Dept. for Speech, Music and Hearing
Quarterly Progress and Status Report

Violin timbre and the picket fence

Jansson, E. V.

journal: STL-QPSR
volume: 31
number: 2-3
year: 1990
pages: 089-095

VIOLIN TIMBRE AND THE PICKET FENCE

Erik V. Jansson

Abstract

On a bicycle ride along a picket fence, the author made the common observation that the picket fence became transparent at a sufficient speed. Objects on the other side became clearly visible. The idea came up that perhaps the same effect is working for acoustical signals and hearing. Such an effect could explain why the timbre of one violin is fairly independent of the played tones but different to the same tones played on another violin. The played tones could be regarded as the pickets and the "partials" deriving from the body, being the object being at rest. In reading an introductory text-book on hearing, support was found that fission can take place for acoustical events even close in time. The events may fuse, though. Also, from Gestalt theory the two possibilities, fission or fusion, can be predicted. It is suggested that experiments are made after a sketched line which may give an important contribution to the understanding of the violin timbre.

1. INTRODUCTION

Two questions have been spinning around in the head of the author for some time: What are the main parameters of the violin tone and what make different tones from one violin similar but different from those of another violin?

Traditionally spectra of tones are analyzed in form of spectral envelopes implying that the properties of the stable tones are the most important, although it is well known that the initial parts, the starting transients, are important for the perception too. Further, it is known that the violin tones are never steady – they vary all the time. Oscilloscope traces of waveforms are too coarse measures to reveal the perceived tone properties.

During a bicycle ride along a picket fence, some ideas were triggered by a common observation. At sufficient speed of the observer, the fence becomes "transparent", and objects on the other side form a steady and rather clear picture. In addition to the time integration, the human eye can enhance the image by focusing at the correct distance. Narrow slits in the fence are sufficient to give the steady picture. In the cinematograph, the same procedure produces moving pictures with a slit (alternating between closed and open) in a constant position. The observed phenomenon was named the "picket fence effect", and starting from this observation attempts were made to form a hypothesis and to develop theories.

A considerable amount of psycho-acoustical experiments has been made on the relations between the physical properties and the perception of the sound. A summary of fundamental findings was obtained by a holiday-reading of a book on hearing for a music acoustics course (Moore, 1983). From the reading, especially psycho-acoustical findings were sought which could be related to "violin timbre and the picket fence effect" and general clues to the understanding of human perception.

The main findings will hereby be summarized before starting the more formal work. References will be given to important original works even if only read as reviews in the Moore book.

2. IMPLICATIONS OF PSYCHO-ACOUSTICAL EXPERIMENTS

The ear works in some respects as a set of band-pass filters, i.e., the critical bands of hearing. Two tones within the same critical band are treated differently compared to two tones in separate critical bands. The critical bands of hearing enables us to separate the lowest five to six partials of a harmonic spectrum, such as a musical tone, and thus work as a spectrograph, i.e., give measures of the frequencies and intensities of the five to six lowest partials. For the seventh and higher partials, at least two partials fall within the same critical band of hearing.

For time-invariant patterns (analyzed in approximately critical bands of hearing), the timbre of a complex tone can be accounted for to 82% in three dimensions (Plomp, 1970; Plomp, Pols, & van de Geer, 1967). Three dimensions of perception were found to correspond remarkably well with a three-dimensional physical configuration (Pols, van de Kamp, & Plomp, 1969).

The hearing has a threshold for the weakest perceivable sound. The threshold varies between the different critical bands but smoothly between adjacent bands. A strong tone or a noise signal in one critical band influences the hearing threshold of neighbouring bands and gives a masked threshold with its level depending on the level and frequency of the masker.

A sound can thus be masked by another simultaneous stronger sound also at a different frequency (note that with two signals either of the two signals can be regarded as a masker or a probe). Furthermore, non-simultaneous masking is possible with the masker before, after, or surrounding the probe sound. The first two masking effects are most prominent within ±100 ms. The masking increases within 1-20 ms duration of the masker and becomes thereafter constant. Houtgast (1972) found that tone bursts intercepted with noise were perceived as continuous when they were above the noise-masking threshold but still below a higher, well reproducible threshold, the pulsating threshold (in "violin timbre and the picket fence" we are looking for the opposite – the complex tone is the masker). Thus, the masking is a rather complex function in our modelling of the hearing.

Auditory objects can be separated by means of fundamental frequency. Two complex tones fuse into a single voice if all partial frequencies coincide (Broadbent & Ladefoged, 1957). Experiments on onset disparities (Rasch, 1978) show that a higher tone in a pair of complex tones can be perceived down to -20 dB when the tones start simultaneously and have the same envelope. When the higher tone starts 30 ms earlier, it can be perceived down to -60 dB. For different rise times, the high tone can be perceived to very low levels. Rasch also found that the high tone is perceived as continuous even if it is terminated after the onset of the lower tone, and that our auditory system perceives no onset differences in time within 30 ms. Thus, small time-differences in onset or two different sets of partials (such as harmonically related partials of the played tone and the non-harmonically related partials deriving from body resonances) may be separated by the perception. Even so, weak sounds that are hardly detectable in an oscilloscope tracing may be important.

An initial sound may colour the following, for instance, a vowel immediately followed by noise. Correlated changes in amplitude or frequency can produce a perceptual fusion of components, also for components not harmonically related. Thus, there are factors working against the picket fence effect, i.e., fusing different sounds from one source into a single perceived sound.

For the perception of temporal patterns, say 3 to 4 elements per second, it was found that timing is the most important factor. For fast-tone sequences, say 10 per second, the sequences form a stream or several streams if the elements are considerably dissimilar. Another interesting finding is that the perception of temporal order of different sounds is poor; a duration of individual sounds of 200 to 700 ms may be needed for correct ordering but only 2-7 ms duration of components for the identification of tone sequences. The observations indicate that we have different modes of listening and that very short component durations are sufficient to identify tone sequences of sounds.

Thus, psycho-acoustic experiments show that auditory objects can be separated by means of fundamental frequencies and onset disparities which are called fission and provide a possibility for an acoustical "picket fence effect." A single cough works as an acoustical "picket" usually with little influence on perceived music. Music may also be intercepted successively by noise bursts and still be perceived as continuous, even if the music signals are deleted within the noise bursts (van Norden, 1975). On the other hand, the initial sound may fuse with the following sound and produce a single object thus deleting the possibility of the acoustical picket fence effect.

3. PICKET FENCE, VISION, AND HEARING

For vision, the space perception is a major parameter and the eye can focus at different distances. The hearing can also localize sources in space. Differences in intensity and arrival time at the two ears are used for the localization (in practical life enhanced by visual clues). Thereby, a trade between the intensity and time cues is used by the hearing. Especially the first arriving sound is important, it has a precedence effect (the same sound may be amplified and added slightly later without influencing the spatial impression).

In everyday life "the auditory world" is analyzed into discrete sound sources or "auditory objects" and not to single attributes as pitch, loudness, etc. The perception of auditory objects "depends primarily on structures in frequency and time." To distinguish between two musical instruments, the time-varying patterns may be important. Schouten (1968) suggested that the identification depends on: 1) whether the sound is periodic or irregular, 2) whether any aspect of the sound is changing as function of time, and 3) whether preceding and following sounds are alike.

It has been suggested that concepts of "source" and "stream" should be introduced and made distinct (Bregman, 1978; Bregman & Pinker, 1978). A ***source*** gives the acoustic waves, for instance, a played violin. A ***stream*** is the percept of the sound elements from the source; the hearing of the played violin, for instance. Many physical cues may be used to form different streams. This acoustic factoring is described by Bregman & Pinker as ***parsing*** – "the acoustic information is parsed to form separate streams in the same way that visual information falling on the retina is parsed to form objects and backgrounds." Simultaneous frequency components can be grouped together and connection over time can be made.

Gestalt psychology says that several factors are governing the perceptual organization – no single rule will always work but all rules together will give a correct interpretation. The factors are as follows.

Similarity: elements which are similar will be grouped together.

Good continuation: properties of a single source should change smoothly.

Common fate: the different components of a single source "usually varies in a similar way."

Belongingness: "a single component in a sound can only be assigned to one source at a time."

Closure: a sound obscured by a second will be perceived as continuous even if turned off during moments of obscure, and non-acoustic clues may be used for the filling in.

Figure-ground phenomenon: sound can be separated in attended and unattended streams (cf., cocktail parties).

Thus, one may conclude that the use of simple visual phenomena to give an idea of acoustical ones is not novel and it has been helpful. Our ***source***, the played violin, should give a ***stream***, the percept of the sound elements of the played violin. The stream of tones from one violin seems to fall in line with Gestalt psychology: similarity, good continuation, closure, etc. But a weak non-harmonic spectrum deriving from the resonances of the violin body may form a second ***stream*** and seems also to fall in line with the factors of Gestalt. Possibly, the effects of the initial transients can be perceived as continuous, somewhat similar to chopped sound above the pulsation threshold. A later portion of the transient is perceived as continuous but masked in the "steady" portion of the tone.

Thus, we should expect that when the auditory world (for example the played violin) is split into two sets of streams, then the acoustical picket fence effect may play an important role. The stream indicating a "constant source" (the resonance properties of the violin body) may stand out perceptually between the stream of fence pickets (the tones in the time domain and the partials in the frequency domain).

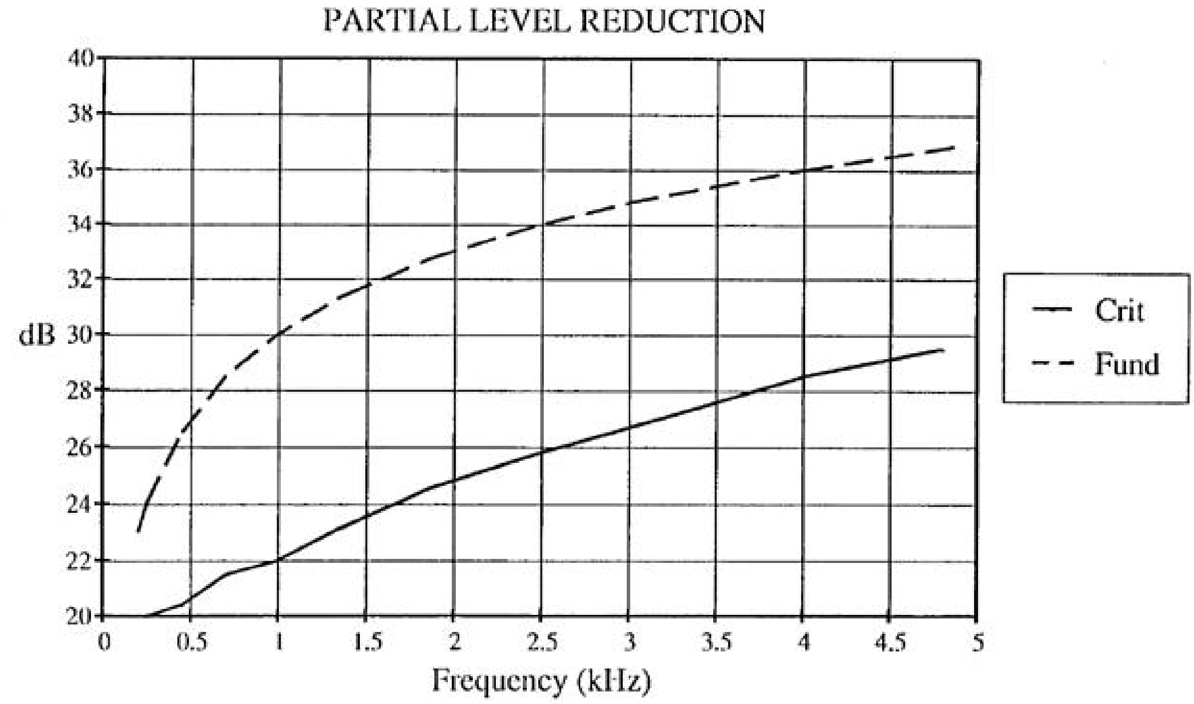

Fig. 1. Minimum level difference tone-to-noise for tone detection assuming 1 Hz bandwidth of the analysis and the total noise level set by the square root of critical bandwidths (full line). The total noise level is also plotted for analysis bands equal to partial separation (equal to fundamental frequency, broken line). The second noise level is adaptable when the partial separations are smaller than the critical bands. Detection threshold is assumed to be reached when noise and tone levels are equal.

4. CONCLUSIONS FROM PHYSICAL ACOUSTICAL PROPERTIES OF THE VIOLIN

In an investigation on "Timbre cues and the identification of musical instruments", Saldanha & Corso (1964) found that "initial transients and a short steady state" gave the best identification for musical instruments, i.e., the musical instruments are indeed not identified from steady state properties only.

Experiments by the author with guitar tones gave that the lowest resonances added a perceptually important thump sound initially to played tones (Jansson, 1983). Similar investigations of the piano tone revealed that there is a most important initial thump sound in the airborne tone which is not a part of the string vibrations (Reinholdt, Askenfelt, & Jansson, 1987).

The played violin string gives rise to a number of phenomena related to the resonances of the body. When a string partial is close to or at a resonance peak, the partial becomes stronger ("resonance amplification"). If the body resonance is easily driven then clear phenomena of coupling between string resonance and body resonance can result (for instance a wolf note). Further, the (transversal) direction of the string vibrations in relation to the body can be important. Strong coupling in one direction but small in the perpendicular result in slightly different resonance frequencies for the two vibration directions (a polarization effect). The mentioned phenomena are, however, limited to narrow frequency ranges and cannot answer the two introductory questions in general and other explanations must be sought.

The picket fence observation suggest that we should look at what happens in the time slits between the played notes (i.e., the pickets), i.e., at the starting transients. Let us assume a bandwidth of 1 Hz (the bandwidths vary somewhat around this value for free strings, see Jansson, 1990) and 10 Hz for the body (for more details see Alonso & Jansson, 1982). This means that the body resonance has reached 63% (-4 dB) of its final steady value after 0.03 s for step excitation. Corresponding time for the string vibration is 0.3 s, i.e., considerably later, and the vibrations of the body resonances should reach their peak values well before the string resonances. The experiments by Rasch showed that an asynchrony of 30 ms in the start between two tones can make the earlier tone detected even with a level 60 dB below the later one. If the later tone after its onset masks the first tone, then the first tone might be heard as continuous through the later and stronger tone.

The ear has a frequency resolving capacity and can be regarded as a bank of filters with bandwidths corresponding to the critical bandwidths of hearing. In a frequency band with noise, the total noise level is set by the width of the frequency band (the noise amplitude in a frequency band equals squareroot of power spectral density times bandwidth). This means that one should expect that a tone must be approximately 20 dB above the noise level (the power spectral density level) to be heard up to 1 kHz, 25 dB at 2 kHz, and 27 dB at 4 kHz, see Fig. 1. From the sixth to the seventh partial of a harmonic spectrum, the effective noise level is set by the frequency distance between the partials. The frequency difference equals the fundamental frequency.

A simple experiment at 1 kHz by the author gave that the tone of an electric sine generator should be 20 dB or more over the noise level to be heard when the tone was centred in the noise of critical bandwidth. With the tone 50 dB above the noise level, the noise was masked. Thus, it is clear that a critical band signal has more information than the level only. The close correspondence with predicted values is accidental but the numbers are still reasonable to use for a first approximation estimate (in the experiment, the displayed noise level was 6 dB below the power spectral density level).

The idea of a picket fence can also be used in the frequency domain by regarding partials as masking pickets. For the played violin, the lowest fundamental is close to 200 Hz (open G-string 196 Hz). This means that at least every second critical band is empty in which low noise-excited resonance peaks may be perceivable. Masking effects are likely to reduce their contribution to the percept, though.

From a series of experiments, a typical open G-string recorded at the left ear of the player was found to have the spectrum shown in Fig. 2. The data in Fig. 1 have been plotted and give an estimate for maximum allowed noise levels without partial maskings (the example should be regarded as a demonstration – the noise level in the diagram is 8 dB above the power spectrum density level). Thereby, it is found that the weak fundamental is likely to be masked by noise and that the partials above 3.5 kHz are only slightly above the "effective" noise level. For the frequency range in-between, the partials dominate. The noise should not mask the partials but is likely to be perceptible. Only the third partial is sufficiently strong to mask the corresponding critical band noise.

5. GENERAL CONCLUSIONS - SUGGESTIONS FOR FUTURE EXPERIMENTAL TESTS

The presented information suggests that the importance of initial transients should be evaluated using information on the string and body resonances. The analysis of the wave-forms as oscilloscope traces (the time domain) seems not to result in a direct answer. In the traces, important perceptual facts can be hidden such as quickly activated but weak body transients immediately preceding and interwoven with the strong string signals. The body transients may be perceived as starts for continuous sounds if their offsets are masked by the string sound.

The traditionally used envelopes of partial levels is also a coarse over-simplification for the frequency domain. Simple experiments indicate that both partials and noise may be detectable but seldom partials alone. The spectrum envelope may even lie below the noise level after correction for noise levels in critical bandwidths. In the noise, the resonance properties of the violin body can be mapped.

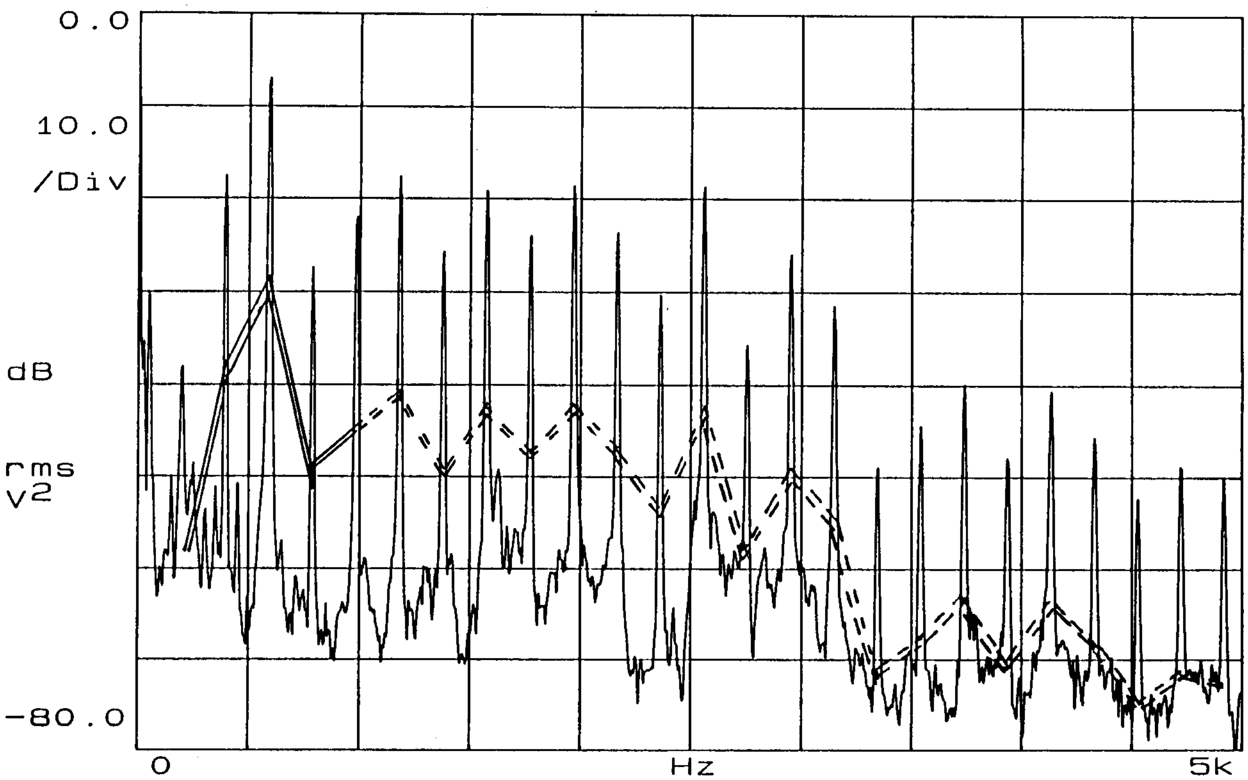

Fig. 2. Spectrum of a played open G-string (196 Hz) measured at the left ear of the player with maximum allowed noise levels not to mask partials (assuming an analysis bandwidth of 1 Hz). Full lines mark the region where the critical band noise level is adaptable and broken lines where the partial separation noise level is adaptable.

If our hearing acts as the picket fence for vision, it may use information both in the time and frequency domains. In the time domain, it may use the information available in the time slits between the notes (the pickets), and in the frequency domain the information in the frequency slits (between partial-pickets). Thereby, it may give a steady "picture" summed over a number of time and frequency slits. It is possible that our hearing picks out an acoustical stream of components that is one major part of the perceived violin tone and represents an acoustical fingerprint of the violin.

Thus, starting from the picket fence effect, it is suggested first to investigate the perception of the initial transient part of the violin tones by analysis and synthesis experiments. Secondly, it should be tested whether this transient part is received as present in the "steady" tone part. The two investigations are related to the picket fence effect in the time domain. Thirdly, it is suggested that it is tested whether the noise in the "steady" tone part is perceived and gives information of the violin properties - the picket fence effect in the frequency domain. Such tests may answer the two leading questions asked in the introduction.

References

Alonso, J. & Jansson, E. (1982): "Eigenmodes, input admittance and the function of the violin," *Acustica* **50**, pp. 329-337.

Bregman, A.S. (1978): "The formation of auditory streams," pp. 63-75 in (ed. J Requin) *Attention and Performance, Vol. 7*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Bregman, A.S. & Pinker, S. (1978): "Auditory streaming and the building of timbre," *Can.J.Psychol.* **32**, pp. 19-31.

Broadbent, D.E. & Ladefoged, P. (1957): "On the fusion of sounds reaching different sense organs," *J.Acoust.Soc.Am.* **29**, pp. 708-710.

Houtgast, T. (1972): "Psychophysical evidence for lateral inhibition in hearing," *J.Acoust.Soc.Am.* **51**, pp 1885-1894.

Jansson, E. (1983): "Acoustics for the guitar player," pp. 7-26 in (ed. E. Jansson) *Function, Construction and Quality of the Guitar*, Publ. issued by the Royal Swedish Academy of Music, no 38, Stockholm.

Jansson, E. (1990): "Experiments with the violin string and bridge," *Appl.Acoustics* **30**, pp. 133-146.

Moore, B.C.J. (1983): *An Introduction to the Psychology of Hearing* (2nd ed.), Academic Press, St Edmunds, Suffolk.

van Norden, L.P.A.S. (1975): *Temporal Coherence in the Perception of Tone Sequences*, Thesis, Technical University, Eindhoven.

Plomp, R. (1970): "Timbre as a multidimensional attribute of complex tones," pp. 397-414 in (eds. R. Plomp & G.F. Smoorenburg) *Frequency Analysis and Periodicity Detection in Hearing*, Sijthoff, Leiden.

Plomp, R., Pols, L.C.W., & van de Geer, J.P. (1967): "Dimensional analysis of vowel spectra," *J.Acoust.Soc.Am.* **41**, pp. 707-712,

Pols, L.C.W., van de Kamp, L.J.Th., & Plomp, R. (1969): "Perceptual and physical space of vowel sounds," *J.Acoust.Soc.Am.* **46**, pp. 458-467.

Rasch, R.A. (1978): "The perception of simultaneous notes such as in polyphonic music," *Acustica* **40**, pp. 21-33.

Reinholdt, A., Askenfelt, A., & Jansson, E. (1987): "Analysis and synthesis of piano tone," *J.Acoust.Soc.Am.* **81**, Suppl. 1, p. S1.

Saldanha, E.L. & Corso, J.F. (1964): "Timbre cues and the identification of musical instruments," *J.Acoust.Soc.Am.* **36**, pp. 2021-2036.

Schouten, J.F. (1968): "The perception of timbre," abstract GP-6-2 in *Reports 6th Int.Congr. on Acoustics, Tokyo, Vol 1.*