Consonance. §2: Psychoacoustic factors

2. Psychoacoustic factors.

'Sensory consonance' refers to the immediate perceptual impression of a sound as being pleasant or unpleasant; it may be judged for sounds presented in isolation (without a musical context) and by people without musical training. 'Musical consonance' is related to judgments of the pleasantness or unpleasantness of sounds presented in a musical context; it depends strongly on musical experience and training, as well as on sensory consonance. These two aspects of consonance are difficult to separate, and in many situations judgments of consonance depend on an interaction of sensory processes and musical experience.

Historically, some theorists have argued that the basis of perceived consonance is physiological or sensory (Helmholtz, 1863), while others have attributed it to the learning of relatively arbitrary cultural patterns (Lundin, 1947). However, one should not regard these theories as mutually exclusive. The relative importance of sensory factors and learning in a particular musical culture will depend on the types of sound being presented, on the instructions given and on the musical experience of the listeners. Psychoacoustic studies have usually emphasized sensory consonance, and tried to explain it in terms of the physical nature of the sounds and the way the sounds are analysed in the peripheral auditory system.

An interval that plays an important role in all scale systems is the octave, in which the frequency of the higher note is double that of the lower one. Consequently, two complete cycles of the higher note occur for each cycle of the lower note. Sensory theories of consonance are based on this fact, and on the fact that other common musical intervals – at least in Western, Indian, Chinese and Arab-Persian music (see Burns and Ward, 1982) – correspond to relatively simple ratios of frequencies (e.g. perfect 5th, 3:2; perfect 4th, 4:3; major 3rd, 5:4). Generally, when pairs of tones are played together, intervals giving simple ratios are heard as consonant, while intervals with complex ratios (such as minor 2nd, 16:15) are heard as dissonant. As a rough rule of thumb, ratios involving integers greater than 6 are heard as dissonant, while intervals involving ratios less than 6 are heard as consonant.

Sensory theories offer two (not mutually exclusive) explanations for this preference for simple ratios. The first is connected with the fact that when two sinusoids (pure tones) with similar frequencies are presented together, the total sound fluctuates in amplitude, an effect called 'beats'. The beats occur as the tones move alternately in phase (the peaks in the two tones coinciding) and out of phase (the peaks in one coinciding with the dips in the other). Beats occur at a rate equal to the difference in frequency between two sinusoidal tones, and in the case of complex tones they also occur between the harmonics (overtones). When two complex tones have fundamental frequencies in a simple ratio, such as 2:1, the harmonics of the upper tone always coincide in frequency with harmonics of the lower tone. Hence, no beats are audible. The more the fundamental frequencies depart from a simple ratio, the greater will be the tendency for beats between the harmonics. Intervals may be preferred that minimize audible beats between harmonics of the two notes (Helmholtz, 1863); this point is expanded later.

The second explanation is connected with the fact that action potentials (nerve impulses,
‘firings’ or ‘spikes’) in the auditory nerve tend to be synchronized to a particular phase of the stimulating wave in the cochlea or inner ear (see Hearing and psychoacoustics); for example, the impulses may occur close to the peaks of the wave. As a result, the time intervals between successive nerve impulses are close to integer multiples of the period of the stimulus (the time taken for one complete cycle). Thus, if the stimulus is a sinusoidal tone with a frequency of 500 Hz, then the intervals between successive nerve impulses cluster around values of 2 milliseconds, 4 milliseconds, 6 milliseconds and so on. Pairs of tones presented together may sound consonant when the intervals between nerve impulses share common values for the two tones (Meyer, 1898; Boomsliter and Creel, 1961).

Beats may be perceived differently depending on their rate (Plomp and Levelt, 1965). If two sinusoids are presented simultaneously, and their frequency separation is very small, then the beats are heard as slow fluctuations in loudness, but the sound is still consonant. If the two tones are moved apart in frequency, then the beats become faster and the percept is of rapid loudness fluctuations which are somewhat unpleasant. For still greater frequency separations, the loudness of the sound becomes steady, but the sound has a rough, unpleasant quality. Finally, when the frequency separation is sufficient, two separate tones are heard, and the overall sound appears to be consonant. At this point the tones are separated by about one quarter of the critical bandwidth, which corresponds to a frequency separation of about 3–4% (somewhat less than one semitone).

Plomp and Levelt and others (Kameoka and Kuriyagawa, 1969) have proposed methods for calculating the consonance of complex tones from the consonances of the simple-tone combinations that they contain. Generally, the calculation requires summation of the dissonance of all combinations of neighbouring harmonics (or partials in the case of non-harmonic complex tones, such as those produced by gongs or bells). The consonance is then inversely related to the total dissonance calculated in this way. The results of the calculations correspond reasonably well with subjective judgments. According to this theory, two simultaneous harmonic complex tones having fundamentals with simple ratios sound consonant because the lower harmonics of the two tones are either widely separated in frequency or coincide. The lower harmonics are usually the more important in determining the overall impression of consonance, because, for most musical instruments, they are more intense than the higher harmonics. If the frequency ratio is less simple (if the ratio cannot be expressed by integers less than about 6), then there will be a number of harmonics from the two tones that differ only a little in frequency, and these will give rise to beats and to dissonance.

Explanations based on beats cannot explain all aspects of the perception of consonance. For example, beats do not occur between successive notes, yet a certain amount of dissonance can be experienced for such notes. Also, beats do not occur when the two notes of an interval are presented one to each ear, yet some dissonance may be experienced. Both of these effects might be the result of learning. A given dissonant interval will have become familiar under conditions where the tones are presented simultaneously to the same ear. Hence that interval is automatically associated with a sensation of dissonance, and this sensation may persist when the tones are presented sequentially or to opposite ears.

The second sensory explanation of the preference for simple ratios – attributing consonance
to the coincidence of neural firing over a period of time – was put forward by Meyer (1898), and has since been supported by Boomsliter and Creel (1961), among others. Boomsliter and Creel emphasized that there is no sensation of pitch when the sounds are very short. Generally, periodic sounds containing fewer than about ten cycles appear click-like rather than tone-like, and the tonality progressively increases as the duration increases (Doughty and Garner, 1948; Moore, 1973). Boomsliter and Creel’s ‘long pattern hypothesis of harmony and hearing’ arose from this fact. They argued that both pitch and the consonance between notes require the synchronization of neural firing to individual cycles of the sounds, and that the synchronization needs to persist for a certain time for it to be analysed. Similar arguments have been advanced more recently to explain the perception of pitches based on the temporal synchrony of neural firing to the stimulus waveform (Plack and Carlyon, 1995; Moore and Sek, 1996).

An explanation in terms of the synchrony of neural impulses to individual cycles of the sounds is supported by the observation that both our sense of musical pitch and our ability to make octave matches largely disappear above 5 kHz (Ward, 1954; Attneave and Olson, 1971), the frequency at which neural synchrony no longer appears to operate (Palmer and Russell, 1986). Furthermore, the highest note (fundamental) for instruments in the orchestra lies just below 5 kHz. One could argue from this that the disappearance of musical pitch at high frequencies is a result of a lack of exposure to tones at these frequencies. However, notes produced by musical instruments do contain harmonics above 5 kHz, so that if the learning of associations between harmonics were the only factor involved, there would be no reason for the change at 5 kHz.

It is of interest that the equal-temperament scale in general use today does not consist of notes in exact simple ratios. This might appear to undermine sensory theories based on either beats or neural synchrony. However, the deviations from a simple ratio scale are small. For example, the interval of a perfect 5th corresponds to a frequency ratio of 3:2; on the equal-temperament scale the ratio is 2.9966:2. This deviation may produce a small increase in beating between the upper harmonics of complex tones, but the effect is not very noticeable.

A different explanation of the preference for simple ratios between the frequencies of musical tones is that humans learn about octave relationships and other musical intervals by exposure to harmonic complex sounds (usually speech sounds) from the earliest moments in life. For example, the first two harmonics in a periodic sound have a frequency ratio 2:1, the 2nd and 3rd have a ratio 3:2, the 3rd and 4th 4:3, and so on. Thus by exposure to these sounds we learn to associate harmonics with particular frequency ratios. Terhardt (1974) suggested that such a learning process could account for the perception of the pitch of complex sounds, and especially the perception of the pitch of tones with ‘missing fundamentals’ (see Hearing and psychoacoustics). Judgments of similarity and of consonance or dissonance may depend upon a similar learning process (Terhardt, 1974); we learn which musical intervals ‘belong’ together by learning the intervals between the lower harmonics in complex tones. Terhardt referred to this as ‘harmony’ and distinguished it from sensory consonance.

While simple ratios may be preferable for simultaneously presented tones, it is not clear whether this is the case for tones presented successively. A number of experiments investigating preferred notes in performances on stringed instruments of various kinds have shown that there is no simple answer. Some workers have found preferences for simple ratios, while others have found that the preferred scale corresponds fairly closely to equal temperament, except that notes higher than the tonic or keynote tend to be sharpened relative to that note. Musicians were asked to play familiar tunes on a monochord, a one-
stringed instrument with continuously variable tuning; while subjects consistently chose the same tuning for a given note within a given tune, they chose different tunings, for what is ostensibly the same note, in different melodies and in different parts of the same melody (Boomsliter and Creel, 1963). However, the chosen patterns formed a structure of small whole-number ratios to the tonic and to additional reference notes linked by small whole-number ratios to the tonic. Thus within small groups of notes simple ratios are preferred, although the ‘reference’ point may vary as the melody proceeds. Others have concluded, in contrast, that ‘there is no evidence … that suggests that the performers tend to play intervals corresponding to exact small-integer ratios, … for either melodic or harmonic situations’ (Burns and Ward, 1982, p.259).

There have been few cross-cultural studies of the perception of consonance. This makes it difficult to assess the relative importance of learning and of innate sensory/perceptual processes; moreover, few cultures remain that are accessible to experimenters but unaffected by Western music. One cross-cultural study found no meaningful differences in preferences for musical intervals between American and Japanese students (Butler and Dalston, 1968). However, aesthetic preferences seem to be distinguishable from psychoacoustic judgments. Consistent results across listeners for judgments of ‘roughness’, have been reported (Taylor, 1965); however, the judgments became very inconsistent when the criterion was changed to ‘pleasantness’. A factor-analytic study of the determinants of consonance judgments yielded two psychoacoustic factors, ‘pitch’ and ‘fusion’, and one aesthetic or ‘evaluative’ one (van de Geer, Levetl and Plomp, 1962).

Some studies have shown that preferences for musical intervals change with age, at least among schoolchildren. One found that British children under nine years of age did not show distinct preferences for tones with simple frequency ratios (as compared to complex ratios), whereas children over 12 did (Valentine, 1913). Such changes may reflect the development of sensory and perceptual skills, or they may reflect increasing familiarity with the ‘grammar’ of the local musical idiom and a development of concepts like the resolution of tension in music. How much the perception of consonance and dissonance is due to basic sensory and perceptual factors and how much to learnt ones remains unresolved.

See also Psychology of music, §II, 1.
Consonance.

BIBLIOGRAPHY

history


S. Gut 'La notion de consonance chez les théoriciens du Moyen Age', *AcM*, xlvi (1976), 20–44


psychoacoustic factors

H. von Helmholtz *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (Brunswick, 1863, 6/1913/R; Eng. trans, 2/1885/R as *On the Sensation of Tone*)

M. Meyer 'Zur Theorie der Differenztöne und der Gehörsempfindungen überhaupt', *Beiträge zur Akustik und Musikwissenschaft*, ii (1898), 25–65

C.W. Valentine 'The Aesthetic Appreciation of Musical Intervals Among Schoolchildren and Adults', *British Journal of Psychology*, vi (1913), 190–216


C. Taylor *The Physics of Musical Sounds* (New York, 1965)


A. Kameoka and M. Kuriyagawa 'Consonance Theory I: Consonance of Dyads', *JASA*, xliv (1969), 1451–9; 'II: Consonance of Complex Tones and its Calculation Method', 1460–69


© Oxford University Press 2007