

REFERENCES

- Abe, M., Nakamura, S., Shikano, K., and Kuwabara, H. (1988). Voice conversion through vector quantization. *Proc. ICASSP*, New York, 655-658.
- Agrawal, A., and Lin, W. C. (1975). Effect of voiced speech parameters on the intelligibility of pb words. *J. Acoust. Soc. Am.*, **57**:217-222.
- Alku, P. (1992). Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Comm.*, **11**:109-118.
- Allen, J., Hunnicutt, M. S., and Klatt, D. (1987). *From Text to Speech*. Cambridge University Press, Cambridge.
- Allen, D. R., and Strong, W. J. (1985). A model for the synthesis of natural sounding vowels. *J. Acoust. Soc. Am.*, **78**(1):58-69.
- The American Heritage Dictionary of the English Language*, William Morris (Ed.), American Heritage Publishing Co. and Houghton Mifflin Co., Boston, 1969.
- Ananthapadmanabha, T. V. (1982). Intelligibility Carried by Speech-Source Functions: Implication for Theory of Speech Perception, STL-QPSR, Royal Institute of Technology, Stockholm, Sweden, 4:49-64.
- Ananthapadmanabha, T. V., and Fant, G. (1982). Calculation of true glottal flow and its components. *Speech Comm.*, **1**:167-184.
- Anderson, V. A. (1977). *Training the Speaking Voice*. Oxford University Press, Oxford.
- Atal, B. S., Chang, J. J., Mathews, M. V., and Tukey, J. W. (1978). Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. *J. Acoust. Soc. Am.*, **63**:1535-1555.
- Atal, B. S., Cuperman, V., and Gersho, A., Eds. (1991). *Advances in Speech Coding*. Kluwer Academic Publishing, Boston.
- Atal, B. S., Cuperman, V., and Gersho, A., Eds. (1993). *Speech and Audio Coding for Wireless and Network Applications*. Kluwer Academic Publishing, Boston.
- Atal, B. S., and David, N. (1979). On synthesizing natural sounding speech by linear prediction. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.* 44-47.
- Atal, B. S., and Hanauer, S. L. (1971). Speech analysis and synthesis by linear prediction of the speech wave. *J. Acoust. Soc. Am.*, **50**(2):637-655.
- Atal, B. S., and Remde, J. R. (1982). A new model of LPC excitation for producing natural-sounding speech at low bit rates. *Proc. ICASSP*, 614-617.
- Badin, P. (1989). Acoustics of voiceless fricatives: Production theory and data. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 3:33-55.
- Badin, P. (1991). Fricative consonants: Acoustic and x-ray measurements, *J. Phonetics*, **19**:397-408.
- Badin, P., and Fant, G. (1984). Notes on vocal tract computation. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 2-3:53-108.
- Baer, T. (1981a). Investigation of the phonatory mechanism. *Proceedings of the Conference on the Assessment of Vocal Pathology*, ASHA Reports, **11**:38-47.
- Baer, T. (1981b). Observations of vocal fold vibration: Measurement of excised larynges. In *Vocal Fold Physiology*, K. N. Stevens, and M. Hirano, Eds. University of Tokyo Press, Tokyo, Japan.
- Baer, T., Gore, J. C., Gracco, L. C., and Nye, P. W. (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *J. Acoust. Soc. Am.*, **90**:799-828.
- Bailly, G., and Benoit, C., Eds. (1992). *Talking Machines*. Elsevier Science Publishing, North-Holland, Amsterdam.
- Baken, R. J. (1987). *Clinical Measurement of Speech and Voice*. College-Hill Press, San Diego, CA.
- Barnwell, T. P. (1979). Objective measures for speech quality testing. *J. Acoust. Soc. Am.*, **66**:1658-1663.
- Bavegard, M. (1996). Towards an articulatory speech synthesizer: Model development and simulations, *TMH-QPSR*, 1:1-15.
- Båvegård, M., and Högberg, J. (1992). Using artificial neural networks to relate articulation to acoustic features of the vocal tract. *STL-QPSR*,

- Royal Institute of Technology, Stockholm, Sweden, 1:33-48.
- Båvegård, M., and Högberg, J. (1993). Using artificial neural nets to compare different vocal tract models. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 2-3:7-13.
- Beker, H. J., and Piper, F. C. (1985). *Secure Speech Communications*. Academic Press, Orlando.
- Benvenuto, N., Marchesi, M., and Uncini, A. (1992). Applications of simulated annealing to the design of special digital filters. *IEEE Trans. Signal Processing*, 40(2):323-332.
- Bocchieri, E. L. (1983). *An Articulatory Speech Synthesizer*. Ph.D. dissertation, University of Florida.
- Bocchieri, E. L., and Childers, D. G. (1984). An animated interactive graphics editor for the study of speech articulation. *Speech Technology*, 2:10-14.
- Bogert, B. P., Heally, M. J., and Tukey, J. W. (1963). The quefrency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking. In *Time Series Analysis*, M. Rosenblatt, Ed. Wiley, New York, pp. 209-243.
- Bohachevsky, I. O., Johnson, M. E., and Stein, M. L. (1986). Generalized simulated annealing for function optimization. *Technometrics*, 28(3):209-217.
- Bonder, L. J. (1983). The n-tube formula and some of its consequences. *Acustica*, 52:216-226.
- Bonomi, E., and Lutton, J.-L. (1984). The n-city travelling salesman problem: Statistical mechanics and the Metropolis algorithm. *SIAM Review*, 26(4):551-568.
- Boone, D. R. (1971). *The Voice and Voice Therapy*. Prentice-Hall Inc., Englewood Cliffs, NJ.
- Borden, G. J., and Harris, K. S. (1980). *Speech Science Primer: Physiology, Acoustics, and Perception*. Williams & Wilkins, Baltimore.
- Bristow, G. (Ed.) (1984). *Electronic Speech Synthesis*. McGraw-Hill Book Co., New York.
- Broad, D. J. (1977a). Acoustics of speech production. In *Topics in Speech Science*, D. J. Broad, Ed., Speech Communications Research Laboratory, Inc., Los Angeles, CA, pp. 10-68.
- Broad, D. J. (1977b). Theory of vocal fold vibration. In *Topics in Speech Science*, D. J. Broad, Ed., Speech Communications Research Laboratory, Inc., Los Angeles, CA, pp. 158-232.
- Brownman, C. P., and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6:201-251.
- Brownman, C. P., and Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *J. Phonetics*, 18:299-320.
- Burrus, C. S., et al. (1994). *Computer-Based Exercises for Signal Processing Using MATLAB*. Prentice Hall, Englewood Cliffs.
- Bush, M. A., Kopec, G. E., and Zue, V. W. (1983). Selecting acoustic features for stop consonant identification. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, 742-745.
- Carnevali, P., Coletti, L., and Patarnello, S. (1985). Image processing by simulated annealing. *IBM Journal of Research and Development*, 29(6):569-579.
- Carhart, R. T. (1940). Air-flow through the larynx. *Quart. J. Spch.*, 26:606-614.
- Carlson, R. (1993). Models of Speech Synthesis, *STL-QPSR*, 1, 1-14.
- Carlson, R., Granstrom, B., and Klatt, D. H. (1979). Some notes on the perception of temporal patterns in speech. In *Frontiers of Speech Communication Research*, B. Lindblom and S. Ohman, Eds. Academic Press, New York, 233-243.
- Cater, J. P. (1983). *Electronically Speaking: Computer Speech Generation*. Howard W. Sams, Indianapolis IN
- Chan, K. (1989). *Modeling of Vocal Fold Vibration*. Ph.D. dissertation, University of Florida, Gainesville.
- Charpentier, F. (1984). Determination of the vocal tract shape from the formants by analysis of the articulatory-to-acoustic nonlinearities. *Speech Comm.* 3(4):291-308.
- Chiba, T., and Kajiyama, M. (1941). *The Vowel, Its Nature and Structure*. Tokyo-Kaiseikan, Tokyo, Japan.
- Childers, D. G. (1977). Laryngeal pathology detection. *Critical Reviews in Bioengineering*, 2:375-426.
- Childers, D. G. (1991). Signal processing methods for the assessment of vocal disorders. *Med. Life Sci. Engineering*, 13:117-130.
- Childers, D. G. (1995). Glottal source modeling for voice conversion. *Speech Comm.*, 16:127-138.
- Childers, D. G. (1997). *Probability and Random Processes Using MATLAB*. McGraw-Hill, Inc., New York.
- Childers, D. G., and Ahn, C. T. (1995). Modeling the glottal volume-velocity waveform for three voice types. *J. Acoust. Soc. Am.*, 97:505-519.
- Childers, D. G., and Ding, C. (1991). Articulatory

- synthesis: Nasal sounds and male and female voices. *J. Phonet.*, **19**:453–464.
- Childers, D. G., and Hu, T. H. (1994). Speech synthesis by glottal excited linear prediction. *J. Acoust. Soc. Am.*, **96**(4):2026–2036.
- Childers, D. G., and Krishnamurthy, A. K. (1985). A critical review of electroglottography. *CRC Critical Reviews in Bioengineering*, **12**:131–161.
- Childers, D. G., and Larar, J. N. (1984). Electroglottography for laryngeal function assessment and speech analysis. *IEEE Trans. on Biomed. Eng.*, **31**(12):807–817.
- Childers, D. G., and Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *J. Acoust. Soc. Am.*, **90**(5):2394–2410.
- Childers, D. G., and Wong, C. F. (1994). Measuring and modeling vocal source-tract interaction. *IEEE Trans. Biomed. Eng.*, **41**:663–671.
- Childers, D. G., and Wu, K. (1990). Quality of speech produced by analysis–synthesis. *Speech Comm.*, **9**:97–117.
- Childers, D. G., and Wu, K. (1991). Gender recognition from speech. Part II: Fine analysis. *J. Acoust. Soc. Am.*, **90**(4):1841–1856.
- Childers, D. G., Hicks, D. M., Moore, G. P., and Alsaka, Y. A. (1986). A model for vocal fold vibratory motion, contact area, and the electroglottogram. *J. Acoust. Soc. Am.*, **80**:1309–1320.
- Childers, D. G., Hicks, D. M., Moore, G. P., Eskenazi, L., and Lalwani, A. L. (1990). Electroglottography and vocal fold physiology. *J. Speech and Hearing Research*, **33**:245–254.
- Childers, D. G., Skinner, D. P., and Kemerait, R. C. (1977). The cepstrum: A guide to processing. *Proc. IEEE*, **65**:1428–1443.
- Childers, D. G., Smith, A. M., and Moore, G. P. (1984). Relationship between electroglottograph, speech, and vocal cord contact. *Folia Phoniatria*, **36**:105–118.
- Childers, D. G., Wu, K., Hicks, D. M., and Yegnanarayana, B. (1989). Voice conversion. *Speech Comm.*, **8**:147–158.
- Coker, C. H. (1967). Synthesis by rule from articulatory parameters. *Proc. 1967 Conf. Speech Comm. Process.*, Boston, MA, pp. 52–53.
- Coker, C. H. (1976). A model of articulatory dynamics and control. *Proc. IEEE*, **64**(4):452–460.
- Coker, C. H., and Fujimura, O. (1966). Model for specification of the vocal-tract area function. *J. Acoust. Soc. Am.*, **40**:1271.
- Cole, R. A., and Cooper, W. E. (1975). Perception of voicing in English affricates and fricatives. *J. Acoust. Soc. Amer.*, **58**(6):1280–1287.
- Colton, R. H., and Estill, J. A. (1981). Elements of voice quality: Perceptual, acoustic and physiologic aspects. In *Speech and Language*, 5, N. J. Lass, Ed. Academic Press, NY, pp. 311–403.
- Cook, P. R. (1991). *Identification of Control Parameters in an Articulatory Vocal Tract Model, with Applications to the Synthesis of Singing*. Ph.D. dissertation, Stanford University, CA.
- Cook, P. R. (1993). SPASM, a real-time vocal tract physical model controller; and singer, the companion software synthesis system. *Computer Music Journal*, **17**(1):30–44.
- Cooper, M. (1987). Human factor aspects of voice input/output. *Speech Technology*, **3**:82–86.
- Corana, A. C., Marchesi, M., Martini, C., and Ridella, S. (1987). Minimizing multimodal functions of continuous variables with the simulated annealing algorithm. *ACM Transactions on Mathematical Software*, **13**(3):262–280.
- Cranen, B., and Boves, L. (1987). On subglottal formant analysis. *J. Acoust. Soc. Am.*, **81**(3):734–746.
- Davis, S. B., and Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans. on Acoust., Speech, and Signal Processing*, **28**(4):357–366.
- Davis, L., and Ritter, F. (1987). Schedule optimization with probabilistic search. *Proceedings of the 3rd IEEE Conf. on Artificial Intelligence Applications*, pp. 231–236.
- de Haan, H. J. (1982). The relationship of estimated comprehensibility to the rate of connected speech. *Perception and Psychophysics*, **32**(1):27–31.
- Deller, J. R., Proakis, J. G., and Hansen, J. H. L. (1993). *Discrete-Time Processing of Speech Signals*. Macmillan, New York.
- Denes, P. B. (1963). On the statistics of spoken English. *J. Acoust. Soc. Am.*, **35**:892–896.
- Denes, P. B., and Pinson, E. N. (1993). *The Speech Chain*, 2nd ed. W. H. Freeman, New York.
- Dudley, H. (1936). Synthesis speech. *Bell Labs. Record*, **15**:98–102.
- Dunn, H. K. (1950). The calculation of vowel resonances, and an electrical vocal tract. *J. Acoust. Soc. Am.*, **22**:740–753.

- Edwards, H. T. (1992). *Applied Phonetics: The Sounds of American English*. Singular Publishing Co., San Diego.
- Eggen, J. H. (1992). *On the Quality of Synthetic Speech: Evaluation and Improvements*. Ph.D. dissertation, University of Eindhoven, The Netherlands.
- Erler, K., and Deng, L. (1993). Hidden Markov model representation of quantized articulatory features for speech recognition. *Computer Speech and Language*, 7:265-282.
- Eskenasi, L. (1988). *Acoustic Correlates of Voice Quality and Distortion Measure for Speech Processing*. Ph.D. dissertation, University of Florida.
- Eskenasi, L., Childers, D. G., and Hicks, D. M. (1990). Acoustic correlates of vocal quality. *J. Speech and Hearing Res.*, 33:298-306.
- Espy-Wilson, C. Y. (1986). A phonetically based semivowel recognition system. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, pp. 2775-2778.
- Fairbanks, G. (1958). Test of phonemic differentiation: The rhyme test. *J. Acoust. Soc. Am.*, 33:596-600.
- Fairbanks, G., Everitt, W. L., and Jaeger, R. P. (1954). Method for time or frequency compression-expansion of speech. *Trans. of Inst. of Radio Engineers, Prof. Group on Audio*, AU2(1):7-12.
- Fallside, F., and Woods, W. A., Eds. (1985). *Computer Speech Processing*. Prentice Hall, Englewood Cliffs.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. Moreton & Co, Gravenhage, Sweden
- Fant, G. (1973). *Speech Sounds and Features*. M. I. T. Press, Cambridge.
- Fant, G. (1979). Glottal source and excitation analysis. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 4:85-107.
- Fant, G. (1985). The vocal tract in your pocket calculator. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 2-3:1-19.
- Fant, G. (1986). Glottal flow: Models and interaction. *J. Phonetics*, 14:393-399.
- Fant, G. (1988). *Glottal Flow Parameters from the Frequency Domain*, The second symposium on advanced man-machine interface through spoken language, Hawaii.
- Fant, G. (1993). Some problems in voice source analysis. *Speech Comm.*, 13:7-22.
- Fant, G. (1995). The LF model revisited: Transformations and frequency domain analysis. *STL-QPSR*, 4:1-13.
- Fant, G., and Lin, Q. G. (1988). Frequency domain interpretation and derivation of glottal flow parameters. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 2-3:1-21.
- Fant, G., and Lin, Q. G. (1989). Comments on glottal flow modelling and analysis. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 1:1-8.
- Fant, G., and Lin, Q. G. (1991). Comments on glottal flow modeling and analysis. In *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*. J. Gauffin and B. Hammarberg, Eds. Singular Publishing Co., San Diego, pp. 1-13.
- Fant, G., Liljencrants, J., and Lin, Q. G. (1985). A four parameter model of glottal flow. *STL-QPSR*, 2-3:119-156.
- Fettweis, A., and Meerkötter, K. (1975). On adaptors for wave digital filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 23(6): 516-525.
- Flanagan, J. L. (1957). Note on the design of terminal analog speech synthesizers. *J. Acoust. Soc. Am.*, 29:306-310.
- Flanagan, J. L. (1972a). *Speech Analysis, Synthesis, and Perception*. Springer-Verlag, New York.
- Flanagan, J. L. (1972b). Voices of men and machines. *J. Acoust. Soc. Am.*, 51:1375-1387.
- Flanagan, J. L. (1982). Talking with computers: Synthesis and recognition of speech by machines. *IEEE Trans. Biomed. Engr.*, BME-29:223-232.
- Flanagan, J. L., and Cherry, L. (1968). Excitation of vocal-tract synthesizers. *J. Acoust. Soc. Am.*, 45(3):764-769.
- Flanagan, J. L., and Golden, R. M. (1966). Phase vocoder. *The Bell System Tech. J.* 45:1493-1509.
- Flanagan, J. L., and Ishizaka, K. L. (1976). Automatic generation of voiceless excitation to a vocal cord-vocal tract speech synthesizer. *IEEE Trans. Acoust., Speech, Signal Process.*, 24(2):163-170.
- Flanagan, J. L., and Ishizaka, K. (1978). Computer model to characterize the air volume displaced by the vibrating vocal cords. *J. Acoust. Soc. Am.*, 63:1559-1565.

- Flanagan, J. L., and Landgraf, I. L. (1968). Self-oscillating source for vocal tract synthesizers. *IEEE Trans. Audio and Electroacoustics*, **16**:57-64.
- Flanagan, J. L., and Rabiner, L. R. (Eds) (1973). *Speech Synthesis*. Dowden, Hutchinson, and Ross, Stroudsburg, Pennsylvania.
- Flanagan, J. L., Coker, C. H., and Bird, C. M. (1962). Computer simulation of a formant vocoder synthesizer. *J. Acoust. Soc. Am.*, **35**:2003(A).
- Flanagan, J. L., Coker, C. H., Rabiner, L. R., Schafer, R. W., and Umeda, N. (1970). Synthetic voices for computers. *IEEE Spectrum*, **7**:22-45.
- Flanagan, J. L., Ishizaka, K., and Shipley, K. L. (1975). Synthesis of speech from a dynamic model of the vocal cords and vocal tract. *Bell System Tech. J.*, **54**(3):485-505.
- Flanagan, J. L., Ishizaka, K. L., and Shipley, K. L. (1980). Signal models for low bit-rate coding of speech. *J. Acoust. Soc. Am.*, **68**(3):780-791.
- Fletcher, H. (1929). *Speech and Hearing*. Van Nostrand, Princeton, NJ, pp. 292-294.
- Fletcher, H., and Steinberg, J. C. (1929). Articulation testing methods. *Bell System Tech. J.*, **8**:848-852.
- Foulke, E., and Sticht, T. G. (1969). Review of research on the intelligibility and compression of accelerated speech. *Psychol. Bull.* **72**(1):50-62.
- French, N. R., and Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *J. Acoust. Soc. Am.*, **19**:90-114.
- Fry, D. (1979). How can you say that? *Human Nature*, February, pp. 38-43.
- Fujisaki, H. (1983). Dynamic characteristics of voice fundamental frequency in speech and singing. In *The Production of Speech*, P. F. MacNeilage, Ed. Springer, New York, pp. 39-55.
- Fujisaki, H., and Ljungqvist, M. (1986). Proposal and evaluation of models for the glottal source waveform. *Proc. IEEE Int. Conf., Acoust., Speech, Signal Process.*, 1605-1608.
- Furui, S. (1989). *Digital Speech Processing, Synthesis, and Recognition*. Marcel Dekker, New York.
- Furui, S., and Sondhi, M. M., Eds. (1992). *Advances in Speech Signal Processing*. Marcel Dekker, New York.
- Garvey, W. D. (1949). Durational factors in speech intelligibility. In *Time-Compressed Speech: An Anthology and Bibliography*. Sam Duker, Ed. Scarecrow Press, 1974.
- Garvey, W. D. (1951). An experimental investigation of the intelligibility of speeded speech. In *Time-Compressed Speech: An Anthology and Bibliography*. Sam Duker, Ed. Scarecrow Press, 1974.
- Garvey, W. D. (1953a). The intelligibility of speeded speech. *J. Exp. Psychology*, **45**:102-108.
- Garvey, W. D. (1953b). The intelligibility of abbreviated speech patterns. *Quarterly Journal of Speech*, **39**:296-306.
- Geman, S., and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**(6):721-741.
- Glass, J. R., and Zue, V. W. (1986). Detection and recognition of nasal consonants in American English. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, pp. 2767-2770.
- Glass, J. R., and Zue, V. W. (1988). Multi-level acoustic segmentation of continuous speech. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, pp. 429-432.
- Gobl, C. (1988). Voice source dynamics in connected speech. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, **1**:123-159.
- Goffe, W. L., Ferrier, G. D., and Rogers, J. (1992). Simulated annealing: An initial application in econometrics. *Computer Science in Economics and Management*, **5**:133-146.
- Goffe, W. L., Ferrier, G. D., and Rogers, J. (1994). Global optimization of statistical functions with simulated annealing. *J. Econometrics*, **60**:65-99.
- Gold, B., and Rabiner, L. R. (1968). Analysis of digital and analog formant synthesizers. *Trans. Audio Electroacoustics*, **AU-16**:81-94.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison Wesley, Reading, MA.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. Academic Press, New York, NY.
- Goldstein, H. (1940). Reading and listening comprehension at various controlled rates. In *Time-Compressed Speech: An Anthology and Bibliography*. Sam Duker, Ed. Scarecrow Press, 1974.
- Gopinath, B., and Sondhi, M. M. (1970). Determination of the shape of human vocal tract from acoustic measurements. *Bell System Tech. J.*, **49**:1195-1214.

- Guo, Q., and Milenkovic, P. (1993). Optimal estimation of vocal tract area functions from speech signal constrained by x-ray microbeam data. *Proc. ICASSP*, 2:538-541.
- Gupta, S. K., and Schroeter, J. (1993). Pitch-synchronous frame-by-frame and segment-based articulatory analysis by synthesis. *J. Acoust. Soc. Am.*, 94(5):2517-2530.
- Gupta, V., Lennig, M., Mermelstein, P., Kenny, P., Seitz, P. F., and O'Shaughnessy, D. (1992). Use of minimum duration and energy contour for phonemes to improve large vocabulary isolated-word recognition. *Computer Speech and Language*, 6:345-359.
- Hamlet, S. L. (1981). Ultrasound assessment of phonatory function. *ASHA Reports*, 11:128-140.
- Hawley, M. E. (Ed.) (1977). *Speech Intelligibility and Speaker Recognition*. Dowden, Hutchinson, and Ross, Stroudsburg, PA.
- Hegerl, G. C., and Höge, H. (1991). Numerical simulation of the glottal flow by a model based on the compressible Navier-Stokes equations. *Proc. ICASSP*, 477-480.
- Heike, G. (1979). Articulatory measurement and synthesis—Methods and preliminary results. *Phonetica*, 36:294-301.
- Heinz, J. M., and Stevens, K. N. (1961). On the properties of voiceless fricatives. *J. Acoust. Soc. Am.*, 33:589-596.
- Heinz, J. M., and Stevens, K. N. (1964). On the derivation of area functions and acoustic spectra from cineradiographic films of speech. *J. Acoust. Soc. Am.*, 36:1037.
- Hermansky, H. (1990). Perceptual linear predictive (PLP) analysis of speech. *JASA*, 87(4):1738-1752.
- Hess, W. (1983). *Pitch Determination of Speech Signals: Algorithms and Devices*. Springer-Verlag, New York, NY.
- Hildebrand, B. H. (1976). *Vibratory Patterns of the Human Vocal Cords During Variations of Frequency and Intensity*. Ph.D. dissertation, University of Florida, Gainesville.
- Hilgers, F. J. M., and Schouwenburg, P. F. (1990). A new low-resistance, self-retaining prosthesis (Provox™) for voice rehabilitation after total laryngectomy. *Laryngoscope*, 100:1202-1207.
- Hinton, G. E., and Sejnowski, T. J. (1983). Analyzing cooperative computation. *Proceedings of the Fifth Annual Conference of the Cognitive Science Society*, Rochester, NY.
- Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica*, 26:89-94.
- Hirano, M. (1975). Phonosurgery: Basic and clinical investigations. *Otologia (Fukuoka)*, 21:239-440.
- Hirano, M. (1977). Structure and vibratory behavior of the vocal folds. In *Dynamic Aspects of Speech Production*, M. Sawashima, and F. S. Cooper, Eds. University of Tokyo Press, Tokyo, Japan, pp. 13-27.
- Hirano, M., and Kakita, Y. (1985). Cover-body theory of vocal cord vibration. In *Speech Science*. R. G. Daniloff, Ed. College-Hill Press, San Diego, pp. 1-46.
- Hirano, M., Kurita, S., and Nakashima, T. (1981). The structure of vocal folds. In *Vocal Fold Physiology*, K. N. Stevens, and M. Hirano, Eds. University of Tokyo Press, Tokyo, Japan, pp. 33-41.
- Hirano, M., and Sato, K. (1993). *Histological Color Atlas of the Human Larynx*. Singular Publishing, San Diego, CA.
- Hirano, M., Matsuo, K., Kakita, Y., Kawasaki, H., and Kurita, S. (1983). Vibratory behavior versus the structure of the vocal fold. In *Vocal Fold Physiology*, I. R. Titze, Ed. The Denver Center for the Performing Arts, Iowa City, pp. 26-40.
- Hogden, J., Lofqvist, A., Gracco, V., Zlokarnik, I., Rubin, P., and Saltzman, E. (1996). Accurate recovery of articulator positions from acoustics: New conclusions based on human data. *J. Acoust. Soc. Am.*, 100:1819-1834.
- Holmes, J. N. (1973). The influence of the glottal waveform on the naturalness of speech from a parallel formant synthesizer. *IEEE Transactions on Audio & Electroacoustics*, 21:298-305.
- Holmes, J. N. (1976). Formant excitation before and after glottal closure. *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Processing*, 39-42.
- Holmes, J. N. (1983). Formant synthesizers: Cascade or parallel. *Speech Comm.*, 2(4):251-274.
- Holmes, J. N. (1987). Mechanisms and models of human speech production. In *Speech Synthesis and Recognition* Van Nostrand Reinhold, Wokingham, U.K., pp. 12-39.
- Holmes, J. N. (1988). *Speech Synthesis and Recognition*. Van Nostrand Reinhold, Wokingham, U.K.
- Horii, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. *J. Speech and Hearing Research*, 22:5-19.
- Horii, Y. (1980). Vocal shimmer in sustained

- phonation. *J. Speech and Hearing Research*, **23**:202–209.
- Hoover, J., Reichle, J., van Tassel, D., and Cole, D. (1987). The intelligibility of synthesized speech: Echo II versus Votrax. *J. Speech and Hearing Research*, **30**:425–431.
- House, A. S., Williams, C. E., Hecker, M. H. L., and Kryter, K. D. (1965). Articulation testing methods: Consonated differentiation with a closed-response set. *J. Acoust. Soc. Am.*, **40**:158–166.
- Hsieh, Y. F. (1994). *A Flexible and High Quality Articulatory Speech Synthesizer*. Ph.D. dissertation, University of Florida.
- Hsiao, Y. -S. (1996). *Speech Synthesis Algorithms for Voice Conversion*. Ph.D. dissertation, University of Florida.
- Hsiao, Y. S., and Childers, D. G. (1996). *A New Approach to Formant Estimation and Modification Base on Pole Interaction*. 30th Asilomar Conf. On Signals, Systems and Computers.
- Hu, H. T. (1993). *An Improved Source Model for a Linear Prediction Speech Synthesizer*. Ph.D. dissertation, University of Florida.
- Huang, X. D., Ariki, Y., and Jack, M. A. (1990). *Hidden Markov Models for Speech Recognition*. Edinburgh University Press, Edinburgh.
- Huggins, A. W. F., and Nickerson, R. S. (1985). Speech quality evaluation using phoneme-specific sentences. *J. Acoust. Soc. Am.*, **77**:1896–1906.
- Iijima, H., Miki, N., and Nagai, N. (1992). Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis. *IEEE Transactions on Signal Processing*, **40**(9):2125–2135.
- Ince, A. N. (1992). *Digital Speech Processing: Speech Coding, Synthesis, and Recognition*. Kluwer Academic Publishing, Boston.
- Ishizaka, K., and Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell System Tech. J.*, **51**(6):1233–1268.
- Ishizaka, K., and Flanagan, J. L. (1977). Acoustic properties of longitudinal displacement in vocal cord vibration, *Bell System Tech. J.*, **56**(6): 889–918.
- Ishizaka, K., French, J., and Flanagan, J. L. (1975). Direct determination of the vocal tract wall impedance. *IEEE Trans. on Acoust., Speech, and Signal Proc.*, **23**(4):370–373.
- Ishizaka, K., Matsudaira, M., and Kaneko, T. (1976). Input acoustic impedance measurement of the subglottal system. *J. Acoust. Soc. Am.*, **60**:190–197.
- Itoh, K., Kitawaki, N., and Takehi, K. (1984). Objective quality measure for speech waveform coding systems. *Rev. Electr. Commun. Lab.*, **32**, pt. 1:220–228.
- Iwahashi, N., and Sagisaka, Y. (1995). Speech spectrum conversion based on speaker interpolation and multi-functional representation with weighting by radial basis function networks. *Speech Comm.*, **16**:139–151.
- Jayant, N. S., and Noll, P. (1984). *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, NJ, Appendix F.
- Johansson, C., Sundberg, J., Wilbrand, H., and Ytterbergh, C. (1983). From sagittal distance to area. *STL-QPSR*, Royal Institute of Technology, Stockholm, **4**:39–49.
- Juang, B. -H. (1984). On using the Itakura-Saito measures for speech coder performance evaluation. *AT&T Technical Journal*, **63**:1477–1498.
- Juang, B. J., Wong, D. Y., and Gray Jr., A. H. (1982). Distortion performance of vector quantization for LPC voice coding. *IEEE Trans. Acoust., Speech, Signal Proc.*, **30**:294–304.
- Kaburagi, T., and Honda, M. (1996). A model of articulatory trajectory formation based on the motor tasks of vocal-tract shapes. *J. Acoust. Soc. Am.*, **99**:3154–3170.
- Kahn, M., and Garst, P. (1983). The effects of five voice characteristics on LPC quality. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, 531–534.
- Kasuya, H., Ogawa, S., and Kikuchi, Y. (1986). An acoustic analysis of pathological voice and its application to evaluation of laryngeal pathology. *Speech Comm.*, **5**:171–181.
- Kasuya, H., Ogawa, S., Mashima, K., and Ebihira, S. (1986). Normalized noise energy as an acoustic measure to evaluate pathological voice. *J. Acoust. Soc. Am.*, **80**(5):1329–1334.
- Kay, S. M. (1988). *Modern Spectral Estimation*. Prentice-Hall, Englewood Cliffs, NJ.
- Kelly, J. K., Jr., and Lochbaum, C. C. (1962). Speech synthesis. *Proc. Fourth Intern. Congr. Acoust.*, **G42**:1–4.
- Kent, R. D., and Read, C. (1992). *The Acoustic Analysis of Speech*. Singular Publishing Co., San Diego.
- Kirkpatrick, S., Gelatt, C. D., Jr., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, **220**(4598):671–680.

- Kitawaki, N. (1992) Quality assessment of coded speech. In S. Furui and M. M. Sondhi, Eds. *Advances in Speech Signal Processing*. Marcel Dekker, Inc., New York, NY, pp. 357-385.
- Kitzing, P. (1986). Glottography: The electrophysiological investigation of phonatory biomechanics. *Acta Oto-Rhino-Laryngologica*, **40**:863-878.
- Klatt, D. H. (1977). Review of the ARPA Speech Understanding Project. *J. Acoust. Soc. Amer.*, **62**(6):1345-1366.
- Klatt, D. H. (1979). Synthesis by rule of segmental durations in English sentences. In *Frontiers of Speech Communication Research*, B. Lindblom and S. Ohman, Eds. Academic Press, London.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *J. Acoust. Soc. Am.*, **67**(3):971-995.
- Klatt, D. H. (1987). Review of text-to-speech conversion for English. *J. Acoust. Soc. Am.*, **82**(3):737-793.
- Klatt, D. H., and Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.*, **87**(2):820-857.
- Klatt, D. H., and Stevens, K. N. (1973). On the automatic recognition of continuous speech: Implications from a spectrogram-reading experiment. *IEEE Trans. on Audio and Electroacoustics*, **AU-21**(3):210-217.
- Koizumi, T., Taniguchi, S., and Hiromitsu, S. (1985). Glottal source-vocal tract interaction. *J. Acoust. Soc. Am.*, **78**:1541-1547.
- Koizumi, T., Taniguchi, S., Hiromitsu, S. (1987). Two-mass models of the vocal cords for natural sounding voice synthesis. *J. Acoust. Soc. Am.*, **82**:1179-1192.
- Koljonen, J., and Karjalainen, M. (1984). Use of computational psychoacoustical models in speech processing: Coding and objective performance evaluation. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, **1**:1.9.1-1.9.4.
- Kryter, K. D. (1970). Masking and speech communication. In *Noise, The Effects of Noise on Man*, Academic Press, New York, NY, pp.
- Kuwabara, H. (1984). A pitch-synchronous analysis/synthesis system to independently modify formant frequencies and bandwidths for voiced speech. *Speech Comm.*, **3**:211-220.
- Kuwabara, H., and Sagisaka, Y. (1995). Acoustic characteristics of speaker individuality: Control and conversion. *Speech Comm.*, **16**:165-173.
- Lalwani, A. L., and Childers, D. G. (1991a). A flexible formant synthesizer. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, **2**:777-780.
- Lalwani, A. L., and Childers, D. G. (1991b). Modeling vocal disorders via formant synthesis. *Proc. ICASSP*, **1**:505-508.
- Larar, J. N., Schroeter, J., and Sondhi, M. M. (1988). Vector quantization of the articulatory space. *IEEE Trans. Acoust., Speech, and Signal Processing*, **36**:1812-1818.
- Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge University Press, Cambridge.
- Lawson, S., and Mirzai, A. R. (1990). *Wave Digital Filters*. E. Horwood, New York, NY.
- Lee, C. K. (1988). *Voice Quality: Analysis and Synthesis*. Ph.D. dissertation, University of Florida.
- Lee, F. F. (1972). Time compression and expansion of speech by the sampling method. *J. Audio Eng. Society*, **20**(9):738-742.
- Lee, M. (1996). *Analysis and Synthesis of Speech Based on an Human Auditory Model*. Ph.D. dissertation, University of Florida.
- Levinson, S. E., and Schmidt, C. E. (1983). Adaptive computation of articulatory parameters from the speech signal. *J. Acoust. Soc. Am.*, **74**(4):1145-1154.
- Licklider, J. C. R. and Pollack, I. (1948). Effects of differentiation, integration, and infinite, peak clipping on the intelligibility of speech. *J. Acoust. Soc. Am.*, **20**:42-51.
- Liljencrants, J. (1985). Dynamic line analogs for speech synthesis. *STL-QPSR*, **1**:1-14.
- Lin, Q. (1992). Vocal-tract computation: How to make it more robust and faster. *STL-QPSR*, **4**:29-42.
- Lin, Q. G. (1990). *Speech Production Theory and Articulatory Speech Synthesis*. Ph.D. dissertation, Royal Institute of Technology, Stockholm, Sweden.
- Lindqvist-Gauffin, J. and Sundberg, J. (1976). Acoustic properties of the nasal tract, *Phonetica*, **33**:161-168.
- Lingard, R. (1985). *Electronic Synthesis of Speech*. Cambridge University Press, Cambridge, New York.
- Leung, H. C., Chigier, B., and Glass, J. R. (1993). A comparative study of signal representations and classification techniques for speech recognition. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, 680-683.
- Logan, J. S., Greene, B. G., and Pisoni, D. B. (1989).

- Segmental intelligibility of synthetic speech produced by rule. *J. Acoust. Soc. Am.*, **86**:566–581.
- Lowry, L. D. (1981). Artificial larynges: A review and development of a prototype self-contained intraoral artificial larynx. *Laryngoscope*, **91**:1332–1355.
- Luce, P. A., Feustel, T. C., and Pisoni, D. B. (1983). Capacity demands in short-term memory for synthetic and natural speech. *Human Factors*, **25**:17–32.
- Mack, M. A., and Gold, B. (1985). The intelligibility of nonvocalized and vocalized semantically anomalous sentences. *MIT Tech. Rept.*, No. 703, July 28, ADA 160 401.
- Mackenzie, J., Laver, J., and Hiller, S. (1983). Structural pathologies of the vocal folds and phonation. *Edinburgh University Dept. of Linguistics*, **16**:80–116.
- Maeda, S. (1974). A characterization of fundamental frequency contours of speech. *QPR-RLE, MIT*, **114**:193–211.
- Maeda, S. (1977). On a simulation method of dynamically varying vocal tract reconsideration of the Kelly–Lochbaum model. In *Articulatory Modeling and Phonetics*, R. Carre, R. Descout, and M. Wajskop, Eds. Groupe de Communication Parlee, Grenoble, France, pp. 281–288.
- Maeda, S. (1982a). A digital simulation method of the vocal-tract system. *Speech Comm.*, **1**(3–4):199–229.
- Maeda, S. (1982b). The role of the sinus cavities in the production of nasal vowels. *Proc. ICASSP*, 911–914.
- Makhoul, J., Roucos, S., and Gish, H. (1985). Vector quantization in speech coding. *Proc. IEEE*, **73**:1551–1588.
- Makhoul, J., Viswanathan, R., and Russell, W. (1976). A framework for the objective evaluation of vocoder speech quality. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, 103–106.
- Malah, D. (1979). Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals. *IEEE Trans. on Acoust., Speech, and Signal Processing*, **27**(2):121–133.
- Markel, J. D., and Gray, A. H. (1974). A linear prediction vocoder simulation based upon the autocorrelation method. *IEEE Trans. Acoust., Speech, Signal Process.*, **22**:124–134.
- Markel, J. D., and Gray, A. H. (1976). *Linear Prediction of Speech*. Springer-Verlag, New York, NY.
- Mattingly, I. G. (1966). Synthesis by rule of prosodic features. *Lang. Speech*, **9**:1–13.
- McAulay, R. J., and Quatieri, T. F. (1986). Speech analysis/synthesis based on a sinusoidal representation. *IEEE Trans. Acoust., Speech, Signal Proc.*, **ASSP-34**:744–754.
- McCandless, S. S. (1974). An algorithm for automatic formant extraction using linear prediction spectra. *IEEE Trans. Acoust., Speech, Signal Process.*, **22**(2):135–141.
- McGowan, R. S. (1994). Recovering articulatory movement from formant frequency trajectories using task dynamics and a genetic algorithm: Preliminary model tests. *Speech Comm.*, **14**:19–48.
- McGowan, R. S., and Lee, M. (1996). Task dynamic and articulatory recovery of lip and velar approximations under model mismatch conditions. *J. Acoust. Soc. Am.*, **99**:595–608.
- Meng, H. M., and Zue, V. W. (1991). Signal representation comparison for phonetic classification. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, 285–288.
- Mermelstein, P. (1967). Determination of the vocal tract shape from measured formant frequencies. *J. Acoust. Soc. Am.*, **41**(5):1283–1294.
- Mermelstein, P. (1973). Articulatory model for the study of speech production. *J. Acoust. Soc. Am.*, **53**(4):1070–1082.
- Mermelstein, P. (1977). On detecting nasals in continuous speech. *J. Acoust. Soc. Amer.*, **61**(2):581–587.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chemical Physics*, **21**:1087–1092.
- Meyer, P., Schroeter, J., and Sondhi, M. M. (1991). Design and evaluation of optimal cepstral lifters for accessing articulatory codebooks. *IEEE Trans. Acoust., Speech, and Signal Processing*, **39**(7):1493–1502.
- Meyer, P., and Strube, H. W. (1984). Calculations on the time-varying vocal tract. *Speech Comm.*, **3**:109–122.
- Meyer, P., Wilhelms, R., and Strube, H. W. (1989). A quasiarticulatory speech synthesizer for German language running in real time. *J. Acoust. Soc. Am.*, **86**(2):523–539.
- Milenkovic, P. H. (1984). Vocal tract area functions from two-point acoustic measurements with formant frequency constraints. *IEEE Trans.*

- Acoust., Speech, and Signal Processing*, 32(6):1122–1135.
- Milenkovic, P. H. (1987). Acoustic tube reconstruction from noncausal excitation. *IEEE Trans. Acoust., Speech, and Signal Processing*, 35(8):1089–1100.
- Milenkovic, P. H. (1993). Voice source model for continuous control of pitch period. *J. Acoust. Soc. Am.*, 93(2):1087–1096.
- Miller, J. D. (1959). Nature of the vocal cord wave. *J. Acoust. Soc. Am.*, 31:667–677.
- Mizuno, H., and Abe, M. (1995). Voice conversion algorithm based on piecewise linear conversion rules for formant frequencies and spectrum tilt. *Speech Comm.*, 16:153–164.
- Moore, G. P. (1936). *A Stroboscopic Study of Vocal Fold Movement*. Ph.D. dissertation, Northwestern University, IL.
- Moore, G. P. (1971). *Organic Voice Disorders*. Prentice-Hall Inc., Englewood Cliffs, NJ.
- Moore, G. P. (1975). Ultra-high speed photography in laryngeal research. *Canadian J. Otolaryngol.*, 4:793–799.
- Moore, G. P., and Childers, D. G. (1983). Glottal area (real size) related to voice production. Transcripts of the Twelfth Symposium Care of the Professional Voice, The Voice Foundation, New York, pp. 92–96.
- Moore, G. P., and von Leden, H. (1958). Dynamic variations of the vibratory pattern in the normal larynx. *Folia Phoniatria*, 10:205–238.
- Morgan, N. (1984). *Talking Chips*, McGraw-Hill Book Co., New York, NY.
- Morse, P. M. (1948). *Vibration and Sound*. McGraw-Hill, New York, NY.
- Morse, P. M., and Ingard, K. U. (1968). *Theoretical Acoustics*. McGraw-Hill, New York, NY.
- Morris, W., Ed. (1969). *The American Heritage Dictionary of the English Language*. American Heritage Publishing Co. and Houghton Mifflin Co., Boston.
- Moulines, E., and Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Comm.*, 9:453–467.
- Moulines, E., and Laroche, J. (1995). Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Comm.*, 16:175–205.
- Murry, T., and Singh, S. (1980). Multidimensional analysis of male and female voices. *J. Acoust. Soc. Am.*, 66:1294–1300.
- Myrick, R., and Yantorno, R. (1993). Vocal tract modeling as related to the use of an artificial larynx. *Proceedings of the 19th Northeast Bioengineering Conference*, pp. 75–77.
- Naik, J. K. (1984). *Synthesis and Evaluation of Natural-Sounding Speech Using the Linear Predictive Analysis-Synthesis Scheme*. Ph.D. dissertation, University of Florida.
- Nakatsui, M., and Mermelstein, P. (1982). Subjective speech-to-noise ratio as a measure of speech quality for digital waveform coders. *J. Acoust. Soc. Am.*, 72:1136–1144.
- Narendranath, M., Murthy, H. A., Rajendran, S., and Yegnanarayana, B. (1995). Transformation of formants for voice conversion using artificial neural networks. *Speech Comm.*, 16:207–216.
- Neuburg, E. P. (1978). Simple pitch-dependent algorithm for high-quality speech rate changing. *J. Acoust. Soc. Amer.*, 63(2):624–625.
- Ning, T., and Whiting, S. (1990). Power spectrum estimation is a orthogonal transformation, *Proc. IEEE Conf. Acoust., Speech, Signal Process.*, 2523–2526.
- Nocerino, N., Soong, F. K., Rabiner, L. R., and Klatt, D. H. (1985). Comparative study of several distortion measures for speech recognition. *Speech Comm.*, 4:317–331.
- Nooteboom, S. G. (1983). The temporal organization of speech and the process of spoken-word recognition. *IPO (Eindhoven) Progress Rept.*, No. 18, pp. 32–36.
- Olive, J. (1971). Automatic formant tracking in a Newton-Raphson technique. *J. Acoust. Soc. Am.*, 50:661–670.
- Olive, J. P. (1992). Mixed spectral representation—Formant and linear predictive coding (LPC). *J. Acoust. Soc. Am.*, 92(4):1837–1840.
- Olive, J. P., Greenwood, A., and Coleman, J. (1993). *Acoustics of American English Speech*. Springer-Verlag, New York, NY.
- Oppenheimer, A. V., and Schaffer, R. W. (1975). *Digital Signal Processing*. Prentice-Hall, Inc., Englewood Cliffs, NJ.
- O'Shaughnessy, D. (1987). *Speech Communication*. Addison-Wesley, Reading, MA.
- Paget, R. (1930). *Human Speech*. Routledge and Kegan Paul, Ltd., London.
- Paige, A., and Zue, V. W. (1970). Computation of

- vocal tract area functions. *IEEE Trans. Audio Electroacoustics*, **18**(1):7-18.
- Papamichalis, P. E. (1987). *Practical Approaches to Speech Coding*. Prentice-Hall, Englewood Cliffs, NJ.
- Papcun, G., Hochberg, J., Thomas, T. R., Laroche, F., Zacks, J., and Levy, S. (1992). Inferring articulation and recognizing gestures from acoustics with a neural network trained on x-ray microbeam data. *J. Acoust. Soc. Am.*, **92**(2):688-700.
- Parsons, T. (1986). *Voice and Speech Processing*. McGraw-Hill, New York, NY.
- Parthasarathy, S., and Coker, C. H. (1990). Phoneme-level parameterization of speech using an articulatory model. *Proc. ICASSP*, 337-340.
- Parthasarathy, S., and Coker, C. H. (1992). On automatic estimation of articulatory parameters in a text-to-speech system. *Computer Speech and Language*, **6**:37-75.
- Peterson, G. E., and Barney, H. L. (1952). Control methods used in a study of the vowels. *J. Acoust. Soc. Am.*, **24**:175-184.
- Pinto, N. B., Childers, D. G., and Lalwani, A. L. (1989). Formant speech synthesis: Improving production quality. *IEEE Trans. Acoust., Speech, Signal Process.*, **37**(12):1870-1887.
- Pisoni, D. B. (1982). Perception of speech: The human listener as a cognitive interface. *Speech Technology*, **1**:10-23.
- Pisoni, D. B., and Hunnicut, S. (1980). Perceptual evaluation of MITalk: The MIT unrestricted text-to-speech system. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 572-575.
- Pisoni, D. B., Nusbaum, H. C., and Green, B. G. (1985). Perception of synthetic speech generated by rule. *Proc. IEEE*, **73**:1665-1676.
- Pitas, I. (1994). Optimization and adaptation of discrete-valued digital filter parameters by simulated annealing. *IEEE Trans. on Signal Processing*, **42**(4):860-866.
- Pohlmann, K. C. (1991). *Advanced Digital Audio*. SAMS, Carmel.
- Pohlmann, K. C. (1995). *Principles of Digital Audio*, 3rd Ed. McGraw-Hill, Inc., New York, NY.
- Portnoff, M. R. (1973). *A Quasi-One-Dimensional Digital Simulation for the Time-Varying Vocal Tract*. MS thesis, MIT, Cambridge, MA.
- Portnoff, M. R. (1981). Time-scale modification of speech based on short-time fourier analysis. *IEEE Trans. on Acoust., Speech, and Signal Process*, **29**(3):374-390.
- Potter, R. K., Kopp, G. A., and Kopp, H. G. (1966). *Visible Speech*. Dover, New York, NY.
- Prado, P. P. L. (1991). *A Target-Based Articulatory Synthesizer*. Ph.D. dissertation, University of Florida.
- Prado, P. P. L., Shiva, E. H., and Childers, D. G. (1992). Optimization of acoustic-to-articulatory mapping. *Proc. ICASSP*, **2**:33-36.
- Quackenbush, S. R., Barnwell, T. P., and Clements, M. A. (1988). *Objective Measures of Speech Quality*. Prentice-Hall, Englewood Cliffs, NJ.
- Quatieri, T. F., and McAulay, R. J. (1992). Shape invariant time-scale and pitch modification of speech. *IEEE Trans. on Signal Processing* **40**(3):497-510.
- Rabiner, L. R. (1968). Digital-formant synthesizer for speech synthesis studies. *J. Acoust. Soc. Am.*, **43**:822-828.
- Rabiner, L. R., and Juang, B. H. (1993). *Fundamentals of Speech Recognition*. Prentice-Hall, Englewood Cliffs, NJ.
- Rabiner, L. R., and Shafer, R. W. (1978). *Digital Processing of Speech Signals*. Prentice-Hall, Englewood Cliffs, NJ.
- Rahim, M. G., Goodyear, C. C., Kleijn, W. B., Schroeter, J., and Sondhi, M. M. (1993). On the use of neural networks in articulatory speech synthesis. *J. Acoust. Soc. Am.*, **93**(2):1109-1121.
- Rayleigh, J. W. S. (1945). *The Theory of Sound*. Dover, New York, NY.
- Riegelberger, E. L., and Krishnamurthy, A. K. (1993). Glottal Source Estimation: Methods of applying the LF-model to inverse filtering. *Proc. ICASSP*, **2**:542-545.
- Rose, R. C., and Barnwell, T. P. III (1990). Design and performance of an analysis-by-synthesis class of predictive speech coders. *IEEE Trans. Acoust. Speech Signal Process.*, **38**(9):1489-1503.
- Rose, K., Gurewitz, E., and Fox, G. C. (1992). Vector quantization by deterministic annealing. *IEEE Transactions on Information Theory*, **38**(4):1249-1257.
- Rosenberg, A. E. (1971). Effect of glottal pulse shape on the quality of natural vowels. *J. Acoust. Soc. Am.*, **49**(2):583-590.
- Rosenberg, A. E., Calson, R., Granstrom, B., and Gaufin, J. (1975). A three-parameter voice source for speech synthesis. *Speech Commun.*, **2**:235-243.
- Rothausser, E. H., Urbanek, G. E., and Pachi, W. P.

- (1968). The preference method for speech evaluation. *J. Acoust. Soc. Am.*, **44**:408–418.
- Rothausler, E. H., et al. (1969). IEEE recommended practice for speech quality measurements, *IEEE Trans. Audio and Electroacoust.*, **17**:225–246.
- Rothausler, E. H., et al. (1971). A comparison of preference measurement methods. *J. Acoust. Soc. Am.*, **49**, pt. 2:1297–1308.
- Rothenberg, M. R. (1973). A New inverse filtering technique for deriving the glottal air flow waveform during voicing. *J. Acoust. Soc. Am.*, **53**:1632–1645.
- Rothenberg, M. R. (1981). An interactive model for the voice source, *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, **4**:1–17.
- Rubin, P., Baer, T., and Mermelstein, P. (1981). An articulatory synthesizer for perceptual research. *J. Acoust. Soc. Am.*, **70**(2):321–328.
- Rye, J. M., and Holmes, J. N. (1982). A versatile software parallel formant speech synthesizer. *JSRU Research Report*, no. 1016. Joint Speech Research Unit, Eastcote, England.
- Saito, S. (1992). *Speech Science and Technology*. IOS Press, Amsterdam.
- Saito, S., and Nakata, K. (1985). *Fundamentals of Speech Signal Processing*. Academic Press, Orlando.
- Saito, S., Fukuda, H., Isogai, Y., and Ono, H. (1981). X-ray stroboscopy. In *Vocal Fold Physiology*, K. N. Stevens, and M. Hirano, Eds. University of Tokyo Press, Tokyo, Japan.
- Saltzman, E. L. (1986). Task-dynamic coordination of the speech articulators: A preliminary research. *Experimental Brain Research*, **15**:129–144.
- Saltzman, E. L., and Kelso, J. A. S. (1987). Skilled actions: A task-dynamic approach. *Psychological Review*, **94**(1):84–106.
- Saltzman, E. L., and Munhall, K. G. (1989). A dynamic approach to gestural patterning in speech production. *Ecological Psychology*, **14**:333–382.
- Savic, M., and Nam, L. H. (1991). Voice personality transformation. *Digital Signal Process.*, **1**:107–110.
- Scherer, R. C. (1981). *Laryngeal Fluid Mechanics: Steady Flow Considerations Using Static Models*. Ph.D. dissertation, University of Iowa, Iowa City.
- Schneiderman, C. R. (1984). *Basic Anatomy and Physiology in Speech and Hearing*. College-Hill Press, San Diego, CA.
- Schroeder, M. R. (1967). Determination of the geometry of the human vocal tract by acoustic measurements. *J. Acoust. Soc. Am.*, **41**(4), p 2:1002–1010.
- Schroeder, M. R. (1993). A brief history of synthetic speech. *Speech Comm.*, **13**:231–237.
- Schroeder, M. R., and Atal, B. S. (1985). Code-excited linear prediction (CELP): High-quality speech at very low bit rates. *Proc. IEEE Int. Conf., Acoust., Speech, Signal Process.*, 937–940.
- Schroeder, M. R., and Strube, H. W. (1979). Acoustic measurements of articulator motions. *Phonetica*, **36**:302–313.
- Schroeter, J., and Sondhi, M. M. (1994). Techniques for estimating vocal-tract shapes from the speech signal. *IEEE Trans. Speech and Audio Processing*, **2**(1):133–150.
- Schroeter, J., Larar, J. N., and Sondhi, M. M. (1987). Speech parameter estimation using a vocal tract/cord model. *Proc. IEEE Int. Conf., Acoust., Speech, Signal Process.*, 308–311.
- Schroeter, J., Larar, J. N., and Sondhi, M. M. (1988). Multi-frame approach for parameter estimation of a physiological model of speech production. *Proc. IEEE Int. Conf., Acoust., Speech, Signal Process.*, **88**(S2.6):83–86.
- Schroeter, J., Meyer, P., and Parthasarathy, S. (1990). Evaluation of improved articulatory codebooks and codebook access distance measures. *Proc. ICASSP*, 393–396.
- Schwab, E. G., Nusbaum, H. C., and Pisoni, D. B. (1985). Some effects of training on the perception of synthetic speech. *Human Factors*, **27**:395–408.
- Schwartz, R., and Makhoul, J. (1975). Where the phonemes are: Dealing with ambiguity in acoustic-phonetic recognition. *IEEE Trans. on Acoust., Speech, and Signal Processing*, **23**(1):50–53.
- Shadle, C. H. (1991). The effect of geometry on source mechanisms of fricative consonants. *J. Phonetics*, **19**:409–424.
- Shearer, W. M. (1979). *Illustrated Speech Anatomy*. Charles C Thomas Publisher, Springfield, IL.
- Shirai, K. (1993). Estimation and generation of articulatory motion using neural networks. *Speech Comm.*, **13**:45–51.
- Shue, Y.-J. (1995). *A Formant-Based Linear Prediction Speech Synthesis/Analysis System*. Ph.D. dissertation, University of Florida.
- Singh, S., and Murry, T. (1978). Multidimensional classification of normal voice qualities. *J. Acoust. Soc. Am.*, **64**:81–87.

- Singhal, S., and Atal, B. S. (1989). Amplitude optimization and pitch prediction in multipulse coders. *IEEE Trans. on Acoust., Speech, Signal Process.*, **37**(3):317–327.
- Smith, M. E., Robinson, K. E., and Strong, W. J. (1981). Intelligibility and quality of linear predictor and eigenparameter coded speech. *IEEE Trans. Acoust., Speech, Signal Proc.*, **29**:391–394.
- Sondhi, M. M. (1974). Model for wave propagation in a lossy vocal tract. *J. Acoust. Soc. Am.*, **55**(5):1070–1075.
- Sondhi, M. M. (1975). Measurement of the glottal waveform. *J. Acoust. Soc. Am.*, **57**(1):228–232.
- Sondhi, M. M. (1979). Estimation of vocal-tract areas: The need for acoustical measurements. *IEEE Trans. Acoust., Speech, and Signal Processing*, **27**(3):268–273.
- Sondhi, M. M. (1986). Resonances of a bent vocal tract. *J. Acoust. Soc. Am.*, **79**(4):1113–1116.
- Sondhi, M. M., and Resnick, J. R. (1983). The Inverse problem for the vocal tract: Numerical methods, acoustical experiments, and speech synthesis. *J. Acoust. Soc. Am.*, **73**(3):985–1022.
- Sondhi, M. M., and Schroeter, J. (1986). A nonlinear articulatory speech synthesizer using both time-and frequency-domain elements. *Proc. ICASSP*, 1999–2002.
- Sondhi, M. M., and Schroeter, J. (1987). A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Trans. Acoust., Speech, and Signal Processing*, **35**(7):955–967.
- Sorokin, V. N. (1994). Inverse problem for fricatives. *Speech Comm.*, **14**:249–262.
- Steeneken, H. J. M., and Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *J. Acoust. Soc. Am.*, **67**:318–326.
- Stevens, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants: Static considerations. *J. Acoust. Soc. Am.*, **50**(4):1180–1192.
- Stevens, K. N. (1993a). Modelling affricate consonants. *Speech Comm.*, **13**:33–43.
- Stevens, K. N. (1993b). Models for the production and acoustics of stop consonants. *Speech Comm.*, **13**:367–375.
- Stevens, K. N. (1998). *Acoustic Phonetics*, The MIT Press, Cambridge.
- Stevens, K. N., and House, A. S. (1955). Development of a quantitative description of vowel articulation. *J. Acoust. Soc. Am.*, **27**(3):484–493.
- Stevens, K. N., and Klatt, D. H. (1974). Current models of sound sources for speech. In *Ventilatory and Phonatory Control Systems*, B. Wyke, Ed. Oxford University Press, New York, pp. 279–292.
- Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., and Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *J. Acoust. Soc. Am.*, **91**:2979–3000.
- Stevens, K. H., Kasowski, S., and Fant, G. (1953). An electrical analog of the vocal tract. *J. Acoust. Soc. Am.*, **25**(4):734–742.
- Story, B. H., and Titze, I. R. (1995). Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Am.*, **97**(2):1249–1260.
- Story, B. H., and Titze, I. R. (1996). Vocal tract functions from magnetic resonance imaging. *J. Acoust. Soc. Am.*, **100**:537–554.
- Strube, H. W. (1982). Time-varying wave digital filters and vocal-tract models. *Proc. ICASSP*, 923–926.
- Sundberg, J., Johansson, C., Wilbrand, H., and Ytterbergh, C. (1987). From sagittal distance to area: A study of transverse, vocal tract cross-sectional area. *Phonetica*, **44**:76–90.
- Tetschner, W. (1993). *Voice Processing* 2nd Ed. Artech House, Boston.
- Therrien, C. W. (1992). *Discrete Random Signals and Statistical Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ.
- Thomas, I. B. (1968). The influence of first and second formants on the intelligibility of clipped speech. *J. Acoust. Eng. Soc.*, **16**:182–185.
- Thomas, T. J. (1986). A finite element model of fluid flow in the vocal tract. *Computer Speech and Language*, **1**:131–151.
- Tiffany, W. R., and Bennett, D. N. (1961). Intelligibility of slow-played speech. *J. Speech Hearing Research*, **4**:248–258.
- Timcke, R., von Leden, H., and Moore, P. (1958). Laryngeal vibrations: Measurements of the glottic wave, Part I: The normal vibratory cycle. *Arch. Otolaryngo*, **68**:1–19.
- Titze, I. R. (1973). The human vocal cords: A mathematical model, Part 1. *Phonetica*, **28**:129–170.
- Titze, I. R. (1974). The human vocal cords: A mathematical model, Part 2. *Phonetica*, **29**:1–21.
- Titze, I. G. (1976). On the mechanics of vocal-fold vibration. *J. Acoust. Soc.*, **60**(6):1366–1380.
- Titze, I. R. (1982). Synthesis of sung vowels using a

- time-domain approach. *Transactions of the Eleventh Symposium: Care of the Professional Voice*, 90-98.
- Titze, I. R. (1984). Parameterization of the glottal area, glottal flow, and vocal fold contact area. *J. Acoust. Soc. Am.*, **75**(2):520-580.
- Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *J. Acoust. Soc. Am.*, **83**(4):1536-1552.
- Titze, I. R. (1989). A four-parameter model of the glottis and vocal fold contact area. *Speech Comm.*, **8**:191-201.
- Titze, I. R. (1994). *Principles of Voice Production*. Prentice-Hall Inc., Englewood Cliffs, NJ.
- Titze, I. R., and Strong, W. J. (1975). Normal modes in vocal cord tissues. *J. Acoust. Soc. Am.*, **57**(3):736-744.
- Titze, I. R., and Talkin, D. T. (1979a). A theoretical study of the effects of the various laryngeal configurations on the acoustics of phonation. *J. Acoust. Soc. Am.*, **66**(1):60-74.
- Titze, I. R., and Talkin, D. T. (1979b). Simulation and interpretation of glottographic waveforms. *ASHA Reports*, **11**:48-53.
- Umeda, N. (1977). Consonant duration in American English. *J. Acoust. Soc. Amer.*, **61**(3):846-858.
- Valbret, H., Moulines, E., and Tubach, J. P. (1992). Voice transformation using PSOLA technique. *Speech Commun.*, **11**:175-187.
- van den Berg, J. W., Zantema, J. T., and Doorneball, P. (1957). On the air resistance and the Bernoulli effect of the human larynx. *J. Acoust. Soc. Am.*, **29**(5):626-631.
- van den Berg, J., and Tan, T. S. (1959). Results of experiments with human larynxes. *Practica Oto-Rhino-Laryngologica*, **21**(6):425-450.
- Vanderbilt, D., and Louie, S. G. (1984). A monte carlo simulated annealing approach to optimization over continuous variables. *J. Computational Physics*, **56**:259-271.
- Van Santen, J., Sproat, R. W., Olive, J. P., and Hirschberg, J. (1997). *Progress in Speech Synthesis*, Springer, New York, NY.
- van Riper, C., and Irwin, J. V. (1958). *Voice Articulation*. Prentice-Hall Inc., Englewood Cliffs, NJ.
- Vecchi, M. P., and Kirkpatrick, S. (1983). Global Wiring by simulated annealing. *IEEE Trans. on Computer-Aided Design*, **2**:215-222.
- Velhuis, R. (1998). A computationally efficient alternative for the Liljencrants-Fant model and its perceptual evaluation. *J. Acoust. Soc. Am.*, **103**(1):566-571.
- Viswanathan, V. R., Russell, W. H., and Huggins, A. W. F. (1984). *Objective Speech Quality Evaluation of Real-Time Speech Coders*. Bolt Beranek Newman, Cambridge, MA, 137, and AD-A141 194/1, 17.
- Voiers, W. D. (1968). The present state of digital vocoding technique: A diagnostic evaluation. *IEEE Trans. Audio Electroacoust.*, **16**:275-279.
- Voiers, W. D. (1977a). Diagnostic acceptability measure for speech communication systems. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 204-207.
- Voiers, W. D. (1977b). Diagnostic evaluation of speech intelligibility. In *Speech Intelligibility and Speaker Recognition*, M. E. Hawley, Ed. Dowden, Hutchinson, and Ross, Stroudsburg, PA, pp. 374-387.
- Voiers, W. D. (1983). Evaluating processed speech using the diagnostic rhyme test. *Speech Technology*, Jan./Feb., 30-39.
- von Leden, H., and Moore, P. (1961). The mechanics of the cricoarytenoid joint. *Arch. Otolaryngo.*, **73**:541-550.
- von Kempelen—Not in additional references?
- Waibel, A. (1988). *Prosody and Speech Recognition*. Pitman, London.
- Wakita, H. (1973). Direct estimation of the vocal-tract shape by inverse filtering of acoustic speech waveforms. *IEEE Trans. Audio Electroacoustics*, **21**(5):417-427.
- Wakita, H. (1979). Estimation of vocal-tract shapes from acoustical analysis of the speech wave: The state of the art. *IEEE Trans. Acoust., Speech, and Signal Process.*, **27**(3):281-285.
- Wakita, H., and Fant, G. (1978). Toward a better vocal tract model. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, 1:9-29.
- Wakita, H., and Gray, H. (1975). Numerical determination of the lip impedance and vocal-tract area functions. *IEEE Trans. Acoust., Speech, Signal Processing*, **23**(6):574-580.
- Waterworth, J. A. (1983). Effect of intonation form and pause duration of automatic telephone number announcements on subjective and memory performance. *Applied Ergonomics*, **14**:39-42.
- Weinstein, C. J., McCandless, S. S., Mondschein, L. F., and Zue, V. W. (1975). A system for

- acoustic-phonetic analysis of continuous speech. *IEEE Trans. on Acoust., Speech, and Signal Processing*, **23**(1):54-67.
- White, J. M. (1995). *A System for Time Modification of Synthesized Speech*, Ph.D. dissertation, University of Florida.
- Witten, I. H. (1982). *Principles of Computer Speech*. Academic Press, Orlando.
- Wong, D. Y. (1980). On understanding the quality problems of LP speech. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, 725-728.
- Wong, D. Y., and Markel, J. D. (1977). An intelligibility evaluation of several linear prediction vocoder modifications. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc.*, 208-211.
- Wong, D. Y., Markel, J. D., and Gray, A. H. (1979). Least squares glottal inverse filtering from the acoustic speech waveform. *IEEE Trans. on Acoust., Speech, and Signal Processing*, **27**(4):350-356.
- Wright, G. T. H., and Owens, F. J. (1993). An optimized multirate sampling technique for the dynamic variation of vocal tract length in the Kelly-Lochbaum speech synthesis model. *IEEE Trans. on Speech and Audio Process.*, **1**(1):109-113.
- Wu, C. J. (1996). *A Flexible 3-D Model of Vocal Fold Vibrations*. Ph.D. dissertation, University of Florida.
- Wu, C. (1996). *Articulatory Speech Synthesizer*. Ph.D. dissertation, University of Florida.
- Xue, Q., Hu, Y. H., and Milenkovic, P. (1990). Analyses of the hidden units of the multi-layer perceptron and its application in acoustic-to-articulatory mapping. *Proc. ICASSP*, 869-872.
- Yu, Z. (1993). A method to determine the area function of speech based on perturbation theory. *STL-QPSR*, Royal Institute of Technology, Stockholm, Sweden, **4**:77-95.
- Zemlin, W. R. (1964). *Speech and Hearing Science*. Stipes, Champaign, IL.
- Zemlin, W. R. (1988). *Speech and Hearing Science: Anatomy and Physiology*. Prentice Hall Inc., Englewood Cliffs, NJ.
- Zue, V. W., Glass, J., Phillips, M., Seneff, S. (1989). Acoustic segmentation and phonetic classification in the Summit system. *Proc. Int'l. Conf. on Acoust., Speech, and Signal Processing*, 389-392.