

ture, such as telephone numbers, some success has been achieved in concatenating "vocoded" whole words. This is because it is possible to smooth vocoder parameters at word boundaries, modify durations, and impose a sentence fundamental frequency contour on the word string (Rabiner *et al.*, 1971; Olive and Nakatani, 1974). Also, Cooper *et al.*, (1984) describe an early plan to concatenate recorded words in a reading machine for the blind application, where the motivation of the listeners might overcome weaknesses of the presentation, but the approach was subsequently abandoned in favor of synthesis by rule.

²Postvocalic devoicing and flapping are actually late rules, occurring after vowel durations are computed. The proper ordering of rules is an important issue in the design of text-to-speech systems.

³Looking at the same data, we might not agree with their intuitions.

⁴The parameter k was assumed to be 0.5 by Delattre *et al.* (1955) and by Holmes *et al.* (1964).

⁵Actually, Peterson *et al.* (1958) proposed the term "diad" as a set of diphones all having essentially the same articulatory trajectory from the middle of one segment to the middle of the next, but differing in prosodic values such as duration and fundamental frequency contour. Hank Truby was the first to use the term "diphone" by separating out prosody as independent variables in synthesis, and calling the remaining phonetic transition (as represented by synthesizer control data) a "diphone." As the term diphone has spread in usage, some authors allow it to refer to larger synthesis units such as consonant clusters when needed to maximize synthesis fidelity (Dixon and Maxey, 1968), but we will restrict the term here to mean a transition between adjacent phonetic segments.

⁶For example, English vowels can be divided into tense inherently long vowels and lax short vowels (House, 1961).

⁷The number of distinguishable stress levels at the lexical and phrasal levels continues to be an area of linguistic dispute; see Vanderslice and Ladefoged (1971) and Coker *et al.* (1973) for extremal positions.

⁸In phonological theory, there is usually a distinction made between a rule that changes a feature or segment discretely, and a feature implementation algorithm that is subject to low-level physiological constraints, contextual influences, and graded behavior. Thus a parameter adjustment rule needed in speech synthesis probably should correspond to the feature implementation level of description (e.g., voice onset time is slightly longer for high versus low vowels even though glottal timing commands might be the same in two situations), whereas allophone selection rules should correspond to actual rule-governed changes to motor commands, as reflected by a change to some segmental feature.

⁹Not all phonological simplifications preserve boundary information; for example [h] deletion and flapping result in an inability to distinguish between "but her" and "butter."

¹⁰If errors were independent, words correct would be approximately equal to phonemes correct to the sixth or seventh power, times the probability of getting the stress correct.

¹¹It is perhaps unfair to evaluate this system against a random sample of words because it was intended to be used in the context of a large morpheme dictionary, and therefore would be activated only for rare words—words that may be more regular in their pronunciation.

¹²Use of the solid curve is equivalent to assuming that another million-word text sample will contain exactly the same 50 000 words, whereas it is likely that a different set of rare words will be found in the new text.

¹³It is surprising how outdated this corpus has become if the goal is to obtain a lexicon representative of modern textual material; Allen and Finkel removed more than 15% of the items as outmoded or too parochial when they were collecting morphemes by hand. We would all benefit from a modern replication of the Kučera and Francis task, especially now that it is practical to examine much larger data bases than only a million words.

¹⁴In theory, every time a new rule was added to the morph decomposition process, it was necessary to go back and check the entire lexicon for accidental incorrect decompositions.

¹⁵An even less sensitive test is the diagnostic rhyme test (Voiers, 1983) which involves a single pair of alternative responses for each familiar CVC word.

¹⁶Most of these "errors" can be attributed to problems with phonemic symbolization; phonetically trained listeners typically perform at better than 99% correct on the same task (Rabiner, 1969).

¹⁷Multipulse linear prediction was designed to make possible the detailed modeling of the voicing source waveform, but in fact it is simply a method of introducing zeros into the representation of any speech sound. It appears that multipulse has little advantage for voiced segments in text-to-speech systems because the rule system imposes an f_0 contour different

from that observed in the original natural speech recording. However, multipulse may be able to better approximate, e.g., the coherent release of plosive bursts (Maeda, 1987).

¹⁸Pisoni and Koen (1981) obtained similar results, although the difference between natural and synthetic speech was greater, perhaps because the MITalk system that they used is not quite as intelligible.

¹⁹Carlson and Granström (1976) had noted the same kind of listener adaptation without feedback in an earlier experimental evaluation. With feedback, listeners can improve considerably in performance on intelligibility tests, even with poor quality synthetic speech (Schwab *et al.*, 1986).

²⁰For example, Xerox Corp. has retrofitted a number of Kurzweil Reading Machines for the blind that are located in public libraries with the more intelligible Prose-2000 text-to-speech board. Digital Equipment Corporation has offered a special price for DECtalk units sold to handicapped individuals and manufacturers of handicapped devices, resulting in a more than one million dollar price reduction on units sold to this population.

Aho, A., and Ullman, J. (1972). *The Theory of Parsing, Translation and Computing* (Prentice-Hall, New York).

Ainsworth, W. A. (1973). "A System for Converting English Text into Speech," *IEEE Trans. Audio Electroacoust.* AU-21, 288–290.

Allen, D. R., and Strong, W. J. (1985). "A Model for the Synthesis of Natural Sounding Vowels," *J. Acoust. Soc. Am.* 78, 58–69.

Allen, J. (1976). "Synthesis of Speech from Unrestricted Text," *Proc. IEEE* 64, 422–433.

Allen, J., Hunnicutt, S., and Klatt, D. H. (1987). *From Text to Speech: The MITalk System* (Cambridge U.P., Cambridge, UK).

Allen, J., Hunnicutt, S., Carlson, R., and Granström, B. (1979). "MITalk-79: The MIT Text-to-Speech System," *J. Acoust. Soc. Am. Suppl.* 1 65, S130.

Ananthapadmanabha, T. V. (1984). "Acoustic Analysis of Voice Source Dynamics," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, QPSR* 2-3, 1–24.

Anderson, M., Pierrehumbert, J., and Liberman, M. (1984). "Synthesis by Rule of English Intonation Patterns," *Proc. Int. Conf. Acoust., Speech Signal Process. ICASSP-84*, 2.9.1–2.9.4.

Anthony, J., and Lawrence, W. (1962). "A Resonance Analogue Speech Synthesizer," *Proc. 4th Int. Cong. Acoust., Copenhagen, Denmark*.

Armstrong, L. E., and Ward, I. C. (1931). *A Handbook of English Intonation* (Cambridge U.P., Cambridge, England), 2nd ed.

ASHA (1981). *Position Paper of the Ad Hoc Committee on Communication Processes for Nonspeaking Persons*, American Speech and Hearing Association, Rockville, MD.

Atal, B. S., and Hanauer, S. L. (1971). "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," *J. Acoust. Soc. Am.* 50, 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. Int. Conf. Acoust., Speech Signal Process. ICASSP-82*, 614–617.

Atal, B. S., and Schroeder, M. R. (1975). "Recent Advances in Predictive Coding: Applications to Speech Synthesis," in *Speech Communication*, edited by G. Fant (Almqvist and Wiksell, Uppsala, Sweden), Vol. I, pp. 27–31.

Atkinson, R. C. (1972). "Teaching Children to Read Using a Computer," *Am. Psychol.* 27, 169–178.

Baer, T. (1981). "Observation of Vocal Fold Vibration: Measurement of Excised Larynges," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo, Japan), pp. 119–136.

Barnwell, T. P. (1971). "An Algorithm for Segment Duration in a Reading Machine Context," *Research Laboratory of Electronics, Tech. Report* 479, M.I.T., Cambridge, MA.

Bassak, G. (1980). "Phoneme-Based Speech Chip Needs Less Memory," *Electronics* 53, 43–44.

Bell, T. (1983). "Talk to Me," *Personal Computing* 7, 120–206 (September 1983).

Bernstein, J. (1986). "Voice Identity and Attitude," *Proceedings Speech Tech '86* (Media Dimensions, New York), pp. 213–215.

- Bernstein, J., and Baldwin, G. (1985). "Spontaneous versus Prepared Speech," *J. Acoust. Soc. Am. Suppl.* 1 78, S37.
- Bernstein, J., Becker, R., Bell, D., Murveit, H., Poza, P., and Stevens, G. (1984). "Telephone Communication between Deaf and Hearing Persons," *Proc. Int. Conf. Acoust., Speech Signal Process. ICASSP-84*, 26.7.1-26.7.4.
- Bernstein, J., and Nessler, L. (1981). "Performance Comparison of Component Algorithms for the Phonemization of Orthography," *Proc. 19th Annu. Assoc. Computational Linguistics*, Stanford Univ., pp. 19-22.
- Bernstein, J., and Pisoni, D. B. (1980). "Unlimited Text-to-Speech System: Description and Evaluation of a Microprocessor-Based Device," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-80*, 576-579.
- Bernstein, L. (Ed.) (in press). *The Vocally Impaired* (Academic, New York).
- Bickley, C. (1982). "Acoustic Analysis and Perception of Breathless Vowels," *Speech Communication Group Working Papers I*, MIT, Cambridge, MA, pp. 71-82.
- Bilger, R., Nuetzel, J., Rabinowitz, W., and Rzezchowski, C. (1984). "Standardization of a Test of Speech Perception in Noise," *J. Speech Hearing Res.* 27, 32-48.
- Bladon, A., and Lindblom, B. (1981). "Modeling the Judgement of Vowel Quality Differences," *J. Acoust. Soc. Am.* 69, 1414-1422.
- Bloomfield, L. (1933). *Language* (Henry Holt, New York).
- Bolinger, D. (1951). "Intonation: Levels versus Configurations," *Word* 7, 199-210.
- Bolinger, D. (1972). "Accent Is Predictable if You Are a Mind-Reader," *Language* 48, 633-644.
- Broad, D. J. (1979). "The New Theories of Vocal Fold Vibration," in *Speech and Language: Advances in Basic Research and Practice*, edited by N. Lass (Academic, New York), Vol. 2, pp. 203-256.
- Broad, D. J., and Fertig, R. H. (1970). "Formant Frequency Trajectories in Selected CVC Syllable Nuclei," *J. Acoust. Soc. Am.* 47, 1572-1582.
- Browman, C. P. (1980). "Rules for Demisyllable Synthesis using Lingua, a Language Interpreter," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-80*, 561-564.
- Bruckert, E., Minow, M., and Tetschner, W. (1983). "Three-Tiered Software and VLSI Aid Developmental System to Read Text Aloud," *Electronics* (21 April 1983).
- Bush, M. A., and Kopec, G. E. (1986). "Network-Based Connected Digit Recognition Using Explicit Acoustic-Phonetic Modeling," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-86*, 1097-1100.
- Carlisle, J. F. (1985). "The Relationship between Knowledge of Derivational Morphology and Spelling Ability in Fourth, Sixth and Eighth Graders," *Status Report on Speech Research SR-82/83*, Haskins Laboratories, New Haven, CT, pp. 151-174.
- Carlson, R., Galyas, K., Granström, B., Pettersson, M., and Zachrisson, G. (1980). "Speech Synthesis for the Non-Vocal in Training and Communication," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR-1*, 13-27.
- Carlson, R., Galyas, K., Granström, B., Hunnicutt, S., Larsson, B., and Neovius, L. (1981). "A Multi-Language Portable Text-to-Speech System for the Disabled," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR 2/3*, 8-16.
- Carlson, R., and Granström, B. (1975). "A Phonetically Oriented Programming Language for Rule Description of Speech," in *Speech Communication*, edited by G. Fant (Almqvist and Wiksell, Uppsala, Sweden), Vol. 2, pp. 245-253.
- Carlson, R., and Granström, B. (1976). "A Text-to-Speech System Based Entirely on Rules," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-76*, 686-688.
- Carlson, R., Granström, B., and Hunnicutt, S. (1982a). "A Multi-Language Text-to-Speech Module," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-82*, 1604-1607.
- Carlson, R., Granström, B., and Hunnicutt, S. (1982b). "Bliss Communication with Speech or Text Output," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-82*, 747-750.
- Carlson, R., Granström, B., and Klatt, D. H. (1979). "Some Notes on the Perception of Temporal Patterns in Speech," 9th International Congress of Phonetic Sciences, Copenhagen, Denmark.
- Carlson, R., Granström, B., and Larssen, K. (1976). "Evaluation of a Text-to-Speech System as a Reading Machine for the Blind," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR 2-3*, 9-13.
- Carlson, R., Granström, B., and Pauli, S. (1972). "Perceptual Evaluation of Segmental Cues," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR 1/1972*, 18-24.
- Carrell, T. D. (1984). "Contributions of Fundamental Frequency, Formant Spacing, and Glottal Waveform to Talker Identification," *Speech Research Laboratory Progress Report 5*, Indiana Univ., Bloomington, IN.
- Catford, J. G. (1977). *Fundamental Problems in Phonetics* (Indiana U.P., Bloomington, IN).
- Chial, M. (1985). "Tutorial on Speech Synthesis," *American Speech-Language-Hearing Foundation 1985 Computer Conference*, New Orleans, LA.
- Chiba, T., and Kajiyama, M. (1941). *The Vowel: Its Nature and Structure* (Tokyo-Kaiseikan, Tokyo, Japan).
- Chomsky, N., and Halle, M. (1968). *Sound Pattern of English* (Harper and Row, New York).
- Church, K. (1983). "Phrase-Structure Parsing: A Method for Taking Advantage of Allophonic Constraints," unpublished Ph.D. thesis, M.I.T., Cambridge, MA.
- Church, K. W. (1985). "Stress Assignment in Letter-to-Sound Rules for Speech Synthesis," *Proc. 23rd Meeting Assoc. Comp. Ling.*, 246-253.
- Clements, G. N. (1985). "The Geometry of Phonological Features," *Phonology* 2, 223-252.
- Cochran, P. S. (1986). "Comparing the Intelligibility of Speech Synthesizers," *Proc. Am. Voice Input/Output Soc. AVIOS-86*, 464-474.
- Cohen, A., and Hart, J. (1967). "On the Anatomy of Intonation," *Lingua* 19, 177-192.
- Cohen, P., and Mercer, R. (1974). "The Phonological Component of an Automatic Speech Recognition System," in *Speech Recognition: Invited Papers Presented at the 1974 IEEE Symposium*, edited by R. Reddy (IEEE, New York), pp. 275-320.
- Coker, C. H. (1968). "Speech Synthesis with a Parametric Articulatory Model," *Speech Symposium*, paper A-4, Kyoto, reprinted in *Speech Synthesis*, edited by J. L. Flanagan and L. R. Rabiner (Dowden, Hutchinson and Ross, Stroudsburg, PA), pp. 135-139.
- Coker, C. H. (1976). "A Model of Articulatory Dynamics and Control," *Proc. IEEE* 64, 452-459.
- Coker, C. H. (1985). "A Dictionary-Intensive Letter-to-Sound Program," *J. Acoust. Soc. Am. Suppl.* 1 78, S7.
- Coker, C. H., and Umeda, N. (1975). "The Importance of Spectral Detail in Initial-Final Contrasts of Voiced Stops," *J. Phonetics* 3, 63-68.
- Coker, C. H., Umeda, N., and Browman, C. P. (1973). "Automatic Synthesis from Ordinary English Text," *IEEE Trans. Audio Electroacoust.* AU-21, 293-297.
- Conroy, D., and Vitale, T. (1986). Personal communication.
- Conroy, D., Vitale, T., and Klatt, D. H. (1986). "DECTalk DTC03 Text-to-Speech System Owner's Manual," EK-DTC03-OM-001, Educational Services of Digital Equipment Corporation, P.O. Box CS2008, Nashua, NH 03061.
- Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman, L. J. (1952). "Some Experiments on the Perception of Synthetic Speech Sounds," *J. Acoust. Soc. Am.* 24, 597-606.
- Cooper, F. S., Gaitenby, J. H., Mattingly, I. G., Nye, P. W., and Sholes, G. N. (1973). "Audible Outputs of Reading Machines for the Blind," *Status Report on Speech Research SR-35/36*, Haskins Laboratories, New Haven, CT, pp. 117-120.
- Cooper, F. S., Gaitenby, J. H., and Nye, P. W. (1984). "Evolution of Reading Machines for the Blind: Haskins Laboratories' Research as a Case History," *J. Rehab. Res. Develop.* 21, 51-87.
- Cooper, F. S., Liberman, A. M., and Borst, J. M. (1951). "The Interconversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech," *Proc. Natl. Acad. Sci. (US)* 37, 318-325.
- Cooper, W. A., Paccia, J. M., and Lapointe, S. G. (1978). "Hierarchical Coding in Speech Timing," *Cognitive Psychol.* 10, 154-177.
- Cooper, W. E., and Sorenson, J. (1981). *Fundamental Frequency in Sentence Production* (Springer, New York).
- Dedina, M. J., and Nusbaum, H. C. (1986). "Pronounce: A Program for Pronunciation by Analogy," *Speech Research Laboratory Progress Report 12*, Indiana University, Bloomington, IN, pp. 335-348.
- Davidson-Nielsen, N. (1974). "Syllabification of English Words with Medial sp, st, sk," *J. Phonetics* 2, 15-45.
- Delattre, P., Liberman, A. M., and Cooper, F. S. (1955). "Acoustic Loci and Transitional Cues for Consonants," *J. Acoust. Soc. Am.* 27, 769-774.
- Dixon, N. R., and Maxey, H. D. (1968). "Terminal Analog Synthesis of

- Continuous Speech Using the Diphone Method of Segment Assembly," *IEEE Trans. Audio Electroacoust.* AU-16, 40-50.
- Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (1977). "Stop Consonant Recognition: Release Bursts and Formant Transitions as Functionally Equivalent Context-Dependent Cues," *Percept. Psychophys.* 22, 109-122.
- Dubno, J. R., and Levitt, H. (1981). "Predicting Consonant Confusions from Acoustic Analysis," *J. Acoust. Soc. Am.* 69, 249-261.
- Dudley, H. (1939). "The Vocoder," *Bell Labs Rec.* 17, 122-126.
- Dudley, H., Riesz, R. R., and Watkins, S. A. (1939). "A Synthetic Speaker," *J. Franklin Inst.* 227, 739-764.
- Dudley, H., and Tarnoczy, T. H. (1950). "The Speaking Machine of Wolfgang Kempelen," *J. Acoust. Soc. Am.* 22, 151-166.
- Dunn, H. K. (1950). "The Calculation of Vowel Resonances, and Electrical Vocal Tract," *J. Acoust. Soc. Am.* 22, 740-753.
- Egan, J. P. (1948). "Articulation Testing Methods," *Laryngoscope* 58, 955-991.
- Elovitz, H., Johnson, R., McHugh, A., and Shore, J. (1976). "Letter-to-Sound Rules for Automatic Translation of English Text to Phonetics," *IEEE Trans. Acoust. Speech Signal Process.* ASSP-24, 446-459.
- Erber, N. P. (1979). "An Approach to Evaluating Auditory Speech Perception Ability," *Volta Rev.* 80, 344-350.
- Estes, S. E., Kerby, H. R., Maxey, H. D., and Walker, R. M. (1964). "Speech Synthesis from Stored Data," *IBM J. Res. Develop.* 8, 2-12.
- Fairbanks, G. (1958). "Test of Phonemic Differentiation: the Rhyme Test," *J. Acoust. Soc. Am.* 30, 596-600.
- Fant, G. (1953). "Speech Communication Research," *Ing. Vetenskaps Akad. Stockholm, Sweden* 24, 331-337.
- Fant, G. (1956). "On the Predictability of Formant Levels and Spectrum Envelopes from Formant Frequencies," in *For Roman Jakobson*, edited by M. Halle (Mouton, 's-Gravenhage, The Netherlands), pp. 109-120.
- Fant, G. (1959). "Acoustic Analysis and Synthesis of Speech with Applications to Swedish," *Ericsson Technics* 15, 3-108.
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, 's-Gravenhage, The Netherlands).
- Fant, G. (1973). "Stops in CV Syllables," in *Speech Sounds and Features*, edited by G. Fant (MIT Press, Cambridge, MA), pp. 110-139.
- Fant, G. (1975). "Non-Uniform Vowel Normalization," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm QPSR* 2-3, 1-19.
- Fant, G. (1979). "Glottal Source and Excitation Analysis," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 1, 85-107.
- Fant, G. (1982). "Preliminaries to Analysis of the Human Voice Source," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 4, 1-25.
- Fant, G., and Ananthapadmanabha, T. V. (1982). "Truncation and Superposition," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 2-3, 1-17.
- Fant, G., Lin, Q. G., and Gobl, C. (1985). "Notes on Glottal Flow Interaction," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm QPSR* 2-3, 21-45.
- Fant, G., and Martony, J. (1962). "Speech Synthesis," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 2, 18-24.
- Fant, G., Stalhammer, U., and Karlsson, I. (1974). "Swedish Vowels in Speech Materials of Various Complexity," in *Speech Communication*, edited by G. Fant (Almqvist and Wiksell, Uppsala, Sweden), Vol. 2, pp. 139-147.
- Fischer, F. W., Shankweiler, D., and Liberman, I. Y. (1985). "Spelling Proficiency and Sensitivity to Word Structure," *J. Memory Lang.* 24, 423-441.
- Fischer-Jorgensen, E. (1954). "Acoustic Analysis of Stop Consonants," *Miscel. Phonetica* 2, 42-59.
- Flanagan, J. L. (1957). "Note on the Design of Terminal Analog Speech Synthesizers," *J. Acoust. Soc. Am.* 29, 306-310.
- Flanagan, J. L. (1958). "Some Properties of the Glottal Sound Source," *J. Speech Hearing Res.* 1, 99-116.
- Flanagan, J. L. (1972). *Speech Analysis Synthesis and Perception* (Springer, New York).
- Flanagan, J. L. (1976). "Computers that Talk and Listen: Man-Machine Communication by Voice," *Proc. IEEE* 64, 405-415.
- Flanagan, J. L. (1981). "Synthesis and Recognition of Speech: How Computers Talk," *Bell Labs. Rec.* 59 (April), 123-130.
- Flanagan, J. L., and Ishizaka, K. (1976). "Automatic Generation of Voiceless Excitation to a Vocal Cord-Vocal Tract Speech Synthesizer," *IEEE Trans. Acoust. Speech Signal Process.* ASSP-24, 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, 1558-1563.
- 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 Syst. Tech. J.* 54, 485-506.
- Flanagan, J. L., and Landgraf, L. (1968). "Self-Oscillating Source for Vocal-Tract Synthesizers," *IEEE Trans. Audio Electroacoust.* AU-16, 57-64.
- Flanagan, J. L., and Rabiner, L. R. (1973). *Speech Synthesis* (Dowden, Hutchinson and Ross, Stroudsburg, PA).
- Flanagan, J. L., Rabiner, L. R., Schafer, R. W., and Denman, J. D. (1972). "Wiring Telephone Apparatus from Computer Generated Speech," *Bell System Tech. J.* 51, 391-397.
- Francis, W. N., and Kucera, H. (1982). *Frequency Analysis of English Usage* (Houghton Mifflin, Boston, MA).
- Friedman, R. H. (1986). Private communication.
- Fromkin, V. (1971). "The Non-anomalous Nature of Anomalous Utterances," *Language* 47, 27-52.
- Fry, D. B. (1958). "Experiments in the Perception of Stress," *Lang. Speech* 1, 126-152.
- Fry, D. B. (1965). "The Dependence of Stress Judgements on Vowel Formant Structure," in *Proceedings of the 6th International Congress of Phonetic Sciences*, edited by E. Zwirner and W. Bethge (Karger, Basel), pp. 306-311.
- Fry, D. B. (1974). "Phonetics in the Twentieth Century," in *Current Trends in Linguistics*, edited by T. S. Sebeok (Mouton, 's-Gravenhage, The Netherlands), Vol. 12, pp. 2201-2239.
- Fujimura, O. (1960). "Spectra of Nasalized Vowels," *Research Laboratory of Electronics QPR* 62, MIT, Cambridge, MA, pp. 214-218.
- Fujimura, O. (1962). "Analysis of Nasal Consonants," *J. Acoust. Soc. Am.* 34, 1865-1875.
- Fujimura, O. (1968). "An Approximation to Voice Aperiodicity," *IEEE Trans. Audio Electroacoust.* AU-16, 68-72.
- Fujimura, O., and Kakita, Y. (1979). "Remarks on Quantitative Description of the Lingual Articulation," in *Frontiers of Speech Communication Research*, edited by B. Lindblom and S. Ohman (Academic, New York), pp. 17-24.
- Fujimura, O., and Lovins, J. (1978). "Syllables as Concatenative Phonetic Elements," in *Syllables and Segments*, edited by A. Bell and J. B. Hooper (North-Holland, New York), pp. 107-120.
- Fujisaki, H., and Nagashima, S. (1969). "Synthesis of Pitch Contours of Connected Speech," *Annual Report Engineering Res. Inst. Univ. Tokyo*, Tokyo, Japan 28, 53-60.
- Fujisaki, H. (1986). "Proposal and Evaluation of Models for the Glottal Source Waveform," *Proc. Int. Conf. Acoust. Speech Signal Process.* ICASSP-86, 1609-1612.
- Gagnon, R. T. (1978). "Votrax Real Time Hardware for Phoneme Synthesis of Speech," *Proc. Int. Conf. Acoust. Speech Signal Process.* ICASSP-78, 175-178.
- Gaitenby, J. (1965). "The Elastic Word," *Status Report on Speech Research SR-2*, Haskins Laboratories, New Haven, CT, pp. 1-12.
- Gaitenby, J. H., Sholes, G. N., and Kuhn, G. M. (1972). "Word and Phrase Stress by Rule for a Reading Machine," *Status Report on Speech Research SR-29/30*, Haskins Laboratories, New Haven, CT, pp. 105-110.
- Gerstman, L. (1957). "Perceptual Dimensions for the Fricative Noise Portions of Certain Speech Sounds," unpublished Ph.D. dissertation, New York University, New York.
- Glushko, R. J. (1981). "Principles for Pronouncing Print: The Psychology of Phonography," in *Interactive Processes in Reading*, edited by A. M. Lesgold and C. A. Perfetti (Erlbaum, Hillsdale, NJ).
- Gold, B., and Rabiner, L. R. (1968). "Analysis of Digital and Analog Formant Synthesizers," *IEEE Trans. Audio Electroacoust.* AU-16, 81-94.
- Goldhor, R. S., and Lund, R. T. (1983). "University-to-Industry Advanced Technology Transfer: A Case Study," *Res. Policy* 12, 121-152.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech* (Academic, New York).
- Goldstein, L., and Browman, C. P. (1986). "Representation of Voicing Contrasts Using Articulatory Gestures," *J. Phonetics* 14, 339-342.
- Groner, G. F., Bernstein, J., Ingber, E., Pearlman, J., and Toal, T. (1982). "A Real-Time Text-to-Speech Converter," *Speech Technol.* 1, 73-76.

- Haggard, M. (1973). "Abbreviation of Consonants in English Pre- and Post-Vocalic Clusters," *J. Phonetics* 1, 9-24.
- Halle, M. (1985). "Speculations about the Representation of Words in Memory," in *Phonetic Linguistics*, edited by V. Fromkin (Academic, New York), pp. 101-114.
- Halle, M., Hughes, G. W., and Radley, J. P. A. (1957). "Acoustic Properties of Stop Consonants," *J. Acoust. Soc. Am.* 20, 107-116.
- Halle, M., and Keyser, S. J. (1971). *English Stress: Its Form, its Growth and its Role in Verse* (Harper and Row, New York).
- Halle, M., and Vergnaud, J. R. (1980). "Three-Dimensional Phonology," *J. Ling. Res.* 1, 83-105.
- Halliday, M. A. K. (1970). "Functional Diversity in Language as seen from a Consideration of Modality and Mood in English," *Found. Lang.* 6, 322-361.
- Harris, C. M. (1953). "A Study of the Building Blocks of Speech," *J. Acoust. Soc. Am.* 25, 962-969.
- Harris, K. (1958). "Cues for Discrimination of American English Fricatives in Spoken Syllables," *Lang. Speech* 1, 1-17.
- Harris, M. O., Umeda, N., and Bourne, J. (1981). "Boundary Perception in Fluent Speech," *J. Phonetics* 9, 1-18.
- 't Hart, J., and Cohen, A. (1973). "Intonation by Rule: A Perceptual Quest," *J. Phonetics* 1, 309-327.
- Hecker, M. H. L. (1962). "Studies of Nasal Consonants with an Articulatory Speech Synthesizer," *J. Acoust. Soc. Am.* 34, 179-188.
- Heffner, R. S. (1969). *General Phonetics* (Univ. Wisconsin, Madison, WI).
- Heinz, J. M., and Stevens, K. N. (1961). "On the Properties of Voiceless Fricative Consonants," *J. Acoust. Soc. Am.* 33, 589-596.
- Henderson, L. (1982). *Orthography and Word Recognition in Reading* (Academic, New York).
- Henke, W. L. (1967). "Preliminaries to Speech Synthesis Based upon an Articulatory Model," *Proc. IEEE Conf. Speech Commun. Process.* 170-182.
- Hertz, S. (1982). "From Text to Speech with SRS," *J. Acoust. Soc. Am.* 72, 1155-1170.
- Hertz, S., Kadin, J., and Karplus, K. (1985). "The Delta Rule Development System for Speech Synthesis from Text," *Proc. IEEE* 73, 1589-1601.
- Hiki, S. (1970). "Control Rule of the Tongue Movement for Dynamic Analog Speech Synthesis," *J. Acoust. Soc. Am. Suppl.* 1 47, S85.
- Hill, K., and Nessly, L. (1973). "Review of the Sound Pattern of English," *Linguistics* 106, 57-101.
- Hirsh, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eldest, E., and Benson, R. W. (1952). "Development of Materials for Speech Audiometry," *J. Speech Hear. Disord.* 17, 321-337.
- Hirshberg, J., and Pierrehumbert, J. (1986). "The Intonational Structuring of Discourse," *Proc. 24th Meeting Assoc. Comp. Linguistics*, Columbia Univ., New York, pp. 136-144.
- Hoard, J. E. (1966). "Juncture and Syllable Structure in English," *Phonetica* 15, 96-109.
- Hoard, J. E. (1971). "Aspiration, Tenseness and Syllabification in English," *Language* 47, 133-140.
- Hockett, C. (1955). *Manual of Phonetics* (Waverley, Baltimore, MD).
- Holbrook, A., and Fairbanks, G. (1962). "Diphthong Formants and their Movements," *J. Speech Hear. Res.* 5, 38-58.
- Holmes, J. N. (1961). "Research on Speech Synthesis Carried out during a Visit to the Royal Institute of Technology, Stockholm, from November 1960 to March 1961," *Joint Speech Research Unit Report JU 11.4*, British Post Office, Eastcote, England.
- Holmes, J. N. (1973). "The Influence of the Glottal Waveform on the Naturalness of Speech from a Parallel Formant Synthesizer," *IEEE Trans. Audio Electroacoust.* AU-21, 298-305.
- Holmes, J. N. (1983). "Formant Synthesizers: Cascade or Parallel," *Speech Commun.* 2, 251-273.
- Holmes, J. N. (1984). "Speech Technology in the Next Decades," *Proceedings of the Tenth International Congress of Phonetic Sciences*, edited by M. P. R. Van den Broecke and A. Cohen (Foris, Dordrecht, The Netherlands), pp. 125-139.
- Holmes, J. N., Mattingly, I. G., and Shearme, J. N. (1964). "Speech Synthesis by Rule," *Lang. Speech* 7, 127-143.
- House, A. S. (1961). "On Vowel Duration in English," *J. Acoust. Soc. Am.* 33, 1174-1178.
- House, A. S., and Fairbanks, G. (1953). "The Influence of Consonantal Environment upon the Secondary Acoustical Characteristics of Vowels," *J. Acoust. Soc. Am.* 25, 105-113.
- House, A. S., Williams, C. E., Hecker, M. H. L., and Kryter, K. D. (1965). "Articulation-Testing Methods: Consonant Differentiation with a Closed-Response Set," *J. Acoust. Soc. Am.* 37, 158-166.
- Huggins, A. W. F., Houde, R. A., and Colwell, E. (1986). "Vidvox Human Factors Investigation," Bolt Beranek and Newman Inc. Report No. 6187, Cambridge, MA, p. 15.
- Hunnicutt, S. (1976). "Phonological Rules for a Text-to-Speech System," *Am. J. Comp. Ling. Microfiche* 57, 1-72.
- Hunnicutt, S. (1980). "Grapheme-to-Phoneme Rules, A Review," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 2-3, 38-60.
- Hunnicutt, S. (1984). "Bliss Symbol-to-Speech Conversion: Blisstalk," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 1, 58-77.
- Hunnicutt, S. (1985). "Lexical Prediction for a Text-to-Speech System," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden QPSR* 2-3, 47-55.
- IEEE (1969). "Recommended Practice for Speech Quality Measurements," *IEEE Trans. Audio Electroacoust.* AU-17, 225-246.
- Ingemann, F. (1978). "Speech Synthesis by Rule using the FOVE Program," *Status Report on Speech Research SR-54*, Haskins Laboratories, New Haven, CT, 165-173.
- Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of Voiced Sounds from a Two-Mass Model of the Vocal Cords," *Bell Syst. Tech. J.* 51, 1233-1268.
- Ishizaka, K., and Matsudaira, M. (1968). "What Makes the Vocal Cords Vibrate?," in *The Sixth International Congress on Acoustics*, edited by Y. Kohasi (Elsevier, New York), Vol. II, pp. B9-B12.
- Itakura, F., and Saito, S. (1968). "Analysis-Synthesis Telephony based on the Maximum Likelihood Method," *Proc. 6th International Congress Acoustics*, Tokyo, paper C-5-5.
- Joos, M. A. (1948). "Acoustic Phonetics," *Language Suppl.* 24, 1-136.
- Kahn, M. (1975). "Arabic Emphatics: The Evidence for Cultural Determinants of Phonetic Sex-Typing," *Phonetica* 31, 38-50.
- Kalikow, D., Stevens, K. N., and Elliott, L. (1977). "Development of a Test of Speech Intelligibility in Noise Using Sentence Materials with Controlled Word Predictability," *J. Acoust. Soc. Am.* 61, 1337-1351.
- Kaplan, G., and Lerner, E. J. (1985). "Realism in Synthetic Speech," *IEEE Spectrum* 22 (April), 32-37.
- Kaplan, R., and Bresnan, J. (1982). "Lexical-Functional Grammar: A Formal System for Grammatical Representation," in *The Mental Representation of Grammatical Relations*, edited by J. Bresnan (MIT, Cambridge, MA).
- Kelly, J., and Gerstman, L. (1961). "An Artificial Talker Driven from Phonetic Input," *J. Acoust. Soc. Am. Suppl.* 1 33, S35.
- Kelly, J., and Gerstman, L. (1962). "Digital Computer Synthesizes Human Speech," *Bell Labs. Rec.* 40, 216-217.
- Kelly, J., and Gerstman, L. (1964). "Synthesis of Speech from Code Signals," U.S. Patent 3,158,685.
- Kelly, J., and Lochbaum, C. (1962). "Speech Synthesis," *Proc. Fourth Int. Congr. Acoust.* paper G42, 1-4 (Reprinted in Flanagan and Rabiner, 1973).
- Kewley-Port, D. (1982). "Measurement of Formant Transitions in Naturally Produced Stop Consonant-Vowel Syllables," *J. Acoust. Soc. Am.* 72, 379-389.
- Klatt, D. H. (1970). "Synthesis of Stop Consonants in Initial Position," *J. Acoust. Soc. Am. Suppl.* 1 47, S93.
- Klatt, D. H. (1972). "Acoustic Theory of Terminal Analog Speech Synthesis," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-72*, 131-135.
- Klatt, D. H. (1973a). "Durational Characteristics of Prestressed Word-Initial Consonant Clusters in English," *Research Laboratory of Electronics QPR* 108, MIT, Cambridge, MA, pp. 253-260.
- Klatt, D. H. (1973b). "Interaction between Two Factors that Influence Vowel Duration," *J. Acoust. Soc. Am.* 54, 1102-1104.
- Klatt, D. H. (1974). "The Duration of [S] in English Words," *J. Speech Hear. Res.* 17, 51-63.
- Klatt, D. H. (1975a). "Vowel Lengthening is Syntactically Determined in a Connected Discourse," *J. Phonetics* 3, 129-140.
- Klatt, D. H. (1975b). "Voice Onset Time, Frication, and Aspiration in Word-Initial Consonant Clusters," *J. Speech Hear. Res.* 18, 686-706.
- Klatt, D. H. (1976a). "Linguistic Uses of Segmental Duration in English: Acoustic and Perceptual Evidence," *J. Acoust. Soc. Am.* 59, 1208-1221.
- Klatt, D. H. (1976b). "Structure of a Phonological Rule Component for a

- Speech Synthesis by Rule Program," IEEE Trans. Acoust. Speech Signal Process. **ASSP-24**, 391-398.
- Klatt, D. H. (1979a). "Synthesis by Rule of Segmental Durations in English Sentences," in *Frontiers of Speech Communication Research*, edited by B. Lindblom and S. Öhman (Academic, New York), pp. 287-300.
- Klatt, D. H. (1979b). "Synthesis by Rule of Consonant-Vowel Syllables," Speech Communication Group Working Papers 3, MIT, Cambridge, MA, pp. 93-104.
- Klatt, D. H. (1980). "Software for a Cascade/Parallel Formant Synthesizer," J. Acoust. Soc. Am. **67**, 971-995.
- Klatt, D. H. (1981). "A Text-to-Speech Conversion System," Proc. AFIPS Office Automation Conference, pp. 51-61.
- Klatt, D. H. (1982a). "The Klattalk Text-to-Speech System," Proc. Int. Conf. Acoust. Speech Signal Process. **ICASSP-82**, 1589-1592.
- Klatt, D. H. (1982b). "A Strategy for the Perceptual Interpretation of Durational Cues," Speech Communication Group Working Papers 1, MIT, Cambridge, MA, pp. 83-91.
- Klatt, D. H. (1982c). "Prediction of Perceived Phonetic Distance from Critical-Band Spectra: A First Step," Proc. Int. Conf. Acoust. Speech Signal Process. **ICASSP-82**, 1278-1281.
- Klatt, D. H. (1986a). "Representation of the First Formant in Speech Recognition and in Models of the Auditory Periphery," in Proc. Montreal Satellite Symposium on Speech Recognition, edited by P. Mermelstein, Twelfth Int. Cong. Acoustics, Toronto, Canada, pp. 5-7.
- Klatt, D. H. (1986b). "Detailed Spectral Analysis of a Female Voice," J. Acoust. Soc. Am. Suppl. **1 80**, S97.
- Klatt, D. H. (1987). "How Klattalk became DECTalk: An Academic's Experiences in the Business World," Speech Tech. **87**, 293-294.
- Klatt, D. H., and Aoki, C. (1984). "Synthesis by Rule of Japanese," J. Acoust. Soc. Am. Suppl. **1 76**, S2.
- Klatt, D. H., and Shipman, D. W. (1982). "Letter-to-Phoneme Rules: A Semi-Automatic Discovery Procedure," J. Acoust. Soc. Am. Suppl. **1 72**, S48.
- Koenig, W. H., Dunn, H. K., and Lacey, L. Y. (1946). "The Sound Spectrograph," J. Acoust. Soc. Am. **18**, 19-49.
- Kučera, H., and Francis, W. N. (1967). *Computational Analysis of Present day American English* (Brown U.P., Providence, RI).
- Kurzweil, R. (1976). "The Kurzweil Reading Machine: A Technical Overview," in *Science, Technology and the Handicapped*, edited by M. R. Redden and W. Schwandt (American Association for the Advancement of Science, Report 76-R-11, Washington, DC), pp. 3-11.
- Labov, W. (1986). "Sources of Inherent Variation in the Speech Process," in *Invariance and Variability in Speech Processes*, edited by J. Perkell and D. H. Klatt (Erlbaum, Hillsdale, NJ), pp. 402-425.
- Ladd, D. R. (1983). "Phonological Features of Intonational Peaks," Language **59**, 721-759.
- Ladefoged, P. (1973). "The Features of the Larynx," J. Phonetics **1**, 73-83.
- Lamel, L., and Zue, V. (1984). "Properties of Consonant Sequences within Words and across Word Boundaries," Proc. Int. Conf. Acoust. Speech Signal Process. **ICASSP-84**, 42.3.1-43.2.4.
- Lawrence, W. (1953). "The Synthesis of Speech from Signals which have a Low Information Rate," in *Communication Theory*, edited by W. Jackson (Butterworths, London, England), pp. 460-469.
- Lee, D., and Lochovsky, F. (1983). "Voice Response Systems," ACM Computing Surveys **15**, 351-374.
- Lee, F. F. (1969). "Reading Machine: From Text to Speech," IEEE Trans. Audio Electroacoust. **AU-17**, 275-282.
- Lehiste, I. (1959). "An Acoustic-Phonetic Study of Internal Open Juncture," Suppl. to *Phonetica* **5**, 1-55.
- Lehiste, I. (1962). "Acoustical Characteristics of Selected English Consonants," Univ. Michigan Speech Research Lab. Report **9**, 1-219.
- Lehiste, I. (1964). "Juncture," in *Proceedings of the Fifth International Congress of Phonetic Sciences*, edited by E. Zwirner and W. Bethge (Karger, Basel, Switzerland), pp. 172-200.
- Lehiste, I. (1967). *Readings in Acoustic Phonetics* (MIT Press, Cambridge, MA).
- Lehiste, I. (1970). *Suprasegmentals* (MIT Press, Cambridge, MA).
- Lehiste, I. (1975a). "Some Factors Affecting the Duration of Syllabic Nuclei in English," *Proceedings of the First Salzburg Conference on Linguistics*, edited by G. Drachman (Verlag Gunter, Narr), pp. 81-104.
- Lehiste, I. (1975b). "The Phonetic Structure of Paragraphs," in *Structure and Process in Speech Perception*, edited by A. Cohen and S. Nooteboom (Springer, Heidelberg, Germany), pp. 195-206.
- Lehiste, I. (1977). "Isochrony Reconsidered," J. Phonetics **5**, 253-263.
- Lehiste, I. Olive, J. P., and Streeter, L. A. (1976). "The Role of Duration in Disambiguating Syntactically Ambiguous Sentences," J. Acoust. Soc. Am. **60**, 1199-1202.
- Lehiste, I., and Peterson, G. E. (1959). "Linguistic Considerations in the Study of Speech Intelligibility," J. Acoust. Soc. Am. **31**, 280-287.
- Lehiste, I., and Peterson, G. E. (1961). "Some Basic Considerations in the Analysis of Intonation," J. Acoust. Soc. Am. **33**, 419-425.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). "Perception of the Speech Code," Psychol. Rev. **74**, 431-461.
- Liberman, A. M., Delattre, P., and Cooper, F. S. (1958). "Some Cues for the Distinction between Voiced and Voiceless Stops in Initial Position," Lang. Speech **1**, 153-167.
- Liberman, A. M., Delattre, P., Cooper, F. S., and Gerstman, L. J. (1954). "The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants," Psychol. Monogr. **68**, 1-13.
- Liberman, A. M., Ingemann, F., Lisker, L., Delattre, P., and Cooper, F. (1959). "Minimal Rules for Synthesizing Speech," J. Acoust. Soc. Am. **31**, 1490-1499.
- Liberman, A. M., and Mattingly, I. G. (1985). "The Motor Theory of Speech Perception Revisited," Cognition **21**, 1-36.
- Liberman, M. Y. (1979). "Phonemic Transcription, Stress, and Segment Durations for Spelled Proper Names," J. Acoust. Soc. Am. Suppl. **1 64**, S163.
- Lieberman, P. (1967). *Intonation, Perception and Language* (MIT Press, Cambridge, MA).
- Liljencrants, J. (1969). "Speech Synthesizer Control by Smoothed Step Functions," Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden **QPSR-4**, 43-50.
- Liljencrants, J. (1985). "Speech Synthesis with a Reflection-Type Line Analog," unpublished Ph.D. thesis, Dept. Speech Commun. and Musical Acoust., Royal Inst. of Tech., Stockholm, Sweden.
- Lindblom, B. (1963). "Spectrographic Study of Vowel Reduction," J. Acoust. Soc. Am. **35**, 1773-1781.
- Lisker, L. (1957). "Minimal Cues for Separating /wrl/ in Intervocalic Position," Word **13**, 256-267.
- Lisker, L. (1978). "Rapid vs. Rabid: A Catalog of Acoustic Features that may Cue the Distinction," Status Report on Speech Research **54**, Haskins Laboratories, New Haven, CT, pp. 127-132.
- Lisker, L., and Abramson, A. S. (1967). "Some Effects of Context on Voice Onset Time in English Stops," Lang. Speech **10**, 1-28.
- Logan, J. S., and Pisoni, D. B. (1986). "Preference Judgements Comparing Different Synthetic Voices," J. Acoust. Soc. Am. Suppl. **1 79**, S24.
- Logan, J. S., Pisoni, D. B., and Greene, B. G. (1986). "Measuring the Segmental Intelligibility of Synthetic Speech: Results from Eight Text-to-Speech Systems," submitted to J. Acoust. Soc. Am.
- Lucassen, J. M., and Mercer, R. L. (1984). "An Information Theoretic Approach to the Automatic Determination of Phonemic Base Forms," Proc. Int. Conf. Acoust. Speech Signal Process. **ICASSP-84**, 42.5.1-42.5.4.
- Luce, P., Feustel, T., and Pisoni, D. (1983). "Capacity Demands in Short-Term Memory for Synthetic and Natural Speech," Human Factors **25**, 17-31.
- MacNeilage, P. F., and DeClerk, J. L. (1969). "On the Motor Control of Coarticulation of CVC Syllables," J. Acoust. Soc. Am. **45**, 1217-1233.
- Maeda, S. (1974). "A Characterization of Fundamental Frequency Contours of Speech," Research Laboratory of Electronics QPR **114**, MIT, Cambridge, MA, pp. 193-211.
- Maeda, S. (1987). "On the Generation of Sounds in Stop Consonants," Speech Communication Group Working Papers V, MIT, Cambridge, MA, pp. 1-14.
- Magnusson, L., Blomberg, M., Carlson, R., Elenius, K., and Granström, B. (1984). "Swedish Speech Researchers Team Up with Electronic Venture Capitalists," Speech Technol. **2**, 15-24.
- Makhoul, J. (1973). "Spectral Analysis of Speech by Linear Prediction," IEEE Trans. Audio Electroacoust. **AU-21**, 140-148.
- Malecot, A. (1956). "Acoustic Cues for Nasal Consonants: An Experimental Study Involving a Tape-Splicing Technique," Language **32**, 274-284.
- Malme, C. I. (1959). "Detectability of Small Irregularities in a Broadband Noise Spectrum," Research Lab. of Electronics Q.P.R. **52**, Mass. Inst. Tech., pp. 139-141.
- Manous, L. M., Pisoni, D. B., Dedina, M. J., and Nusbaum, H. C. (1985). "Comprehension of Natural and Synthetic Speech Using a Sentence Verification Task," Speech Research Laboratory Progress Report **11**, In-

- diana University, Bloomington, IN, pp. 33–58.
- Marcus, M. (1980). *A Theory of Syntactic Recognition for Natural Language* (MIT Press, Cambridge, MA).
- Markel, J. D. (1972). "Digital Inverse Filtering: A New Tool for Formant Trajectory Estimation," *IEEE Trans. Audio Electroacoust.* AU-20, pp. 129–137.
- Martin, J. G., and Bunnell, H. T. (1982). "Perception of Anticipatory Coarticulation Effects in Vowel-Stop Consonant-Vowel Sequences," *J. Exp. Psychol.: Human Percept. Perform.* 8, 473–488.
- Mathews, M. V., Miller, J. E., and David, E. E. (1961). "Pitch-Synchronous Analysis of Voiced Sounds," *J. Acoust. Soc. Am.* 33, 179–186.
- Matsui, E., Suzuki, T., Umeda, N., and Omura, H. (1968). "Synthesis of Fairy Tales using an Analog Vocal Tract," *Proc. 6th International Congress on Acoustics*, Tokyo, Japan, B159–162.
- Mattingly, I. G. (1966). "Synthesis by Rule of Prosodic Features," *Lang. Speech* 9, 1–13.
- Mattingly, I. G. (1968). "Synthesis-by-Rule of General American English," Supplement to Status Report on Speech Research, Haskins Laboratories, New Haven, CT, pp. 1–223.
- Mattingly, I. G. (1974). "Speech Synthesis for Phonetic and Phonological Models," in *Current Trends in Linguistics*, edited by T. S. Sebeok, Mouton, 's-Gravenhage, The Netherlands, Vol. 12, pp. 2451–2487.
- Maxey, H. D. (1963). "Terminal Analog Synthesis of Voiced Fricatives," *J. Acoust. Soc. Am.* 35, 1890 (A).
- McIlroy, M. D. (1974). "Synthetic Speech by Rule," unpublished Technical Memo, Bell Laboratories, Murray Hill, NJ.
- McGillivray, R. (1983). "Terminals with Voice Output," in *Aids and Appliances Review* 9 (Carroll Center for the Blind, 770 Centre St., Newton, MA 02158), pp. 32–41.
- Mermelstein, P. (1973). "Articulatory Model for the Study of Speech Production," *J. Acoust. Soc. Am.* 53, 1070–1082.
- Miller, R. L. (1959). "Nature of the Vocal Cord Wave," *J. Acoust. Soc. Am.* 31, 667–677.
- Minow, M., and Klatt, D. H. (1983). "DECTalk DTC01 Text-to-Speech System Owner's Manual," EK-DTC01-OM-001, Educational Services of Digital Equipment Corporation, P.O. Box CS2008, Nashua, NH 03061.
- Monsen, R. B., and Engebretson, A. M. (1977). "Study of Variation in the Male and Female Glottal Wave," *J. Acoust. Soc. Am.* 62, 981–993.
- Morris, L. R. (1979). "A Fast Fortran Implementation of the NRL Algorithm for Automatic Translation of English Text to Votrax Parameters," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-79*, 907–913.
- Munson, W., and Karlin, J. (1962). "Isopreference Method for Evaluating Speech Transmission Circuits," *J. Acoust. Soc. Am.* 27, 1213–1219.
- Nakata, K., and Mitsuoka, T. (1965). "Phonemic Transformation and Control Aspects of Synthesis of Connected Speech," *J. Radio Res. Labs.* 12, 171–186.
- Nakatani, L. H., and Dukes, K. D. (1973). "Sensitive Test of Speech Communication Quality," *J. Acoust. Soc. Am.* 53, 1083–1092.
- Nakatani, L. H., and Dukes, K. D. (1977). "Locus of Segmental Cues for Word Juncture," *J. Acoust. Soc. Am.* 62, 714–719.
- Nakatani, L. H., Egan, D., Ruedisueli, L., and Hawley, P. (1986). "A Talking Tutor and Trainer for Teaching the Use of Interactive Computer Systems," *Proc. Conf. on Human Factors in Computing Systems (ACM, New York)*, pp. 29–34.
- Nakatani, L., and O'Connor-Dukes, K. (1979). "Phonetic Parsing Cues for Word Perception," unpublished Technical Memorandum, Bell Laboratories, Murray Hill, NJ.
- Nakatani, L. H., and Schaffer, J. (1978). "Hearing Words without Words: Prosodic Cues for Word Perception," *J. Acoust. Soc. Am.* 63, 234–244.
- NCHS (1977). "Pamphlet from the National Center for Health Statistics," Washington, DC.
- Nixon, C. W., Anderson, T. R., and Moore, T. J. (1985). "The Perception of Synthetic Speech in Noise," *Armstrong Aerospace Medical Research Laboratories Report*, Wright-Patterson AFB, OH 45433.
- Nusbaum, H. C., Dedina, M. J., and Pisoni, D. B. (1984). "Perceptual Confusions of Consonants in Natural and Synthetic CV Syllables," *Speech Research Laboratory Progress Report 10*, Indiana Univ., Bloomington, IN, pp. 409–422.
- Nusbaum, H. C., Pisoni, D. B., and Schwab, E. C. (1984). "Subjective Evaluation of Synthetic Speech: Measuring Preference, Naturalness and Intelligibility," *Speech Research Laboratory Progress Report 10*, Indiana Univ., Bloomington, IN, pp. 391–408.
- Nye, P., and Gaitenby, J. (1974). "The Intelligibility of Synthetic Monosyllable Words in Short, Syntactically Normal Sentences," *Status Report on Speech Research SR-37/38*, Haskins Laboratories, New Haven, CT, pp. 169–190.
- Nye, P., Hankins, J., Rand, T., Mattingly, I., and Cooper, F. (1973). "A Plan for the Field Evaluation of an Automated Reading System for the Blind," *IEEE Trans. Audio Electroacoust.* AU-21, 265–268.
- Ochsmann, R. B., and Chapanis, A. (1974). "The Effects of 10 Communication Modes on the Behavior of Teams during Cooperative Problem Solving," *Int. J. Man-Machine Studies* 6, 579–619.
- O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. G., and Cooper, F. S. (1957). "Acoustic Cues for the Perception of Initial /w,j,r,l/ in English," *Word* 13, 24–36.
- Öhman, S. E. G. (1966). "Coarticulation in VCV Utterances: Spectrographic Measurements," *J. Acoust. Soc. Am.* 39, 151–168.
- Öhman, S. E. G. (1967). "Word and Sentence Intonation: A Quantitative Model," *Speech Transmission Laboratory, Royal Institute of Technology, Stockholm QPSR* 2-3, 20–54.
- Olive, J. P. (1977). "Rule Synthesis of Speech from Diadic Units," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-77*, 568–570.
- Olive, J. P., and Liberman, M. Y. (1979). "A Set of Concatenative Units for Speech Synthesis," in *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America*, edited by J. J. Wolf and D. H. Klatt (American Institute of Physics, New York), pp. 515–518.
- Olive, J. P., and Liberman, M. Y. (1985). "Text-to-Speech—An Overview," *J. Acoust. Soc. Am. Suppl.* 1 78, S6.
- Olive, J. P., and Nakatani, L. H. (1974). "Rule Synthesis of Speech by Word Concatenation: A First Step," *J. Acoust. Soc. Am.* 55, 660–666.
- Olive, J. P., and Spickenagle, N. (1976). "Speech Resynthesis from Phoneme-Related Parameters," *J. Acoust. Soc. Am.* 59, 993–996.
- Oller, D. K. (1973). "The Effect of Position in Utterance on Speech Segment Duration in English," *J. Acoust. Soc. Am.* 54, 1235–1247.
- Olson, R., Foltz, G., and Wise, B. (1985). "Reading Instruction and Remediation with the Aid of Computer Speech," oral presentation, Meeting of the Society for Computers in Psychology, Boston, MA.
- O'Malley, M. H., Larkin, D. K., and Peters, E. W. (1986). "Beyond the Reading Machine: What the Next Generation of Intelligent Text-to-Speech Systems Should Do for the User," in *Proceedings Speech Tech '86* (Media Dimensions, New York), pp. 216–219.
- O'Shaughnessy, D. (1977). "Fundamental Frequency by Rule for a Text-to-Speech System," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-77*, 571–574.
- O'Shaughnessy, D. (1979). "Linguistic Features in Fundamental Frequency Patterns," *J. Phonetics* 7, 119–145.
- O'Shaughnessy, D., and Allen, J. (1983). "Linguistic Modality Effects on Fundamental Frequency in Speech," *J. Acoust. Soc. Am.* 74, 1155–1171.
- Oshika, B., Zue, V., Weeks, R., Neu, H., and Aurbach, J. (1975). "The Role of Phonological Rules in Speech Understanding Research," *IEEE Trans. Acoust. Speech Signal Process. ASSP-23*, 104–110.
- Peck, C. (1969). "An Acoustic Investigation of the Intonation of American English," *Nat. Lang. Studies I*, Univ. Michigan, Ann Arbor, MI, pp. 36–47.
- Peterson, G. E., and Barney, H. L. (1952). "Control Methods Used in a Study of the Vowels," *J. Acoust. Soc. Am.* 24, 175–184.
- Peterson, G. E., and Lehiste, I. (1960). "Duration of Syllabic Nuclei in English," *J. Acoust. Soc. Am.* 32, 693–703.
- Peterson, G., Wang, W., and Sivertsen, E. (1958). "Segmentation Techniques in Speech Synthesis," *J. Acoust. Soc. Am.* 30, 739–742.
- Pierrehumbert, J. (1981). "Synthesizing Intonation," *J. Acoust. Soc. Am.* 70, 985–995.
- Pike, K. L. (1945). *The Intonation of American English* (Univ. Michigan Press, Ann Arbor, MI).
- Pisoni, D. B. (1982). "Perception of Speech: The Human Listener as a Cognitive Interface," *Speech Technol.* 1, 10–23.
- Pisoni, D. B., and Hunnicutt, S. (1980). "Perceptual Evaluation of MITalk: The MIT Unrestricted Text-to-Speech System," *Proc. Int. Conf. Acoust. Speech Signal Process. ICASSP-80*, 572–575.
- Pisoni, D. B., and Koen, E. (1981). "Some Comparisons of Intelligibility of Synthetic and Natural Speech at Different Speech-to-Noise Ratios," *Speech Research Laboratory Progress Report 7*, Indiana Univ., Bloomington, IN, pp. 243–254.
- Pisoni, D. B., Nusbaum, H. C., and Greene, B. G. (1985). "Perception of Synthetic Speech Generated by Rule," *Proc. IEEE* 73, 1665–1676.
- Pols, L. C. W., and Olive, J. P. (1983). "Intelligibility of Consonants in CVC Utterances Produced by Diadic Rule Synthesis," *Speech Commun.*

- Portnoy, K. (Ed.) (1979-present). *Communication Outlook* (Artificial Language Lab, Michigan State Univ., Lansing, MI).
- Potter, R. K. (1946). "Introduction to Technical Discussions of Sound Portrayal," *J. Acoust. Soc. Am.* **18**, 1-3. (See also five related articles that follow.)
- Potter, R. K., Kopp, G. A., and Green, H. C. (1947). *Visible Speech* (van Nostrand, New York).
- Quarby, D. J., and Holmes, J. N. (1984). "Implementation of a Parallel-Formant Speech Synthesizer using a Single-Chip Programmable Signal Processor," *IEEE Proceed.* **131**, 563-569.
- Rabiner, L. R. (1968). "Speech Synthesis by Rule: An Acoustic Domain Approach," *Bell Syst. Tech. J.* **47**, 17-38.
- Rabiner, L. R. (1969). Private communication.
- Rabiner, L. R., Schafer, R. W., and Flanagan, J. L. (1971). "Computer Synthesis of Speech by Concatenation of Formant-Coded Words," *Bell System Tech. J.* **50**, 1541-1558.
- Repp, B. H. (1986). "Perception of the [m]-[n] Distinction in CV Syllables," *J. Acoust. Soc. Am.* **79**, 1987-1999.
- Rodet, X. (1984). "Time-Domain Formant-Wave-Function Synthesis," *Comput. Music J.* **8**, 9-14.
- Rosen, G. (1958). "A Dynamic Analog Speech Synthesizer," *J. Acoust. Soc. Am.* **30**, 201-209.
- Rosenberg, A. (1971). "Effect of Glottal Pulse Shape on the Quality of Natural Vowels," *J. Acoust. Soc. Am.* **49**, 583-590.
- Rothenberg, M., Carlson, R., Granström, B., and Gauffin, J. (1975). "A Three-Parameter Voice Source for Speech Synthesis," in *Speech Communication*, edited by G. Fant (Almqvist and Wiksell, Uppsala, Sweden), Vol. 2, pp. 235-243.
- Schwab, E. C., Nusbaum, H. C., and Pisoni, D. B. (1986). "Effects of Training on the Perception of Synthetic Speech," *Human Factors* **37**, 395-408.
- Schwartz, M. F. (1967). "Transitions in American English /s/ as Cues to the Identity of Adjacent Stop Consonants," *J. Acoust. Soc. Am.* **42**, 897-899.
- Sejnowski, T. J., and Rosenberg, C. R. (1986). "NETalk: A Parallel Network that Learns to Read Aloud," *Elec. Engr. Comp. Sci. Tech. Report JHU/EECS-86/01*, Johns Hopkins University, Baltimore, MD, pp. 1-23.
- Shearme, J. N., and Holmes, J. N. (1961). "An Experimental Study of the Classification of Sounds in Continuous Speech According to their Distribution in the Formant 1-Formant 2 Plane," *Proc. Fourth Int. Cong. Phonetic Sci.*, pp. 234-240.
- Sherwood, B. A. (1979). "The Computer Speaks," *IEEE Spectrum*, August, 18-25.
- Sherwood, B. A. (1981). "New Technology Provides Computer Voices for Education," *Speech Technol.* **1**, 24-29.
- Shoup, J. E. (1980). "Phonological Aspects of Speech Recognition," in *Trends in Speech Recognition*, edited by W. A. Lea (Prentice-Hall, New York), pp. 125-138.
- Sivertsen, E. (1961). "Segment Inventories for Speech Synthesis," *Lang. Speech* **4**, 27.
- Social Security Administration (1985). *Report of Distribution of Surnames in the Social Security Number File, September 1, 1984* (Department of Health and Human Resources, Washington, DC), Social Security Administration Pub. No. 42-004.
- Stevens, K. N. (1972). "The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data," in *Human Communication: A Unified View*, edited by E. E. David and P. B. Denes (McGraw-Hill, New York), pp. 51-66.
- Stevens, K. N. (1977). "Physics of Larynx Behavior and Larynx Modes," *Phonetica* **34**, 264-279.
- Stevens, K. N. (1987). Book in preparation.
- Stevens, K. N., Fant, G., and Hawkins, S. (1987). "Some Acoustical and Perceptual Correlates of Nasal Vowels," in *Festschrift for Ilse Lehiste*, edited by R. Channon and L. Shockey (Foris, Dordrecht, The Netherlands), pp. 241-254.
- Stevens, K. N., and House, A. S. (1955). "Development of a Quantitative Description of Vowel Articulation," *J. Acoust. Soc. Am.* **27**, 484-493.
- Stevens, K. N., and House, A. S. (1961). "An Acoustical Theory of Vowel Production and Some of Its Implications," *J. Speech Hear. Res.* **4**, 303-320.
- Stevens, K. N., and House, A. S. (1963). "Perturbation of Vowel Articulation by Consonantal Context: An Acoustical Study," *J. Speech Hear. Res.* **6**, 111-128.
- Stevens, K. N., Kasowski, S., and Fant, G. (1953). "An Electrical Analog of the Vocal Tract," *J. Acoust. Soc. Am.* **25**, 734-742.
- Stewart, J. Q. (1922). "An Electrical Analogue of the Vocal Organs," *Nature* **110**, 311-312.
- Streeter, L. (1978). "Acoustic Determinants of Phrase Boundary Perception," *J. Acoust. Soc. Am.* **64**, 1582-1592.
- Sundberg, J., and Gauffin, J. (1979). "Waveform and Spectrum of Glottal Voice Source," in *Frontiers of Speech Communication Research*, edited by B. Lindblom and S. Ohman (Academic, London) pp. 301-322.
- Suppes, P. (1979). "Current Trends in Computer Assisted Instruction," in *Advances in Computers* (Academic, New York).
- Suppes, P. (Ed.) (1981). *University-Level Computer-Assisted Instruction at Stanford: 1968-1980* (Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, CA).
- Svensson, S. G. (1974). "Prosody and Grammar in Speech Perception," *Monographs from the Institute of Linguistics* (University of Stockholm, Stockholm, Sweden).
- Teranishi, R., and Umeda, N. (1968). "Use of Pronouncing Dictionary in Speech Synthesis Experiments," Reports of the Sixth International Congress on Acoustics, Tokyo, Vol. II, pp. 155-158.
- Titze, I. R. (1974). "The Human Vocal Cords: a Mathematical Model," *Phonetica* **29**, 1-21.
- Titze, I. R., and Talkin, D. (1979). "A Theoretical Study of the Effects of the Various Laryngeal Configurations on the Acoustics of Phonation," *J. Acoust. Soc. Am.* **66**, 60-74.
- Uldall, E. (1960). "Attitudinal Meanings Conveyed by Intonation Contours," *Lang. Speech* **3**, 223-234.
- Umeda, N. (1975). "Vowel Duration in American English," *J. Acoust. Soc. Am.* **58**, 434-445.
- Umeda, N. (1976). "Linguistic Rules for Text-to-Speech Synthesis," *Proc. IEEE* **64**, 443-451.
- Umeda, N. (1977). "Consonant Duration in American English," *J. Acoust. Soc. Am.* **61**, 846-858.
- Umeda, N. (1981). "Influence of Segmental Factors on Fundamental Frequency in Fluent Speech," *J. Acoust. Soc. Am.* **70**, 350-355.
- Umeda, N., and Coker, C. H. (1974). "Allophonic Variation in American English," *J. Phonetics* **2**, 1-5.
- Umeda, N., Matsui, E., Suzuki, T., and Omura, H. (1968). "Synthesis of Fairy Tales using an Analog Vocal Tract," *Proc. 6th International Congress on Acoustics, Tokyo, Japan*, pp. B159-162.
- Umeda, N., and Quinn, A. M. S. (1981). "Word Duration as an Acoustic Measure of Boundary Perception," *J. Phonetics* **9**, 19-28.
- Umeda, N., and Teranishi, R. (1975). "The Parsing Program for Automatic Text-to-Speech Synthesis Developed at the Electrotechnical Laboratory in 1968," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-23**, 183-188.
- Vanderheiden, G. C. (1978). *Nonvocal Communication Resource Book* (University Park, Baltimore, MD).
- Vanderheiden, G. C. (1985). "1985 Updates for the Nonvocal Communication Resource Book," *Trace Center of the Waisman Center, University of Wisconsin, Madison, WI*.
- Vanderslice, R. (1968). "Synthetic Elocution," *Working Papers in Phonetics, UCLA, Los Angeles, CA* **8**, 1-131.
- Vanderslice, R., and Ladefoged, P. (1971). "Binary Suprasegmental Features and Transformational Word-Accentuation Rules," *Language* **48**, 819-839.
- Venezky, R. L. (1965). *A Study of English Spelling-to-Sound Correspondences on Historical Principles* (Ann Arbor Press, Ann Arbor, MI).
- Venezky, R. L. (1970). *The Structure of English Orthography* (Mouton, 's-Gravenhage, The Netherlands).
- Voiers, W. (1983). "Evaluating Processed Speech using the Diagnostic Rhyme Test," *Speech Technol.* **1**, 30-39.
- Wang, W., and Peterson, G. E. (1958). "Segment Inventory for Speech Synthesis," *J. Acoust. Soc. Am.* **30**, 743-746.
- Warrick, A., Nelson, P. J., Cossaltor, J. G., Cote, C., McGillis, J., and Charbonneau, J. R. (1977). "Synthesized Speech as an Aid to Communication and Learning for the Non-Verbal," *Proc. Workshop on Communication Aids for the Handicapped, Ottawa, Canada*, pp. 120-135.
- Werner, E., and Haggard, M. (1969). "Articulatory Synthesis by Rule," in *Speech Synthesis and Perception* (Psychological Laboratory, Univ. Cambridge, Cambridge, England), Vol. 1, pp. 1-35.
- Wiggins, R. (1980). "An Integrated Circuit for Speech Synthesis," *IEEE Int. Conf. Acoust. Speech Signal Process.* **ICASSP-80**, 398-401.

- Wijk, A. (1969). In *Alphabets for English*, edited by W. Haas (Manchester Univ. Press, Manchester, England).
- Woods, W. A. (1970). "Transition Network Grammars for Natural Language Analysis," *Commun. ACM* 13, 591-606.
- Wright, C. E., Altom, M. J., and Olive, J. P. (1986). "Diagnostic Evaluation of a Synthesizer's Acoustic Inventory," *J. Acoust. Soc. Am. Suppl.* 1 79, S25.
- Wright, J. T., Malsheen, B. J., and Peet, M. (1986). "Comparison of Segmental Intelligibility and Pronunciation Accuracy for Two Commercial Text-to-Speech Systems," *Proc. Am. Voice Input/Output Soc. AVIOS-86*, 235-261.
- Young, R. W. (1948). "Review of U.S. Patent 2,432,321, Translation of Visual Symbols, R. K. Potter, assignor (9 December 1947)," *J. Acoust. Soc. Am.* 20, 888-889.
- Young, S. J., and Fallside, F. (1979). "Speech Synthesis from Concept: A Method for Speech Output from Information Systems," *J. Acoust. Soc. Am.* 66, 685-695.
- Zue, V. W., and Laferriere, M. (1979). "An Acoustic Study of Medial /t,d/ in American English," *J. Acoust. Soc. Am.* 66, 1039-1050.
- Zue, V. W., and Shattuck-Hufnagel, S. (1979). "The Palatalization of Alveolar Fricatives in American English," *Proc. 9th Int. Cong. Phonetics, Copenhagen, Denmark*, pp. 215-216.