibrato simply adds a "vibrato quality" to the tone.

There are rather narrow limits for the rate and extent of an acceptable vibrato. A rate slower than 5 undulations per second tends to sound unacceptably slow, and vibrato rates exceeding 8 undulations per second tend to sound nervous. Similarly, vibrato rates exceeding ±2 semitones tend to sound bad. Particularly when combined with a slow rate, such wide extents tend to remind of a singer who possessed a wonderful voice many years ago.

## 2. Pitch

Normally, the fundamental frequency determines the pitch we perceive. In the case of vibrato tones this fundamental frequency is not steady but continuously moves up and down. What pitch do we hear in such cases? Another question raised by vibrato tones is whether the pitch we hear from a vibrato tone is as accurately and precisely perceived as the pitch perceived from vibrato-free tones. If this is the case, singers would have a good reason for using the vibrato: it would mean that the demands for accuracy with respect to phonation frequency would be reduced as soon as there is a vibrato.

Both these questions were studied in an experiment where musically trained subjects adjusted the fundamental frequency of a vibrato-free tone so that its pitch agreed with the pitch of a preceding vibrato tone (Sundberg, 1977). In order to gain complete control over the stimuli, synthetic sung tones were used throughout. The results showed that the subjects tuned the frequency of the response tone so that it agreed, within a few cents, with the linear average of the vibrato tone's undulating fundamental frequency, as is illustrated in Figure 10. Essentially the same result was found by Shonle & Horan (1980) and Iwamiya et al. (1983). However, Shonle & Horan found that it is actually not the linear but the logarithmic average that corresponds to the pitch we hear. For typical vibrato extents the difference between these averages is small.

PITCH PERCEIVED OF VIBRATO TONES

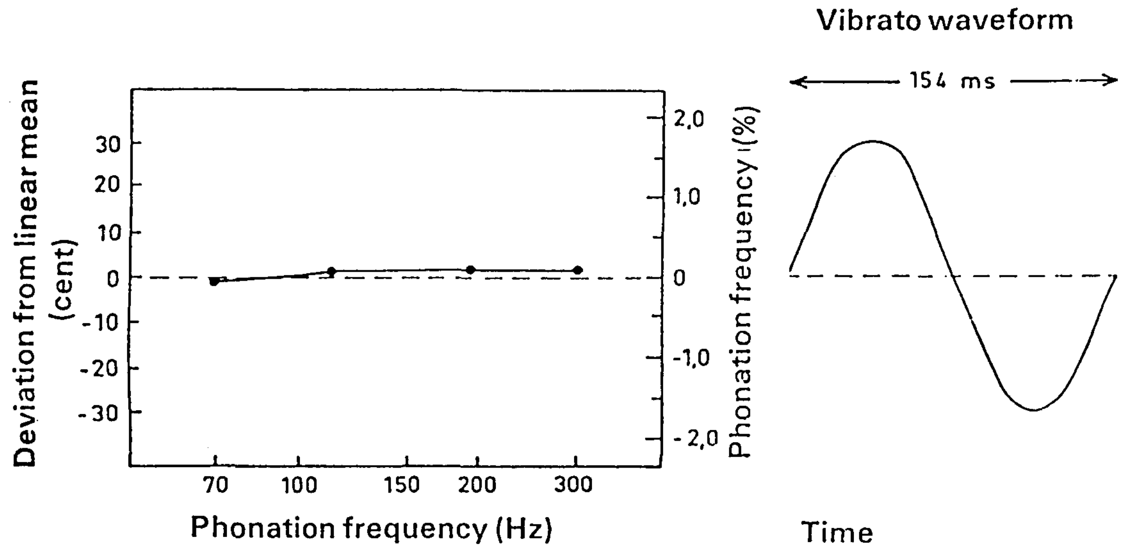


*Figure 10. Fundamental frequency corresponding to the pitch perceived from a vibrato tone. From (Sundberg 1977).*

In my investigation, the entire experiment was repeated twice; the subjects first matched the pitch of vibrato tones, and then the same subjects matched the pitch of vibrato-free tones. In this way it was possible to find out to what extent the accuracy of the pitch perceived was affected by the vibrato. The subjects' matchings of the pitch of the same stimulus tones were almost entirely equally accurate, regardless of the presence of the vibrato. Apparently, the vibrato did not reduce the certainty with which the subjects perceived the pitch.

d'Alessandro & Castellengo (1991) recently studied the relations between fundamental frequency gestures of very short durations corresponding to fractions of a vibrato period, i.e. tones of about 80 ms duration. Interestingly they found that the rising half of a vibrato cycle, when presented alone, was perceived as a higher pitch than the falling half, correspondingly presented. Thus, the pitch perceived of the rising curve was 15 cents higher than the mean, while the pitch of the falling was 11 cents below the mean, as illustrated in this Figure 11. The authors assume that this asymmetry results from the ears' paying more attention to the final portion of such short stimuli. In other words, their results can probably be accounted for in terms of a weighting function which pays more attention to recent than to less recent information.

There are often good reasons for distinguishing between physical and perceptual measures. The pitch of vibrato tones offers a striking example. The pitch perceived of a vibrato tone is as well defined as that perceived from a vibrato-free tone. Thus, the pitch of a vibrato tone can be regarded as being constant. Yet, the associated fundamental frequency varies. Moreover, the pitch of a vibrato tone remains constant only if the vibrato rate is not too slow; for rates near 4 Hz and lower, the pitch is clearly undulating. Thus, while fundamental frequency varies, the pitch is either constant or varying, depending on the rate of variation.

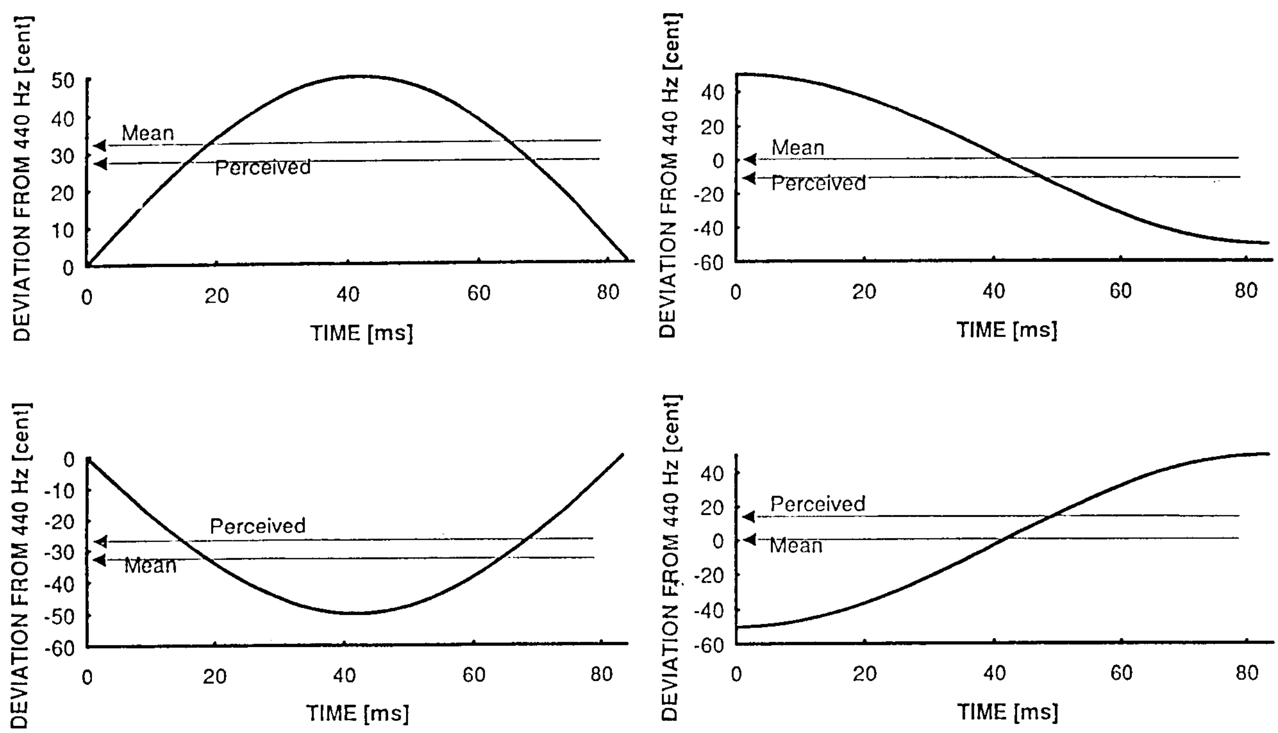


*Figure 11. Fundamental frequencies corresponding to the pitches perceived from short tones with a fundamental frequency changing according to a half a sine wave curve. The linear mean and the frequency corresponding to the perceived pitch are shown. Data from d'Alessandro & Castellengo (1991).*

## 3. Amplitude variations

The perceptual significance of the amplitude vibrato is often overestimated. On the other hand, no formal investigation seems to have been made of this issue. Therefore, I will here rely on my own experiences from synthesising singing.

The main perceptual effect of the vibrato is dependent upon the frequency modulation and it is generally quite hard to focus one's attention on the amplitude fluctuations. For instance, the perceptual significance of different phase relations between frequency and amplitude modulation discussed above seems small; the auditory image of a perfectly normal vibrato tends to persist, even if the amplitude vibrato is twice as fast as the frequency vibrato, a case occurring when the frequency of the strongest spectrum partial agrees with the frequency of the first formant. Also, if the amplitude vibrato is very small - which happens when the strongest partial is far away from the first formant - the impression is almost the same as when the amplitude vibrato is great. i.e., when the strongest partial is close to the first formant.

## 4. Vibrato phase and pitch change

It has often been observed that pitch changes occur in accordance with the phase of the vibrato waveform. Thus, ascending intervals are initiated during the rising phase and descending intervals during the negative phase. Myers & Michel (1987) tested the validity of this observation in a formal experiment where singer subjects alternated between two given pitches at different tempos. A great number of cases were observed where there was no synchronisation between pitch change and vibrato phase; the vibrato period was often disrupted by the pitch change.

This observation does not rule out the possibility that singers strive to synchronise pitch changes with the vibrato cycle, and manage to do so only after practising. In other words, it seems necessary to systematically study real performances from this point of view.

## 5. Special perceptual effects of the vibrato

### *a. Gluing partials together*

The sound of a voice fills the air with a number of partials, i.e. a chord of sine tones of different frequencies. Such sets of tones are generally lumped by our perceptual systems and heard as units called voiced sounds. It is in fact quite difficult for us to hear any of these partials as an autonomous tone, even in cases when a particular partial is much stronger than all other partials. Spectrum partials seem soaked with a glue that make them stick together.

However, under special conditions this glue can be dissolved, so that a vowel spectrum is perceived as a collection of individual, autonomous partials. One way of dissolving the glue is to have them stay perfectly constant in frequency, i.e. a total absence of a vibrato. This is an interesting but yet poorly explored property of the perceptual system. In short, the ear tends to switch from a pasting, holistic listening mode to an analysing listening mode when the fundamental frequency is constant.

A sound example quoted from Chowning (1980) offers a striking example of the vibrato's ability of gluing partials together. It presents first a cluster of sine tones, arriving independently of one another, all of constant frequencies. Then, after a little while, they split into three groups. Each group partials then starts to move all its partials in its own, private synchrony. All of a sudden the grey cluster of sine tones is transformed into three voices singing a chord.

### *b. Increasing vowel intelligibility*

As mentioned before, the fundamental is not the only spectrum partial that undulates in frequency because of the vibrato. The voice source spectrum comprises a series of harmonic partials, so that partial number *n* always has the frequency of *n* times the fundamental frequency. In other words, the frequency of each partial changes in synchrony with the frequency of the fundamental. The amplitude of a spectrum partial is determined by two factors. One is its number: the higher the number of a partial, the weaker its amplitude in the voice source spectrum. The second factor is the frequency distance between the partial and the formants. When a partial approaches a nearby formant because of the vibrato, it gains in amplitude and vice versa, as was mentioned before.

It seems reasonable to assume that this phenomenon is used by the auditory system to gain information as to the frequencies of the formants. If a loud partial increases its amplitude considerably, when the fundamental frequency is rising in the vibrato cycle, there must be a formant just above the frequency of that partial. In other words, if the frequency and amplitude fluctuations of a partial are in phase, the frequency of the nearest formant is higher than that of the undulating partial, as illustrated by partial number 1 in Figure 12. And, conversely, if a loud partial decreases in amplitude as the fundamental frequency is rising in the vibrato cycle - that is, if they vary in counterphase - there must be a formant just below the frequency of that partial, as in the case of partial number 2 in the figure. Theoretically, there must be more information on formant frequencies in a vibrato tone than in a nonvibrato tone.

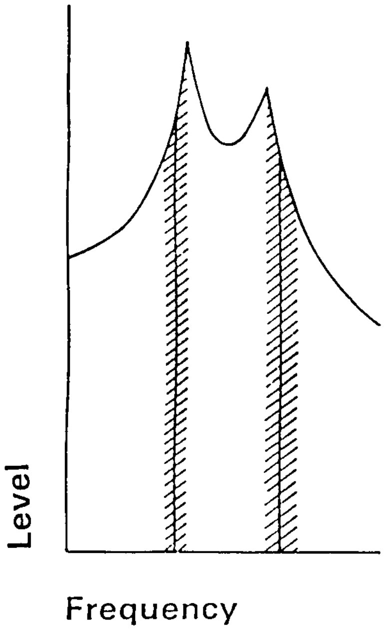


*Figure 12. As the vibrato induces an undulation of the frequencies of the spectrum partials, their amplitudes vary according to their position relative to the formants. In the case illustrated here the amplitude of the first partial varies in phase with the fundamental frequency, thus suggesting that the first formant is slightly higher than this partial. For the second formant counterphase conditions applies indicating that there is a formant slightly lower than the second partial. Thus, the amplitude variations of the partials provide information on the frequency locations of the formants and should therefore be expected to facilitate vowel intelligibility at high pitches.*

Often it is rather easy to determine the formant frequencies by looking at the spectrum: they are represented by peaks in the spectrum envelope, as illustrated in the left part of Figure 13. But when phonation frequency is very high, difficulties occur. If there are four or five partials and about the same number of formants, then it is hardly possible to guess where the formant frequencies are. This situation is illustrated in the right part of the same figure. A reasonable hypothesis would be that the vibrato makes guessing easier in such cases.

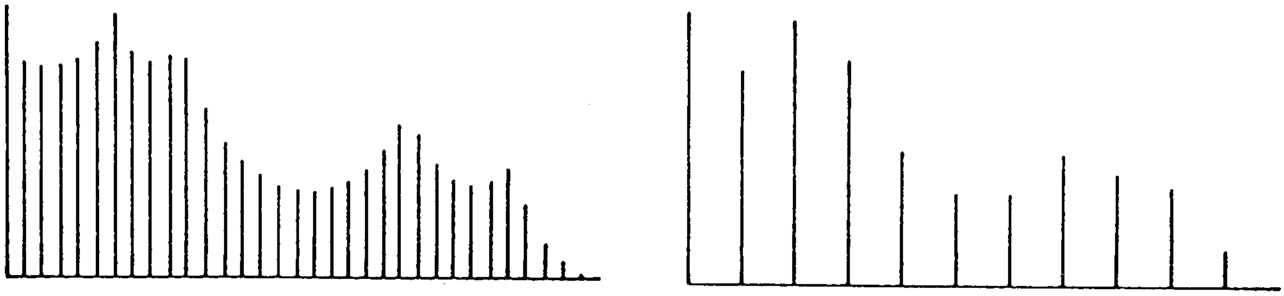


*Figure 13. Formant appear as spectrum peaks which are easy to discern under conditions of low pitch, as the spectrum partials are then densely packed along the frequency axis (left graph). When pitch is high, the spectrum partials appear with large intervals along the frequency axis, and the possibilities to identfy the formant frequencies are small; in fact, the formant frequencies are the same in both graphs as indicated by the fact that the same spectrum envelope fits both spectra.*

I investigated this hypothesis by means of an experiment with synthetic vowels (Sundberg, 1977). Phonation frequency was varied between 260 and 1040 Hz. The formant frequencies were those used by a female professional soprano when she sang the vowels at the pitch of C4, approximately 260 Hz. The vowels were presented with and without vibrato to a group of phonetically trained listeners. Their task was to identify the vowel they heard. The results revealed that the agreement in the listeners' interpretations was not clearly affected by the presence of the vibrato. The conclusion must be that the vibrato does not help the listener to identify the vowel sung.

As this conclusion is clearly counterintuitive, McAdams & Rodet (1988) made a different experiment to test it. To four subjects they presented tones that were equal when there was no vibrato and different when presented with a vibrato. Figure 14 shows the formant patterns they used for obtaining this effect. A perceptible effect was revealed. However, the task of the subjects was difficult. Some needed no less than 35 training sessions to note the difference. On the other hand, once they had learnt it, the subjects could readily distinguish and also learn to identify the two different stimuli. This experiment shows that it is possible to recognise the effect but also that it does not play an important role in vowel identification.

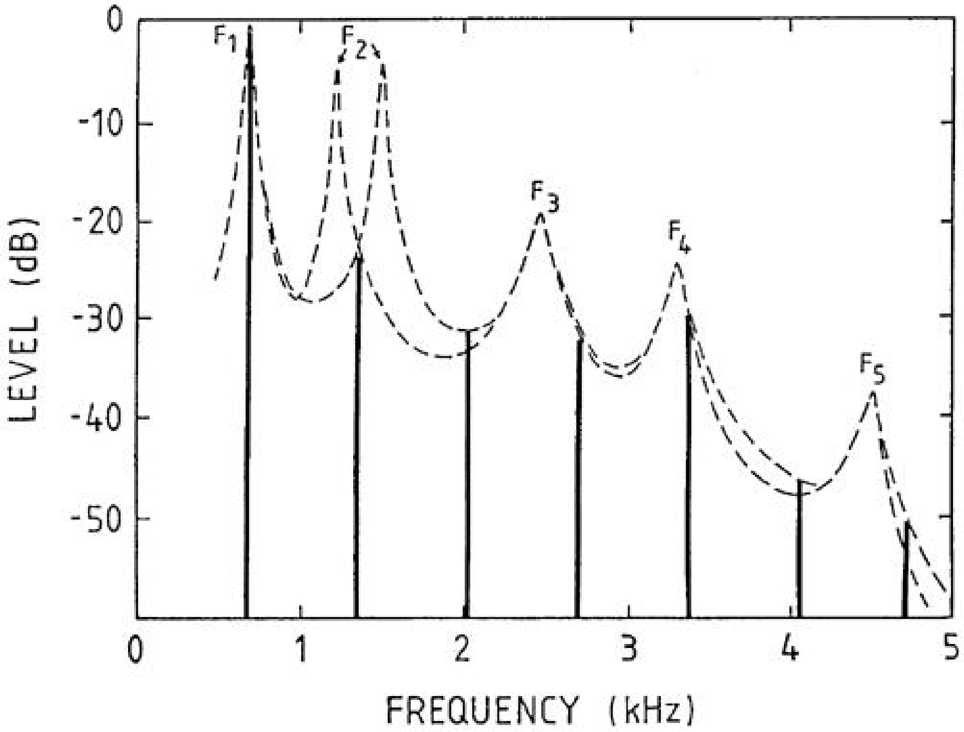


*Figure 14. The two spectra used in the experiment by McAdams & Rodet (1988). The two spectra are almost identical although they differ with respect to the second formant frequency. However, a vibrato will reveal this difference. When the spectra were presented with vibrato, subjects could learn, after proper training, to hear a difference. After McAdams & Rodet, 1988.*

### *c. Arpeggio*

The vibrato is related to a rather clear phenomenon associated with pitch glides. Figure 15 offers an example. It shows a synthesised fundamental frequency pattern which is heard as a sung arpeggio. First, the frequency changes by a glissando between two target pitches, one octave apart. Then a vibrato is superimposed on this fundamental frequency pattern. As a result, the frequency fall and rise are interrupted by small wiggles. These wiggles seem to cause the ear to perceive discrete pitches, or arpeggio tones rather than a glissando. Singers seem to use this effect when singing coloratura passages, i.e., sequences of short tones sung legato (Sundberg, 1989). Under these conditions the strategy seems to be that the fundamental frequency curve makes a wiggle around the target frequencies.

## 6. Why do singers use vibrato?

Above we have seen that the vibrato does not seem to reduce the demands on pitch accuracy; the pitch of a vibrato tone is practically as well defined as the pitch of a vibrato-free tone. Furthermore, the vibrato does not seem to increase vowel intelligibility. Why do singers sing with vibrato then? Two possibilities can be imagined.

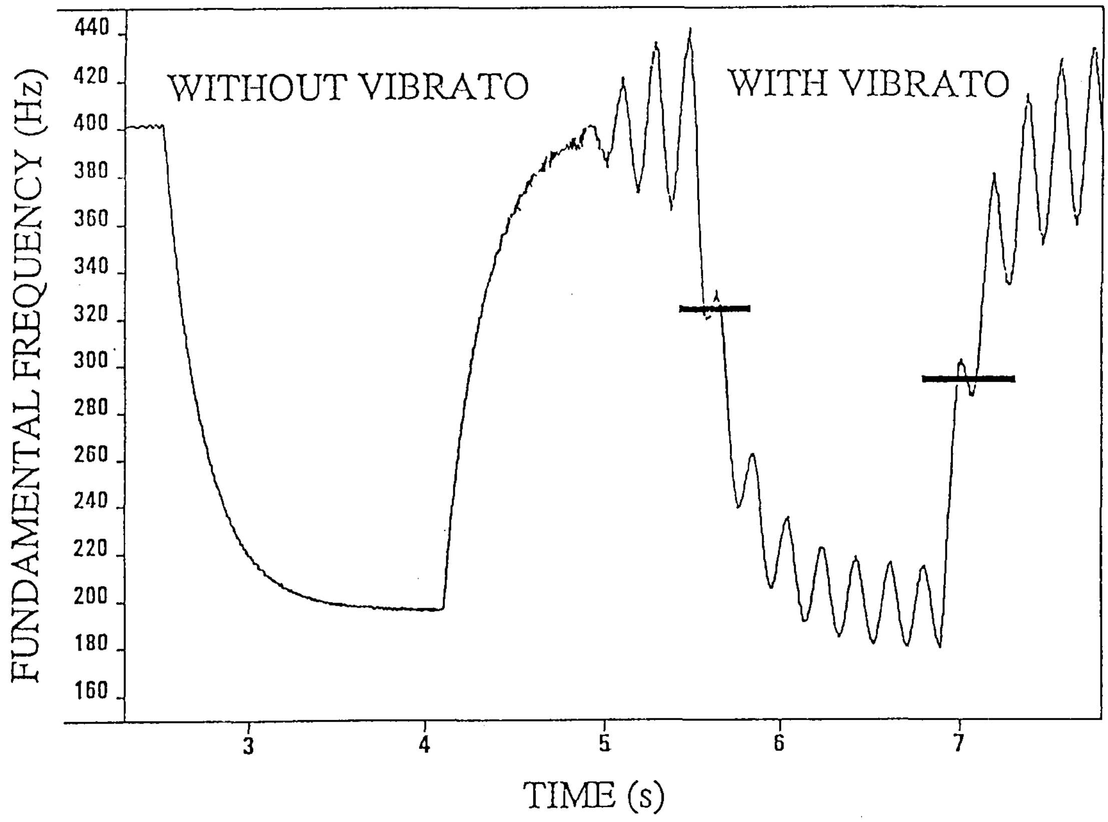


*Figure 15. Synthesised fundamental frequency pattern producing performance of a sung arpeggio. First the frequency changes by a glissando between the two target pitches, one octave apart. When a vibrato is added, the frequency fall and rise are interrupted by a little wiggle which causes the perception of an arpeggio.*

First, the experiment showing that the vibrato did not reduce accuracy of the pitch perceived, concerned successive tones; first the subjects heard the stimulus tone, and thereafter (not simultaneously) they heard the vibrato-free response tone. In our Western music culture, this is a rare case. Mostly we hear several tones at the same time which form consonant intervals such as thirds, fifths, and octaves. In such cases a failing accuracy in tuning is revealed to us not only in terms of the pitch perceived but also in terms of beats. If such intervals are sung slightly out of pitch, beats occur. The greater the pitch error, the quicker the beats, and eventually they give rise to roughness.

However, if the notes producing the consonant interval are sung with vibrato, no beats will occur. This means that the vibrato eliminates those beats which reveal that the consonant interval was not in accordance with the pure tuning. It should be noted, that pure tuning is not generally used in music practise, because it tends to sound queer in many contexts (Sundberg, 1991). Certainly this is a very good acoustical argument for using vibrato in singing as well as in other types of music. The singer's freedom is increased regarding the choice of fundamental frequency so that it can be used artistically, for expressive purposes. In other words, the vibrato gives the singer access to fine tuning as a means of musical expression.

It is also possible that the vibrato makes the singer's voice easier to discern against the background of a loud orchestral accompaniment. The vibrato causes the partial to vary in amplitude, and it seems quite likely that a signal with strong high partials which vary in amplitude is more readily detected than a signal with a constant spectrum. The same may also apply to the frequency undulations. These hypotheses, sometimes heard in discussions at voice conferences, would be possible to test by formal experiments with synthesised stimuli.

A third possible *raison d'être* for the vibrato is the following. Findings by Sundberg & Askenfelt (1983) suggest that the vibrato is often missing when the singer runs into phonatory problems, although, of course, the vibrato may be absent for artistic reasons as well. Large & Iwata (1971) showed that vibrato-free tones consumed less air flow than vibrato tones. This suggests that vibrato tones are produced with lesser degree of glottal adduction than nonvibrato tones (Gauffin & Sundberg, 1989). We may then speculate that in a vibrato singing, phonation is somewhat further away from overadducted, or pressed phonation than in nonvibrato singing. Perhaps the singer uses the vibrato to signal to the audience that phonation is far from pressed. In other words, the vibrato might be used in order to inform the listeners that the singer is solving a difficult vocal task without a struggle. It is certainly a basic condition for creating an aesthetically and artistically satisfactory result that difficult tasks are solved without apparent difficulty.

## Acknowledgement

The author is indepted to Eric Prame for valuable comments on the manuscript.

## References

Bennett, G. (1981). "Singing synthesis in electronic music". In *Research Aspects of Singing*, Stockholm: Royal Swedish Academy of Music, Publ. No. 33.

Björklund, A. (1961). "Analysis of soprano voices". *J. Acoust. Soc. Amer.* 33, pp. 575-582.

Castellengo, M. & Colas, D. (1991). "Etude acoustique de la réalisation et de la perception du trill vocal". Paper presented at the Voice Conference in Besançon, France.

Chowning, J. M. (1980). "Computer synthesis of the singing voice". In *Sound Generation in Winds, Strings, Computers* (J. Sundberg, ed.), Stockholm: Royal Academy of Music, Publ. No. 29.

d'Alessandro, C. & Castellengo, M. (1991). "Etude, par la synthese, de la perception du vibrato vocal dans les transition de notes". Paper given at the International Voice Conference in Besançon.

d'Alessandro, C. & Castellengo, M, (1991), "Etude de la perception des notes courtes chantées en presence de vibrato". Paper given at the 12th International Congress of Phonetic Sciences, Aix-en-Provence.

Damsté, H., Reinders, A. & Tempelaars, S. (1982). "Why should voices quiver?". In *Vox Humana.* Studies presented to Aatto Sonninnen, Publ. from the Institute of Finnish Langu. & Comm., Univ. Jyväskylä, pp. 26-34.

Fant, G. (1960). *Acoustic Theory of Speech Production*, The Hague: Mouton.

Gauffin, J. & Sundberg, J. (1989). "Spectral correlates of glottal voice source wave form", *J. Speech and Hear. Res.* 32, pp. 556-565.

Gramming, P. & Sundberg, J. (1988). "Spectrum factors relevant to phonetogram measurement", *J. Acoust. Soc. Amer.* 83, pp. 2352-2360.

Hakes, J., Doherty, T. & Shipp, T. (1987). "Acoustic properties of straight tone, vibrato, trill and trillo", *J. Voice*, 1, pp. 148-156.

Hakes, J., Doherty, T. & Shipp, T. (1990). "Trillo rates exhibited by professional early music singers", *J. Voice*, 4, pp. 305-308.

Horii, Y. (1989). "Acoustic analysis of vocal vibrato: A theoretical interpretation of data", *J. Voice* 3, pp. 36-43.

Imaizumi, S., Saida, H., Shimura, Y. & Hirose, H. (1993). "Harmonic analysis of the singing voice: Acoustic characteristics of vibrato". Paper given at Stockholm Music Acoustics Conf. 1993 (SMAC 93).

Iwamiya, S., Kosygi, K. & Kitamura, O. (1983). "Perceived pitch of vibrato tones", *J. Acoust. Soc. Japan* 4:2, pp. 73-82.

Keidar, A., Titze, I. & Timberlake, C. (1984). "Vibrato characteristics of tenors singing high C's". In *Transcripts of the 13th Symposium Care of the Professional Voice* (V. Laurence, ed.), Part I, pp. 105-110.

Large, J. & Iwata, S. (1971). "Aerodynamic study of vibrato and voluntary 'straight' pairs in singing", *Fol. Phoniatr.* 23, pp. 50-65.

McAdams, S. & Rodet, X. (1988). "The role of FM-induced AM in dynamic spectral profile analysis". In *Basic Issues in Hearing* (H. Duifhuis, J. Horst & H. Wit, eds.), London: Academic Press, pp. 359-369.

Michel, J. & Myers, D. (1991). "The effects of crescendo on vocal vibrato", *J. Voice* 5, pp. 292-298.

Myers, D. & Michel, J. (1987). "Vibrato and pitch transitions", *J. Voice* 1, pp. 157-161.

Ramig, L. & Shipp, T. (1987). "Comparative measures of vocal tremor and vocal vibrato", *J. Voice* 1, pp. 162-167.

Rothenberg, M., Miller, D. & Molitor, R. (1988). "Aerodynamic investigation of sources of vibrato", *Fol. Phoniatr.* 40, pp. 244-260.

Schultz Coulon, H. (1978). "The neuromuscular phonatory control system and vocal function", *Acta Otolaryngol.* 86, pp. 142-153.

Schultz Coulon, H. (1976). "Zur Bedeutung der kinästetisch-reflektorischen Phonationskontrolle für die Genauigkeit der Stimme", *Fol. Phoniatr.* 28, pp. 335-348.

Schultz Coulon, H. & Battmer, R. (1981). "Die quantitative Bewertung des Sängervibratos", *Fol. Phoniatr.* 33, pp. 1-14.

Seashore, C. (1938). *Psychology of Music, new printing 1967.* New York: Dover Publications.

Shipp, T., Leanderson, R. & Sundberg, J. (1980). "Some acoustic characteristics of vocal vibrato", *J. Research in Singing* IV:1, pp. 18-25.

Shonle, J. & Horan, K. (1980). "The pitch of vibrato tones", *J. Acoust. Soc Amer.* 67, pp. 246-252.

Sundberg, J. (1975). "Formant technique in a professional female singer", *Acustica* 32, pp. 89-96.

Sundberg, J. (1977). "Vibrato and vowel identification", *Arch. of Acoustics* 2, pp. 257-266.

Sundberg, J. (1978). "Effects of the vibrato and the 'singing formant' on pitch", pp. 51-69 in *Musica Slovaca* VI, 1978, Bratislava; also in *J. Research in Singing* 5:2 (1977), pp. 5-17.

Sundberg, J. (1989). "Synthesis of singing by rule". In *Current Directions in Computer Music Research* (M. Mathews & J. Pierce, eds), Cambridge, MA: The MIT Press, pp. 45-55 and pp. 401-403.

Sundberg, J. (1991). *The Science of Musical Sounds*, San Diego: Academic Press.

Sundberg, J. & Askenfelt, A. (1983). "Larynx height and voice source: a relationship?". In *Voice Physiology* (J. Abbs & D. Bless, eds), Houston TX: Collegehill, pp. 307-316.

Sundberg, J. (1984) "Using acoustic research for understanding various aspects of the singing voice". In *Transcripts of the 13th Symposium Care of the Professional Voice* (V. Laurence, ed.), Part I, pp. 90-104.

Ternström, S. & Sundberg, J. (1988). "Intonation precision of choir singers", *J. Acoust. Soc. Amer.* 84, pp. 59-69.

Winckel, F. (1967). *Music Sound and Sensation. A Modern Exposition.* New York: Dover Publications.

Winckel, F. (1953). "Physikalische Kriterien für objektive Stimmbeurteilung", *Fol. Phoniatr., Separatum*, pp. 231-252.

Winckel, F. (1974). "Acoustical cues in the voice for detecting laryngeal diseases and individual behavior". In *Ventilatory and Phonatory Control Systems* (B. Wyke, ed.), London: Oxford University Press.