Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles

DAVID HURON

Music Perception 2001, 19, 1-64
Voice Leading

The manner in which individual parts or voices move from tone to tone to create perceptually independent musical lines.
Rules of voice leading

1. **Registral Compass Rule**: Write in the region between $F_2$ and $G_5$

2. **Textural Density Rule**: Harmony should be written using three or more concurrent voices.

3. **Chord Spacing Rule**.
   - soprano - alto $<=$ octave
   - alto-tenor $<=$ octave
   - tenor - bass no restriction.

4. **Avoid Unisons Rule**: Avoid concurrent unison

5. **Common Tone Rule**: Pitches common to consecutive chords should be retained in the same voice.
6. **Nearest Chordal Tone Rule.** Move to the nearest available pitch.

7. **Conjunct Motion Rule.** If possible by a diatonic step. Or, expressed in reverse:

8. **Avoid Leaps Rule.** Avoid large melodic intervals.

9. **Part-Crossing Rule.** Pitch contours should not cross.

10. **Part Overlap Rule.** No part should move to a more extreme pitch higher than the immediately preceding pitch in another part.
11. *Parallel Unisons, Fifths, and Octaves Rule*. No two voices should move in parallel octaves, fifths, or unisons. Some even say:

12. *Consecutive Unisons, Fifths, and Octaves Rule*. None even if not parallel.

13. *Exposed (or Hidden or Direct) Octaves (and Fifths) Rule*. Bass and soprano voices should not approach octaves by similar motion between two parts unless at least one of the parts moves by diatonic step.
Perceptual principles

1. *Toneness Principle*. Strong auditory images are evoked when tones exhibit a high degree of toneness or virtual pitch weight eg Harmonic complex tones between F2 and G5.
Fig. 1. Changes of maximum pitch weight versus pitch for complex tones from various natural and artificial sources; calculated according to the method described by Terhardt, Stoll, and Seewann (1982a, 1982b). Solid line: pitch weight for tones from C to C, having a sawtooth waveform (all fundamentals at 60 dB SPL). Dotted lines: pitch weights for recorded tones spanning the entire ranges for harp, violin, flute, trumpet, violoncello, and contrabassoon. Pitch weights were calculated for each tone where the most intense partial component of the tone was the fundamental.

In order to evoke strong auditory streams, use continuous or recurring rather than brief or intermittent sound sources. Intermittent sounds should be separated by no more than roughly 800 ms of silence in order to ensure the perception of continuity.

But avoid overlapping sounds as in normal piano or gamelan (*tutupan*) performance.
3. Minimum Masking Principle

Approximately equivalent amounts of spectral energy should fall in each critical band / auditory filter. For typical complex harmonic tones, this generally means that simultaneously sounding notes should be more widely spaced as the register descends.
Fig. 3. Approximate size of critical bands represented using musical notation. Successive notes are separated by approximately one critical bandwidth = roughly 1-mm separation along the basilar membrane. Notated pitches represent pure tones rather than complex tones. Calculated according to the revised equivalent rectangular bandwidth (ERB) (Glasberg & Moore, 1990).
3. Minimum Masking Principle

**Fig. 4.** Average spacing of tones for sonorities having various bass pitches from C₄ to C₂. Calculated from a large sample (>10,000) of four-note sonorities extracted from Haydn string quartets and Bach keyboard works (Haydn and Bach samples equally weighted). Bass pitches are fixed. For each bass pitch, the average tenor, alto, and soprano pitches are plotted to the nearest semitone. (Readers should not be distracted by the specific sonorities notated; only the approximate spacing of voices is of interest.) Note the wider spacing between the lower voices for chords having a low mean tessitura. Notated pitches represent complex tones rather than pure tones.
4. Tonal Fusion Principle

To maintain independence avoid intervals that promote tonal fusion eg:

unisons, octaves, perfect fifths
Fusion and consonance

Fig. 5. Comparison of sensory consonance for complex tones (line) from Kaestner (1909) with interval prevalence (bars) in the upper two voices of J.S. Bach’s three-part Sinfonias (BWVs 787–801). Notice especially the discrepancies for P1 and P8. Reproduced from Huron (1991b).

The coherence of an auditory stream is maintained by close pitch proximity in successive tones within the stream.

Pitch-based streaming is assured when pitch movement is within van Noorden’s fission boundary (region 1 - normally 2 semitones or less for tones less than 700 ms in duration).

When pitch distances are large, it may be possible to maintain the perception of a single stream by reducing the tempo.
5. *Pitch Proximity Principle*

![Graph showing pitch separation and inter-onset time](image)

**Fig. 6.** Influence of interval size and tempo on stream fusion and segregation (van Noorden, 1975, p. 15). Upper curve: temporal coherence boundary. Lower curve: fission boundary. In Region 1, the listener necessarily hears one stream (small interval sizes and slow tempos). In Region 2, the listener necessarily hears two streams (large interval sizes and fast tempos).
Fig. 8. Frequency of occurrence of melodic intervals in notated sources for folk and popular melodies from 10 cultures (n = 181). African sample includes Pondo, Venda, Xhosa, and Zulu works. Note that interval sizes correspond only roughly to equally tempered semitones.
5. Pitch Proximity Principle

Fig. 9. Distribution of note durations in 52 instrumental and vocal works. Dotted line: note durations for the combined upper and lower voices from J. S. Bach’s two-part Inventions (BWVs 772–786). Dashed line: note durations in 38 songs (vocal lines only) by Stephen Foster. Solid line: mean distribution for both samples (equally weighted). Note durations are corrected for any duration prolongation and duration reduction on the harpsichord.

"Bach notes shorter since harpsichord not continuous?"
5. *Pitch Proximity Principle*

Fig. 7. Schematic illustration of two possible perceptions of intersecting pitch trajectories. “Bounced” perceptions (right) are more common for stimuli consisting of discrete pitch sequences, when the timbres are identical.

Failure to perceive crossing explained by pitch proximity.

demo on Bregman CD track 17

The perceptual union of concurrent tones is encouraged when pitches move together maintaining constant ratios.

Darwin, Ciocca & Sandell, 1994 JASA showed that a mistuned harmonic continued to contribute to pitch with greater mistuning when coherent FM than when none.

Bregman micro-mod track 24.
7. *Onset Synchrony Principle.*

To keep parts independent, avoid note onsets of less than 100ms.

Fig. 13. Results of a comparative study of tonally fused intervals in polyphonic and (predominantly) homophonic repertoires by J. S. Bach. Only perfect harmonic intervals are plotted (i.e., fourths, fifths, octaves, twelfths, etc.). As predicted, in polyphonic textures, most perfect intervals are approached with asynchronous onsets (one note sounding before the onset of the second note of the interval). In the more homophonic chorale repertoire, most perfect intervals are formed synchronously; that is, both notes tend to begin sounding at the same moment. By contrast, the two repertoires show no differences in their approach to imperfect and dissonant intervals. From Huron (1997).
8. *Principle of limited density*

To keep parts independent, have 3 or fewer parts.

Fig. 14. Voice-tracking errors while listening to polyphonic music. Solid columns: mean estimation errors for textural density (no. of polyphonic voices); data for 130 trials from five expert musician subjects. Shaded columns: unrecognized single-voice entries in polyphonic listening; data for 263 trials from five expert musician subjects (Huron, 1989b). The data show that when listening to polyphonic textures that use relatively homogeneous timbres, tracking confusions are common when more than three voices are present.
9. *Timbral differentiation*

To keep parts independent, have unique timbres

See also:
Bregman CD demo 9

10. *Spatial differentiation*

To keep parts independent, have them spatially separated.

Bregman demo of streaming by spatial location in East African xylophone music. Demo 41.