Hearing Research 174 (2002) 101–106

HEARING RESEARCH

www.elsevier.com/locate/heares

Temporal pitch in electric hearing

Fan-Gang Zeng *

Departments of Otolaryngology, Anatomy and Neurobiology, Biomedical Engineering and Cognitive Sciences, University of California, Irvine, CA 92697, USA

Received 1 July 2002; accepted 15 August 2002

Abstract

Both place and temporal codes in the peripheral auditory system contain pitch information, however, their actual use by the brain is unclear. Here pitch data are reported from users of the cochlear implant, which provides the ability to change the temporal code independently from the place code. With fixed electrode stimulation, both frequency discrimination and pitch estimate data show that the cochlear implant users can only discern differences in pitch for frequencies up to about 300 Hz. An integration model can predict pitch estimation from frequency discrimination, reinforcing Fechner's hypothesis relating sensation magnitude to stimulus discriminability. The present results suggest that 300 Hz is the upper boundary of the temporal code and that the absolute place information should be included in the present pitch models. They further suggest that future cochlear implants need to increase the number of independent electrodes to restore normal pitch range and resolution.

Key words: Pitch; Temporal pitch; Frequency discrimination; Electric hearing; Cochlear implant; Fechner's model

1. Introduction

Pitch is a fundamental attribute of auditory perception allowing sounds to be ordered from low to high (Stevens, 1954; ANSI, 1973). Neurophysiological studies in animals have shown that pitch can be encoded by either the place of excitation as a function of stimulus frequency along the cochlea, i.e. the 'place code' (Johnstone and Boyle, 1967; Rhode, 1978) or the temporal structure of discharges that is phase-locked to the frequency, i.e. the 'temporal code' (Kiang et al., 1965; Rose et al., 1967). Because the place and temporal codes usually co-vary with stimulus frequency in acoustic hearing, their relative contribution to pitch perception has been a topic of continuous debate for over 100 years but still remains unsettled (Helmholtz, 1877; Wever, 1948; Licklider, 1951; Evans, 1978; Srulovicz and Goldstein, 1983; Young and Sachs, 1979; Chatterjee and Zwislocki, 1997; Moller, 1999).

Burns and Viemeister (1976) used sinusoids of different frequencies to amplitude-modulated white noise (SAM noise) to study the temporal coding of pitch in acoustic hearing. The SAM noise has a sinusoidal temporal envelope but a noise spectrum, forcing listeners to rely on the temporal envelope cue to perform pitch-related tasks. Their result showed that listeners could derive pitch by following the temporal envelope frequency to about 500 Hz. Similarly, Formby (1985) found that the difference limen in modulation frequency of the SAM noise increases dramatically from about 3 to 122 Hz for modulation frequency at 80 and 400 Hz, respectively. This large increase in modulation frequency discrimination contrasts greatly the 1–2 Hz difference in tonal frequency discrimination within the same 80- to 400-Hz range. The results from the SAM noise experiments imply that the upper boundary of temporal pitch is 300–500 Hz. However, this implication has been controversial because SAM noise only elicits weak pitch sensations in normal-hearing listeners.

Here we use cochlear implants to delineate the relative contribution of place and temporal cues to pitch perception. The cochlear implant directly stimulates the auditory nerve through a fixed electrode array inserted into the cochlea, providing a unique opportunity to

---

* Tel.: +1 (949) 824-1539; Fax: +1 (949) 824-5907.
E-mail address: fzeng@uci.edu (F.-G. Zeng).

PII: S0378-5955(02)00644-5

study the temporal code independent of the place code. We systematically changed the stimulus frequency on a single electrode pair and measured the frequency difference limen and pitch–frequency function in four cochlear implant users. We found that these implant users could only detect differences in pitch for frequencies up to roughly 300 Hz. Correspondingly, they could perceive increased pitch with frequency until the same 300-Hz frequency was reached. These results suggest that 300 Hz is the upper boundary of the temporal code for pitch perception.

2. Materials and methods

2.1. Subjects

Four post-lingually deafened adults (one female and three males) participated in this study. Three subjects were cochlear implant users of the Ineraid device (Eddington et al., 1978) and one subject (JM) was a Nucleus-22 user (Clark et al., 1987). The subjects had used the implant for 5–20 years and were average to 'star' users (from occasional to regular telephone use). They all had extensive experience in psychophysical and speech tests. An informed consent was approved by a local Institutional Review Board and obtained from each individual subject after the nature and possible consequences of the study were explained.

2.2. Stimuli

Biphasic (100 or 200 μs/phase) pulses of different frequencies were used in the experiments. All stimuli were 300 ms in duration and delivered through a constant current source to the most apical or basal electrode. The distance between the most apical and the most basal was about 20 mm for Ineraid users (AM and MK), was undetermined for the third Ineraid user (DC) (see fig. 5A in Eddington et al., 1978), and was 12.75 mm (17×0.75 mm) for the Nucleus user. The reference electrode was outside the cochlea in Ineraid users and adjacent (BP+1) to the stimulating electrode in the Nucleus user. Customized research interfaces were used to deliver the electric stimulation (for Ineraid devices, see Zeng and Shannon, 1994; for Nucleus devices see Shannon, 1992).

2.3. Procedure

A balancing procedure was used to measure the comfortable loudness level for all stimulus frequencies and electrodes (Zeng and Shannon, 1994). All stimuli were presented at the same comfortable loudness level. A two-down, one-up, three-alternative, forced-choice, adaptive procedure was then used to measure frequency discrimination (Levitt, 1971). An absolute magnitude estimation procedure was used to estimate the numerical value of the pitch for all stimulus frequencies and electrodes (Zwislocki and Goodman, 1980). The sequence of frequency and electrode was randomized and the results averaged over four to six estimates.

3. Results

Fig. 1 shows that, for both apical (circles) and basal (triangles) electrodes, the frequency difference limen (y-axis) increases monotonically as a function of the standard frequency (x-axis). The average difference limen was 12 Hz at the standard frequency of 50 Hz, gradually increased to 40 Hz at 200 Hz, and dramatically elevated to 135 Hz or more at 300 Hz and above. This trend in the frequency difference limen can be accurately ($R^2 = 0.92$–1.00) described by an exponential function (solid lines for apical electrodes and dashed lines for basal electrodes).

Fig. 2 shows pitch estimates (y-axis) as a function of frequency (x-axis) from both apical (circles) and basal (triangles) electrodes in the same cochlear implant subjects. Note first that all subjects showed a saturated pitch–frequency function. Pitch monotonically increased with frequency up to roughly 300 Hz, but saturated with further increases in frequency. Second, note the interaction between stimulus frequency and electrode position. Pitch was dominated by the temporal code at low frequencies, demonstrated by similar pitch perception at these frequencies between apical and basal electrodes (no significant difference in pitch estimates was found at 50 Hz: 2.5 vs. 6.1; t-test, $P > 0.1$). On the other hand, pitch was dominated by the place code at the highest frequencies, demonstrated by the constant difference between apical and basal electrode stimulation at these frequencies (significantly lower pitch was obtained with the apical stimulation than the basal stimulation at the highest frequency: 79.0 vs. 89.1; t-test, $P < 0.01$).

Following Fechner's (1966) original proposal that the perceptual difference in sensation is constant, we can derive the pitch–frequency function (Fig. 2) from the frequency difference limen data (Fig. 1). Fechner's proposal can be written as:

$$\Delta P = C \tag{1}$$

where ΔP is the perceptual difference in pitch and C is a constant. We showed in Fig. 1 that the frequency difference limen (ΔF) is described by an exponential function:

$$\Delta F = ae^{bF} \tag{2}$$

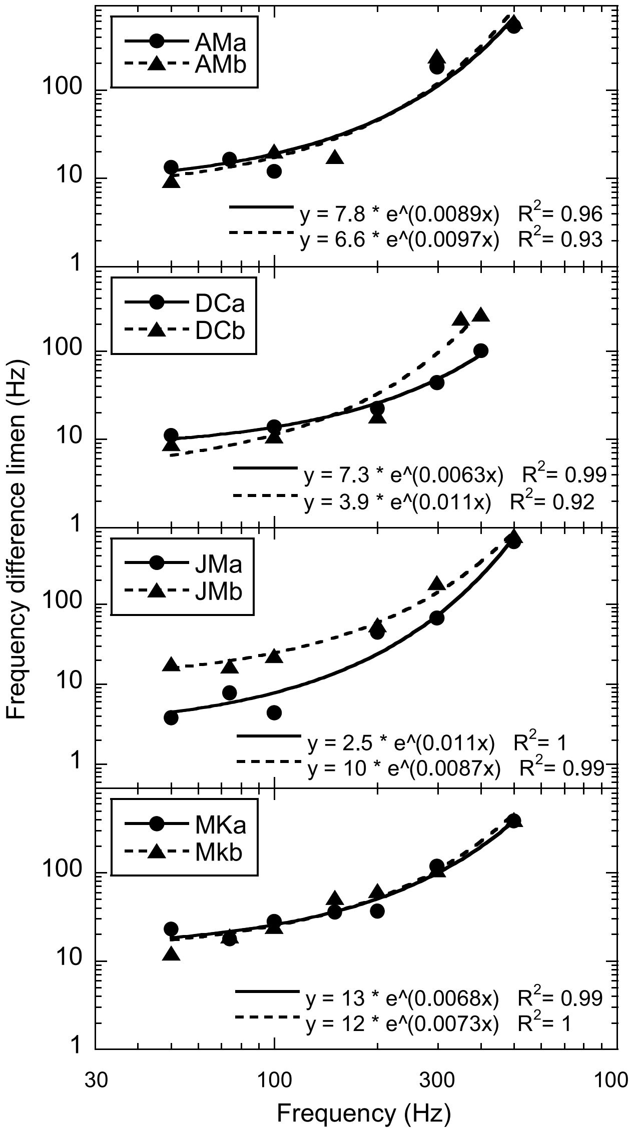

Fig. 1. Frequency discrimination as a function of frequency in four cochlear implant subjects. The *x*-axis is the standard frequency of biphasic pulse trains and the *y*-axis is the difference limen, both in Hz. Circles represent data collected with apical electrodes and triangles represent data collected with basal electrodes. The solid and dashed lines represent the fit of an exponential function (see inserted equations in each panel) to the data collected with apical and basal electrodes, respectively.

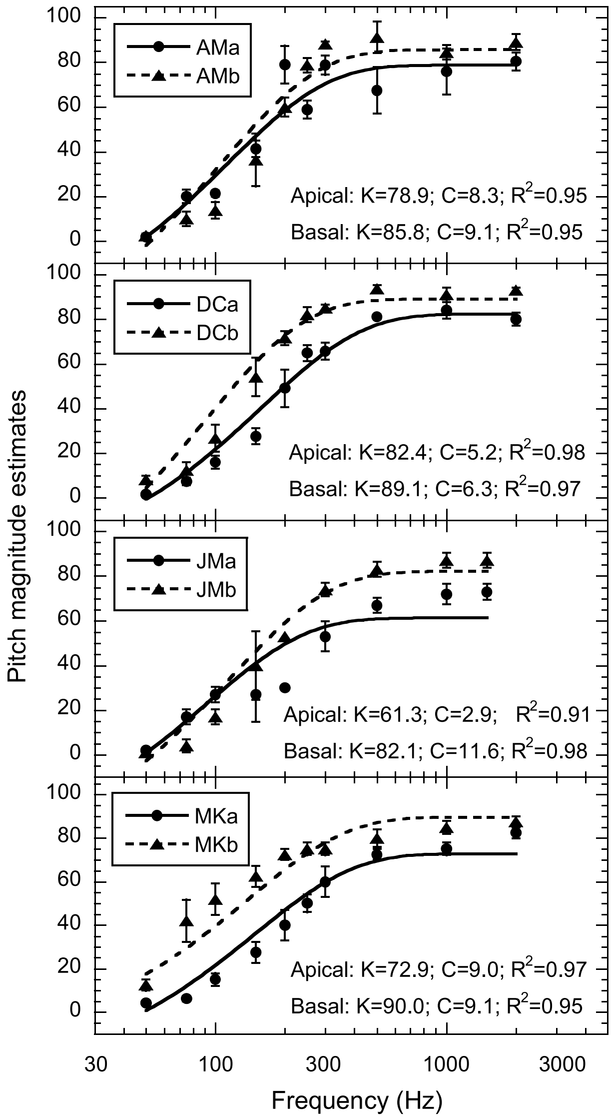

Fig. 2. Pitch magnitude estimates as a function of frequency in four cochlear implant subjects. The *x*-axis is the frequency of a biphasic pulse train and the *y*-axis is its estimated subjective pitch. Circles represent data collected with apical electrodes and triangles represent data collected with basal electrodes. Error bars represent ± standard errors. The solid and dashed lines represent the predicted pitch function (Eq. 3 in the text) from the frequency discrimination data for apical and basal electrodes, respectively. The data points at 50 Hz were weighted four times more than the rest of the data points to reflect the dominant temporal pitch at this frequency. *K* depicts the saturation pitch value and *C* depicts the constant perceptual difference in pitch (Eq. 3 in the text).

where a and b are depicted for each data set in Fig. 1. Dividing Eq. 1 by Eq. 2 and integrating the combined equation, we obtain the predicted pitch–frequency function:

$$P = K - \frac{C}{abe^{bF}} \tag{3}$$

where K depicts the saturation pitch determined by the electrode position.

The solid lines in Fig. 2 represent the predicted pitch–frequency functions from the apical electrode and the dashed lines represent that from the basal electrode ($R^2 = 0.91$–0.98). The mean saturation pitch for apical electrodes ($K = 73.9$; S.D. $= 9.3$) is significantly lower than that for basal electrodes ($K = 87.0$; S.D. $= 3.6$; *t*-test, $P < 0.05$). The constant perceptual difference in pitch (C) is not statistically different between apical and basal electrodes (6.4 vs. 9.0; *t*-test, $P > 0.1$). These values indicate roughly 10 discriminable steps ($=K/C$) for temporal pitch in electric hearing.

4. Discussion

Pitch estimate and rate discrimination in electric hearing have been mostly examined separately in previous studies (Bilger, 1977; Eddington et al., 1978; Simmons et al., 1981; Hochmair-Desoyer et al., 1983; Shannon, 1983; Clark et al., 1987; Townshend et al., 1987; Pijl and Schwarz, 1995; McDermott and McKay, 1997; Wilson et al., 1997; Fearn et al., 1999). The results from these studies were mostly consistent with the present data. One exception was that one of the five subjects (SR2) in the Wilson et al.'s study and two subjects in the Hochmair-Desoyer's study showed no sign of saturation in temporal pitch for rates up to 500–1000 Hz. It is not clear why these subjects could discern pitch differences at these high rates. One possibility is that these subjects had an unusually high degree of nerve survival. On the other hand, the relationship is loose at best between temporal pitch and nerve survival. For example, most cochlear implant users have significant nerve loss, but their pitch performance is highly consistent with the temporal pitch data obtained using either the sinusoidally amplitude-modulated noise (Burns and Viemeister, 1976; Formby, 1985) or the filtered click trains (McKay and Carlyon, 1999) in normal-hearing listeners. Together, the present available data suggest that the upper boundary using temporal code for pitch is likely much lower than that suggested in previous neurophysiological studies (300 vs. 3000 Hz).

However, the fact that cochlear implant users cannot discriminate temporal pitch at higher than 300-Hz frequencies does not necessarily mean that they cannot process temporal information at these frequencies. Physiological studies have found much higher synchronization in the auditory nerve's response to electric stimulation than acoustic stimulation (Dynes and Delgutte, 1992; Litvak et al., 2001). Psychophysical data show that the implant users could detect temporal fluctuations at frequencies as high as 4000 Hz (Shannon, 1992). Similarly, speech performance improved significantly in single-channel Vienna implant users when the signal bandwidth was increased from 300 to 900 Hz (Hochmair-Desoyer et al., 1985). This use of higher than 300-Hz frequency boundary is in line with the phase-locking capability of the auditory nerve. However, the percept evoked by the high-rate modulation or stimulation may reflect different degrees of 'roughness' in timbre rather than 'highness' in pitch.

Second, the accurate prediction of pitch estimation from frequency discrimination by a simple integration model is strikingly similar to the close coupling between intensity discrimination and loudness estimation in electric hearing (Zeng and Shannon, 1994, 1999). These results reinforce Fechner's classic hypothesis relating stimulus discriminability to sensation magnitude, indicating that this relationship is not coincidental but possibly a consequence of the sensory and nervous systems' structure and function (Rushton, 1961; Stevens, 1970).

Third, the present result bears important considerations for the design of future cochlear implants. Current cochlear implants typically have 12 to 22 physical electrodes, but have a reduced number of effective functional channels (four to eight, see Fishman et al., 1997) with a relatively small perceptual change in place pitch from apical to basal electrodes (fig. 2 of this report and see also fig. 4 in Eddington, 1980). Therefore, the total pitch range that current cochlear implants may cover is limited to at most 1200–2400 Hz (300 Hz × four or eight channels). Future cochlear implant needs to increase both the density and range (towards both the apical and basal ends) of independent electrodes than presently available to restore the normal pitch resolution and range. This requirement may be met by (1) increasing the physical number of electrodes; (2) introducing virtual channels between electrodes (McDermott and McKay, 1994); and (3) restoring stochastic electric responses (Rubinstein et al., 1999; Zeng et al., 2000).

Finally, the present results obtained with simple stimuli can be used to test the validity of contemporary pitch models of complex sounds. To predict the perception of periodicity pitch of complex sounds, modern models have all combined time and place information (Srulovicz and Goldstein, 1983; Lyon, 1984; Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Meddis and O'Mard, 1997; Shamma and Klein, 2000). For example, the popular autocorrelation theory (e.g. Meddis and O'Mard, 1997) first computes a temporal autocorrelation function in each frequency channel and then sums the autocorrelation function across all frequency channels to form the basis for pitch estimation. We now consider a simple case in acoustic and electric stimulation to test the validity of this autocorrelation model.

For simplicity, let us assume only four frequency bands ranging from low to high frequency (or apical to basal) in the model (y-axis labeled with 'Place' in Fig. 3). In acoustic stimulation with intact cochlear filters, a high-frequency tone will activate the high-frequency channel and produce a short-latency peak in the autocorrelation function (dashed line in A). The summed autocorrelation function (solid line on the bottom) will have a short latency, indicating a high pitch. Similarly, a low-frequency tone will activate the low-

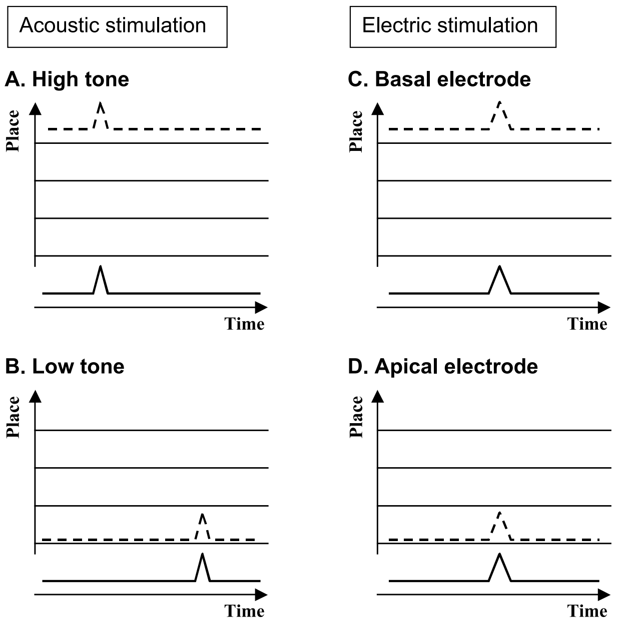

Fig. 3. Schematic representation of the autocorrelation pitch model in acoustic (A and B) and electric (C and D) stimulation. A shows the autocorrelation function for a high tone and B shows that for a low tone. Notice the inseparable relationship between the place and time codes in acoustic stimulation (the high-frequency channel is associated with a short latency while the low-frequency channel is associated with a long latency). C shows basal electrode stimulation and D shows apical electrode stimulation with the same periodic waveform. Note the identical autocorrelation function with basal and apical electrode stimulation, which would result in an identical summed autocorrelation function (bottom trace) and presumably the same pitch according to the summed autocorrelation model.

frequency channel and produce a long-latency peak in the autocorrelation function, indicating a low pitch (B). With acoustic stimulation, the place (the frequency channel being activated) and the time (latency) are intrinsically correlated. However, with electric stimulation, identical temporal information can be provided to different channels by independently stimulating the basal (C) and apical (D) electrodes. The summed autocorrelation function should be identical between basal and apical stimulation, predicting an independence of rate pitch on the place of stimulation. This prediction is clearly inconsistent with the observed joint dependence of pitch perception on rate and place (Fig. 2; see also Eddington et al., 1978; Eddington, 1980; Tong et al., 1983; Townshend et al., 1987; Nelson et al., 1995). Other pitch models (Srulovicz and Goldstein, 1983; Shamma and Klein, 2000) have incorporated place as a central template for pitch matching, but only used the place information in a relative sense to calculate pitch. The present data suggest that these modern pitch models of complex sounds should take absolute place information into account.

Acknowledgements

I thank Len Kitzes, Ying-Yee Kong, Ginger Stickney, Brian Moore, and two anonymous reviewers for their helpful comments on an earlier version of the manuscript. This work is supported by National Institutes of Health (NIDCD RO1 DC02267).

References

ANSI, 1973. Psychoacoustical Terminology. American National Standard Institute, New York.

Bilger, R.C., 1977. Psychoacoustic evaluation of present prostheses. Ann. Otol. Rhinol. Laryngol. 86 (3 Pt 2), 92–104.

Burns, E.M., Viemeister, N.F., 1976. Nonspectral pitch. J. Acoust. Soc. Am. 60, 863–869.

Chatterjee, M., Zwislocki, J.J., 1997. Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch. Hear. Res. 111, 65–75.

Clark, G.M. et al., 1987. The University of Melbourne-Nucleus Multi-Electrode Cochlear Implant. Karger, Basel.

Dynes, S.B., Delgutte, B., 1992. Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. Hear. Res. 58, 79–90.

Eddington, D.K., 1980. Speech discrimination in deaf subjects with cochlear implants. J. Acoust. Soc. Am. 68, 885–891.

Eddington, D.K., Dobelle, W.H., Brackmann, D.E. et al., 1978. Auditory prostheses research with multiple channel intracochlear stimulation in man. Ann. Otol. Rhinol. Laryngol. 87 (6 Pt 2), 1–39.

Evans, E.F., 1978. Place and time coding of frequency in the peripheral auditory system: some physiological pros and cons. Audiology 17, 369–420.

Fearn, R., Carter, P., Wolfe, J., 1999. The perception of pitch by users of cochlear implants: possible significance for rate and place theories of pitch. Acoust. Australia 27, 41–43.

Fechner, G.T., 1966. Elements of Psychophysics. Vol. I. Holt, Rinehart&Winston, New York.

Fishman, K.E., Shannon, R.V., Slattery, W.H., 1997. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J. Speech Lang. Hear. Res. 40, 1201–1215.

Formby, C., 1985. Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners. J. Acoust. Soc. Am. 78, 70–77.

Helmholtz, H.L.F., 1877. On the Sensation of Tones. Reproduced in 1954 by Dover Publication Inc., New York.

Hochmair-Desoyer, I.J., Hochmair, E.S., Burian, K., Stiglbrunner, H.K., 1983. Percepts from the Vienna cochlear prosthesis. Ann. N. Y. Acad. Sci. 405, 295–306.

Hochmair-Desoyer, I.J., Hochmair, E.S., Stiglbrunner, H.K., 1985. Psychoacoustic temporal processing and speech understanding in cochlear implant patients. In: Schindler, R.A., Merzenich, M.M. (Eds.), Cochlear Implants. Raven Press, New York, pp. 291–304.

Houtsma, A.J., Smurzynski, J., 1990. Pitch identification and discrimination for complex tones with many harmonics. J. Acoust. Soc. Am. 87, 304–310.

Johnstone, B.M., Boyle, A.J.F., 1967. Basilar membrane vibration examined with the Mossbauer technique. Science 158, 389–390.

Kiang, N.Y.S., Watanabe, T., Thomas, E.C. et al., 1965. Discharge Patterns of Single Fibres in the Cat's Auditory Nerve. MIT Press, Cambridge, MA.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.

Licklider, J.C.R., 1951. A duplex theory of pitch perception. Experientia 7, 128–134.

Litvak, L., Delgutte, B., Eddington, D., 2001. Auditory nerve fiber responses to electric stimulation: modulated and unmodulated pulse trains. J. Acoust. Soc. Am. 110, 368–379.

Lyon, R.F., 1984. Computational models of neural auditory processing. IEEE Proc. 36, 1–4.

McDermott, H.J., McKay, C.M., 1994. Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea. J. Acoust. Soc. Am. 96, 155–162.

McDermott, H.J., McKay, C.M., 1997. Musical pitch perception with electrical stimulation of the cochlea. J. Acoust. Soc. Am. 101, 1622–1631.

McKay, C.M., Carlyon, R.P., 1999. Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains. J. Acoust. Soc. Am. 105, 347–357.

Meddis, R., O'Mard, L., 1997. A unitary model of pitch perception. J. Acoust. Soc. Am. 102, 1811–1820.

Moller, A.R., 1999. Review of the roles of temporal and place coding of frequency in speech discrimination. Acta Otolaryngol. 119, 424–430.

Nelson, D.A., Van Tasell, D.J., Schroder, A.C. et al., 1995. Electrode ranking of 'place pitch' and speech recognition in electrical hearing. J. Acoust. Soc. Am. 98, 1987–1999.

Pijl, S., Schwarz, D.W., 1995. Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. J. Acoust. Soc. Am. 98, 886–895.

Rhode, W.S., 1978. Some observations on cochlear mechanics. J. Acoust. Soc. Am. 64, 158–176.

Rose, J.E., Brugge, J.F., Anderson, D.J., Hind, J.E., 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J. Neurophysiol. 30, 769–793.

Rubinstein, J.T., Wilson, B.S., Finley, C.C. et al., 1999. Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. Hear. Res. 127, 108–118.

Rushton, W.A., 1961. Peripheral coding in the nervous system. In: Rosenblith, W. (Ed.), Sensory Communication. MIT Press, Cambridge, MA, pp. 169–181.

Shackleton, T.M., Carlyon, R.P., 1994. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. J. Acoust. Soc. Am. 95, 3529–3540.

Shamma, S., 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.

Shannon, R.V., 1983. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. Hear. Res. 11, 157–189.

Shannon, R.V., 1992. Temporal modulation transfer functions in patients with cochlear implants. J. Acoust. Soc. Am. 91, 2156–2164.

Simmons, F.B., White, R.L., Walker, M.G., Mathews, R.G., 1981. Pitch correlates of direct auditory nerve electrical stimulation. Ann. Otol. Rhinol. Laryngol. 90 (2 Pt 3), 15–18.

Srulovicz, P., Goldstein, J.L., 1983. A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. J. Acoust. Soc. Am. 73, 1266–1276.

Stevens, S.S., 1954. Pitch discrimination, mels, and Kock's contention. J. Acoust. Soc. Am. 26, 1075–1077.

Stevens, S.S., 1970. Neural events and the psychophysical law. Science 170, 1043–1050.

Tong, Y.C., Dowell, R.C., Blamey, P.J., Clark, G.M., 1983. Two-component hearing sensations produced by two-electrode stimulation in the cochlea of a deaf patient. Science 219, 993–994.

Townshend, B., Cotter, N., Van Compernolle, D. et al., 1987. Pitch perception by cochlear implant subjects. J. Acoust. Soc. Am. 82, 106–115.

Wever, E.G., 1948. Theory of Hearing. Wiley, New York.

Wilson, B., Zerbi, M., Finley, C., Lawson, D., van den Honert, C., 1997. Relationships between temporal patterns of nerve activity and pitch judgments for cochlear implant patients. In: The 8th Quarterly Progress Report of NIH Contract N01-DC-5-2103, Speech Processors for Auditory Prostheses.

Young, E.D., Sachs, M.B., 1979. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J. Acoust. Soc. Am. 66, 1381–1403.

Zeng, F.G., Fu, Q.J., Morse, R., 2000. Human hearing enhanced by noise. Brain Res. 869, 251–255.

Zeng, F.G., Shannon, R.V., 1994. Loudness-coding mechanisms inferred from electric stimulation of the human auditory system. Science 264, 564–566.

Zeng, F.G., Shannon, R.V., 1999. Psychophysical laws revealed by electric hearing. NeuroReport 10, 1931–1935.

Zwislocki, J.J., Goodman, D.A., 1980. Absolute scaling of sensory magnitudes: a validation. Percept. Psychophys. 28, 28–38.