The Hypothesis of Self-Organization for Musical Tuning Systems

Jean-Julien Aucouturier

Grad. School of Arts and Sciences
The University of Tokyo, Japan
aucouturier@gmail.com

Abstract. Musical tuning systems are found in intriguing diversity in human cultures over the world and over the history of human music making, from the western hegemony of 12-tone equal temperament (C, $C\sharp$, $D\ldots$) to e.g. the inharmonic indian 22 *shruti* system. Traditional justifications for the adoption of such musical systems consider tuning as an algorithmic optimization of consonance. However, it is unclear how this can be implemented in a realistic evolutionary process, with no central entity in charge of optimization. Inspired by the methodology of artificial language evolution, we propose that tuning systems can emerge as the result of local musical interactions in a population. We show with computer simulations that such interactional mechanisms are capable of generating coherent artificial tunings that resemble natural systems, sometimes with a diversity and complexity unaccounted by previous theoretical justifications. However, the self-organization of realistic tuning systems is found here to require non-trivial environmental and cultural constraints. Notably, advanced musical activities such as primitive harmonic accompaniment (drone tones) and using different types of instruments simultaneously seem to be necessary ingredients.

1 Introduction

Considerable attention has been paid to the emergence of linguistic abilities in human evolutionary history, with cross-disciplinary effort over the fields of linguistics, anthropology, neurophysiology or artificial life (see e.g. [Christiansen and Kirby, 2003]). However, its musical counterpart has been virtually ignored until recently [Mithen, 2006], with only emerging awareness that the cognitive abilities of music making and listening may be more than secondary outcomes of language capacities, but rather central driving forces in the evolution of the modern human mind [Cross, 2007]. Music making is a multifaceted behavior, which can be jointly analysed from a physical point of view (describing the properties of the musical stimuli, e.g. frequency, periodicity, etc.), but also a cognitive (starting from the basic sensations of e.g. pitch, rhythm, etc. evoked by the former) and cultural one (music is grounded in social interaction, as noted e.g. in [Blacking, 1973]). This paper is concerned with only one aspect of the above, namely the question of the evolution of musical tuning systems.

A vast majority of human music[1] is based on discrete sound events in time, or *tones*, each with a distinct height or *pitch*. Tone height is a continuous physical parameter (related to the oscillation frequency, in Hertz), but it typically uses only a discrete and finite set of pitch values (e.g. one of $\{C, C\sharp, D, D\sharp, ...\}$ in the western tradition). The choice of the number and spacing of such frequency values can be defined as a *tuning system*. Tuning systems are generally found periodic, with tone frequencies having a ratio of 2:1 referred to as the same pitch class, one *octave* apart.

Most of music in the European/North American world relies on so-called "12-tone equal-tempered" tuning system, a set of 12 pitch values per octave, which are equally spaced logarithmically. This for instance is the standard system for tuning a piano, where the smallest interval between 2 keys (e.g. between C and $C\sharp$) is called a *semi-tone* (an interval also measured as 100 *cents*). However, the predominance in our culture of 12-tone equal temperament should not occlude the great variety of other possible tuning systems found in the many examples of indigenous music over the world and over the history of human music making. Equal (or "fairly equal") tempered systems using tone numbers other than 12 are found e.g. in central Javanese gamelan (*sléndro* tuning with five notes to an octave [Lindsay, 1979]) or medieval 19-tone ("one-third-comma meantone") and 31-tone ("one-fourth-comma-meantone") temperaments [Alves, 1989]. Non-equally spaced tuning systems are also common, e.g. in the 22-tone indian musical system where octave is divided into 22 *shrutis*, built with two types of semitones (90 and 114 cents) [Kolinski, 1961]. In other words, while any tone frequency between 261.626 Hz and 277.183 Hz would only sound to a western listener as either a high-tuned C or a low $C\sharp$, other cultures may very well have a word for it, and make it a distinct and well-defined pitch class.

The mechanisms by which various human cultures have come to internally decide on conventions to discretize a continuous space such as tone frequency are intriguing and reminiscent of other features of human cultural and cognitive evolution, such as:

- basic color terms (where a space of *visual* frequencies is divided into classes for "blue","green", "red", etc. [Berlin and Kay, 1969])
- vowel sounds in human phonologic systems (from a space of continuous motor commands [De Boer, 2000])
- and even words (lexicon) in human language (which may be seen as indexing an arguably continuous and infinite symbolic space [Steels, 2003])

The construction of musical tuning systems is by no means arbitrary, and notably obeys to strong physiological constraints. Various pitch combinations will sound more or less "natural" (or "pleasant", "restful") when used in combination, and this sensation of *consonance* has been found quite universal among human cultures, notably for its maximum manifestation in octave (2:1) and perfect fifth (3:2) intervals [Burns and Ward, 1978]. A plausible model of tone consonance was derived in [Plomp, 1965], where the dissonance of intervals of pure

[1] although there are notable exceptions, with continuous drones and non-discrete pitch classes, e.g. in Indian Carnatic ragams

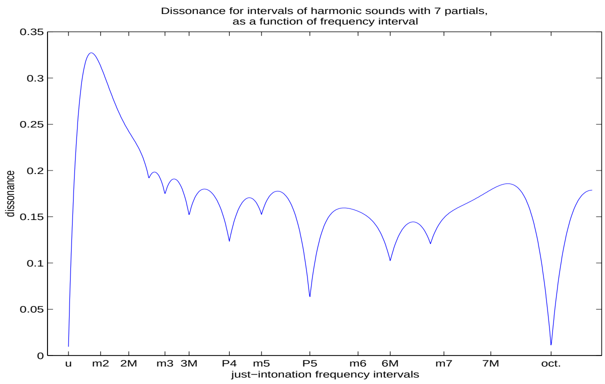

Fig. 1. Dissonance curve for intervals of harmonic tones with 7 partials, as a function of frequency interval. The curve is generated using the mathematical parameterization of [Sethares, 1993]. We observe that minima of the dissonance curve correspond to many of the degrees of the "just intonation" tuning system.

sine waves was found to be a simple exponentially increasing-then-decreasing function of their frequency ratio

$$d(f_1, f_2) = \exp^{-a|f_2 - f_1|} - \exp^{-b|f_2 - f_1|} \tag{1}$$

When considering intervals of complex harmonic sounds, made of a series of n sinusoidal partials, dissonance can be calculated as the sum of the dissonances of all pairs of partials:

$$d = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} d(f_i, f_2) \tag{2}$$

Figure 1 shows the dissonance curve for intervals of harmonic tones with 7 partials (with geometrically decreasing amplitude by factor 0.8), as a function of distance between the tones' fundamental frequencies. It appears that many of the local minima of the curve correspond to pitch classes of the "just intonation" tuning system (a system based on harmonic ratios, which are approximated by the 12-tone equal temperament). Notably, the most consonant intervals are the unison (1:1), then the octave (2:1), then the perfect fifth (3:2), perfect fourth (4:3), etc.

Such physiological evidence provide apparently straightforward explanations to the evolution and creation of tuning systems, which have been viewed notably as

- Solutions of an optimisation problem: The degrees of musical scales should correspond to local minima of the dissonance curve [Sethares, 1993].
- Outcomes of an algorithmic procedure: Start with a tone, and iteratively add the octave and fifth interval above and below any tone that was previously determined ("Cycle of fifths" construction); Start with octave, fifth

and fourth, and add the pitch classes corresponding to their 3 first harmonics [Fink, 2003].

Such explanations indeed promisingly account for many of the properties of actual musical tuning systems. Notably, as dissonance curves depend on instrument timbres (i.e. amplitude and frequency ratio - not necessarily harmonic - of partials), scales are naturally found to adapt to different instruments, thus explaining their cultural diversity. Also, the fact that typical harmonic series cannot be heard easily after the second or third partial can be used to explain the predominance of 7-tone and 9-tone scales in ancient music.

However, it is unclear how such explanations can be implemented in a realistic evolutionary process, where no central entity can be expected to do the optimization. Although recent musical constructions were indeed derived through pure mathematical thinking (see e.g. [Honingh and Bod, 2005]), this seems an implausible process to explain many of the world's indigenous tuning systems, which rather emerged through local interactions in a musically inclined population. I have a guitar, and you have too: let's just vaguely tune in, and play.

In this context, even apparently trivial observations such as the sheer existence of shared pitch categories do require a second look. Under which interactional processes can a population agree on a common set of pitches, and not collapse in extreme scenarios where only one pitch is used (the maximally consonant unison interval), or each individual uses its own distinct set. Moreover, there are many properties of human tuning systems which cannot be explained by traditional algorithmic approaches:

- There is great diversity in the number of pitch classes, from the 4 notes of the world's arguably oldest instrument [Fink, 2003], 5-tone Javanese *slendro*, up to 22-tone Indian system and 25-tone classical Arab tuning. Dissonance curves as shown in Figure 1 only account for limited number of degrees, mainly a function of the number of partials (which does not scales to 20 and more).
- Even within the western 12-tone system, many intervals such as minor second (C-$C\sharp$) or augmented fifth/minor sixth (C-$G\sharp$) are not easily explained by dissonance-based algorithms.
- Tones forming very dissonant intervals (sometimes less than a semi-tone) are often found to co-exist in a given tuning system (and do not fusion into one another), etc.

This paper proposes to examine a few of these aspects with the help of computer simulations in which a population of adaptive agents evolve tuning systems through only local interactions. Inspired by the ideas of artificial language evolution [Steels, 2003], we try to evolve artificial tuning systems with properties similar to the natural ones, through musical equivalents of language games.

2 Methods

We simulate the evolution of artificial tuning systems using the methodology of language evolution research in the lines of [Steels, 2003]: A population of agents, each equipped with a simple cognitive architecture (here a model of tone consonance perception, and a tuning system, i.e. a set of tones), engage in one-to-one musical games (i.e. "try to play music together"). Game after game, the agents adapt their internal tuning system to increase the success of further games. We study how shared tone categories can emerge among the total population, and how their properties depend on the distributed activities of the agents. This mechanism, when successful, strongly opposes the previous algorithmic justifications for the emergence of human musical tuning systems: in our framework, there is no central agency that controls how agents are supposed to act.

2.1 Agent architecture

The agents A_i are equipped with simple cognitive structures, each of which is inaccessible to other agents:

- A tuning system $ts_i = \{t_i^1, \ldots, t_i^{N_i}\}$, i.e. a set of N_i tones. Tones are described by a fundamental frequency $f(t_i^k)$, and a timbre $\mathcal{T}(t_i^k)$ which is given as part of the agents environment (see Section 2.3). The fundamental frequencies of each agent's tones are initialized at random. In all the experiments reported here, all agents have tuning systems of the same size, i.e. $N_i = N \forall i$
- A tuning procedure, i.e. the ability to change the fundamental frequency of part or all of the tones in ts_i
- A restricted model of human hearing simulating the perception of tonal consonance. The model uses the mathematical parameterization of the experimental Plomb-Levelt curves [Plomp, 1965] given in [Sethares, 1993]: it computes a value of dissonance $d(t_1, t_2) = d(f(t_1), \mathcal{T}(t_1); f(t_2), \mathcal{T}(t_2))$ for two tones t_1 and t_2 given by their fundamental frequencies $f(t_1), f(t_2)$ and their timbres $\mathcal{T}(t_1), \mathcal{T}(t_2)$.

Agents are able to produce sounds for interaction with other agents, by simply playing a tone or a series of tones from their internal tuning system. In the current experiment, the agents are capable of monophony only, i.e. tones are only played once at a time.

2.2 Interaction protocol

In the experiments reported here, we investigate two different musical games. Both games, which can be regarded as imitation games as in [De Boer, 2000], are designed to allow the agents to align their internal tuning systems with that of other agents, in a way that enables realistic musical interaction. At each new game (or iteration), two agents, a *player* P and a *tuner* T are drawn at random from the population. Their interaction depends on the type of game:

- **Game 1: Single Tone Shift**
  This game corresponds to the very simple musical situation where two agents play a single tone at unison, and one of them (the tuner) try to adapt its

tone to generate a consonant interval with the other agent's. The player P and the tuner T draw one single tone at random from their internal tuning systems: $t_P \in ts_P$ and $t_T \in ts_T$. The tuner agent T shifts the fundamental frequency of its tone t_T so as to minimize dissonance with t_P, i.e.

1. finds the frequency value f^* corresponding to the nearest dissonance minimum around $f(t_T)$, i.e.

$$f^* = \arg\min_{f_m} |f_m - f(t_T)|; f_m \in \mathcal{M} \tag{3}$$

   where $\mathcal{M}$ is the set of all frequencies f corresponding to local minima of $d(f, \mathcal{T}(t_T); f(t_P), \mathcal{T}(t_P))$.
2. deletes t_T from tuning system ts_T
3. and inserts new tone t_T^* in ts_T having $f(t_T^*) = f^*$ and $\mathcal{T}(t_T^*) = \mathcal{T}(t_T)$

– **Game 2: Drone Shift**
In this game, a single note is continuously sounded by agent T, in harmony with all the tones in agent P's tuning system. Such accompaniment is called a *drone* note, and is systematic in many forms of traditional music - often parts of musical instruments are designed to produce such drone notes without requiring the attention of the player, e.g. sympathetic strings in indian *sitar* or drone pipes in Irish *Uilleann pipes* [Sadie, 1984]. More precisely, the tuner T draw one single tone $t_T \in ts_T$ at random from its internal tuning systems. The player P plays all the tones in its tuning system. Agent T shifts the fundamental frequency of its tone t_T so as to minimize the sum of the dissonance with all tones in ts_P, i.e.

1. finds the frequency value f^* corresponding to the nearest dissonance minimum around $f(t_T)$, i.e.

$$f^* = \arg\min_{f_m} |f_m - f(t_T)|; f_m \in \mathcal{M} \tag{4}$$

   where $\mathcal{M}$ is the set of all frequencies f corresponding to local minima of $\sum_{t_i \in ts_P} d(f, \mathcal{T}(t_T); f(t_i), \mathcal{T}(t_i))$.
2. deletes t_T from tuning system ts_T
3. and inserts new tone t_T^* in ts_T having $f(t_T^*) = f^*$ and $\mathcal{T}(t_T^*) = \mathcal{T}(t_T)$

2.3 Environment

The environmental constraints imposed on the agents mainly concern how tones are implemented. Here, we define a tone t as a series of N_t sinusoidal oscillators, or partials, each with a frequency in Hz and an amplitude, i.e. $t = \{f_i, a_i\}_{i=1:N_t}$. The frequency of a tone's first partial is referred to as its fundamental frequency $f(t) = f_1$. The parameters of a tone's partials for $i \geq 2$ are determined by its timbre $\mathcal{T}(t)$, which can be understood as representing a given musical instrument. In this work, we consider two alternative timbres:

– **Timbre 1: Harmonic timbre**
The represents the ideal harmonic sounding body (i.e. a string), where the

partials frequencies are integer multiples of the fundamental frequency, with a geometrically decreasing amplitude:

$$f_i = i f_1 \quad (5)$$
$$a_i = \alpha^i a_1 \quad (6)$$
$$(7)$$

where we take $\alpha = 0.8$ and $a_1 = 1$.

– **Timbre 2: Compressed timbre**
  This represents an inharmonic timbre with partial frequencies not integer multiples on a fundamental, but rather more narrowly spaced according to a geometrical law (we take here the formulation of [Sethares, 1993]). Such timbres are typical of certain bells, such as the 2500-year-old Chinese Zheng bells [Shen, 1987].

$$f_i = A^{\log_2(i)} f_1 \quad (8)$$
$$a_i = \alpha^i a_1 \quad (9)$$
$$(10)$$

where we take $A = 1.90$, $\alpha = 0.8$ and $a_1 = 1$.

All tones are generated using $N_t = 7$ partials. Fundamental frequencies f_i for all agents are constrained to be in the range [262 Hz,523 Hz] corresponding to the octave between $C4$ and $C5$ in the 12-tone equal temperament system (this includes notorious diapason $A4 = 440$ Hz), both at initialization, and during updates in musical games. All frequencies are quantized to the nearest 10 cents (i.e. a tenth of a semi-tone), which corresponds to the frequency resolution achieved by professional piano tuners.

2.4 Measures

In the experiments reported here, we mainly focus on two properties of the evolved artificial tuning systems:

– Number of distinct tones: This measures the number of distinct fundamental frequencies found in the total population, at a given iteration n. Upon (random) initialization, the number can be as high as $\sum_{i=1}^{N_a} N_i$, where N_a is the size of the population and N_i the size of the tuning system of agent A_i. Low counts of distinct tones after n iterations, if observed, show that a limited, shared set of tones has been evolved through local interactions; actual numbers can be compared to natural tuning systems.
– Distribution of intervals: This measures the histogram of the intervals in the tuning systems of each agent in a population, at a given iteration n. Intervals in an agent's tuning system ts are measured as differences of fundamental frequency of each tone to the lowest tone is ts. A 5-tone tuning system

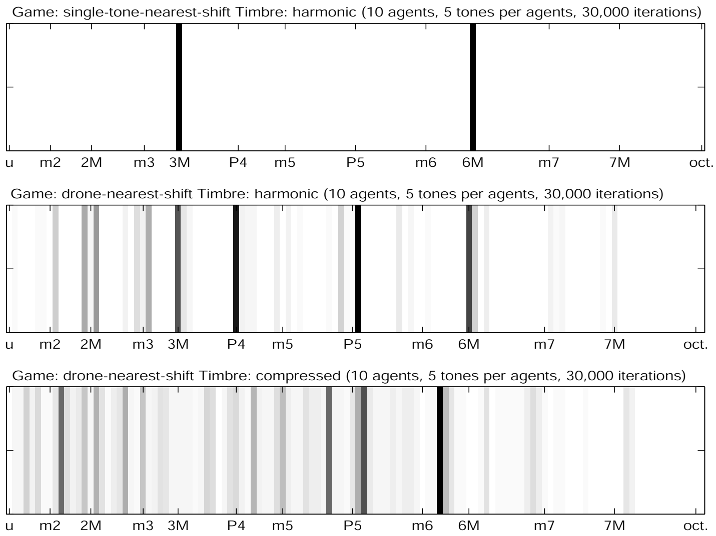

Fig. 2. Interval histograms in tuning systems evolved after 30,000 iterations for different simulation settings. Single-tone-shift games do not reproduce the diversity and complexity of natural systems as well as drone-shift games.

$\{t_1, \ldots, t_5\}$ (ordered by increasing fundamental frequency $f(t_i)$) yields 4 intervals $\delta_i = f(t_{i+1}) - f(t_1) \forall i \in [1, 4]$. The global histogram shows the proportion of such intervals over the whole population. Upon initialization, the interval histogram over the whole population shows a random distribution. Peaky and sparse histograms, if observed after n iterations, show that tuning systems of individual agents have evolved to share the same structure.

3 Results

3.1 The conditions for self-organization are non-trivial

Simulations involving simple games don't converge to anything resembling natural tuning systems. Figure 2 compares the histograms of tone intervals in the global population of agents for different simulation settings, after 30,000 iterations, averaged over 50 runs. We observe that interactions based on the single-tone-shift game systematically produce unrealistically simple tuning systems with 3 tones. The tones are separated by a major third (3M) and a perfect fourth (P4), i.e. span a major sixth interval (6M). This forms a minimalistic-yet-usable musical system, but doesn't reproduce the diversity and complexity of natural systems. Moreover, the convergence under such an interaction protocol seems to be independent from the number N of initial tones in each individual agent. Figure 3-left shows the convergence profile (in number of distinct tones against number of iterations) for $N \in \{3, 5, 8, 10, 15, 20\}$. All settings (averaged over 10 runs) converge to the same set of 3 tones, at a speed correlated with N.

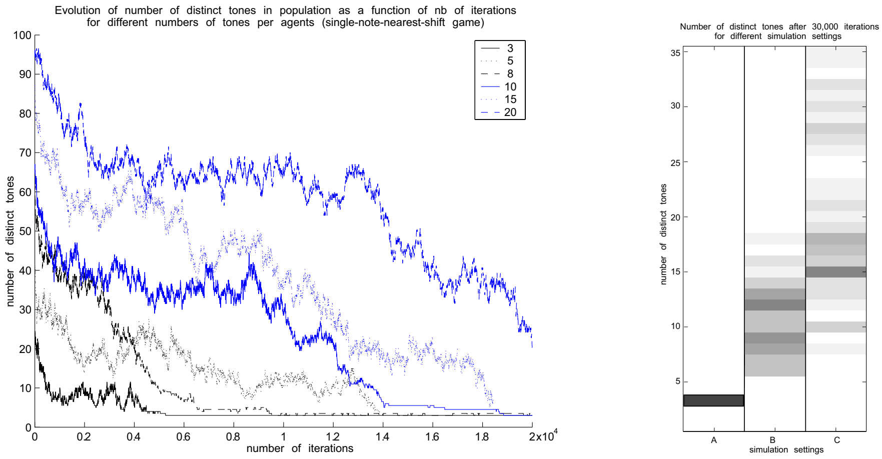

Fig. 3. Left: Convergence profile (in number of distinct tones against number of iterations), for single-tone-shift game and varying number of tones per agents $N \in \{3, 5, 8, 10, 15, 20\}$. All settings converge to (the same) 3 tones, with speed correlated to N. Results averaged over 10 runs. Right: Comparison of the histograms of number of distinct tones after 30,000 iterations for different simulation settings. A: single-tone-shift game, 10 tones per agents, 10 agents, harmonic timbre. B: drone-shift game, 10 tones per agents, 10 agents, harmonic timbre. C: drone-shift game, 10 tones per agents, 10 agents, compressed timbre. Results averaged over 50 runs.

3.2 Specific musical constraints are needed

Other more complex musical situations, as implemented with e.g. drone-shift game, do favor the emergence of more realistic tuning systems, notably:

- Emergence of shared pitch class in a given population, but not the same ones for different runs.
- Increased diversity of system size: Figure 3-right shows that, at constant simulation settings, tone numbers form a bell-like distribution with notable variance.
- Realistic sizes: between 5 and 15 tones for harmonic timbres (Figure 3-right:B), which is in accordance with many natural systems found the world over.
- Interval complexity: emerging systems include strong components at very consonant intervals (P5, P4, 3M, 6M), in accordance with algorithmic constructions based on the Plomb-Levelt consonance curves. However, typical systems also incorporate more dissonant intervals found in 12-tone equal temperament, but not well explained by previous hypotheses, notably minor second (m2) and major seventh (7M). Moreover, the observed interval distributions also suggest the possibility of emergence of rare non-harmonic intervals, such as low minor thirds (m3) and in between P5 and m6.

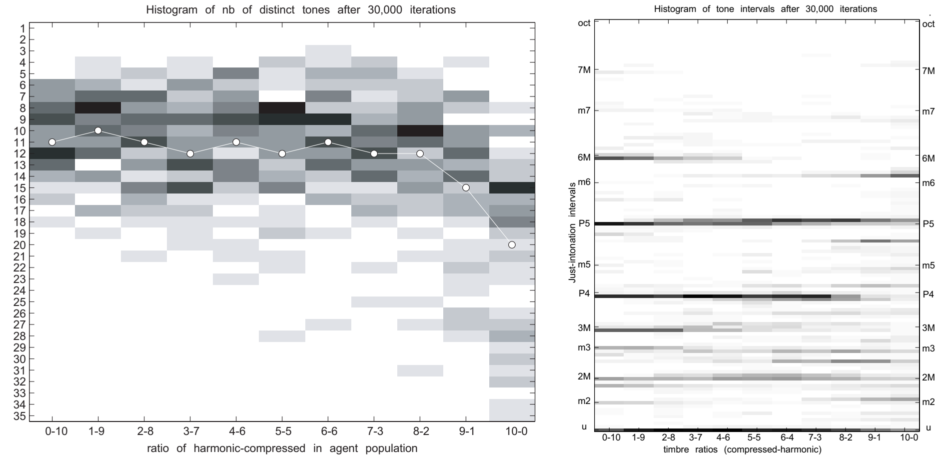

Fig. 4. Left: Histogram of number of distinct tones in tuning systems evolved after 50,000 iterations with heterogeneous populations of agents of varying timbre ratios (foremost left: purely harmonic, foremost right: purely compressed). White circles identify distribution means. (settings: population of 10, 10 tones per agents, drone-shift games, results averaged over 50 runs). Right: Interval histograms of tuning systems evolved after 50,000 iterations with heterogeneous populations of agents of varying timbre ratios, same settings.

On the whole, these observations suggest that certain properties of musical activities, such as the primitive harmonic accompaniment translated with one-to-many drone tuning, are needed to explain the emergence of natural tuning systems.

3.3 Instruments are a clear evolutionary factor

It also appears (in accordance with [Sethares, 1993]) that the choice of a given timbre (or instrument) influences the emergence of certain intervals as well as the size of tuning systems. Figure 3-right:B&C shows that, all other conditions held equal, compressed timbres encourage the emergence of tuning systems with more distinct tones than harmonic timbres, and also generate greater size diversity over several runs. The structure of systems evolved with compressed timbre is also quite different from the harmonic ones (Figure 2): strong perfect fifth (P5), no major third (3M) nor perfect fourth (P4), and clear non-harmonic intervals (between m5 & P5, m6 & 6M, m2 & 2M).

3.4 Population heterogeneity is another factor for diversity

Finally, it appears that environmental factors leading to heterogeneity in the population of agents favor the emergence of even more complexity in tuning systems. We investigate populations of agents made of two groups of varying relative sizes, each with a different timbre (harmonic or compressed). We find that intermediate settings between between harmonic-only and compressed-only generate hybrid, interpolated properties in the evolved tuning systems, notably:

- Hybrid number of distinct tones: Figure 4-left shows the histograms of number of distinct tones for different timbre ratios, after 50,000 drone-shift iterations, averaged over 50 runs: ratios intermediate between $0-10$ (harmonic-only) and $10-0$ (compressed-only) generate intermediate numbers of tones (notably at high ratios), as well as intermediate variance values.
- Hybrid intervals: Figure 4-right shows the histogram of intervals in the tuning systems converged with different timbre ratios, after 30,000 iterations, averaged over 50 runs. It appears that timbre hybridation generates systems with mixtures of harmonic and inharmonic intervals:
  - inharmonic intervals (characteristics of compressed timbres) such as high-tuned m6, high P4, high m2 and low P5 appear as early as (4-6),(5-5),(7-3) and (8-2) resp.
  - harmonic intervals such as m2, 6M, 3M and P4 disappear when ratios increase over (2-8), (4-6), (5,5) and (8-2) resp.

Further, at intermediate ratios, hybridation also favors the emergence of alien intervals which belong neither to the purely harmonic nor compressed systems, and stably co-exist with neighboring intervals, e.g. high-tuned 2M from (2-8) to (8-2), and high-tuned 3M from (2-8) to (6-4).

4 Conclusion

The results of these simulations show that coherent musical tuning systems can emerge as the results of local musical interactions between the members of a population. The artificial systems that emerge show properties similar to the ones found in natural tuning systems, such as number of distinct tones and intervalic structure. Moreover, such self-organizing systems are evolved with a diversity and complexity which is not easily explained by previous theoretical justifications based on algorithmic procedures. For instance, under certain conditions, systems are found that include 20 tones or more, with inharmonic ratios that are alien to the consonance profiles of the timbres used in the population. This is in remarkable accordance with the structure of much-debated tunings such as the ancient Indian 22 *shrutis* system. On the whole, this makes self-organization a promising hypothesis to explain some of the properties of natural tuning systems.

However, our simulations also show that the implementation of such mechanisms into a realistic interactional framework requires non-trivial environmental and cultural constraints. Simple musical games, in which tones are adapted on a one-to-one basis, are insufficient to generate tuning systems with diversity and complexity resembling that of real-world systems. Advanced musical concepts, such as harmony (e.g. with drone tones), and non-homogeneity in the population (e.g. making music with different types of instruments simultaneously) seem to be necessary ingredients. This suggests directions for future research, e.g. incorporating more advanced musical concepts such as melodic interestingness (optimally consonant phrases are boring), or the need for key modulation (which is one of the reasons for preferring the imperfect equal-temperament over just-intonation systems).

References

[Alves, 1989] Alves, B. (1989). The just intonation system of Nicola Vicentino. *1/1: Journal of the Just Intonation Network*, 5(2):8–13.

[Berlin and Kay, 1969] Berlin, B. and Kay, P. (1969). *Basic Color Terms.* Berkeley: University of California Press.

[Blacking, 1973] Blacking, J. (1973). *How Musical is a Man?* Seattle: University of Washington Press.

[Burns and Ward, 1978] Burns, E. and Ward, W. D. (1978). Categorical perception: Phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 63:456–468.

[Christiansen and Kirby, 2003] Christiansen, M. H. and Kirby, S. (2003). *Language Evolution.* Oxford: Oxford University Press.

[Cross, 2007] Cross, I. (2007). *Music and cognitive evolution*, pages 649–667. Oxford: Oxford University Press. In Robin Dunbar & Louise Barrett (Eds.) Handbook of Evolutionary Psychology.

[De Boer, 2000] De Boer, B. (2000). Self organization in vowel systems. *Journal of Phonetics*, 28(4):441–465.

[Fink, 2003] Fink, R. (2003). *The Neanderthal flute and origin of the scale: fang or flint? A response*, pages 83–87. Rahden: Verlag Marie Leidorf. In Ellen Hickmann, Anne Draffkorn Kilmer and Ricardo Eichmann (Eds.) Studies in Music Archaeology III.

[Honingh and Bod, 2005] Honingh, A. and Bod, R. (2005). Convexity and the well-formedness of musical objects. *Journal of New Music Research*, 34(3):293–303.

[Kolinski, 1961] Kolinski, M. (1961). The origin of the indian 22-tone system. *Studies in Ethnomusicology*, 1:3–19.

[Lindsay, 1979] Lindsay, J. (1979). *Javanese Gamelan.* Oxford: Oxford University Press.

[Mithen, 2006] Mithen, S. (2006). *The Singing Neanderthals: The Origins of Music, Language, Mind and Body.* Cambridge, MA: Harvard University Press.

[Plomp, 1965] Plomp, R. and Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 38:548–560.

[Sadie, 1984] Sadie, S. e. (1984). *The New Grove Dictionary of Musical Instruments.* New York: Grove's Dictionaries of Music.

[Sethares, 1993] Sethares, W. A. (1993). Local consonance and the relationship between timbre and scale. *Journal of the Acoustical Society of America*, 94(3):1218–1228.

[Shen, 1987] Shen, S. (1987). Acoustics of ancient chinese bells. *Scientific American*, 256:94.

[Steels, 2003] Steels, L. (2003). The evolution of communication systems by adaptive agents. In *Adaptive Agents and Multi-Agent Systems: Adaptation and Multi-Agent Learning. LNAI 2636*, pages 125–140.