J Hear Sci, 2019; 9(3): 39–49
DOI: 10.17430/1003501


Contributions:
A Study design/planning
B Data collection/entry
C Data analysis/statistics
D Data interpretation
E Preparation of manuscript
F Literature analysis/search
G Funds collection

# A NATURAL THEORY OF MUSIC BASED ON MICROMECHANICAL RESONANCES BETWEEN COCHLEAR SENSING CELLS

**Andrew Bell**[ABCDEF]

John Curtin School of Medical Research, The Australian National University, Canberra, Australia

**Corresponding author:** Andrew Bell, Eccles Institute of Neuroscience, John Curtin School of Medical Research, 131 Garran Road, The Australian National University, Canberra, ACT 2601, Australia; email: andrew.bell@anu.edu.au

## Abstract

The origins of music remain obscure. Here it is pointed out that the outer hair cells in the cochlea lie approximately in a regular, hexagonal array, so it appears significant that important musical ratios – in particular the semitone, octave, perfect fifth, and major third – appear in the relative distances between adjacent cells. The speculation made here is that if the intercell distances are inherently tuned in this way, then incoming sound could initiate reverberating activity between the cells, and a musical ratio could be signaled by simultaneous standing waves in one cell–cell distance and in another which adjoins it. Essentially, the spacings between the cells might correspond to the lengths of miniature, musically tuned cavities.

This speculative model of cell–cell interaction can explain recent observations that the human cochlea spontaneously emits low-level sound at intervals close to a semitone, and that the hearing thresholds of some subjects exhibit a whole "keyboard" of semitone-like intervals. These recent findings are striking, and appear more than coincidence. They prompt the key question, why? A possible answer may lie, it is suggested, with the distinct 2-dimensional geometry of the outer hair cells in the plane of the basilar membrane, which commonly exhibits a 19° alignment. This angle corresponds to a relative distance of 1.06, which is close to a semitone. It is pointed out that the same geometry which generates a natural semitone also produces other musically significant ratios. Perhaps, then, music might be an innate property of the human auditory system – and hence that there might be a natural basis for preferred musical intervals.

Natural theories have often been criticised, with some saying that music is a learned faculty and depends only on culture. However, evidence has accumulated that there do seem to be musical universals, and therefore that music might indeed have a natural basis, most commonly thought to be via some neural processing in the brain. The explanatory model put forward here as the basis for further exploration suggests that musical analysis might actually begin in the periphery – in the cochlea itself.

**Key words:** semitone • cochlea • spontaneous otoacoustic emission • hearing threshold • outer hair cell • neurophysiology

## UNA TEORÍA NATURAL DE MÚSICA BASADA EN LAS RESONANCIAS MICROMECÁNICAS ENTRE LAS CÉLULAS SENSORIALES DE LA CÓCLEA

### Resumen

El origen de la música sigue sin estar claro. En este artículo inicialmente se quiere remarcar que las células auditivas externas de la cóclea están en una disposición hexagonal casi regular, por lo que parece importante que los intervalos musicales básicos - en particular el semitono, la octava, la quinta justa y la tercera mayor - aparezcan en las distancias correspondientes entre las células adyacentes. Esto, nos lleva a lanzar la hipótesis que, si las distancias entre las células se ajustan inherentemente de esta manera, entonces el sonido entrante puede iniciar una reverberación entre las células y, un intervalo de música dado puede señalizarse mediante ondas estacionarias simultáneas entre una célula y la otra y en células adyacentes. Básicamente, el espacio entre las células puede corresponder a las longitudes de los resonadores en miniatura musicalmente sincronizados.

Este modelo especulativo de la interacción célula-célula puede explicar las observaciones recientes de que la cóclea humana emite espontáneamente un sonido de bajo nivel cerca de los semitonos y que los umbrales de audición en algunas personas muestran todo el "teclado" de los intervalos similares a semitonos. Estos descubrimientos recientes son sorprendentes y parecen más que accidentales. Surge la pregunta: ¿por qué? Quizás la respuesta a esta pregunta esté relacionada con la clara geometría bidimensional de las células ciliadas externas en el plano de la membrana basal, que generalmente muestra una disposición de 19 grados. Este ángulo corresponde a una distancia relativa de 1.06, que está cerca de un semitono. Cabe señalar que la misma geometría que crea un semitono natural también crea otros intervalos importantes en la música. Quizás, por lo tanto, la música es una característica innata del sistema auditivo humano y, por lo tanto, puede haber una base natural para los intervalos musicales preferidos.

Las teorías naturales a menudo han sido criticadas y algunos dicen que la música es una habilidad aprendida y que depende solo de la cultura. Sin embargo, se ha acumulado bastante evidencia de que probablemente haya universales musicales y, por lo tanto, que la música puede tener un fondo innato que a menudo se asocia con el procesamiento natural en el cerebro. El modelo presentado en este artículo, que es la base para futuras investigaciones, sugiere que el análisis de la música en realidad puede comenzar en la parte periférica, en la cóclea misma.

**Palabras clave:** semitono • cóclea • emisión otoacústica espontánea • umbral auditivo • células ciliadas externas • neurofisiología

# ЕСТЕСТВЕННАЯ ТЕОРИЯ МУЗЫКИ, ОСНОВАННАЯ НА МИКРОМЕХАНИЧЕСКИХ РЕЗОНАНСАХ МЕЖДУ СЕНСОРНЫМИ КЛЕТКАМИ В УЛИТКЕ


## Аннотация

Происхождение музыки остается не до конца понятным. В этом исследовании отмечается, что внешние слуховые клетки в улитке находятся в почти правильном, шестиугольном расположении, поэтому представляется важным, что основные музыкальные интервалы - полутона, октава, чистая квинта и большая терция - появляются на соответствующих им расстояниях между соседними клетками. По мнению авторов статьи, если расстояние между клетками по своей природе «настроено» именно таким образом, то поступающий звук может вызывать резонанс между клетками, и данный музыкальный интервал может сигнализироваться посредством одновременных волн, стоящих между одной и другой клетками и в соседних клетках. В итоге, расстояние между клетками может соответствовать длине миниатюрных, музыкально настроенных резонаторов.

Эта спекулятивная модель взаимодействия клетка – клетка может объяснить недавние наблюдения о том, что улитка человека самопроизвольно издает низкочастотный звук, близкий к полутонам, и что пороги слышимости у некоторых людей показывают всю «клавиатуру» интервалов похожих на полутона. Данное открытие удивительно и кажется более чем случайным. Возникает вопрос: почему? Возможно, ответ на этот вопрос связан с четкой двумерной геометрией наружных волосковых клеток в плоскости базилярной мембраны, которая обычно расположена под углом 19 градусов. Этот угол соответствует относительному расстоянию 1,06, которое близко к полутону. Следует отметить, что та же самая геометрия, которая создает естественный полутон, также создает другие важные интервалы в музыке. Возможно, поэтому музыка является врожденной особенностью слуховой системы человека и, следовательно, может существовать естественная основа для предпочтительных музыкальных интервалов.

Естественные теории часто подвергаются критике, и существуют утверждения, что музыка - это полученный навык и зависит только от культуры. Однако, есть доказательства того, что, вероятно, существуют музыкальные универсалии, и, таким образом, музыка действительно может иметь врожденную основу, которая чаще всего связана с естественной работой мозга. Модель, представленная в этой статье, которая является основой для дальнейших исследований, предполагает, что музыкальный анализ может фактически начинаться в периферической части, то есть в самой улитке.


**Ключевые слова:** полутон • улитка, спонтанная отоакустическая эмиссия • слуховой порог • наружные волосковые клетки • нейрофизиология.

# NATURALNA TEORIA MUZYKI OPARTA NA REZONANSACH MIKROMECHANICZNYCH MIĘDZY KOMÓRKAMI CZUCIOWYMI W ŚLIMAKU


## Streszczenie

Pochodzenie muzyki pozostaje niejasne. W niniejszym opracowaniu zwraca się uwagę, że zewnętrzne komórki słuchowe w ślimaku leżą w prawie regularnym, sześciokątnym układzie, więc wydaje się istotne, że podstawowe muzyczne interwały – w szczególności półton, oktawa, kwinta czysta i tercja wielka – pojawiają się w odpowiadających im odległościach między sąsiadującymi komórkami. W artykule przyjęto założenie, że jeśli odległości między komórkami są z natury dostrojone w ten sposób, wówczas przychodzący dźwięk może wzbudzić rezonans między komórkami, a dany interwał muzyczny może być sygnalizowany przez jednoczesne fale stojące między jedną a druga komórką i w komórkach do nich przyległych. Zasadniczo odstępy między komórkami mogą odpowiadać długościom miniaturowych, muzycznie zestrojonych rezonatorów.

Ten spekulatywny model interakcji komórka–komórka może wyjaśnić ostatnie obserwacje, świadczące o tym, że ludzki ślimak spontanicznie emituje dźwięk o niskim poziomie bliski półtonowi i że progi słyszenia u niektórych osób pokazują całą „klawiaturę" interwałów podobnych do półtonów. Te ostatnie odkrycia są zadziwiające i wydają się czymś więcej niż przypadkiem. Pojawia się pytanie: Dlaczego? Być może odpowiedź na to pytanie wiąże się z wyraźną dwuwymiarową geometrią zewnętrznych komórek włosowatych w płaszczyźnie błony podstawnej, która zwykle wykazuje ułożenie 19-stopniowe. Kąt ten odpowiada względnej odległości 1,06, która jest bliska półtonowi. Należy zwrócić uwagę, że ta sama geometria, która tworzy naturalny półton, tworzy także inne ważne w muzyce interwały. Być może zatem muzyka jest wrodzoną cechą układu słuchowego człowieka – a zatem może istnieć naturalna podstawa do preferowanych interwałów muzycznych.

Teorie naturalne były często krytykowane, a niektórzy twierdzą, że muzyka jest wyuczoną umiejętnością i jest zależna wyłącznie od kultury. Zgromadzono jednak dowody, że prawdopodobnie istnieją muzyczne uniwersalia, a zatem muzyka może rzeczywiście mieć wrodzone podłoże, które najczęściej wiązane jest z naturalnym przetwarzaniem zachodzącym w mózgu. Przedstawiony w niniejszym artykule model, będący podstawą do dalszych badań, sugeruje, że analiza muzyczna może faktycznie rozpoczynać się w części obwodowej – w samym ślimaku.


**Słowa kluczowe**: półton • ślimak • spontaniczna emisja otoakustyczna • próg słuchu • zewnętrzne komórki włosowate • neurofizjologia

## Introduction

Despite the vastness of the literature on the theoretical underpinnings of music, its origins and evolutionary function remain a profound puzzle [1–4]. Why do humans spend so much time listening to or performing music? There is no solid scientific explanation, but an accessible overview of the core issues can be found in Ball [5] (see also [6–8]).

In this paper, the unconventional proposal is made that music could originate within the ear itself. Existing mechanical models of the cochlea – predominantly involving traveling waves – provide only indirect hints of how musical ratios might arise (e.g., the spiral staircase idea of Shera [9] and the feedforward model of Motallebzah and colleagues [10]). However, a novel resonance model developed by Bell [11] emphasises the micromechanical interactions between rows of outer hair cells, and this surface acoustic wave (SAW) model has the potential to generate multiple resonant frequencies at a single spot on the basilar membrane. The SAW model suggests that minute mechanical resonances may be sustained by the active outer hair cells within the cochlea, and that the sympathetic resonances generated from sound entering the organ allow the hair cell system to immediately detect frequency ratios. The model is reminiscent of the miniature piano strings in the ear imagined by Helmholtz [12], but here the strings are actually the spaces between individual sensing cells, in which case the resonant elements might be more aptly described as hollow cavities.

The implication is that if the distances between the cells are set at musically significant ratios, then so too will be the resonant frequencies. From published micrographs of the outer hair cell pattern and the limited amount of work done on analysing the intercell distances [11], there is reason to think such micromechanical resonances may occur [13], and the present paper integrates this idea with recent audiological findings [14,15].

Outer hair cells in the mammalian cochlea are arranged in a hexagonal lattice (Fig. 1), a configuration which means that each cell has a small number of fixed distances (and therefore ratios) to its nearest neighbours. Another important factor is that the stereocilia of each sensing cell are arranged in two V-shaped arms, a shape which, it is hypothesised, permits the cell to make on-the-spot comparisons of activity – in particular, reverberation of waves – associated with each arm. In this way, it is suggested that musical ratios in a sound can be simply and immediately picked out without the need for complex neural processing.

The trigger for this paper comes from two recent findings. The first is the finding by Bell and Jedrzejczak [14] that the most common frequency ratio between sounds emitted by normal human ears – faint but precise, narrow-band tones called spontaneous otoacoustic emissions (SOAEs) – occurs at a ratio of 1.063 (±0.005), very close to an equal-tempered semitone of 1.059. (See [17] for a recent review of otoacoustic emissions and their relationship to cochlear mechanics). The semitone-like finding is complemented by another by Dewey and colleagues [18] who found regular ripples in hearing thresholds, so that, in one subject at least, one can discern a virtual "keyboard" of semitone-like ratios extending over nearly two octaves. In general, the periodic minima in the subject's hearing thresholds match the frequencies of their SOAEs, suggesting that both phenomena have the same underlying origin. These core findings are revisited in more detail below, noting that the musical significance of the periodicity appears to have so far escaped notice. The message derived from both these recent papers is that the findings are too regular to be coincidental; instead it is concluded that the human cochlea appears to be intrinsically tuned to a basic musical interval that might reasonably be regarded as "a natural semitone".

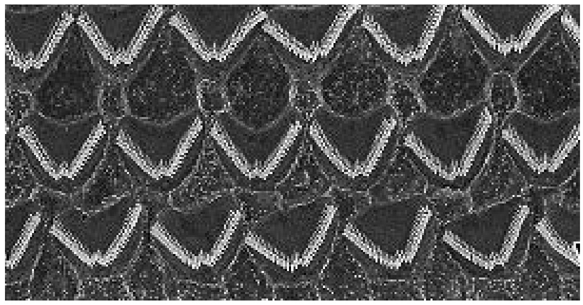


**Figure 1.** The array of outer hair cells of a mouse, showing the regular arrangement of V-shaped stereocilia (reproduced from [16], open access)

This paper elaborates on the musical connections just outlined, describes key findings of the active nature of outer hair cells (OHCs) and their associated otoacoustic emissions, and investigates the strongly musical ratios that arise from the geometry seen in Figure 1.

The core concept set out here is that the natural semitone is a fundamental "atom" of music, at least melodically if not harmonically (see [14] for further discussion of this point). In equating the observed ratio of 1.063 (±0.005) with a natural semitone, it is suggested that the cochlea contains an intrinsic interval that provides the ear with a ready musical template of 1/12th of an octave. In other words, the octave is divided into 12 equal steps because our ears are built that way. As will be described in more detail later, the human ear contains a continuous array of tuned elements, each derived from local feedback between one active cell and its neighbouring cell. Because the feedback resonance can occur over multiple pathways, it includes multiple frequencies – the shortest of which is a principal frequency (the characteristic frequency) and the other, lower frequency ones are derived from slightly longer paths. As will become clear, the closest frequency is 1.06 times lower because, due to geometry, the path length is effectively 1.06 times longer. This is the origin of the semitone, and it therefore becomes a matter of fact, not just theory, why the octave is built up of 12 semitones. Under certain circumstances, the underlying semitone structure becomes apparent (as with the observations of Dewey and colleagues), and we see something that closely resembles a piano keyboard.

Alternative *n*-tet tunings (9, 10, or 11 equal intervals in an octave) have sometimes been entertained by theoreticians (e.g. [19]), even though some people find the results audibly distressing (c.f. [20]). Adopting a naturalistic stance, it is no longer a puzzle why the smallest musically significant interval for humans is $2^{1/12}$ (1.059).

This paper extends the micromechanical resonance model to show that the same hexagonal arrangement of sensing cells which produces the semitone can also manifest other important musical ratios such as the octave, fifth, and major third. In this way, the foundations for a wider natural theory of music might be set in which all the ratios of the diatonic scale arise from reverberation of sound waves between adjacent cochlear hair cells. Indeed, it can be supposed that the entire cochlear tuning curve is built up of all these cell–cell interactions, so that the high-frequency tip (the characteristic frequency) originates from the strongest and shortest pathway, while all the other pathways are longer and weaker and contribute to the low-frequency tail. We conclude that music depends on geometry, a proposition going back to Pythagoras [5,6,8,21]. Once the existence of a natural semitone (and other ratios) in the ear is acknowledged, then the longstanding, but generally out of favour, idea of a natural basis for music takes on new life.

## Data and model

The starting point is the work of Bell & Jedrzejczak [14] who recorded faint, narrow-band sounds – synchronised spontaneous otoacoustic emissions – from 140 ears of 81 normal subjects and found that this objective measure of cochlear function had a peak at a frequency ratio of 1.063 ±0.005, a value remarkably close to an equal tempered semitone of 1.059. In that same publication, attention was drawn to the otoacoustic emissions of one particular subject described by Braun [22] in which there was a run of 10 emissions extending over 10.4 semitones, an average interval of 104 cents, and another run of 11 emissions spanning 11.4 semitones, again an average of 104 cents per interval. These findings were taken as evidence of a "keyboard" in the human cochlea, and the present paper brings forward further evidence for such a remarkable entity.

The work on the fine-structure of hearing thresholds investigated by Dewey and colleagues [18] contributes another important piece of empirical data. These researchers found, from repeated sets of painstaking measurements, that the sensitivity of the human ear to small steps in frequency – its threshold microstructure – has regular peaks and dips. Although the authors did not draw attention to the inherent size and regularity of the frequency steps, it is evident from the data shown in Figure 2 that there is a pattern of repeating ripples stretching across 1.79 octaves. Each ripple is about 89 cents wide, and the subjective threshold minima correspond with objective peaks in the SOAE spectrum, in line with the Braun [22] findings.

Somewhat curiously, the pattern in Fig. 2 is not strictly regular: there are 19 strong threshold minima which are more or less evenly placed (dashed lines in Fig. 2) but in 6 cases the expected minima are missing and interpolations (dotted lines) have been added instead. Four of those 6 instances fall close to a local threshold maximum instead of a minimum. Nevertheless, the regularity is striking, with the 24 marked intervals stretching across 1.79 octaves and generating an average keyboard spacing of 89 cents (or 1.053). Individual spacings vary from 73 to 106 cents. Note that the 89 cent average seen here is smaller than the 104 cent figure evident in the Braun work.

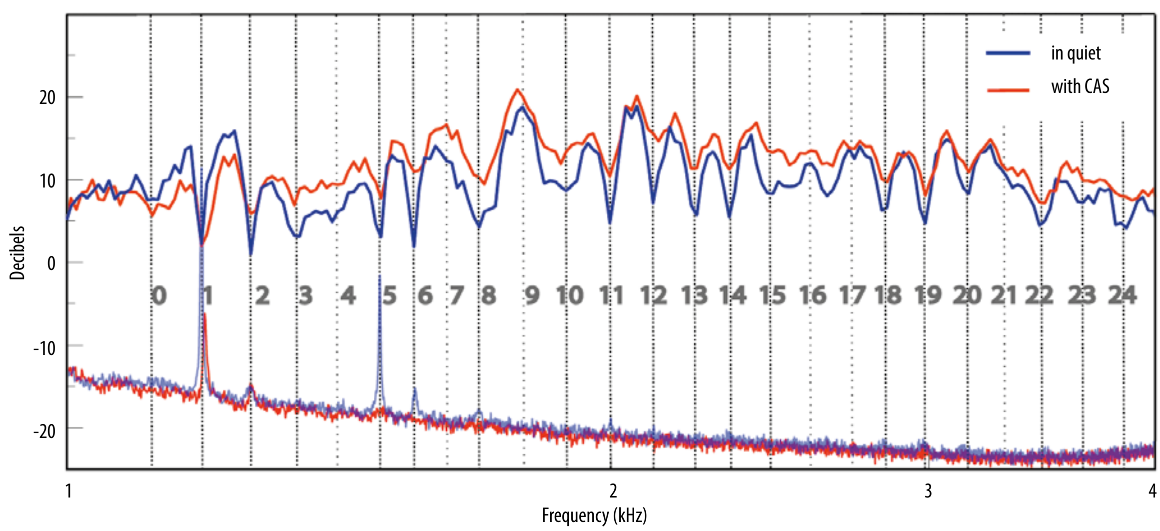



**Figure 2.** A keyboard in the human cochlea. Fine-grain measurements of hearing thresholds in the left ear of a female subject reveal it contains an almost unbroken sequence of semitone-like intervals extending over nearly 2 octaves. At top, the blue curve is the threshold measured in quiet, and red with contralateral noise. Below are spectra of the sound emitted by the ear – its spontaneous otoacoustic emissions. The 19 dashed vertical lines have been added to align with threshold minima, while the 6 dotted lines are interpolations added when the sequence is broken and expected minima are missing. The result in this subject is 24 'notes' ranging over 1.79 octaves, an average separation of 0.89 semitone (89 cents) and equating to a slightly narrow 'natural semitone' of 1.053. Adapted from Fig. 8C of Dewey et al. [15], with permission of Springer Nature, © 2014

In two other subjects tested by Dewey and colleagues, the ripples are much less apparent, suggesting that having a 'musical ear' depends on the particular individual tested. But taken together with the findings of preferred spacings of SOAEs in normal subjects [14], there appears to be intriguing evidence that, at least in a proportion of subjects, the human cochlea is intrinsically tuned to an interval close to a semitone.

Shifting the focus to other musical intervals, there is the peculiar circumstance of the enlarged octave. When measured accurately by careful subjective matching, the octave is found not to be precisely 2:1, as music theory would have it, but is slightly stretched, so that it takes on a value of about 2.03, depending on the subject and frequency [23–26]. Numerically, it may be more than coincidence that an enlarged octave of about 2.03 corresponds to an enlarged semitone of 1.061 (i.e., $2.03^{1/12}$), close to the value found experimentally by Bell & Jedrzejczak [14].

Finally, there is important work on the perception of musical intervals done by Levelt and colleagues [27], work which does not seem to have been repeated or received the attention it deserves. Making no mention of music to their subjects, Levelt and colleagues asked subjects to judge the similarity of different, simultaneously sounded, frequency ratios, and then used multidimensional scaling to discover several prominent auditory landmarks or "reference points" in auditory space (as illustrated in Figure 5 of [14]). Along the dimension of frequency ratio ($x$-axis), Levelt and colleagues found four landmarks, of which the two most prominent were the semitone and the fifth. It would be interesting to repeat the original work and fill in some of the details – for example, whether the first peak falls most closely to the equal-tempered semitone, the just semitone, or the natural semitone. In this context, Hall & Hess [28] report that the category boundary between the unison and semitone (their Fig. 11) is about 5 cents higher than expected from equal temperament, supporting an enlarged subjective semitone.

Bell & Jedrzejczak conjectured that the natural semitone might be a property of the automatic frequency-shift detectors reported by Demany and colleagues [29] and was part of a dual-template mechanism for analysing musical stimuli according to their harmonic and melodic components [30], and this point is considered further in the Discussion below.

What is the origin of the identified natural semitone? One idea put forward by Bell & Jedrzejczak [14] was that the ratio may derive from physical coupling of tuned elements in the ear, and that the behaviour of such coupled oscillators might be similar to the automatic frequency-shift detectors of Demany and colleagues. In this way, a keyboard of 1.06 steps might arise, generating something akin to the 'spiral staircase' in the cochlea that has been described by Shera on the basis of his detailed otoacoustic measurements [9]. There is also recent further evidence that multiple, discrete resonating elements might exist in the mammalian cochlea, each with a step size of about 1.1 [31]. However, rather than being fixed in place, it is possible that the steps might, in line with Demany's suggestion, be dynamic, synchronising automatically to incoming tones and able to register shifts of frequency – melodies – in terms of relative semitone steps.

An alternative speculation for the source of the natural semitone, an idea touched on by Bell and Jedrzejczak and which is explored in more detail here, is the surface acoustic wave (SAW) model of cochlear function. The SAW model is a dynamic model of cochlear tuning which involves reverberating wave activity between rows of outer hair cells, and this activity – likened to the vibrating strings of an 'underwater piano' [11] – has the potential to generate many interesting musical properties. The current paper elaborates on this model, indicating how it can give rise not only to a semitone but to other musical intervals as well.

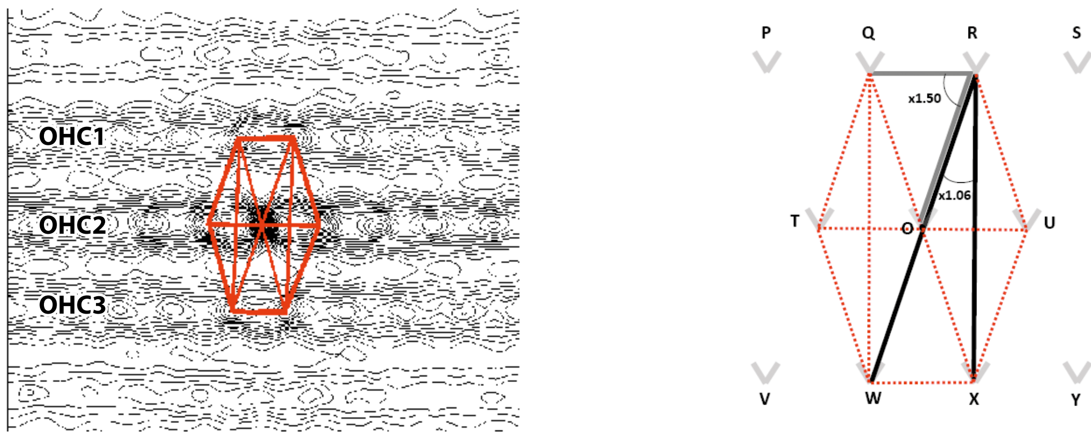
## The cochlea, cell–cell distances, and musical ratios

The precise geometric arrangement of outer hair cells (OHCs) in the mammalian cochlea (Fig. 1) has no accepted explanation. However, one possibility is that the arrangement is a way of splitting sound into its constituent frequencies, in particular its musical components [13]. The SAW model was originally developed as a micromechanical model of how SOAEs might be generated [11], but another notable aspect is that it also provides a possible mechanism for detecting the musical components of a sound. The idea here is that musical properties might be sensed as ratios between multiple acoustic resonances set up between adjacent outer hair cells. The model builds on the regularity with which the cells lie and suggests that wave interactions between neighbouring OHCs – which, uniquely, are both detectors and generators of sound energy [32] – could lead to the formation of multiple standing waves.

The SAW model supposes that in response to a musically rich sound, concerted activity of a local group of cells could, given the right spacings between them, lead to multiple standing waves and, together with a coincidence detector, allow the sound's musical components to be extracted. Figure 3 shows how this could work, given a particular set of observed intercellular distances. The outer hair cells in the cochlea are arranged in a repeating hexagonal lattice (Fig. 1 and Fig. 3a), so that by adjusting the hexagon's aspect ratio – which varies from base to apex – many musically significant frequency ratios can be generated. The geometry illustrated in Fig. 3a is from observations in a monkey, but it is not unlike the few observations that have been made on humans [33].

A distinctive feature of Fig. 1 and Fig. 3a is the offset of the central row (OHC2) from the flanking rows (OHC1 and OHC3), a characteristic which gives rise to diagonal lines that are about 20° from the vertical (observe the directions of red lines in the figure). In other words, the oblique cell–cell distances are 1/cos 20° (= 1.06) times longer than the shortest row-to-row distances, and this ratio, as with some others evident in Fig. 3, is taken to be musically important, as the rest of the text endeavours to make clear. The actual arrangement shown in Fig. 3a is idealised in Fig. 3b to illustrate how multiple musical ratios might arise.

Because the aspect ratio of the hexagon varies from base to apex (see [11]), a range of angles and preferred




(a) (b)

**Figure 3.** Musical possibilities from actual (a) and idealised (b) outer hair cell geometry. Outer hair cells lie in three precise rows (OHC1, OHC2, OHC3), producing a hexagonal array (red lines) whose relative dimensions can give rise to musical ratios.

**(a)** Contour map of spatial autocorrelations of the positions of outer hair cell stereocilia in a monkey (from Bell [11] based on observations of Lonsbury-Martin [34]). Each peak of the contour map is the preferred position of an outer hair cell (compare with Fig. 1). The repeating pattern of intercellular distances is outlined in red, producing a hexagonal geometry. Note that the prominent oblique red lines are tilted by about 20° to the vertical, meaning that they are longer by a factor of 1/cos 20° (= 1.06) compared to the vertical red lines.

**(b)** An idealised version of (a) in which angle WRX is specified to be 19.4°. Trigonometry then shows that the hexagon contains important musical ratios. In this case, the semitone (ratio 1.06, black lines) and perfect fifth (ratio 1.50, grey lines) feature strongly, and other musically significant ratios are also present (Table 1)

**Table 1.** Musical ratios arising from the specific hexagonal geometry and cell–cell distances shown in Fig. 3 (i.e., the aspect ratio of the hexagon's height to its width is 0.707, leading to a relative cell spacing of 0.3533 along the rows and 0.5 between the rows). Musical ratios emerge if wave fronts are exchanged between the cells and their two stereociliar arms act as coincidence detectors. Note that, like a hollow pipe, the length of a cavity is inversely proportional to the standing wave frequency it supports

| Dimensions | Ratio | Name | By inversion | Notes |
|---|---|---|---|---|
| PV:VP, QW:WQ, etc. | 1.000 | unison | unison | In the SAW model, this distance is the resonant cavity between OHC1 and OHC3 in which whole-wavelength standing waves are set up. Such a strong resonance is the 'characteristic frequency' of the basilar membrane |
| PV:VY | 1.060 | semitone | major seventh | Initial setting of aspect ratio (i.e., 1.06/3 = 0.3533) |
| WR:RX, QW:WR, etc. | 1.061 | semitone | major seventh | A second semitone ratio, this time marked as *black lines* in Fig. 3. Angle WRX is 19.4° |
| QR:RO, etc. | 1.501 | perfect fifth | fourth | Marked as *grey lines* in Fig. 3 |
| VO+OW:QW | 1.259 | major third | minor sixth | For clarity, not marked on figure |
| VW+WX:XR | 1.415 | tritone | tritone | For clarity not marked |
| OV+VW:OW | 2.040 | stretched octave? | stretched octave? | For clarity, not marked |
| PQ:QS, etc | 2.000 | octave | octave | Not marked |

ratios is possible (not just the 19° alignment), and this variation provides scope for further investigation. The pattern shown in Fig. 3 is fairly representative of what micrographs commonly show near the mid portion of the basilar membrane (i.e., the 19° alignment, which is considered important in producing a natural semitone). At other places on the basilar membrane, perhaps fifths and octaves might dominate in the interactions between cells.

In summary, the basis of the semitone, and possibly other musical intervals, is a reverberation process created by positive feedback between neighbouring outer hair cells. Effectively, resonance occurs between the cells in response to incoming sound stimulation. In sustaining this process, the stereocilia of each hair cell are both detectors and effectors [32], a facility which allows them to exchange wave energy, and, with the appropriate combination of wave speed and distance between cells, permits a standing wave resonance of fixed frequency to be set up. The process is similar to what happens with a vibrating string, except in the cochlear case the cells sustain the vibration by pumping in energy cycle by cycle. Build-up of wave energy occurs because there is a virtually instantaneous 'kick-back' of the stereocilia that occurs whenever a wavefront is detected, and, like a child building up waves in a bathtub, the process repeats and repeats [35]. The outcome of having multiple possible intercell distances is different sets of standing wave frequencies, and because each cell has two independent stereociliar arms (see Fig. 1) it is able to test for coincidence of standing waves in each arm. If the angle between the cavities is 19°, as illustrated in Figure 3, then the corresponding lengths (and hence frequencies) are in the ratio 1/cos19° = 1.058.

Other important musical ratios can be sensed in a similar way. In each case, an outer cell tests for coexisting standing waves in cavities formed in adjacent stereociliar arms. In effect, the two arms function as coincidence detectors: if the standing wave formed in one arm of a cavity coincides with a standing wave formed in the other, the coincidence is registered as the simultaneous absence of motion in both arms (because each arm will naturally lie at a node). In the SAW model, a cavity describes the space between one cell and its neighbours; however, since the space is filled with cochlear fluid, the waves must be fluid-borne and of short wavelength. Waves with such properties – so-called "squirting waves" – have been described in the subtectorial space [35].

Returning to Fig. 3 and Table 1, it can be seen that for the same intercellular distances in the hexagon which produce a semitone, other musically significant ratios also arise – notably the octave, perfect fifth, and major third – illustrating the virtue of the 19° alignment. The simultaneous emergence of all these ratios reinforces the idea that the cochlea itself may be the instrument for detecting musical ratios. To elaborate, Fig. 3b illustrates how a semitone could arise from a geometry in which the cell spacing WR is 1.06 times longer than XR, an alignment corresponding with the favoured 19° angle. Other musical ratios then arise similarly: a perfect fifth, for example, derives from the coincidence of standing waves in QR and RO. Due to symmetries of the hexagon, both these ratios appear multiple times in different combinations, and so they are expected to generate strong responses in the auditory system. Other comparisons yield further musical ratios, as listed in Table 1. The ratios include the major third (1.259), an enlarged minor third (1.213), minor seventh (1.782), the tritone (1.415), and of course their inversions. As foreshadowed, it is possible to consider the cochlear tuning curve, with its sharp tip and broad tail, as being a superposition of all these feedback responses.

The simultaneous appearance of both the semitone and the fifth is thought to be significant. It is compatible with the findings of Levelt and colleagues [27], who found that these two ratios form outstanding musical landmarks (Fig. 5 of [14]).

The proposed mechanism ties together a number of major strands in auditory science. The SAW model draws upon the distinctive features of OHCs – their regular geometrical arrangement, their simultaneous sensing and motor properties, and their two stereociliar arms – and matches them to some universal features of music, particularly the pivotal roles that the semitone, the fifth, and the octave play in defining auditory space. There are still many unanswered questions, but the following discussion is intended to point a way forward. The aim is to indicate how the SAW model might be used as a foothold to establish a natural theory of music, one which can explain recent audiological findings and which appears to generally align with a range of standard musical features.

## Discussion

### Natural theories of music

There have been recurring efforts to establish a natural theory of music (for context see [3,6,8,36]). It is fair to say that most these attempts have not been greeted with success, with Ball [5] saying that claims of a natural foundation for diatonic scales in mathematical and acoustical principles "have occasionally been comically absurd" (*ibid.*, p. 70). Helmholtz himself thought that "the construction of scales and harmonic tissue is a product of artistic invention, and by no means furnished by the natural formation or natural function of our ears" [12, pp. 365–6]. More recently, Patel [7] made the simple claim that "there are no sonic universals in music" (*ibid.*, p. 12).

The review by Burns & Ward [8] comes to the conclusion that, based on the evidence available at the time, musical intervals are learned rather than a direct result of characteristics of the auditory system. In fact, they regarded the evidence as arguing against the existence of frequency-ratio detectors at any level (*ibid.*, p. 260), although they add the important caveat that their own experiments with *isolated* musical intervals may have little to do with the perception of melodies (c.f. [20]).

Work by Brown & Jordania [36] is notable for its effort to "breakthrough the scepticism" that has surrounded discussion about universals in music, and although the authors are wary of ascribing a biological basis to music, they expressly identify 12 pitches per octave as the lowest common denominator across cultures. Likewise, Killin [4] is drawn to the strong common thread underlying all cultures' musics, but prefers a "social brain hypothesis". A remarkable fact illustrated in Killin's Fig. 6 is that around 500 BCE in China a set of 64 bronze gongs were cast which played 12 semitones to the octave, concrete evidence (still existing today) of the importance of the semitone even in ancient times.

The wide-ranging musicological survey by McDermott and Hauser [1] also points to certain recurring features of music which suggest the presence of some innate "machinery", an in-built mechanism that somehow encodes sound in terms of its relative pitch, most commonly in steps of 1 or 2 semitones. They highlight the difficulty of otherwise explaining relative pitch using the common beating model of consonance and dissonance, which clearly can't work for melodies. However, the real problem as they see it is that the mechanism is confounded by the environmental and cultural overlays to which everyone is exposed. Nevertheless, the McDermott and Hauser survey generally supports the idea that the innate elements might reside within "pre-existing structures in the auditory nervous system" (*ibid*., p.44), and one possibility they suggest is that there are frequency ratio detectors in the brain. The present work conjectures that such ratio detectors might in fact be located peripherally, in the cochlea.

Bidelman [37] approaches the subject of the origins of music by noting that "musical intervals and chords deemed more pleasant sounding by listeners are also more prevalent in tonal compositions", apparently connecting music with a desire for harmony and the avoidance of dissonance. But this seems to overlook that the most beautiful of melodies often involve the tone and the semitone, which are not regarded as harmonious. In fact, in the context of melodies, it doesn't seem appropriate to speak of intervals like the semitone as being dissonant when in fact there is no "dissonance" – a somewhat pejorative term – in sequential tones (Huron [38] defines dissonance as "a low-level auditory irritation"). Bidelman may be closer to the mark when he remarks (p. 7) that "dissonance may challenge the auditory system in ways that simple consonance does not", and goes on to make the case that there is a subcortical pitch processor that may be mirrored in upstream sites, including cortex. That subcortical processor could, of course, be as peripheral as the cochlea. Significantly, Bidelman concludes his 2013 paper by supporting an innate basis for music residing somewhere in the auditory system.

The arguments are far from resolved, but it is hoped that experimental approaches focusing on the cochlea may open up a fresh way of exploring the issue.

## Melody and the natural semitone

Melody is the musical property elicited by the sequential – that is, non-simultaneous – presentation of two tones, and the outstanding feature here is that, rather than the semitone being an inferior ratio remaining after other harmonious rankings have been considered (as commonly viewed), the semitone is itself primary, the "atom" of melody, at least in Western cultures.

The earlier paper by Bell & Jedrzecjzak [14] suggests that harmony and melody are two separate elements of our musical perception, supporting an idea first put forward by Demany & Semal [30]. Harmony is the perception of the consonance or dissonance created when two tones are sounded together (simultaneously, so that both tones enter the cochlea and interact). It is a subjective measure, and what intervals are considered harmonious have varied over the millennia. Pythagoras confined harmony to either the joint sounding of the octave or the fifth, whereas nowadays it has been expanded to include many more intervals including the fourth, major third, minor third, and others [5,39] – but not the semitone, which invariably ends up at the very bottom of consonance ratings [37]. Interestingly, Brown & Jordania [36] point out that harmonising on narrow intervals such as seconds is a practice widely distributed across the globe (see also [40]).

The geometric SAW model under examination here is distinguished by a prominent semitone, and the audiometric evidence points towards this feature. For simplicity, the text here has therefore focused on the semitone and emphasised the melodic aspects of music. We take the semitone to be the essential building block of melody, discounting notions that it should be regarded of lesser importance just because it is antithetical to harmony. Nevertheless, it is true that the music literature focuses heavily on harmony, with much of it dealing with consonance and dissonance. However, rather than dwell on why the semitone appears dissonant in a harmonic setting, the point is simply made that there must be more to music than just the rate of beating of partials [23,30,41,42].

It is also worth noting that even in theoretical calculations of consonance and dissonance – e.g. the curves of Helmholtz and of Plomp & Levelt – the semitone appears very deep and, moreover, very precisely defined. Rakowski [43] measured the accuracy with which music students could tune musical intervals, and the findings were that, after the unison and octave, the semitone was the next most accurately tunable interval. Additionally, Larrouy-Maestri [20] found that, in a melody, very small errors in its tuning were readily perceived, and in this regard the special nature of the semitone in generating melody has not been sufficiently explored. Some beginnings have been made [44], but we are far from understanding what makes the semitone so distinctive.

The explanation offered here for why the semitone stands apart is that it is actually a primary building block of music picked out by the OHC geometry in response to incoming sounds. It is remarkable that 70% of the intervals in all melodies are, in Western music at least, comprised of 0, 1, or 2 semitones (p. 109 of [5]; [45]). That is, without the semitone and the tone, there would be no memorable tunes in the Western canon. It is clearly a principal task of the ear to discriminate two sequential tones differing by either one or two steps. Scientifically, this is a major puzzle, and the usual explanation is that the semitone is picked out by some complicated brain process further on in the auditory system [37].

For non-Western music the scheme is somewhat different (p. 109 of [5]; [2]). In Chinese music, for example, 70% of the intervals are 0, 2, or 3 semitones – in this pentatonic music, a distinct semitone is absent, although it is implicit within the prevailing two- and three-step spacing. In Japanese music, the semitone seems less significant, with two-thirds of the steps being of 0, 1, 2, 3, or 4 semitones, while in African music 80% fall within this range. Of course, in modern times both Japanese and African people are easily able to appreciate the elements of Western-style music, including the semitone, so clearly they possess some mechanism that allows them to readily extract and identify this interval. The mechanism is here supposed to reside in their ears; however, otoacoustic measurements of these non-Western populations would be of great value in backing up this claim.

The music literature is heavily weighted towards Western experience [36], and more studies on Eastern styles of music, as well as the corresponding otoacoustic emission data, are very much needed.

The outstanding role of the semitone – either alone or in combination with the tone – suggests it is a fundamental aspect of many musics, and the reason offered here is that it is a built-in feature of the human cochlea. Krumhansl [46] points out that the semitone is relatively easy to produce, label, and remember, and has considerable power as a cue to key-finding. This view is supported by Jackendoff & Lerdahl [47], who note that half and whole steps, although harmonically rough, are the most common, stable, and effortless components in a melody. Relevant here, perhaps, is that when music is played to monkeys, they only respond when the music contains notes taken from the diatonic scale – that is, it contains tones and semitones [48]. However, more generally, the literature on whether animals perceive music is equivocal [1], although a study of how their cochlear hair cells are arranged might be illuminating.

Against the view that melodic machinery is built in to the auditory system, it can be argued that some musical cultures are totally bereft of the semitone – for example the gamelan orchestras of Bali – but this is hard to establish (Ch. 10 of [19]; [49]) and such instances are left as test cases. A broad comparison of the scales of five of the world's major cultures shows that the semitone appears to play a strong organising role (Fig. 4 of [2]). Although certain cultures use intervals smaller than a semitone in some contexts, these "microtonal" intervals are mostly used ornamentally rather than for melodic purposes [36].

## Synthesis of harmony and melody

Among all the discussion of the roles that harmony and melody play in music, it is significant that the SAW model can give at least a basic account of how both aspects might arise. Indeed, one of the claims of the present model is that it can, given an appropriate geometric arrangement, potentially account for all 12 tones of the diatonic scale, whether sequentially or simultaneously perceived. So then, in terms of the SAW model, what is the effective difference between consonance and dissonance? An attempt to supply an answer is not made here, but a framework from which it might be usefully addressed is the synthesis of harmony and melody attempted by Bell & Jedrzejczak [14]. In this work, they support the ideas of Van Noorden [41] and Demany & Semal [30] and conclude that musical perception is in fact built on these two separate, but complementary templates, one harmonic and the other melodic.

Ball neatly sums up the matter by saying that if melody is the path, harmony is the terrain (p. 165 of [5]). In terms of the SAW model, the suggestion is that there is the same physical mechanism – coincidence of standing waves – behind each of the two distinct templates. The harmonic template could, following Helmholtz, depend on the OHCs detecting the simultaneous coincidence of partials. Melody, on the other hand, might derive from a process, again involving the OHCs, which is able to register (and remember) previous semitone steps. The findings of preferred semitone intervals in SOAEs and in auditory microstructure support the possibility that such processes may exist. When the ear is excited by a certain frequency, a series of small frequency steps is formed (Fig. 4 of [14]), perhaps generated by the 19° tilt in the geometry of the underlying hair cells, and this template – associated with the automatic frequency-shift detector of Demany and Ramos [50] – might be used to measure the number of semitones. Memory is therefore essential for extracting melody, requiring the position of the steps to be marked, either in the cochlea or the brain (or both), and Demany and Ramos found experimental evidence that the memory of frequency shifts had a time-frame of several seconds.

The above text suggests that a large part of musical experience might be based on just two or three musical landmarks – the octave, the semitone, and perhaps the fifth – and from them the entire chromatic scale and sense of harmony might follow. In vision, three components – red, green, and blue – define all of colour space, whereas in auditory space, the speculation made here is that the ratios of 2:1 and 3:2 establish harmony, while the cochlear semitone establishes melody (and dissonance). The two processes work together but somewhat independently, and operate according to Demany and Semal's melodic and harmonic templates.

In all this the integrating roles of central brain processes cannot be neglected [8,37,51–53], and it is acknowledged that culture can indeed shape musical perceptions [54], but the otoacoustic and hearing threshold measurements suggest that there is also some peripheral process at work in our perception of music. According to the proposal put forward here, reverberation between neighbouring hair cells in the cochlea is at the core of that process. Depending on the details of their outer hair cell geometry, some ears are highly musical, while others are less so.

## Conclusions

This paper has pointed to evidence that frequency ratios exist in the cochlea which are very close to a semitone. The empirical data supports the notion that, at least in Western subjects, there is a natural semitone in the ear, suggesting a naturalistic basis for music more generally. It appears that the natural semitone, with a frequency ratio of about 1.06, is a result of some subtle cochlear processing of sound, and might be associated with the regularity with which the outer hair cells in the cochlea are arranged. Together with an automatic frequency-shift detector, both these features would allow the ear to readily perceive musical intervals and melodies. Further work is needed to confirm or discount these speculations.

Establishing the ear itself as the initial site of musical perception would represent a radical change in direction for music theory [55], since recent work has tended to focus on the neurophysiology of the brain as the locus of our musical sense [56,57]. For a long time, experimentalists have focused largely on harmony, measuring it in terms of consonance and dissonance. On this basis, simple integer ratios have become preeminent, and the complex, even irrational semitone has been relegated to an inferior status [21,39]. Yet such a perspective has difficulty explaining how the semitone can be at the core of melody, a quintessential musical dimension.

Of course, it is accepted that the brain is also necessary for accurate music perception. In some cases of amusia, the ear is perfectly intact but certain brain regions are compromised [42,58]. However, in such cases it could still be that the ear is primary, with brain processing occurring in a subsequent step.

This work has looked for a possible mechanism for how a natural semitone might arise, and puts forward a micromechanical model – the surface acoustic wave (SAW) model – for scientific evaluation. The model involves reverberation of sound between adjacent outer hair cells, and relies on the regular geometric arrangement of the cochlear outer hair cells. Its outstanding feature is that it sets out a common physical mechanism through which the semitone, and other important musical ratios, are detected. Although speculative, the proposal has the potential to expand and unify our understanding of many psychophysical observations concerning music, opening up possible new solutions to longstanding puzzles.

From the standpoint of conventional music theory, the existence of natural ratios may be surprising, despite the fact that naturalistic foundations for music have often been proposed [36]. An advantage of the peripheral analysis of sound is that it could be both fast and effective, giving it a strong evolutionary advantage [1,3,59].

Thorough investigation and testing will be needed in order to decide the merits of the ideas raised, but further objective evidence would have major implications. The suspicion raised here is that the human faculty for music might be biologically hard-wired, and that our wondrous musical sense is innate.

## References

1. McDermott J, Hauser M. The origins of music: innateness, uniqueness, and evolution. Music Perception, 2005;23:29–59.
2. Justus T, Hutsler JJ. Fundamental issues in the evolutionary psychology of music: assessing innateness and domain specificity. Music Perception, 2005;23:1–27.
3. Honing H, ten Cate C, Peretz I, Trehub SE. Without it no music: cognition, biology and evolution of musicality. Philos Trans R Soc Lond B Biol Sci, 2015;370:e20140088.
4. Killin A. The origins of music: evidence, theory, and prospects. Music & Science, 2018;1:1–23.
5. Ball P. The Music Instinct: How music works and why we can't do without it. Oxford: Oxford University Press; 2010.
6. Parncutt R. Harmony: A psychoacoustical approach. New York: Springer; 1989.
7. Patel AD. Music, Language, and the Brain. Oxford: Oxford University Press; 2008.
8. Burns EM, Ward WD. Intervals, scales, and tuning. In: Deutsch D, editor. The Psychology of Music. New York: Academic; 1982. p. 241–69.
9. Shera CA. The spiral staircase: tonotopic microstructure and cochlear tuning. J Neurosci, 2015;35:4683–90.
10. Motallebzadeh H, Soons JAM, Puria S. Cochlear amplification and tuning depend on the cellular arrangement within the organ of Corti. Proc Natl Acad Sci USA, 2018;115:5762–7.
11. Bell JA. The Underwater Piano: A Resonance Theory of Cochlear Mechanics. PhD thesis, Australian National University, Canberra 2005.
12. Helmholtz HLF. On the Sensations of Tone as a Physiological Basis for the Theory of Music. London: Longmans, Green; 1875.
13. Bell A. Musical ratios in geometrical spacing of outer hair cells in the cochlea: strings of an underwater piano? 7th International Conference on Music Perception and Cognition; Sydney: Causal Productions, Adelaide; 2002.
14. Bell A, Jedrzejczak WW. The 1.06 ratio in the cochlea: evidence and outlook for a natural musical semitone. PeerJ, 2017;5:e4192.
15. Dewey JB, Dhar S. Wideband profiles of stimulus-frequency otoacoustic emissions in humans. MoH Proceedings, 2014.
16. Sekerkova G, Richter C-P, Bartles JR. Roles of the espin actin-bundling proteins in the morphogenesis and stabilization of hair cell stereocilia revealed in CBA/CaJ congenic jerker mice. PLOS Genetics, 2011;7:e1002032.
17. Shera CA, Charaziak KK. Cochlear frequency tuning and otoacoustic emissions. Cold Spring Harbor Perspectives in Medicine, 2018:10.1101/cshperspect.a033498.
18. Dewey JB, Lee J, Dhar S. Effects of contralateral acoustic stimulation on spontaneous otoacoustic emissions and hearing threshold fine structure. J Assoc Res Otolaryngol, 2014;15:897–914.
19. Sethares WA. Tuning, Timbre, Spectrum, Scale. London: Springer; 1998.
20. Larrouy-Maestri P. "I know it when I hear it": on listeners' perception of mistuning. Music & Science, 2018;1:1–10.
21. McAdams S. Music: a science of the mind? Contemporary Music Review, 1987;2:1–61.
22. Braun M. High-multiple spontaneous otoacoustic emissions confirm theory of local tuned oscillators. SpringerPlus, 2013;2:135.
23. Bonnard D, Micheyl C, Semal C, Dauman R, Demany L. Auditory discrimination of frequency ratios: the octave singularity. J Exp Psychol Hum Percept Perform, 2013;39:788–801.
24. Bonnard D, Dauman R, Semal C, Demany L. Harmonic fusion and pitch affinity: is there a direct link? Hear Res, 2016;333:247–54.
25. Van den Brink G. Octave and fifth settings for pure tones and residue sounds. In: Evans EF, Wilson JP, editors. Psychophysics and Physiology of Hearing. London: Academic Press; 1977, p. 373–9.
26. McKinney MF, Delgutte B. A possible neurophysiological basis of the octave enlargement effect. J Acoust Soc Am, 1999;106:2679–92.
27. Levelt WJM, van de Geer JP, Plomp R. Triadic comparisons of musical intervals. Br J Math Stat Psychol, 1966;19:163–79.
28. Hall DE, Hess JT. Perception of musical interval tuning. Music Perception, 1984;2:166–95.
29. Demany L, Pressnitzer D, Semal C. Tuning properties of the auditory frequency-shift detectors. J Acoust Soc Am, 2009;126:1342–8.
30. Demany L, Semal C. Harmonic and melodic octave templates. J Acoust Soc Am, 1990;88:2126–35.
31. Bell A, Wit HP. Cochlear impulse responses resolved into sets of gammatones: the case for beating of closely spaced local resonances. PeerJ, 2018;6:e6016.
32. Brownell WE. What is electromotility? The history of its discovery and its relevance to acoustics. Acoustics Today, 2017;13(1):20–7.
33. Bredberg G, Ades HW, Engström H. Scanning electron microscopy of the normal and pathologically altered organ of Corti. Acta Otolaryngol, 1972; Suppl. 301:3–48.
34. Lonsbury-Martin BL, Martin GK, Probst R, Coats AC. Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy. Hear Res, 1988;33:69–94.
35. Bell A, Fletcher NH. The cochlear amplifier as a standing wave: "squirting" waves between rows of outer hair cells? J Acoust Soc Am, 2004;116:1016–24.

36. Brown S, Jordania J. Universals in the world's musics. Psychology of Music, 2011;41:229–48.
37. Bidelman GM. The role of the auditory brainstem in processing musically relevant pitch. Frontiers in Psychology, 2013;4:e264.
38. Huron D. Sweet Anticipation: Music and the psychology of expectation. Cambridge, MA: MIT Press; 2006.
39. Pesic P. Hearing the irrational: music and the development of the modern concept of number. Isis, 2010;101:501–30.
40. Jordania J. Choral Singing in Human Evolution. Saarbrücken: Lambert Academic Publishers; 2015.
41. van Noorden L. Two channel pitch perception. In: Clynes M, editor. Music, Mind, and Brain. New York: Plenum; 1982.
42. Cousineau M, McDermott JH, Peretz I. The basis of musical consonance as revealed by congenital amusia. Proc Nat Acad Sci, 2012;109:19858–63.
43. Rakowski A. Intonation variants of musical intervals in isolation and in musical contexts. Psychology of Music, 1990;18:60–72.
44. Patel AD. A new approach to the cognitive neuroscience of melody. In: Peretz I, Zatorre RJ, editors. The Cognitive Neuroscience of Music. Oxford: Oxford University Press; 2003, p. 325–45.
45. Vos PG, Troost JM. Ascending and descending melodic intervals: statistical findings and their perceptual relevance. Music Perception, 1989;6:383–96.
46. Krumhansl CL. Tonal hierachies and rare intervals in music cognition. Music Perception, 1990;7:309–24.
47. Jackendoff R, Lerdahl F. The capacity for music: what is it, and what's special about it? Cognition, 2006;100:33–72.
48. Wright AA, Rivera JJ, Hulse SH, Shyan M, Neiworth JJ. Music perception and octave generalization in rhesus monkeys. J Exp Psychol Gen, 2000;129:291–307.
49. Rahn J. "Equiheptatonic" tuning in Thai classical music: strict propriety and step sizes. Music & Science, 2019;2:1–15.
50. Demany L, Ramos C. On the binding of successive sounds: perceiving shifts in nonperceived pitches. J Acoust Soc Am, 2005;117:833–41.
51. Bidelman GM, Heinz MG. Auditory-nerve responses predict pitch attributes related to musical consonance–dissonance for normal and impaired hearing. J Acoust Soc Am, 2011;130:1488–502.
52. Cousineau M, Demany L, Pressnitzer D. The role of peripheral resolvability in pitch-sequence processing. J Acoust Soc Am, 2010;128:EL236–41.
53. Peretz I, Zatorre RJ. The Cognitive Neuroscience of Music. Oxford: Oxford University Press; 2003.
54. McDermott HJ, Schultz AF, Undurraga EA, Godoy RA. Indifference to dissonance in native Amazonians reveals cultural variation in music perception. Nature, 2016;535:547–50.
55. McDermott JH. What can experiments reveal about the origins of music? Current Directions in Psychological Science, 2009;18:164–8.
56. Bidelman GM, Grall J. Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. Neuroimage, 2014;101:204–14.
57. Bidelman GM, Krishnan A. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. J Neurosci, 2009;29:13165–71.
58. Peretz I, Hyde KL. What is specific to music processing? Insights from congenital amusia. Trends in Cognitive Sciences, 2003;7:362–7.
59. Fitch WT. The biology and evolution of music: a comparative perspective. Cognition, 2006;100:173–215.