Time domain processing in the auditory system

Alain de Cheveigné
CNRS/Université Paris 7, and ATR-HIP

ABSTRACT

This paper reviews models of auditory signal processing that operate in the *time domain*. This is in contrast with the traditional view that the cochlea performs a Fourier transformation, and that subsequent processing is carried out in the frequency domain. We classify models into parameter *estimation*, channel *labelling*, and within-channel *filtering* models. Phenomena accounted for are sound localization, pitch perception, vowel timbre identification, and various aspects of sound segregation ("cocktail-party") effects. Most models may be implemented based on a time-domain stochastic neural representation (spike trains) such as found in the auditory nerve. Processing involves "neural filters" based on delay lines and gating neurons (coincidence counters), that select or delete certain spikes, thereby modifying the statistics of the spike train in a way that strengthens or weakens the representation of parts of the sound environment. In particular "cancellation filters", in which the interaction is mainly inhibitory, are suggested as a likely ingredient for models of auditory processing.

Introduction

The cochlea resembles a bank of filters covering the range of audible frequencies (about 20 Hz to 20 kHz in man), and is often viewed as playing essentially the role of a Fourier transformer: the slowly-varying spectrum from the cochlea is coded as an average discharge rate within auditory nerve fibers, and handed to the auditory nervous system for subsequent processing. Logically, such processing should occur in the *frequency* domain. This idea, that dates back to Helmholtz (1877), still dominates much of Hearing Science.

However there is evidence that some aspects of auditory processing within the nervous system may be carried out in the *time* domain. Unfortunately, time-domain models of auditory processing are less well developed. The purpose of this paper is to review a series of time-domain models that account quite well for many aspects of observed psychophysical performance, and that require physiological assumptions that are not unreasonable.

Most of these models can be implemented using two basic ingredients: *delay lines* and *gating neurons*. The latter are sensitive to the interaction of neural discharges on their inputs, and particularly to their relative timing. Interaction is excitatory or inhibitory: the occurrence of a spike on one pathway allows, or on the contrary disallows, the transmission of a spike on the other pathway.

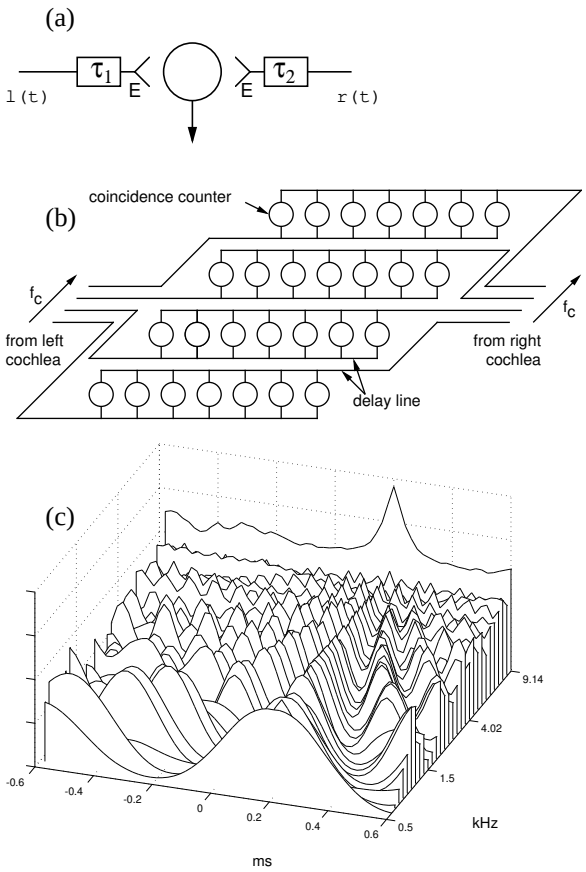

Fig. 1 *Jeffress (1948) model of auditory localization. (a) Elementary delay line and coincidence counting neuron. (b) Array of coincidence counters organized tonotopically and by interaural delay (correlation lag). (c) Pattern evoked by a wideband sound source displaced from the axis. A ridge occurs at the internal delay that compensates for the external interaural delay of 0.2 ms. The topmost curve represents the sum of cross-correlation functions over frequency ("summary cross-correlation function").*

1 Estimation: location, pitch and timbre

1.1 Jeffress's localization model

The "cross-correlation" model of auditory localization was imagined by Jeffress (1948). The circuit of Fig. 1(a) has two inputs that are delayed by different amounts before converging

on a neuron that acts as a coincidence counter. The neuron responds when nerve impulses ("spikes") arrive simultaneously at both synapses, within a time window of coincidence that is short relative to the periods of important stimulus components. Assuming statistical independence between spike-generating processes, the output probability density is proportional to the *product* of the input densities. The model includes an array of such circuits, with delay parameters τ_1 and τ_2 such that $\tau = \tau_1 - \tau_2$ covers the range of possible interaural delays (about $\pm 0.8ms$ in man). This array comes close to calculating the running cross-correlation function of input spike densities:

$$CCF_t(\tau) \quad = \quad \int_{-\infty}^{t} w(\theta - t) l(\theta) r(\theta - \tau) d\theta$$

where w is a windowing function and $l(t)$ and $r(t)$ are spike densities at left and right ears. Every frequency channel has its own array, and together they form a two-dimensional array indexed by frequency (inherited from cochlear filtering) and interaural delay [Fig. 1(b)]. In response to a wide-band source localized to one side of auditory space, a ridge of activity arises where the internal (neural) delay compensates for the external (acoustic) interaural delay [Fig. 1(c)].

Jeffress's model remained speculative for many years before physiological and anatomical evidence was found to support it in the MSO and IC (where MSO projects) of mammals and equivalent centres in birds (Carney et al., 1989; Chan et al. 1989; Konishi et al., 1988; Yin et al., 1984, 1987, 1990). Rarely was a model more prescient.

Note that Jeffress's model would work as well if excitatory-excitatory (EE) interaction were replaced by inhibitory-excitatory (IE) interaction, and the search for a *maximum* by a search for a *minimum*. We come back to this idea later on in the paper.

1.2 Licklider's pitch model

Licklider (1951, 1959) imagined a similar model to explain pitch. As in Jeffress's model, this model involves a 2-D array of delay lines and coincidence counting neurons. Instead of coming from separate ears, inputs originate from the same ear, and the circuit thus calculates the equivalent of a running autocorrelation function within each channel:

$$ACF_t(\tau) \quad = \quad \int_{-\infty}^{t} w(\theta - t) s(\theta) s(\theta - \tau) d\theta$$

Instead of being sensitive to correlation between ears, the model is sensitive to correlations within the stimulus itself, as arise when the stimulus is periodic. In response to a periodic sound, a ridge appears within the two-dimensional pattern (filter channel frequency X autocorrelation delay) at a delay equal to the period of the stimulus. This is the cue to the pitch of the stimulus in Licklider's model.

The model has been "reinvented" in various forms since it was first proposed (see de Cheveigné, 1998 for a review). Cariani et al. (1996a,b) found physiological support for it, in that autocorrelation histograms of auditory-nerve fiber discharges correlate well with the psychophysics of pitch. However, contrary to Jeffress's model, no as yet no one has yet located the neural centers where the processing might occur.

1.3 Meddis and Hewitt's vowel timbre model

Time-domain processing is usually invoked for localization, commonly for pitch, but rarely for *timbre*. The timbre of steady-state sounds is usually attributed to spectral cues extracted in the cochlea and exploited centrally (possibly by pattern matching). Nevertheless, Meddis and Hewitt (1992) proposed a vowel identification model in which the 2D autocorrelation pattern was summed over frequency, to obtain a "summary autocorrelation function". Identification then involved pattern-matching of the "short-lag" portion of this summary autocorrelation function to stored templates. A similar scheme was used with success by de Cheveigné (1997). The fact that ACF patterns support vowel identification is not completely a surprise. The spectral envelope of a vowel, characteristic of its timbre, may be estimated by LPC analysis based on the low-order coefficients of the short-term autocorrelation (or autocovariance) function of the vowel's waveform (Rabiner and Schafer, 1978). The summary ACF resembles the waveform ACF.

1.4 Cochlear filtering: what for?

The three models of this section postulated cochlear filtering as a first step, but did not exploit it. If anything the filter bank seems a handicap, as the same time-domain operations must be reproduced in multiple channels. Cochlear filtering comes to its own in the spectro-temporal models of the next section, as a basis for *segregation* of multiple sources.

2 Channel selection and labeling

2.1 Lyon's binaural segregation model

Jeffress's model was designed to give the azimuth of a *single* source, but can be extended to give the positions of several sources. As long as both sources occupy distinct portions of the spectrum, they may produce distinct "ridges" in the 2D cross-correlation pattern, within the limits of the channels that they dominate. Indeed, the fact that ridges occur at different places in different channels can be used to *label* each channel as belonging to a source. Lyon's (1983-1988) binaural sound-segregation system was based on this idea. Lyon's model was the first of a series of *spectro-temporal* models, in which temporal information is used to label portions of a spectral representation.

2.2 Weintraub's voice segregation scheme

Weintraub (1985) built a system directly inspired from Lyon's ideas, but using periodicity information rather than binaural information. A channel was labelled as belonging to a voice if it was dominated by the periodicity of that voice. Meddis and Hewitt (1992) built a model of concurrent vowel perception based on the same idea. An originality of Meddis and Hewitt's model is that it requires the periodicity of only *one* voice (the stronger

one), whereas Weintraub's and other models usually assume that the periodicity (F_0) of both must be estimated.

2.3 Computational Auditory Scene Analysis

Spectro-temporal channel selection is the basis for a variety of CASA (computation auditory scene analysis) models (Cooke, 1991; Brown, 1992; Ellis, 1996). In addition to binaural or periodicity cues, channels may be labelled to other scene analysis cues such as "common fate", top-down heuristics, etc.. To the extent that these models employ realistic "cochlear" filter banks, and that channel labeling is based on time-domain processing, these models can loosely be included with the time-domain models of auditory processing that are the subject of this paper.

2.4 What to do with shared channels?

In a spectro-temporal model, channels are "atomic": information is divided among channels, but cannot be split further. However, cochlear frequency resolution is not perfect, and so some channels may actually be responding to several sources at the same time. Models diverge in their strategy at this point: (a) assign the channel to one source or the other, (b) assign it to both, (c) assign it to neither, (d) try to split it. Strategy (d) implies *within-channel* time-domain processing. Models of within-channel processing are described in the next section.

3 Within-channel filtering

3.1 Durlach's EC model

The phenomenon of binaural release from masking is well known. The detection threshold of a target signal in noise is *lower* (i.e. masking is less effective) when target and masker have different interaural phase relations. To explain this phenomenon, Durlach (1963) imagined a model in which signals at both ears were *equalized* by scaling and delaying one relative to the other, and then subtracted one from the other to *cancel* the masker. Durlach worked with signals that were narrowband, for which it made no difference whether processing occurred on the raw signal or within a single auditory filter channel. The model may be extended trivially to the wide-band target case by assuming identical processing within channels, with the same parameters (scale and delay) in each.

When the signal-to-noise ratio (SNR) is small, a convenient way to estimate parameters is to search the parameter space for values that produce the minimum cancellation residual. This can be done by a neural circuit similar to that used in Jeffress's model, but involving inhibitory rather than excitatory interaction [Fig. 2(a)]. In response to a source localized to one side, a *valley* appears in the 2D array. The valley serves as a cue to the azimuth of the source, as did the ridge in Jeffress's model. Physiological support for this model has been found in the LSO (Joris and Yin, 1998, and references therein).

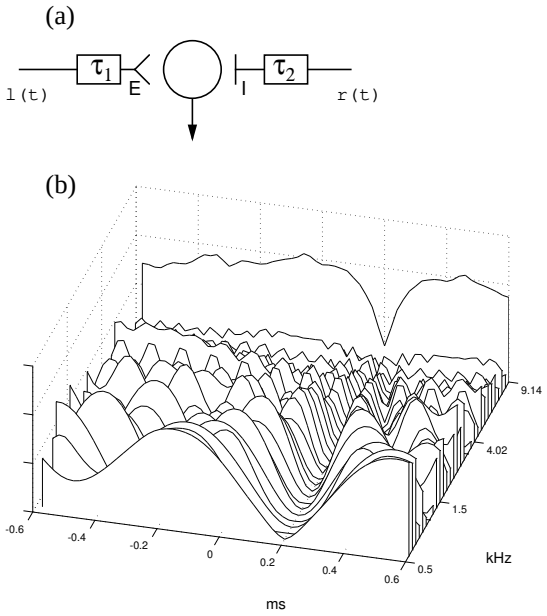

Fig. 2 *Cancellation model of auditory localization, formally equivalent to Jeffress's autocorrelation model. (a) Elementary delay and inhibitory coincidence counting neuron. (b) Pattern evoked by a wideband source with a 0.2 ms interaural time difference.*

Residual activity within the valley (if any) may then be exploited as evidence of a second, weaker source. Such residual activity is usually thought of as a slowly-varying pattern of activity across frequency channels, to be processed in the frequency domain. However nothing in principle would prevent it from being exploited in the time domain. This opens up the perspective of successive stages of time-domain processing, each operating on the residual output of preceding stages.

3.2 Culling and Summerfield's modified EC model

Culling and Summerfield (1995) found that Durlach's EC model does not describe subject's performance in the situation where two concurrent vowels must be segregated on the basis of binaural disparity. More precisely, they found no evidence that the *same* equalization parameters were being applied to different frequency channels. Instead they found evidence of independent processing within channels, with different parameters for each channel. They proposed a "modified EC model", in which parameters are found by searching for a minimum residual, as in Durlach's model, but *independently* for each channel.

Psychoacoustics show that this model is closer to what is going on in the auditory system than Durlach's model, but it poses a rather sticky problem. If processing is independent within each channel, on what basis are channels grouped and assigned to sources? A possible answer is that they are *not*, but that the auditory system makes do with this rather strange "two-valued" pattern. For each channel there are two values: the "raw" value before cancellation, and the residual after cancellation. The first reflects the stronger source *within that channel*,

and the second the weaker source. This information is of course less useful than if channels were unambiguously assigned to sources, but possibly more than the unprocessed, single-valued pattern.

3.3 Cancellation model of vowel segregation

Vowels mixed together are easier to identify when they have different F_0s. The effectiveness of the harmonicity cue, exploited when F_0s are different, has been observed in a large number of so-called "double-vowel" experiments. Various models have been proposed to explain it, among them the channel-selection model of Meddis and Hewitt (1992) mentioned above. Meddis and Hewitt's model partitions the set of peripheral channels into channels dominated by one vowel, and those dominated by the other. It is quite successful for mixtures of equal-amplitude vowels, that each dominate the spectrum near their own formants. However it fails if one vowel is stronger than the other and dominates *all* channels. Segregation effects are nonetheless observed in that case, that Meddis and Hewitt's model cannot explain (de Cheveigné et al., 1997). To explain such effects, one must admit that F_0-guided segregation can occur *within* channels, and not just between channels.

Within-channel segregation can be performed in the time domain, with a "cancellation filter" similar to that proposed for Durlach's EC model but operating on spike train coming from the same rather than different ears (Fig. 3). The filter removes any spike that forms with a preceding spike an interval equal to its delay parameter, τ. Tuned to the period of an interfering sound, the filter modifies the interval statistics in a way that suppresses the representation of that sound in favor of the weaker target. The model has been tested with data recorded in the auditory nerve (de Cheveigné, 1993), and it accounts for psychoacoustic data quite well (de Cheveigné, 1997).

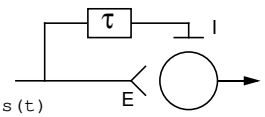

Fig. 3 *Neural cancellation filter for harmonic segregation.*

3.4 Cancellation model of pitch perception

The *multiplication* operation that appears in the autocorrelation function of Licklider's pitch model can be replaced by *subtraction*, and the search for a maximum by the search for a minimum. Subtraction might be implemented by an inhibitory gating or coincidence-counting neuron such as used in the cancellation model of vowel segregation (Fig. 3). To a first approximation, cancellation and autocorrelation pitch models have equivalent behaviors (de Cheveigné, 1998), but they differ in their physiological "signature", and in certain second-order properties such as phase-sensitivity.

3.5 Multiple period estimation

An attractive feature of the cancellation model of pitch perception is that it can be generalized to estimate multiple periods. This is done by searching the parameter space of a *cascade* of cancellation filters (de Cheveigné and Kawahara, 1997, 1998). The search can be sequential: after the period of one sound has been estimated, it is cancelled, and the others are estimated from the cancellation residual. It can also be exhaustive: the multiple dimension parameter space of cascaded filters is searched exhaustively for a minimum output. This model may help explain how, for example, an orchestra conductor can hear out the notes of individual instruments, among many playing simultaneously in the orchestra.

3.6 Ringing suppression

A problem that confronts time-domain models of auditory processing is that of *ringing* of auditory filters. For transient sounds the ringing often dominates the time-domain response of the channel, but it has little to say about the signal. Several auditory models incorporate schemes to reduce the effects of ringing: spectrotemporal adaptation, cross-channel summation, subtraction of the response to noise, etc.. The cancellation filter of Fig. 3, tuned to the ringing period of the channel, is also quite effective in removing the effects of ringing (de Cheveigné, 1998).

4 Discussion

Our review presented a panorama of time-domain auditory processing models. Some are quite old (Jeffress, Licklider, Durlach), others more recent. In some cases (localization, binaural interaction) it is difficult to imagine processing occurring other than in the time domain. In others (pitch, timbre, F_0-guided segregation) the time-domain models are under strong competition from healthy frequency domain ("place") models. However even in these cases many aspects are best explained in the time domain. Delgutte (1984) had proposed a variety of time-domain schemes of processing in the auditory system, as alternatives to frequency-domain processing. Our review is certainly incomplete and biased in favor of our own work on cancellation models of F_0-guided segregation and pitch.

4.1 The issue of delay lines

Most models can be implemented physiologically given two ingredients: delay lines and gating neurons (excitatory or inhibitory). The availability of gating neurons does not seem to be an issue. Short delay lines (on the order of naturally occurring interaural delays) have also been found. However there is as yet little evidence for delay lines that: (a) have delays on the order of pitch periods (10 or 20 ms), (b) are temporally precise, and (c) are "wide-band" (carrying spike trains similar to primary auditory nerve fibers). The requirement (b) of temporal precision implies specialized neural structures, that have been found in pathways that converge on the LSO (Joris and Yin, 1998) but with relatively short delays.

Requirement (c) is that the delay line carries (and thus memorizes) the detailed temporal structure over a relatively long duration. This is a quite severe constraint. A delay line that carries a *single pulse* per period might be easier to implement, as it needs to "memorize" less information. This consideration might constitute an argument in favor of the "strobed auditory integration" (SAI) model of Patterson et al. (1992), that can be interpreted as a cross-correlation between a nerve fiber firing probability pattern and a single "strobe" for each period. The single strobe might be easier to delay accurately than a detailed pattern. Processing might for example be implemented by an array of coincidence counters (as in Licklider's model), on which converge both undelayed primary or primary-like fibers, and delayed fibers carrying a single pulse (or small number of pulses) per period. The array would calculate the function:

$$SAI_t(\tau) \quad = \quad \int_{-\infty}^{t} w(\theta - t)s(\theta)S(\theta - \tau)d\theta$$

where w is a window function, $s(t)$ is the undelayed auditory nerve fiber firing density, and $S(t)$ is the delayed strobe function. This is slightly different from the original proposal of Patterson et al., of a static buffer into which the temporal firing pattern was loaded on each strobe.

4.2 The role of peripheral filtering

Several time-domain models take little or no advantage of peripheral filtering. Jeffress's localization model would work if waveforms from both ears were correlated before filtering rather than after. The same can be said of Licklider's pitch model, Durlach's binaural model, de Cheveigné's cancellation filter, etc.. On first view, filtering seems to make things harder, because the same processing needs to be repeated in all filtered channels. Exquisite mechanical selectivity is a characteristic of cochleas of many species, and it *must* be of use: that a model might do better without cochlear filtering can almost be interpreted as an argument *against* that model.

The paradox may be resolved by the following argument. Peripheral filtering may not be necessary for the *principle* of time-domain models, but it may be essential for their successful implementation. Mechanical-to-neural transduction is not linear, and for this and other reasons neural signal processing is likely to have a limited dynamic range. Parallel processing of band-limited channels, each with its own frequency and phase characteristics, may offer a way to compensate for these defects. Cochlear filtering may thus be seen as the "last linear stage" (Møller, 1983) before time-domain processing takes over within the auditory nervous system.

4.3 The cancellation principle

Several models were based on the principle that the auditory system attempts to *cancel* correlates of a source. Cancellation constitutes a way of estimating the source's characteristics. It also has the effect of "clearing" the internal representation, so that the emergence of new sources is easier to spot. Scene analysis might thus proceed via a succession of estimate-cancel-estimate steps, as elements of the auditory scene are peeled away one after the other (Nakatani et al, 1995). The cancellation principle fits well with recent physiological evidence of inhibitory processing of interaural time and level differences in the LSO (Batra et al., 1997; Joris et al. 1998, Tsuchitani, 1997). It also fits well with the recent idea of *blind separation* of concurrent sources. Cancellation of a source produces a residual that is uncorrelated with that source. Given N sources, cancellation of all combinations N-1 sources would produce N uncorrelated outputs. Blind separation (and related techniques) can be seen as a principled way of estimating the parameters necessary for cancellation.

Conclusion

Models of time-domain auditory processing were reviewed, including recent models based on the principle of cancellation. The models account well for observed performance on psychoacoustic tasks, and their physiological assumptions are relatively easy to satisfy (although direct physiological confirmation is in many cases lacking). Time-domain processing within the auditory nervous system is a reasonable alternative to the more common assumption of frequency-domain processing of slowly-varying spectral patterns, produced in the cochlea and carried by a tonotopic rate code. Auditory processing models provide interesting examples of time-domain processing of the stochastic fine structure of neural spike trains.

References

- Batra, R., Kuwada, S., and Fitzpatrick, D. C. (1997). "Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. Coincidence detection," J. Neurophysiol. 78, 1237-1247.

- Brown, G. J. (1992), "Computational auditory scene analysis: a representational approach," Sheffield, Department of Computer Science unpublished doctoral dissertation.

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

- Cariani, P. A., and Delgutte, B. (1996b). "Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase-invariance, pitch circularity, rate-pitch and the dominance region for pitch," J. Neurophysiol. 76, 1717-1734.

- Carney, L. H., and Yin, T. C. T. (1989). "Responses of low-frequency cells in the inferior colliculus to interaural time differences of clicks: excitatory and inhibitory components," J. Neurophysiol. 62, 144-161.

- Chan, J. C. K., Yin, T. C. T., and Musicant, A. D. (1987). "Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. II. Responses to band-pass filtered noises," J. Neurophysiol. 58, 543-561.
- Cooke, M. P. (1991), "Modeling auditory processing and organisation," Sheffield, Department of Computer Science unpublished doctoral dissertation.
- Culling, J. F., and Summerfield, Q. (1995). "Perceptual segregation of concurrent speech sounds: absence of across-frequency grouping by common interaural delay.," J. Acoust. Soc. Am. 98, 785-797.
- 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.
- de Cheveigné, A. (1997). "Concurrent vowel segregation III: A neural model of harmonic interference cancellation," 101, 2857-2865.
- de Cheveigné, A., Kawahara, H., Tsuzaki, M., and Aikawa, K. (1997). "Concurrent vowel identification I: Effects of relative level and F0 difference," J. Acoust. Soc. Am. 101, 2839-2847.
- de Cheveigné, A., and Kawahara, H. (1997). "Modeling the perception of multiple pitches.", Proc. IJCAI workshop on Computational Auditory Scene Analysis, Nagoya.
- de Cheveigné, A. (1998). "Cancellation model of pitch perception," J. Acoust. Soc. Am. 103, 1261-1271.
- de Cheveigné, A., and Kawahara, H. (1998). "Multiple period estimation and pitch perception model," Speech Communication, submitted.
- Delgutte, B. (1984). "Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds," J. Acoust. Soc. Am. 75, 879-886.
- Durlach, N. I. (1963). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. 35, 1206-1218.
- Ellis, D. (1996), "Prediction-driven computational auditory scene analysis," MIT unpublished doctoral dissertation.
- Helmholtz, H. v. (1877). "On the sensations of tone (English translation A.J. Ellis, 1954)," New York, Dover.
- Jeffress, L. A. (1948). "A place theory of sound localization," J. Comp. Physiol. Psychol. 41, 35-39.
- Joris, P. X., and Yin, T. C. T. (1998). "Envelope coding in the lateral superior olive. III. Comparison with afferent pathways," J. Neurophysiol. 79, 253-269.
- Konishi, M., Takahashi, T. T., Wagner, H., Sullivan, W. E., and Carr, C. E. (1988). "Neurophysiological and anatomical substrates of sound localization in the owl," in "Auditory function - neurobiological bases of hearing," Edited by G. M. Edelman, W. E. Gall and W. M. Cowan, New York, Wiley, 721-745.
- Lyon, R. F. (1983-1988). "A computational model of binaural localization and separation," in "Natural computation," Edited by W. Richards, Cambridge, Mass, MIT Press, 319-327.
- 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.
- Licklider, J. C. R. (1951). "A duplex theory of pitch perception," Experientia 7, 128-134.
- Nakatani, T., Okuno, H. G., and Kawabata, T. (1995). "Residue-driven architecture for computational auditory scene analysis.", Proc. IJCAI, 165-172.
- 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.
- Rabiner, L. R., and Schafer, R. W. (1978). "Digital processing of speech signals," Englewood Cliffs, NJ, Prentice-Hall.
- Meddis, R., and Hewitt, M. J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," J. Acoust. Soc. Am. 91, 233-245.
- Møller, A. R. (1983). "Auditory physiology," New York, Academic Press.
- Tsuchitani, C. (1997). "Input from the medial nucleuas of the trapezoid body to an interaural level detector," HR 105, 211-224.
- Wein85 Weintraub, M. (1985), "A theory and computational model of auditory monaural sound separation," University of Stanford unpublished doctoral dissertation.
- Yin, T. C. T., and Chan, J. C. K. (1990). "Interaural time sensitivity in medial superior olive of cat," J. Neurophysiol. 64, 465-488.
- Yin, T. C. T., Chan, J. C. K., and Carney, L. H. (1987). "Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation," J. Neurophysiol. 58, 562-583.
- Yin, T. C. T., Kuwada, S., and Sujaku, Y. (1984). "Interaural time sensitivity of high-frequency neurons in the inferior colliculus," J. Acoust. Soc. Am. 76, 1401-1410.