Pitch perception models - a historical review

Alain de Cheveigné

CNRS - Ircam, Paris, France
cheveign@ircam.fr

Abstract

This paper analyzes theories and models of pitch from a historical perspective. Pythagoras is credited with the first "psychophysical" model, the monochord, that he used to formulate a law that links a physical quantity (ratio of string lengths) to a psychological quantity (musical interval). The relation between pitch itself and frequency emerged progressively with Aristoxenos, Boethius, Mersenne and Galileo. The anatomist Du Verney first proposed the idea of resonance within the ear, and of a "tonotopic" projection from the ear to the brain. The notion of frequency analysis, formalized mathematically by Joseph Fourier, was developed by Helmholtz into a beautiful theory of hearing that bridged mathematics, physiology and music. Helmholtz unfortunately followed Ohm in postulating that pitch is determined by one particular frequency component, the fundamental, thus sparking a controversy that has drained energy of hearing scientists for decades, opposing tenants of "spectral pitch" to tenants of "temporal pitch". Today the terms of the disagreement have shifted, and the disagreement is now between models based on "pattern matching" (originated by de Boer but already hinted to by Helmholtz) and those based on "autocorrelation" (originated by Licklider, but already implicit in earlier work). Despite the disagreements, there are deep connections between these various theories of pitch, and between them and the many methods that have been proposed for the artificial equivalent of pitch perception: fundamental frequency estimation. Using a historical perspective I will try to make apparent these relations between models and methods. The aim is to help us go beyond the controversies and develop a better understanding of how we perceive pitch.

1. Introduction

The history of yesterday's ideas suggests that today's might not last, and that better ones await us in the future. By looking carefully at theories that did not survive, we may learn to identify the weak points of our own theories and fix them. The historical perspective has other virtues. Among factors that slow down progress in Science, Boring cites the need to conform to the "Zeitgeist", the spirit of the times [1]. Another factor is controversy that may lock progress into sterile argument. History serves as an antidote to these factors. Models are often reincarnations of older ideas, themselves with roots deeper in time. By digging up the roots we can see the commonalities and differences between successive or competing models. Anyone who likes ideas will find many good ones in the history of science.

Some early theories focused on explaining *consonance and musical scales* [2], others on the *physiology of the ear* [3], and others again on the *physics of sound* [4, 5, 6]. Certain thinkers, such as Helmholtz [7], have tried to address all these aspects, others were less ambitious. Music once constituted a major part of Science, and theories of music were theories of the world. Today, music and science go each their own way, and the goal of hearing science is more modestly to explain how we perceive sound. However, music is still an important part of our auditory experience, and, historically, theories of hearing have often been theories of musical *pitch*.

Today, two competing explanations of pitch prevail: *autocorrelation* and *pattern-matching*, that inherit from the rival theories of *place* and *time*, themselves rooted in early concepts of *resonance* and *time interval*. Autocorrelation and pattern-matching each have variants. The historical perspective reveals both their unity and the originalities of each, and suggests directions in which future models might evolve. This paper is a short version of an upcoming chapter on pitch perception models [8].

2. Resonance

2.1. Interval and ratio, pitch and frequency

Pythagoras (6th century BC) is credited for relating *musical intervals* to ratios of string length on a monochord [6]. The monochord consists of a board with two bridges between which a string is stretched. A third bridge divides the string in two parts. Intervals of unison, octave, fifth and fourth arise for length ratios of 1:1, 1:2, 2:3, 3:4, respectively. The monochord can be seen as an early example of a *psychophysical model*, in that it relates the perceptual property of musical interval to a ratio of physical quantities. The physics of the model were quickly occluded by the mathematics or mystics of the *numbers* involved in the ratios [2]. Ratios of numbers between 1 and 4 were taken to govern both musical consonance and the relations between heavenly bodies. Aristoxenos (4th century BC) disagreed with the Pythagoreans that numbers are relevant to music, and instead argued that musical scales should be defined based on what one *hears* [9]. Two millenia later, Descartes made the same objection to Mersenne [2]). In 1581 the role of number was also challenged, from a different perspective, by Vincenzo Galilei (father of Galileo). Using weights to vary the *tension* of a string, he found that the abovementioned intervals arise for ratios of 1:1, 1:4, 4:9, and 9:16 respectively [10, 2]. These ratios are different from those found for length: they are more complex, and don't agree with the importance that the Pythagoreans gave to numbers from 1 to 4. Deciding the respective roles of mathematics, physics, and perception in the "laws" of music is still a problem today.

In addition to interval, the Greeks had the concept of *pitch*, a quantity by which sounds can be ordered from grave to acute [9]. They probably associated it with *rate*, but semantic overlaps between rate (of vibration), speed (of propagation) and force (of excitation) makes this unsure. The relation between ratios of string length and ratios of *vibration frequency* was established by Galileo Galilei [11], whereas Mersenne [12], using strings long enough to count vibrations, determined the actual frequencies of each note of the scale. This provided a relation of *pitch* with number that was firmly grounded in the physics of sound.

2.2. Sympathetic resonance in the ear

A string produces musical sounds, but it can also vibrate in sympathetic *resonance* as noted by Aristotle [10]. The perception of like by like was a common notion, and so the concept of resonance has been used in theories of hearing from antiquity onwards [1, 3, 6].

In 1683, Du Verney proposed that the bony spiral lamina within the cochlea serves as a resonator: "*...being wider at the start of the first turn than the end of the last ...the wider parts can be caused to vibrate while the others do not ...they are capable of slower vibrations and consequently respond to deeper tones ...in the same way as the wider parts of a steel spring vibrate slowly and respond to low tones, and the narrower parts make more frequent and faster vibrations and respond to sharp tones ...according to the various motions of the spiral lamina, the spirits of the nerve which impregnate its substance [that of the lamina] receive different impressions that represent within the brain the various aspects of tones*" [13]. This paragraph concentrates several key concepts of place theory: *frequency-selective response*, *tonotopy*, and *tonotopic projection* to the brain. Subsequent progress of the resonance theory is recounted in [3]. In 1758, Le Cat [14] proposed that the basilar membrane is constituted of *strings* like those of a harpsichord, and Helmholtz later used a similar metaphor.

2.3. Superposition and Ohm's law

Mersenne reported that he could hear within the sound of a string, or a voice, up to five pitches corresponding to the fundamental, the octave, the octave plus fifth, etc. [12]. He knew also that a string can respond sympathetically to higher harmonics, and yet he found it hard to accept that it could vibrate *simultaneously* at all those frequencies. This was easier for a younger mind such as that of Sauveur, who in 1701 coined the terms "fundamental" and "harmonic" [15]. The physics of string vibration were worked out in the 18th century by a succession of physicists: Taylor, Daniel Bernoulli, Lagrange, dAlembert, and Euler [4]. Euler in particular, by introducing the concept of *linear superposition*, made it easy to understand the multiple vibrations of a string that had so troubled Mersenne.

Mersenne and Galileo usually conceived of vibrations as merely being periodic, without regard to their shape, but 18th century physicists found that solutions were often easy to derive if they assumed "pendular" (sinusoidal) vibrations. For linear systems, they could then extend the solutions to any *sum of sinusoids* thanks to Euler's principle. A wide variety of shapes can be obtained in this way, meaning that the method was quite general. That *any* shape can be obtained in this way was proved in 1820 by Fourier [16]. In particular, any periodic wave can be expressed as the *superposition* of sinusoids with periods that are integer fractions of the fundamental period. Fourier's theorem had a tremendous impact on mathematics and physics.

Up to that point, pitch had been closely associated with progress in the physics of periodic vibration, and it seemed obvious that this new tool must somehow be relevant to pitch. In 1843 Ohm formulated a law, later rephrased and clarified by Helmholtz, according to which every pitch corresponds to a *sinusoidal partial* within the stimulus waveform. For Ohm, the presence of a partial was ascertained by applying Fourier's theorem, and Helmholtz proposed that the same operation is approximated by the cochlea[17, 7].

Ohm's law extended the principle of linear superposition to the sensory domain. Just as a complex waveform is the sum of sinusoids, so for Helmholtz the sensation produced by a complex sound such as a musical note was "composed" of simple sensations, each evoked by a partial. In particular, he associated the main pitch of a musical tone to its *fundamental* partial.

2.4. The missing fundamental

Ohm's law is the result of a *choice*. Mersenne had given little attention to the shape of periodic vibrations which he had no means to observe. His law relating frequency to pitch did not mention shape. However, Fourier's theorem now implied that, depending on its shape, a vibration might contain *several* sinusoidal partials, each with a different frequency. This raised an obvious question: does pitch relate (a) to the period of the vibration as a whole, or (b) to the period of one of the partials? If (b) is true, then Fourier analysis is required to determine pitch, if (a) it is unnecessary. Ohm chose (b).

Seebeck had already addressed the question experimentally, using a siren to produce periodic stimuli with several pulses irregularly spaced within a period[1]. Regardless of the number of pulses, pitch followed the fundamental period, consistent with (a). Furthermore, by applying Fourier's theorem to the waveform, Seebeck showed that pitch salience did not depend on the relative amplitude of the fundamental partial, which for some pulse configurations was very small. Since the same pitch was also heard when the stimulus contained *only* that partial, he could conclude that *pitch does not depend on a particular partial*. This contradicted (b). Low pitch in the absence of a fundamental partial was already known from earlier work on beats [18].

Nevertheless, Ohm chose (b) and Helmholtz endorsed this choice. Many authors have puzzled over the Seebeck-Ohm-Helmholtz controversy and the reasons why Helmholtz did not take seriously Seebeck's arguments[1, 19, 20, 21]. One reason was no doubt that, by extending Ohm's law to upper harmonics, Helmholtz could explain the higher pitches that some people (among which Mersenne and himself) occasionally heared. One can speculate that additional reasons were the conviction that a theorem as powerful as Fourier's *must* be relevant, and the desire to ensure that the parts of his monumental theory would fit together.

Helmholtz had three options to address the missing fundamental problem without renouncing his theory, two of which he used. The first was to invoke *nonlinear distortion* in sound-producing apparatus or in the ear. As an explanation of periodicity pitch, that hypothesis was quite weak already at the time, as argued by Helmholtz's translator, Ellis [7]. However it took over sixty years before Schouten and Licklider laid the explanation to rest. With an optical siren, Schouten produced a complex tone that lacked a fundamental. He managed, not only to prove that the distortion product at the fundamental had a very low amplitude, but also to *cancel* it. The absence of a fundamental component was verified by adding a sinusoidal tone with a nearby frequency and checking for absence of a beat. The low pitch was unaffected by removeal of the fundamental partial, as it was unaffected when Licklider masked it with noise [22, 23]. This rules out the distortion product explanation of low pitch. However distortion products *do* exist, and they sometimes *do* affect pitch, so that explanation tends to resurface from time to time.

A second option was Helmholtz's concept of "unconscious inference" that prefigured *pattern matching* (next section)[24]. A third option, that Helmholtz apparently did not use, was to treat cochlear resonators as *strings*.

As Mersenne and others had noticed, a string vibrates sympathetically with sounds tuned to its fundamental mode *and with their harmonics*. Thus it responds to a periodic sound regardless of whether or not it contains a fundamental partial. It is, in essence, a filter tuned to periodicity. Helmholtz had used the bank-of-strings metaphor to describe the cochlea. Nevertheless, he chose to characterize each filter as if it were a *Helmholtz resonator* tuned to a single sinusoidal partial. Had he chosen to treat them as a strings, the missing fundamental problem would have not existed. Of course, a bank of strings does not fit Fourier's theorem, and this is perhaps why he did not choose this option. If he *had* chosen it, the model would have eventually been proven wrong as cochlear filters are not tuned to periodicity.

3. Pattern matching

We are confronted with incomplete patterns everyday, and our brain is good at "reconstructing" perceptually the parts that are missing. Pattern matching models assume that this is how pitch is perceived when the fundamental partial is missing. The idea is thus that the fundamental partial *is* the necessary correlate of pitch, as Ohm claimed, but that it may nevertheless be absent if other parts of the pattern (harmonics often associated with it) are present. This idea was prefigured by Helmholtz's "unconscious inference" and John Stuart Mill's concept of "possibilities" [24, 1]. As a possible mechanism, Thurlow suggested that listeners use their own voice as a "template" to match with incoming patterns of harmonics [25].

In 1956, de Boer described the concept of pattern matching in his thesis [26], but the best-known models are those of Goldstein [27], Wightman [28] and Terhardt [29]. These models are closely related, but each has its characteristic flavor. Goldstein's is probabilistic and performs optimum processing of a set of estimates of partial frequencies (obtained by a process that is not defined, but that could be Helmholtz's cochlear analysis). Wightman takes the limited-resolution profile of activity across the cochlea, and feeds it to a hypothetical internal "Fourier transformer" to obtain a pattern akin to the autocorrelation function. Terhardt follows Ohm in positing for each partial its own sensation of *spectral pitch*, from which an internal template derives a *virtual pitch* that matches that of the (possibly missing) fundamental. That template is learned.

Pattern-matching models are well known and will not be described in greater detail here. There is a close relationship between pattern-matching models and spectrum-based signal-processing methods for *fundamental frequency estimation*, such as subharmonic summation, harmonic sieve, autocorrelation or cepstrum [30, 8]. For the last two, this reflects the fact that the Fourier transform, applied to a spectrum (power spectrum for autocorrelation, log spectrum for cepstrum) is sensitive to the regular pattern of harmonics.

Terhardt's model is distinct in that it requires that templates be *learned* by exposure to harmonically rich stimuli, an idea that is attractive but constraining. It can be argued that learning could just as well occur from exposure to patterns of *subharmonics* (superperiods) of periodic sounds, and that harmonically rich stimuli are thus unnecessary [8]. Shamma and Klein went further and showed that templates may be learned by exposure to *noise* [31]. What this suggests is that the harmonic relations within the template are a mathematical property that needs merely to be *discovered*, not learned. Indeed, other devices embody the pattern-matching properties of a harmonic template without having "learned" them. Examples are the autocorrelation function and the string.

4. Temporal models

Democritus (5th century BC) and Epicurus (4th century BC) are credited with the idea that a sound-producing body emits *atoms* that propagate to the listener's ear, an idea later adopted by Beeckman and Gassendi [2]. Related is the idea that a string "hits" the air repeatedly, and that pitch reflects the rate at which sound pulses hit the ear [32, 2]. If so, it should be a simple matter to measure the interval between *two* consecutive atoms or pulses, rather than wait for a series of pulses to build up sympathetic vibration in a resonator.

The influence of this temporal view of pitch can be observed indirectly in the "coincidence" theories of consonance that developed in the 16th and 17th centuries. Two notes were judged consonant if their vibrations coincided often [2].

Early temporal models assumed that patterns of pulses are handled by the "brain", and thus they tend to be less elaborate than resonance models. Compare for example Anaxagoras (5th century BC) for whom hearing involved *penetration of sound to the brain*, and Alcmaeon of Crotona (5th century BC) for whom *hearing is by means of the ears, because within them is an empty space, and this empty space resounds* [6]. The latter obviously "explains" more. A similar contrast is seen between the monumental resonance theory of Helmholtz [7], and the two-page "telephone theory" that Rutherford opposed to it, according to which the ear is merely a telephone receiver that transmits pulses to the brain [33].

Rutherford is one of several thinkers that opposed the Helmholtz theory [34]. One can speculate that they disapproved of Ohm's choice (Sect. 2.3), objected to the obligatory Fourier analysis, and in general resented the weight of Helmholtz's authority. Some of these theories qualify as "temporal" (e.g. those of Hurst or Bonnier [8]), others were essentially variations on the theme of a resonant cochlea.

Rutherford was aware that the maximum rates he observed in frog or rabbit nerves (352 per second) were insufficient to carry pitch over its full range (up to 4-5 kHz for musical pitch). The need for high firing rates was relaxed in 1930 by Wever and Bray's "volley theory" [35]. Subsequent measurements from the auditory nerve confirmed that the volley principle is essentially valid (in a stochastic form), in that synchrony to temporal features is measurable up to about 4-5 kHz in the auditory nerve [36]. Synchrony is also observed at more central neural relays, but the upper frequency limit decreases as one proceeds.

Temporal and resonance models differ essentially in the *time* required to make a frequency measurement. Resonance involves the build-up of energy by accumulation of successive waves, and this requires time that varies inversely with *frequency resolution*. The relation $\Delta F.\Delta t \geq k$ that constrains temporal and spectral resolution was formalized by Gábor [37], but it was known already to Helmholtz. Helmholtz reasoned that notes occur at a rate of up to 8 per second in music, and from this he calculated the narrowest possible bandwidth for cochlear filters. Frequency resolution was thus dictated by necessary temporal resolution, rather than by constraints related to the implementation of cochlear filters.

In contrast, a time-domain mechanism needs just enough time to measure the interval between two events (plus time to make sure that each event *is* an event, plus time to make sure that they are not both part of a larger pattern). The time required is on the order of *two periods* of the lowest expected frequency; accuracy is limited only by noise or imperfection of the implementation [8]. An explanation of the puzzling fact that a time-domain mechanism can escape Gábor's relation was given by Nordmark [38].

A weakness of temporal models, as described so far, is their reliance on *events*. Events need of course to be extracted from the waveform (or from the neural pattern that it evokes). For simple waveforms this is trivial: one may use peaks or zero-crossings. For complex waveforms the problem is more delicate, as evident from the difficulties encountered by time-domain methods of fundamental frequency estimation [30]. It is hard (perhaps impossible) to find a definition of "event" that allows stable period measurement in every case.

This weakness is evident in the *phase sensitivity* of early temporal models [28]. For example, a mechanism that measures intervals between peaks is confused by waveforms that have several peaks per period. A mechanism that measures intervals between *envelope* peaks is confused by phase manipulations that produce two envelope peaks per period, etc. As pitch is often invariant for such phase manipulations, such phase-sensitive mechanisms cannot hold. The autocorrelation model provided a solution to this problem.

5. Autocorrelation

In the *autocorrelation* (AC) model, each sample of the waveform is used, as it were, as an "event". Each is compared to every other sample, and the inter-event interval that gives the best match (on average) indicates the period. Concretely, comparison is performed by *multiplying* samples and summing the products over a time window. If samples are equal their products tend to be large, and so the *autocorrelation function* (ACF) has a peak at the period (and its multiples). The peak is the cue to pitch. A slightly more straightforward idea is to *subtract* samples and sum the squared differences, as proposed in the *cancellation* model of [39]. The cue to pitch is then a *dip* in the difference function. Cancellation and AC models are formally equivalent [39].

The original formulation of the AC model is due to Licklider [40], although an interesting precursor was proposed by Hurst in 1895 according to which a pulse travels up the basilar membrane, is reflected at the apex, travels down, and meets the next pulse at a position that indicates the period [41]. In Licklider's model, the ACF was calculated within the auditory nervous system, for each channel of the auditory filter bank. The model was reformulated and implemented computationally by Meddis and Hewitt [42], and confronted to autocorrelation statistics of actual nerve recordings by Cariani and Delgutte [43]. A similar model based on first order interspike interval statistics was proposed by Moore [44]. *Cancellation* was cited earlier. Another variant is the *strobed temporal integration* model of Patterson and colleagues, in which patterns are cross-correlated with a *strobe* function consisting of one pulse per period [45]. Yost proposed a simpler predictive model based on waveform autocorrelation [46]. One may cite also a number of "autocorrelation" models in which the ACF was produced by an internal "Fourier transformer" operating on a spectral profile coming from the cochlea [28].

An important theorem, the *Wiener-Khintchine theorem*, say that ACF and power spectrum are Fourier transforms one of the other. In this sense the AC model can be seen as an incarnation of the two steps of *spectral analysis* and *pattern matching*. This implies a relation between these rival approaches, as stressed early on by de Boer [26]. They differ of course in how they might be implemented in the auditory nervous system, in properties such as frequency versus temporal resolution (Sect. 4), and in the way they can be extended to handle mixtures of tones [47, 8].

It is interesting to compare autocorrelation to the *string* which we encountered several times in this review. Implementation of autocorrelation requires a *delay*, associated with a multiplier (e.g. a coincidence-detector neuron). Delayed patterns are *multiplied* with undelayed patterns. The string too consists of a delay that, as it were, feeds upon itself. Delayed patterns are *added* to undelayed patterns, and their sum delayed again. This shows a basic similarity between string and AC. It also shows their difference. In the AC model a pattern is delayed at most once. In the string it is delayed many times, and these multiple delays are necessary for the build-up of resonance that allows the string to be selective.

6. Discussion

Autocorrelation and pattern matching are the two major options for explaining pitch today. Pitch is evoked mainly by stimuli that are periodic, and its value depends on their *period*. The two approaches can be seen as two different ways of extracting the period from the stimulus. Autocorrelation does so directly, and pattern-matching indirectly via a first stage of Fourier transformation. The choice between them corresponds to that made by Ohm, a century and a half ago (Sect. 2.3).

Cochlear frequency resolution, as Helmholtz pointed out, must be limited. Filters are of roughly constant "Q", and thus have difficulty resolving upper harmonics, closely spaced on a logarithmic scale. Pattern-matching depends on frequency resolution, and cannot work for stimuli that contain only partials that are *unresolved*. Indeed, such stimuli tend to have a weak pitch, and this can be interpreted in favor of pattern-matching [8]. On the other hand, the pitch does exist and thus needs explaining, which pattern matching cannot do. This argues in favor of the AC model. The AC model could, in principle, cover both resolved and unresolved stimuli, but the marked behavioral differences between them suggest that there might instead be *two mechanisms* [48, 49, 50, 51]. The superior performance in pitch tasks for conditions in which partials are resolved is strongly suggestive of a pattern-matching mechanism, that breaks down for unresolved conditions. It might nevertheless be due to other factors that co vary with "resolution" [8]. The issue of resolved vs. unresolved is currently a central issue in pitch theory.

Pattern matching and AC models both have many variants. At times discussions may tend to focus on relatively minor differences between rival formulations (e.g. between Terhardt's vs. Goldstein's formulation of pattern-matching, or between first-order and all-order spike statistics for the AC model). The historical approach is useful to widen the perspective, to emphasize the similarities between variants, and possibly even to suggest new, perhaps radically different, directions in which to seek explanations of pitch and hearing.

This author is mainly interested (Licklider would have said "ego-involved" [52]) in a particular variant of the AC model, cancellation. The reason is that it brings together mechanisms of pitch and of *sound segregation* that may be of use in particular to explain perception of multiple pitch [53, 47]. An algorithm based on cancellation has recently proved to be effective for fundamental frequency estimation [54]. The concept of cancellation fits well with the ideas on redundancy and neural metabolism reduction of Barlow [55].

Interest in pitch is fueled by interest in music, a very old activity. Ideas for this review were searched for in sources as ancient and diverse as possible. There are big gaps. Many important sources are known only indirectly from citations of later authors, suggesting that much material of interest has been *lost*. Indeed, there is evidence that some of the knowledge that developed over the last 25 centuries was known long before that, in Sumer, Egypt, China, and possibly even South America. Sources consulted were exclusively in English or French. Those written in Latin, German or other languages were inaccessible for lack of linguistic competence. A more complete review is due to appear shortly [8].

7. Conclusions

The history of models of pitch perception has been reviewed. Modern ideas reincarnate older ideas, and their roots extend as far back as records are available. Models that are in competition today may have common roots. The historical approach allows commonalities and differences to be put in perspective. Hopefully this should help to defuse sterile controversy that is sometimes harmful to the progress of ideas [1]. It also may be of use to newcomers to the field to understand, say, why psychoacousticians insist on studying musical pitch with unresolved stimuli (that sound rather unmusical), why they add low pass noise (which makes tasks even more difficult), etc. The good reasons for these customs are easier to understand with a vision of the debates from which present-day pitch theory evolved.

8. References

[1] Boring EG (1942) Sensation and perception in the history of experimental psychology. New York: Appleton-Century.

[2] Cohen, H. F. (1984). ”Quantifying music,” Dordrecht, D. Reidel (Kluwer).

[3] von Békésy, G., and Rosenblith, W. A. (1948). ”The early history of hearing - observations and theories,” J. Acoust. Soc. Am. 20, 727-748.

[4] Lindsay, R. B. (1966). ”The story of acoustics,” J. Acoust. Soc. Am. 39, 629-644.

[5] Lindsay, R. B. (1973). ”Acoustics: historical and philosophical development,” Stroudsburg, Dowden, Hutchinson and Ross.

[6] Hunt, F. V. (1992 (original: 1978)). ”Origins in acoustics,” Woodbury, New York, Acoustical Society of America.

[7] von Helmholtz, H. (1877). ”On the sensations of tone (English translation A.J. Ellis, 1885, reprinted 1954),” New York, Dover.

[8] de Cheveigné, A. (2004). ”Pitch perception models,” in ”Pitch,” Edited by C. Plack and A. Oxenham, New York, Springer Verlag, in press.

[9] Macran, H. S. (1902). ”The harmonics of Aristoxenus (reprinted 1990, Georg Olms Verlag, Hildesheim),” Oxford, The Clarendon Press.

[10] Wiener, P. P. (1973-1974). ”The Dictionary of the History of Ideas: Studies of Selected Pivotal Ideas,” New York, Charles Scribner’s Sons.

[11] Galilei, G. (1638). ”Mathematical discourses concerning two new sciences relating to mechanicks and local motion, in four dialogues (translated by THO. WESTON, reprinted in Lindsay, 1973, pp 40-61),” London, Hooke.

[12] Mersenne, M. (1636). ”Harmonie Universelle (reproduit 1975, Paris: Editions du CNRS),” Paris, Cramoisy.

[13] Du Verney, J. G. (1683). ”Traité de l’organe de l’ouie, contenant la structure, les usages et les maladies de toutes les parties de l’oreille,” Paris.

[14] Le Cat, C.-N. (1758). ”La Théorie de l’ouie: supplément à cet article du traité des sens,” Paris, Vallat-la-Chapelle.

[15] Sauveur, J. (1701). ”Système général des intervalles du son (translated by R.B. Lindsay as ”General system of sound intervals and its application to sounds of all systems and all musical instruments”, reprinted in Lindsay, 1973, pp 88-94),” Mmoires de l’Acadmie Royale des Sciences 279-300, 347-354.

[16] Fourier, J. B. J. (1820). ”Traité analytique de la chaleur,” Paris, Didot.

[17] Ohm, G. S. (1843). ”On the definition of a tone with the associated theory of the siren and similar sound producing devices (translated by Lindsay, reprinted in Lindsay, 1973, pp 242-247,” Poggendorf’s Annalen der Physik und Chemie 59, 497ff.

[18] Young, T. (1800). ”Outlines of experiments and inquiries respecting sound and light,” Phil. Trans. of the Royal Society of London 90, 106-150 (plus plates).

[19] Schouten, J. F. (1970). ”The residue revisited,” in ”Frequency analysis and periodicity detection in hearing,” Edited by R. Plomp and G. F. Smoorenburg, London, Sijthoff, 41-58.

[20] de Boer, E. (1976). ”On the ”residue” and auditory pitch perception,” in ”Handbook of sensory physiology, vol V-3,” Edited by W. D. Keidel and W. D. Neff, Berlin, Springer-Verlag, 479-583.

[21] Turner, R. S. (1977). ”The Ohm-Seebeck dispute, Hermann von Helmholtz, and the origins of physiological acoustics,” The Britsh Journal for the History of Science 10, 1-24.

[22] Schouten, J. F. (1938). ”The perception of subjective tones,” Proc. Kon. Acad. Wetensch (Neth.) 41, 1086-1094 [reprinted in Schubert Schubert, E. D. (1979). ”Psychological acoustics (Benchmark papers in Acoustics, v 13),” Stroudsburg, Pennsylvania, Dowden, Hutchinson & Ross, Inc].

[23] Licklider, J. C. R. (1954). ”Periodicity pitch and place pitch,” J. Acoust. Soc. Am. 26, 945.

[24] Warren, R. M., and Warren, R. P. (1968). ”Helmholtz on perception: its physiology and development,” New York, Wiley.

[25] Thurlow, W. R. (1963). ”Perception of low auditory pitch: a multicue mediation theory,” Psychol. Rev. 70, 461-470.

[26] de Boer, E. (1956), ”On the ”residue” in hearing,” unpublished doctoral dissertation.

[27] Goldstein, J. L. (1973). ”An optimum processor theory for the central formation of the pitch of complex tones,” J. Acoust. Soc. Am. 54, 1496-1516.

[28] Wightman, F. L. (1973). ”The pattern-transformation model of pitch,” J. Acoust. Soc. Am. 54, 407-416.

[29] Terhardt, E. (1974). ”Pitch, consonance and harmony,” J. Acoust. Soc. Am. 55, 1061-1069.

[30] Hess, W. (1983). ”Pitch determination of speech signals,” Berlin, Springer-Verlag.

[31] Shamma, S., and Klein, D. (2000). ”The case of the missing pitch templates: how harmonic templates emerge in the early auditory system,” J. Acoust. Soc. Am. 107, 2631-2644.

[32] Bower, C. M. (1989). ”Fundamentals of Music (translation of De Institutione Musica, Anicius Manlius Severinus Boethius, d524),” New Haven, Yale University Press.

[33] Rutherford, E. (1886). ”A new theory of hearing,” J. Anat. Physiol. 21, 166-168.

[34] Wever, E. G. (1949). ”Theory of hearing,” New York, Dover.

[35] Wever, E. G., and Bray, C. W. (1930). ”The nature of acoustic response: the relation between sound frequency and frequency of impulses in the auditory nerve,” Journal of experimental psychology 13, 373-387.

[36] Johnson, D. H. (1980). ”The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones,” J. Acoust. Soc. Am. 68, 1115-1122.

[37] Gábor, D. (1947). ”Acoustical quanta and the theory of hearing,” Nature 159, 591-594.

[38] Nordmark, J. O. (1968). ”Mechanisms of frequency discrimination,” J. Acoust. Soc. Am. 44, 1533-1540.

[39] de Cheveigné, A. (1998). ”Cancellation model of pitch perception,” J. Acoust. Soc. Am. 103, 1261-1271.

[40] Licklider, J. C. R. (1951). ”A duplex theory of pitch perception,” Experientia 7, 128-134.

[41] Hurst, C. H. (1895). ”A new theory of hearing,” Proc. Trans. Liverpool Biol. Soc. 9, 321-353 (and plate XX).

[42] Meddis, R., and Hewitt, M. J. (1991). ”Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification,” J. Acoust. Soc. Am. 89, 2866-2882.

[43] Cariani, P. A., and Delgutte, B. (1996). ”Neural correlates of the pitch of complex tones. I. Pitch and pitch salience,” J. Neurophysiol. 76, 1698-1716.

[44] Moore, B. C. J. (1977, 2003). ”An introduction to the psychology of hearing,” London, Academic Press.

[45] Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (1992). ”Complex sounds and auditory images,” in ”Auditory physiology and perception,” Edited by Y. Cazals, K. Horner and L. Demany, Oxford, Pergamon Press, 429-446.

[46] Yost, W. A. (1996). ”Pitch strength of iterated rippled noise,” J. Acoust. Soc. Am. 100, 3329-3335.

[47] de Cheveigné, A., and Kawahara, H. (1999). ”Multiple period estimation and pitch perception model,” Speech Communication 27, 175-185.

[48] Houtsma, A. J. M., and Smurzynski, J. (1990). ”Pitch identification and discrimination for complex tones with many harmonics,” J. Acoust. Soc. Am. 87, 304-310.

[49] Carlyon, R. P., and Shackleton, T. M. (1994). ”Comparing the fundamental frequencies of resolved and unresolved harmonics: evidence for two pitch mechanisms?,” J. Acoust. Soc. Am. 95, 3541-3554.

[50] Carlyon, R. P. (1996). ”Masker asynchrony impairs the fundamental-frequency discrimination of unresolved harmonics,” J. Acoust. Soc. Am. 99, 525-533.

[51] Carlyon, R. P. (1998). ”Comments on ”A unitary model of pitch perception” [J. Acoust. Soc. Am. 102, 1811-1820 (1997)],” J. Acoust. Soc. Am. 104, 1118-1121.

[52] Licklider, J. C. R. (1959). ”Three auditory theories,” in ”Psychology, a study of a science,” Edited by S. Koch, New York, McGraw-Hill, I, 41-144.

[53] de Cheveigné, A. (1993). ”Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing,” J. Acoust. Soc. Am. 93, 3271-3290.

[54] de Cheveigné, A., and Kawahara, H. (2002). ”YIN, a fundamental frequency estimator for speech and music,” J. Acoust. Soc. Am. 111, 1917-1930.

[55] Barlow, H. B. (1961). ”Possible principles underlying the transformations of sensory messages,” in ”Sensory Communication,” Edited by W. A. Rosenblith, Cambridge Mass, MIT Press, 217-234.