Sound-Motion Similarity in Musical Experience

Rolf Inge Godøy
Department of Musicology, University of Oslo
e-mail: r.i.godoy@imv.uio.no

Music, Mind, Motion, Machines

www.fourms.uio.no
Main similarity issues in our sound-motion research:

- *Cross-modal similarity in music*, in particular subjective perception of similarities between sound and body motion

- *Approximate similarity in imitation*: the ability to reproduce salient and readily recognizable features in oral and improvisational musical contexts

- *Similarity in musical translations*: recognizing musical ideas as similar across very different instrumental and/or vocal versions and arrangements

A suspected common factor in these (and several other similarity issues) is *motor cognition*, i.e. our disposition to perceive and think in terms of body motion
Action sensations

Action units by constraints

Multimodal action images

Continuous sound tracing

Continuous sound
Motor theory of perception: Mental simulation and imitation of body motion (including postures):

- *Motor theory* initially in linguistics, but in recent years extended to several other domains
- Perception and cognition as a process of constant mental simulation
- Covert imitation of the body motion of others and the assumed actions behind what we hear
- Neurocognitive research seems to confirm this incessant mental simulation/imitation going on
- Overt imitation can be found in *scat singing* or *beatboxing*, as imitations of various instrumental (and everyday) sounds with the vocal apparatus
Motion shapes in beatboxing:

- Vocal imitation of what are really not vocal sounds
- Attests to the feasibility of the motor theory in auditory perception
- Attests to quite sophisticated motor control of vocal apparatus
- One important feature of the motor theory is *approximate imitation*, as in quasi languages (French, Italian, German)
- Likewise approximate imitation in various so-called *air instrument* performances and *sound-tracing*
Motion shapes in listening

- Attest to quite extensive musical knowledge also among untrained listeners

- Air instrument performance, i.e. motion *as if playing an instrument*: detail annotations suggest global features of onsets, pitch and loudness well rendered, more variability in details

- Sound-tracing: various statistical evaluations seem to confirm pitch and loudness, including fluctuations in these, as the most prominent, with more variability for timbral features, e.g. spectral centroid, filter open-close, etc.

- Need to look at categories of music-related motion
Music-related body motion

- Two main classes of music-related body motion (see Godøy and Leman 2010 for details)
- *Sound-producing* body motion (instrumental or vocal), furthermore including *excitatory*, *modulatory*, *ancillary*, and *communicative* motion, all which can also occur in air instrument performance
Example motion - sound similarity in production:

- Goal-directed trajectories
- Coarticulation (contextual smearing) in both motion and resultant sound
- In this example, a long prefix and a short suffix:
Understanding Coarticulation in Music

like cascade of tones leading up to the goal point of the Bb5, and with the marimba player’s right hand continuing after hitting this Bb5 with a ‘follow through’ motion similar to what can be seen e.g. in golf or tennis.

Figure 3. The marimba performance with four mallets of a rapid prefix trajectory leading up to the goal point of the B♭5. The position, velocity, and acceleration of the four mallets are displayed below the notation of this passage, and at the bottom of the figure we see a ‘cumulative’ trajectory picture of the mallets’ and whole body of the performer by the markers’ ‘tails’.

We hypothesize that the focus on goal points in coarticulation may be related both to the abovementioned Psycholinguistic Refractory Period [4] and to findings on...
Music-related body motion

- *Sound-accompanying* body motion, such as in *dancing, walking, gesticulating*, and *sound-tracing*, rendering some feature(s) of pitch, loudness, timbre, various fluctuations, etc.

- Here, spontaneous sound-accompanying motion trajectories modified in successive repetitions, question of what features are reflected:
Timescales for similarity considerations:

- **Sub-chunk level**: Continuous sound and movement below the chunk level of duration (i.e. below roughly 0.5 seconds)

- **Chunk level**: Holistically perceived fragments of sound and movement roughly in the 0.5 to 5 seconds range as with Pierre Schaeffer’s *sonic objects*, and this is also the timescale of postures and motion

- **Supra-chunk level**: Concatenations of chunks into larger scale units, i.e. into sections, movements, and whole works
Timescales in music production

• Although large span between longest and shortest durations, limits in either direction

• Various biomechanical constraints (maximum speed, need for rests, posture changes, etc.) and attention constraints (with a need to make motion automatic) determine timescales in music

• Body motion also shape sound, e.g. fusion by coarticulation

• And: so-called phase transition thresholds in motion, e.g. between singular strokes and tremolo:
Chunk-level typological categories by motion

- Sound categories suggested by Pierre Schaeffer (1966), and that correspond to biomechanically distinct body motion categories:
  - *Impulsive* = discontinuous effort
  - *Sustained* = continuous effort
  - *Iterative* = rapid series of impulses, i.e. continuous effort but bouncing back and forth such as in a drum roll
  - *Phase-transitions* between categories based on duration and density of events
Some sub-chunk level morphological categories:

- **Grain** = Continuous movement across a rough surface, e.g. the 'brrrrrrr' of a double bass
- **Gait (“Allure”)** = Slower fluctuations in harmonic content, in pitch, in loudness, etc.
- Schaeffer also suggested these categories may apply across different timbres and instruments, e.g. a generic 'brrrrrr' grain from different sources

General point: Sound seems to be a good transducer of motion shape information, and motion shapes could be the basis for similarity differentiation of most perceived sonic features as in Schaeffer's theory:
<table>
<thead>
<tr>
<th>CRITERIA for musical perception</th>
<th>TYPES</th>
<th>CLASSES</th>
<th>GENRES</th>
<th>SPECIES (site and width of the dimensions in the musical field)</th>
<th>DURATION (threshold of recognition for short sounds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2-3) Qualification (4-9) Evaluation</td>
<td>TYPICAL</td>
<td>MUSICAL</td>
<td>MUSICAL</td>
<td>MUSICAL</td>
<td>TYPICAL</td>
</tr>
<tr>
<td>TYPO-MORPHOLOGICAL SUMMARY</td>
<td>MUSICAL</td>
<td>MUSICAL</td>
<td>MUSICAL</td>
<td>MUSICAL</td>
<td>MUSICAL</td>
</tr>
<tr>
<td>MASSES</td>
<td>TONE</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>TONIC type N</td>
<td>TONE</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>COMPLEX</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>VARIABLE Y</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>ANY W.K.T</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>DYNAMICS</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>HARMONIC TIMBRE</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>MELODIC PROFILE</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>VARIATIONS</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>PROFILE OF MASS</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>GRAIN</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>SUSTAIN</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
<tr>
<td>GAIT</td>
<td>X</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
<td>MAJOR</td>
</tr>
</tbody>
</table>
Some main issues of similarity in general:

- Correlations *subjective - signal based* similarities
- Subjective feature differentiation (Schaeffer 1966)
- Intentionality in feature focus (Schaeffer 1966)
- Context of features (Tversky 1977)
- Prototypes (e.g. Rosch 1975)
- *Inter-categorical vs. intra-categorical* variation
- Categorical boundaries, e.g. between 'plucked' and 'bowed' sounds dependent on attack rate
General issues of motion similarity:

- Our sensitivity to motion cues for qualitative features
- Relevant example: BioWalker with quantitative to qualitative mappings [http://www.biomotionlab.ca/Demos/BMLwalker.html](http://www.biomotionlab.ca/Demos/BMLwalker.html)
- *Global similarities*, e.g. quantity of motion (QoM)
- *Local/detail similarities* (cf. air instrument studies)
- A number of salient features, e.g. *acceleration profiles*, *mean square jerk* (MSJ), etc. (cf. Hogan and Sternad 2007)
- Overall shape: compare mocap trajectory shapes for different performances of the same music
Similarities in objects and trajectories

- Approximate similarity
- High-level *motor equivalence*, i.e. motor programs applicable to different effectors, makes for similarity across several variants
- Recognize fairly large variations in effector anatomy
- Likewise large variations in optimal motion modes between musicians
Similarities in objects and trajectories

- Exploring hybrids in music, both at the chunk and the sub-chunk levels
- Timbre identity for one instrument across pitch range
- Timbre identity for one instrument across playing modes
- Timbre similarity in spite of differences, tamtam-piano:
- MDS studies of timbral similarity
- Musical translations (arrangements, orchestrations), literal ('google translate') vs. more idiomatic
Similarities in objects and trajectories

- Salient but approximate similarity
- In still pictures
- In motion pictures
- In postures and trajectory shapes
A motor theory approach to similarity:

- Motor theory (radical versions) suggests that most perception and cognition, hence, most reasoning, understanding, feeling, etc. is based on sensations (real or imagined) of body motion.

- Motor images have the great advantage of being very flexible and adaptable, e.g. by so-called motor equivalence, yet also very robust.

- Motor images may be the basis for categories and prototypes.

- Humans (and other animals) seem to be very sensitive to nuances of body motion (expressivity and affect).
In summary:

• Our claim: experience of similarity in music often linked with sensations of body motion

• Similarity often approximate, cf. Harnad's idea of "approximationism" in categorization

• Extensive challenges of bridging the qualitative-quantitative gap in similarity experience

• An analysis-by-synthesis experimental approach seems promising

• Good processing tools for similarity, e.g. the MIRToolbox, MoCapToolbox, Motiongrams, MoCapgrams, Meinard's work....

• Our own work towards a Musical and Motion Database