

Recording Human Evoked Potentials That Follow the Pitch Contour of a Natural Vowel

Hilmi R. Dajani*, David Purcell, Willy Wong, Hans Kunov, and Terence W. Picton

Abstract—We investigated whether pitch-synchronous neural activity could be recorded in humans, with a natural vowel and a vowel in which the fundamental frequency was suppressed. Small variations of speech periodicity were detected in the evoked responses using a fine structure spectrograph (FSS). A significant response ($P \ll 0.001$) was measured in all seven normal subjects even when the fundamental frequency was suppressed, and it very accurately tracked the acoustic pitch contour (normalized mean absolute error $< 0.57\%$). Small variations in speech periodicity, which humans can detect, are therefore available to the perceptual system as pitch-synchronous neural firing. These findings suggest that the measurement of pitch-evoked responses may be a viable tool for objective speech audiometry.

Index Terms—Bioelectric potentials, biomedical signal processing, speech processing, time-frequency analysis.

I. INTRODUCTION

Objective audiometric tests that use clicks, tone bursts, or steady state tones give only limited information about a subject's ability to perceive speech [1]–[3]. Consequently, there is an interest in developing objective hearing tests in which speech is the stimulus signal. In speech, pitch is the primary determinant of prosody, which includes linguistic, emotional, and speaker-specific information [4]–[6]. The speech signal has a fundamental frequency F_0 , the inverse of its basic periodicity, whose evolution over time is commonly referred to as the F_0 or pitch contour. Previous studies have reported the recording of potentials evoked by natural and synthetic speech [7]–[9], and have described a peak in the spectrum of the response at F_0 . Krishnan *et al.* [10] have recently detected evoked responses that closely followed gross directional changes of around 10–50 Hz in the pitch contours of synthetic lexical tones of Mandarin Chinese, using a periodicity detection short-term autocorrelation algorithm [11].

The ear is very sensitive to small changes in F_0 [4], the difference limen for (synthetic) vowels being in the range of 0.3% to 0.5% F_0 for the male voice [12]. In this paper, we tracked evoked responses that closely followed fine variations of the order of a few Hertz in the F_0 of a natural vowel using a fine structure spectrograph (FSS) [13]. This time-frequency analyzer allows the detection of small envelope and frequency variations in modulated signals. As the physiological mechanisms underlying pitch perception remain uncertain [14], the measurement of pitch-evoked responses in the electroencephalogram

Manuscript received June 25, 2004; revised February 6, 2005. Asterisk indicates corresponding author.

*H. R. Dajani was with the Institute of Biomaterials and Biomedical Engineering and the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada. He is now with the Rotman Research Institute, Baycrest Centre for Geriatric Care, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada (e-mail: h.dajani@utoronto.ca).

D. Purcell is with the Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada.

W. Wong and H. Kunov are with the Institute of Biomaterials and Biomedical Engineering and the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada.

T. W. Picton is with the Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, ON M6A 2E1, Canada.

Digital Object Identifier 10.1109/TBME.2005.851499

(EEG) with natural speech might give a clearer picture of physiological pitch processing. Moreover, it may provide an objective measure of speech perception in normal and hearing-impaired listeners.

II. METHODS

A. Subjects

Seven subjects 22–65 years old (two females) participated in this study. All had hearing thresholds of 20 dB hearing level or less, for 500-, 1000-, 2000-, and 4000-Hz tones in their right ear.

B. Auditory Stimuli and Recording of Evoked Potentials

We used two stimuli in the experiment: a 2-s recording of the vowel /a/ spoken by an adult male, and a version of this vowel in which the signal at the fundamental frequency F_0 (approximately 165 Hz) was suppressed (501 tap FIR high-pass filter, 300-Hz cutoff frequency). The vowel was recorded in a quiet room using a Sound Professionals microphone and preamplifier. The signal from the microphone was digitized at 32 kHz using an Avance AC97 16-bit sound card. During the experiment, the stimuli were generated using a modified MASTER system [15]. Digital-to-analog conversion was performed at 32 kHz using a National Instruments 16 bit 6052E input/output board. Electrical analog signals were routed to a GSI Model 16 audiometer, where their magnitude could be adjusted. The stimuli were presented using an Etymotic ER 2 insert earphone, and were calibrated to 75-dB SPL (average level), with the foam ear insert sealed in a Knowles DB-100 Zwislocki coupler. The level of the signal at F_0 in the unaltered vowel was estimated at 57-dB SPL by filtering the vowel with a bandpass filter centered at 165 Hz (501 tap FIR filter, 130- and 200-Hz cutoff frequencies), and comparing the time domain amplitude of the output signal with that of the vowel. In the high-pass filtered vowel, the level of signal at F_0 was approximately 12-dB SPL, which is below the threshold of hearing in the range of 165 Hz (e.g., [16]).

The subjects were presented 1350 repetitions of each of the stimuli in two 45-min sessions, separated by a short break. The order of the sessions, for the two stimuli, was randomized across subjects. The evoked potentials were recorded from gold-plated Grass electrodes placed at the vertex (Cz) and at the posterior midline of the neck just below the hair line. The ground electrode was placed on the collarbone. All electrode impedances were below 8 k Ω at 10 Hz. The signal was amplified and filtered using a Grass P55 battery-operated amplifier with a gain of 10 000 and a bandpass filter of 1–1000 Hz. The signal was further amplified with a gain of 5 on the acquisition board and digitized at 2 kHz, with a resolution of 16 bits. Measurements were performed in an Industrial Acoustics Company sound-insulated room. The subjects sat in a comfortable reclining chair in the dark, and were allowed to fall asleep.

A control experiment to test if there was any electrical leakage from the sound generating equipment to the electrodes was performed with one subject who exhibited a strong pitch-evoked response. This subject participated in another session under the same experimental conditions but with the insert earphone inserted into a Zwislocki coupler taped to his chest. A blocked earplug was inserted into the subject's ear canal to prevent any perception of the sound. The coupler presents the same acoustic load to the transducer as when the insert earphone is in the ear.

C. Analysis

Hand-labeling of pitch periods in the acoustic waveform is a standard method for obtaining reference pitch contours [4], [17]. In this paper, the boundaries of the pitch periods were identified using a peak detector

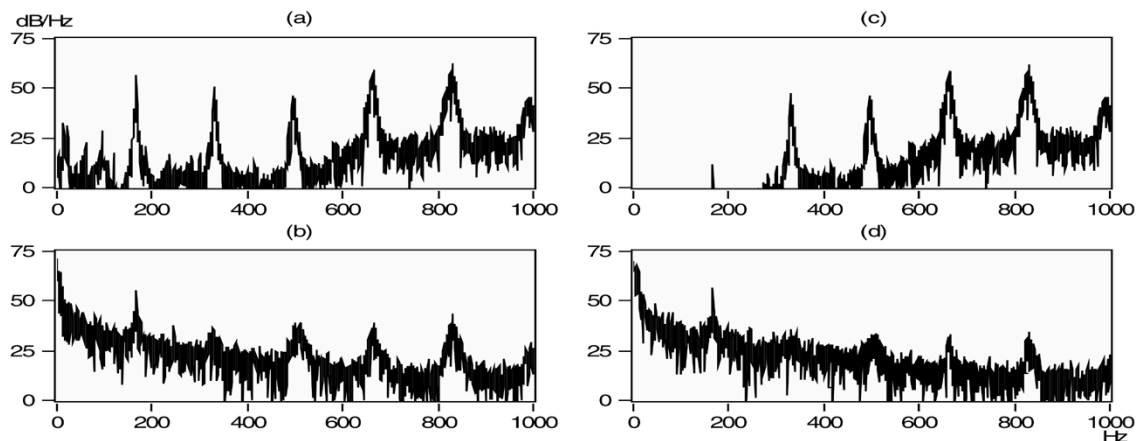


Fig. 1. Power spectra (up to 1000 Hz) of the acoustic stimuli (top row) and grand averaged response waveforms from all subjects (bottom row). Left column is for the /a/ stimulus, while right column is for the high-pass filtered version.

that fits a quadratic polynomial to the signal, implemented on successive 6-ms intervals. The output of the peak detector was inspected visually, and a handful of errors and omissions were corrected manually. The waveform representing the reciprocals of the interpeak intervals, which gives the pitch contour, was interpolated using a cubic spline interpolator, resampled to 2 kHz, and smoothed with a 128-ms moving window averaging function to suppress pitch “jitter.”

To obtain the pitch contour in the EEG, a noise rejection technique was first employed where the mean spectral amplitude between 20 and 400 Hz in each response waveform, or sweep, was tested against the mean spectral amplitude in the same frequency range averaged over the 1350 sweeps in a session. A sweep was rejected if its noise metric exceeded the mean by more than 1.5 standard deviations. Using this algorithm, the number of rejected sweeps ranged from 11–75 (average 33) per session. The remaining recorded sweeps were then synchronously averaged.

The fine structure spectrogram (FSS) of the synchronously averaged sweeps was obtained [13]. The FSS uses a large number of highly overlapping filter/detector (F/D) stages, each consisting of a bandpass filter, followed by a rectifier and a low-pass filter, and with all stages feeding into a detector of local spectral peaks. If the bandpass and low-pass filters are wideband relative to the signal modulation frequency, and the outputs of the F/D stages are probed very frequently, then the FSS can detect small frequency and envelope modulations in synthetic modulated signals and in speech. For the special case when the bandwidth of the bandpass filter equals twice that of the low-pass filter and a square-law rectifier is used, the FSS is equivalent to peak detection on a magnitude-squared short-time Fourier transform (STFT) that is highly over-sampled in time and frequency (in this case implemented using a 256-ms 512-point Hanning window zero padded to 4096 points, with window separation of 0.5 ms, or 1 sample) [13]. The detection of local peaks in the STFT, also called “ridge extraction,” can be used to estimate the instantaneous frequency and amplitude of components in the time-frequency plane [18]–[20]. Over-sampling of the STFT, as done in the FSS, allows the detection of small instantaneous frequency and amplitude variations. To extract the pitch contour from the three-dimensional FSS of the evoked response, the highest point in the FSS was determined at each time instant within a search range from 156 to 174 Hz. This range is three times the range for F0 (162–168 Hz) in the acoustic waveforms. Analysis was performed up to 1.79 s in the signal, in order to avoid end effects in the FSS.

To test for the presence of the pitch-evoked response, the spectrum of the averaged EEG sweep was obtained, and the peak spectral value within the range of 162–168 Hz was compared to 20 frequency bins (9.8 Hz) from the background noise above 174 Hz and 20 frequency bins below 156 Hz. The values of 156 and 174 Hz were chosen as

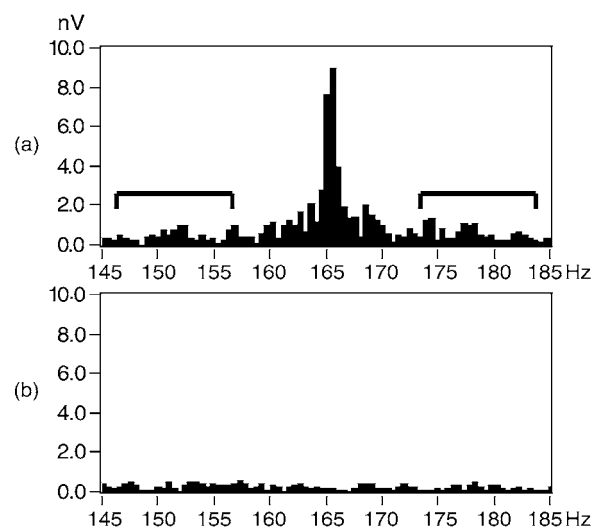


Fig. 2. Amplitude spectra of the synchronously averaged EEG sweeps in one subject (subject 5). (a) When presented with the /a/ stimulus. (b) In the control experiment, with the earphone inserted into the Zwislocki coupler. Horizontal bars indicate the noise estimate region.

TABLE I
AVERAGE RECORDING TIME (AND RANGE ACROSS SUBJECTS) IN SECONDS
REQUIRED TO ACHIEVE SIGNIFICANCE LEVELS OF $P < 0.05$ AND $P < 0.001$

Stimulus	$P < 0.05$	$P < 0.001$
/a/	6 (2 - 12)	77 (8 - 311)
HPFiltered /a/	7 (2 - 12)	114 (20 - 262)

boundaries for the background noise by extending the observed range for F0 as determined from the acoustic waveforms by an equal interval of 6 Hz on either side. The statistical test employed an F-ratio, which gives the probability of the presence of a response within background noise [15].

To estimate the accuracy of the EEG pitch contour, we used the standard measure of the % absolute error (ERR) relative to a reference pitch contour at each time instant [4], [21], [22]

$$\text{ERR} = \frac{100 \times \text{Absolute_error}(\text{Hz})}{\text{Reference_contour_frequency}(\text{Hz})} \quad (1)$$

We also report the mean and standard deviation in hertz of the absolute error relative to the reference contour [22], and the % Normalized Mean absolute Error NME, which is defined as

$$\text{NME} = \frac{100 \times \text{Mean_absolute_error}(\text{Hz})}{\text{Reference_contour_mean_frequency}(\text{Hz})} \quad (2)$$

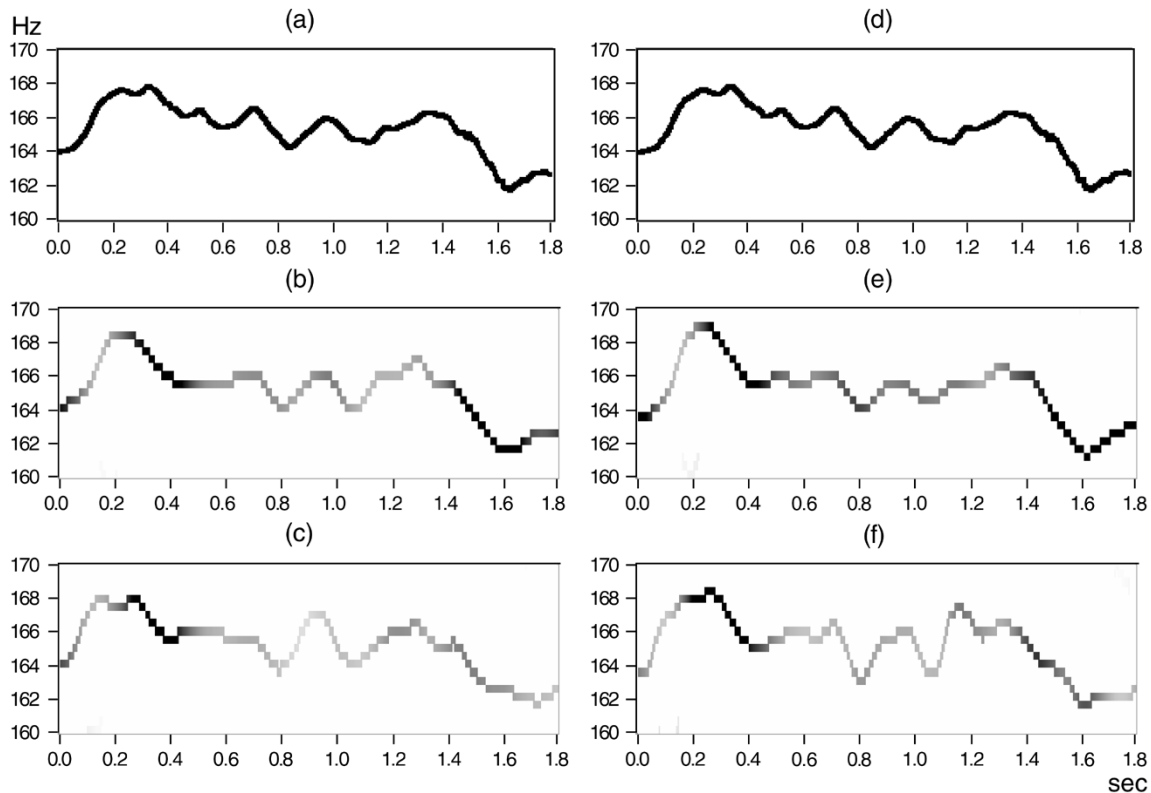


Fig. 3. Comparison of the reference acoustic pitch contours (top row) and the FSS (256-ms 512-point Hanning window zero padded to 4096 points, with window separation of 0.5 ms, or 1 sample) of the pitch-evoked response in subjects 6 and 3 (second and third rows). Left column is for the /a/ stimulus, while right column is for the high-pass filtered version. After smoothing, the pitch contours for both acoustic stimuli are practically identical, except for a very small time delay due to the high-pass filter. The grey scale in the FSS reflects the instantaneous amplitude of the pitch-evoked response.

III. RESULTS

The top row of Fig. 1 shows power spectra of the acoustic stimuli (left column is for the unaltered /a/ stimulus, right column is for the high-pass filtered version), and the bottom row shows power spectra of the grand averaged response waveforms from all subjects. There is a strong peak at the fundamental frequency F_0 (between 160 and 170 Hz) in the response spectra with both stimuli. There are also other peaks in the responses at harmonics of F_0 , the highest being near the frequency of the first formant (approximately 830 Hz). For each subject, with both the unaltered and high-pass filtered vowels, there was a statistically significant pitch-evoked response above the noise ($P \ll 0.001$). In the control experiment (Fig. 2), there was no evidence for a false response due to any interaction between the sound generating equipment and the electrodes ($P = 0.15$). Table I shows the average and range of recording times required to achieve a statistically significant response. Fig. 3 shows the reference acoustic pitch contours (top row), and the second and third rows show the FSS of the pitch-evoked responses in two subjects (left column is for the /a/ stimulus, right column for the high-pass filtered version). The accuracy of the EEG pitch contours is shown in Table II.

IV. DISCUSSION

In this paper, we demonstrated that a bioelectric response that closely follows the pitch contour of a natural vowel can be detected in normal hearing subjects. Since the pitch is not constant, simple frequency transforms are not appropriate and fine structure variations in the pitch-evoked response (of the order of a few Hertz) were tracked using a time-frequency representation referred to as the FSS. The observed pitch-evoked responses show that information about

small variations in the speech pitch frequency, to which humans are very sensitive, is likely available to the perceptual pitch processor in the form of the compound electrical activity of pitch-synchronous neurons. Moreover, the results show that a representation like the FSS may be useful in the analysis of the EEG, when there is interest in frequency and envelope fine structure, and as long as local peaks in the time-varying spectrum track instantaneous changes in amplitude and frequency.

It is likely that the pitch-evoked response originates from the earliest levels of auditory processing. The envelope (and frequency) of voiced speech fluctuates at the fundamental frequency F_0 , as there is typically an abrupt signal increase at the start of a pitch period, after which amplitude decays exponentially due to the interaction between the glottal source and the vocal cavity [23]. In the region of the cochlea where the individual harmonics of F_0 are not resolved, peripheral rectification by the inner hair cells would produce neural responses that fluctuate strongly with the envelope of the stimulus [24], [25]. This has been demonstrated in neural recordings with single and two formant vowel-like stimuli. Peripheral fibers with high characteristic frequencies (>3–4 kHz) had the strongest fluctuations at F_0 in their firing patterns, while fibers with low characteristic frequencies phase locked to the cycle-by-cycle variations of the waveform and showed poorer phase-locking to the envelope, resulting in firing patterns with smaller fluctuations at F_0 [26]–[29]. Beyond the auditory periphery, phase-locking of neural activity to the fundamental frequency of speech stimuli has been found at the level of the brainstem and the cortex (e.g., [30]). In experiments with steady tones whose envelopes fluctuated at a frequency close to the F_0 in our experiment (i.e., around 165 Hz), the main source for the recorded evoked response in humans was located in the brainstem, with smaller contributions from cortical and other stages of the auditory system [24], [31].

TABLE II
MEASURES OF THE ACCURACY OF THE EEG PITCH CONTOURS RELATIVE TO THE REFERENCE ACOUSTIC PITCH CONTOURS, FOR THE STIMULUS AND THE HIGHPASS FILTERED VERSION. MaxERR IS THE MAXIMUM % ABSOLUTE ERROR OVER THE DURATION OF THE PITCH CONTOUR (1). NME IS THE % NORMALIZED MEAN ABSOLUTE ERROR (2)

Stimulus	/a/			
Subject	MaxERR(%)	Mean Abs Error(Hz)	Stdev Abs Error(Hz)	NME(%)
1	1.47	0.7	0.5	0.42
2	1.87	0.95	0.62	0.57
3	1.23	0.76	0.54	0.46
4	1.35	0.76	0.55	0.46
5	1.37	0.72	0.52	0.44
6	1.11	0.66	0.42	0.4
7	1.14	0.69	0.44	0.41
All subjects	Range:1.11-1.87	Mean: 0.75 Stdev: 0.10		Range:0.4-0.57
Stimulus	HPFiltered /a/			
Subject	MaxERR(%)	Mean Abs Error(Hz)	Stdev Abs Error(Hz)	NME(%)
1	1.31	0.64	0.51	0.39
2	2.75	0.88	0.7	0.54
3	1.74	0.86	0.62	0.52
4	1.37	0.83	0.53	0.5
6	1.12	0.62	0.46	0.37
7	1.35	0.71	0.5	0.43
All subjects	Range:1.12-2.75	Mean: 0.76 Stdev: 0.11		Range:0.37-0.54

The pitch of speech can be perceived even when the fundamental frequency is missing from the spectrum of the signal (e.g., in “telephone” speech) [23]. With the high-pass filtered version of the vowel stimulus, the evoked response at the fundamental frequency F0 is as strong as it is for the unaltered vowel, and its contour also very closely follows the acoustic pitch contour. Since F0 is suppressed in the spectrum of the stimulus, the peripheral nonlinearity may also play a role in extracting the fundamental periodicity, even though experiments with complex steady state tones pointed away from a peripheral origin for the evoked response when the fundamental component was missing from the stimulus [32], [33].

While further work with different vowels and other speech patterns is needed, the successful recording of pitch-evoked responses from several synthetic vowels with different pitch contours [10] and the detection of pitch contour fine structure in the responses evoked by natural vowels in this study suggest that speech pitch-evoked responses may be a useful tool for objective speech audiometry. Speech-evoked recordings could provide information on how hearing impairment secondary to cochlear lesions alters neural responses to speech stimuli [9]. Since pitch-evoked responses are probably dependent on synchronous firing in the auditory system, they could also help in the detection of disorders of temporal processing in the elderly [34], [35] and in patients with auditory neuropathy [36]. In a clinical application, it would be desirable to reduce the 45-min recording time per stimulus that we used in our study. If the goal is not to very accurately estimate variations in the pitch-evoked contour but to simply detect the evoked response at a statistical confidence level of $P < 0.05$, then far fewer sweeps are required and very short recording times are possible (Table I).

ACKNOWLEDGMENT

The authors would like to thank P. Van Roon for her assistance with the data collection, and Dr. M. S. John for providing helpful advice.

REFERENCES

[1] J. W. Hall, III, *Handbook of Auditory Evoked Responses*. Boston, MA: Allyn and Bacon, 1992.

[2] D. Ballantyne, *Handbook of Audiological Techniques*. London, U.K.: Butterworth-Heinemann, 1990.

[3] S. A. Gelfand, *Essentials of Audiology*, 2nd ed. New York: Thieme, 2001.

[4] W. Hess, *Pitch Determination of Speech Signals, Algorithms and Devices*. Berlin, Germany: Springer-Verlag, 1983.

[5] —, “Pitch and voicing determination,” in *Advances in Speech Signal Processing*, S. Furui and M. M. Sondhi, Eds. New York: Marcel Dekker, 1992, pp. 3–48.

[6] A. de Cheveigné and H. Kawahara, “YIN, a fundamental frequency estimator for speech and music,” *J. Acoust. Soc. Am.*, vol. 111, pp. 1917–1930, 2002.

[7] G. C. Galbraith, P. W. Arbagey, R. Branski, N. Comerci, and P. M. Rector, “Intelligible speech encoded in the human brain stem frequency-following response,” *Neuroreport*, vol. 6, pp. 2363–2367, 1995.

[8] G. C. Galbraith, S. P. Jhaveri, and J. Kuo, “Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 6, pp. 46–53, 1997.

[9] A. Krishnan, “Human frequency-following responses: Representation of steady-state synthetic vowels,” *Hear. Res.*, vol. 166, pp. 192–201, 2002.

[10] A. Krishnan, Y. Xu, J. T. Gandour, and P. A. Cariani, “Human frequency-following response: Representation of pitch contours in Chinese tones,” *Hear. Res.*, vol. 189, pp. 1–12, 2004.

[11] P. Boersma, “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound,” in *Proc. Inst. Phon. Sci.*, vol. 17, 1993, pp. 97–110.

[12] J. L. Flanagan and M. G. Saslow, “Pitch discrimination for synthetic vowels,” *J. Acoust. Soc. Am.*, vol. 30, pp. 435–442, 1958.

[13] H. R. Dajani, W. Wong, and H. Kunov, “Fine structure spectrography and its application in speech,” *J. Acoust. Soc. Am.*, vol. 117, no. 6, pp. 3902–3918, Jun. 2005.

[14] R. Plomp, *The Intelligent Ear: On the Nature of Sound Perception*. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

[15] M. S. John and T. W. Picton, “MASTER: A windows program for recording multiple auditory steady-state responses,” *Comput. Meth. Programs Biomed.*, vol. 61, pp. 125–150, 2000.

[16] *Acoustics—Reference Zero for the Calibration of Audiometric Equipment—Part 2: Reference Equivalent Threshold Sound Pressure Levels for Pure Tones and Insert Earphones*, ISO 389-2, 1994.

[17] T. F. Quatieri, *Discrete-Time Speech Signal Processing: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall PTR, 2002.

[18] N. Delpat, B. Escudé, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrèsani, “Asymptotic wavelet and gabor analysis: Extraction of instantaneous frequencies,” *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 644–664, Mar. 1992.

- [19] R. Carmona, W. Hwang, and B. Torrésani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Trans. Signal Process.*, vol. 45, no. 10, pp. 2586–2590, Oct. 1997.
- [20] —, "Multiridge detection and time-frequency reconstruction," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 480–492, Feb. 1999.
- [21] A. de Cheveigné and H. Kawahara, "Comparative evaluation of F0 estimation algorithms," in *Proc. Eurospeech 2001*, 2001, pp. 2451–2454.
- [22] C. Wang and S. Seneff, "Robust pitch tracking for prosodic modeling in telephone speech," in *Proc. ICASSP-00*, vol. 3, 2000, pp. 1343–1346.
- [23] D. O'Shaughnessy, *Speech Communications: Human and Machine*. Piscataway, NJ: IEEE, 2000.
- [24] O. G. Lins, P. E. Picton, T. W. Picton, S. C. Champagne, and A. Durieux-Smith, "Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz," *J. Acoust. Soc. Am.*, vol. 97, pp. 3051–3063, 1995.
- [25] B. Delgutte and N. Y. S. Kiang, "Speech coding in the auditory nerve: I. Vowel-like sounds," *J. Acoust. Soc. Am.*, vol. 75, pp. 866–878, 1984.
- [26] B. Delgutte, "Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers," *J. Acoust. Soc. Am.*, vol. 68, pp. 843–857, 1980.
- [27] B. Delgutte and N. Y. S. Kiang, "Speech coding in the auditory nerve: I. Vowel-like sounds," *J. Acoust. Soc. Am.*, vol. 75, pp. 866–878, 1984.
- [28] H. E. Secker-Walker and C. L. Searle, "Time-domain analysis of auditory-nerve-fiber firing rates," *J. Acoust. Soc. Am.*, vol. 88, pp. 1427–1436, 1990.
- [29] P. X. Joris and T. C. T. Yin, "Responses to amplitude-modulated tones in the auditory nerve of the cat," *J. Acoust. Soc. Am.*, vol. 91, pp. 215–232, 1992.
- [30] A. R. Palmer, "Neural signal processing," in *Hearing*, B. C. J. Moore, Ed. San Diego, CA: Academic, 1995, pp. 75–121.
- [31] A. T. Herdman, O. Lins, P. Van Roon, D. R. Stapells, M. Scherg, and T. W. Picton, "Intracerebral sources of human auditory steady-state responses," *Brain Topogr.*, vol. 15, pp. 69–86, 2002.
- [32] S. Greenberg, J. T. Marsh, W. S. Brown, and J. C. Smith, "Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones," *Hear. Res.*, vol. 25, pp. 91–114, 1987.
- [33] G. C. Galbraith, "Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli," *Electroencephalogr. Clin. Neurophysiol.*, pp. 321–330, 1994.
- [34] P. J. Fitzgibbons and S. Gordon-Salant, "Auditory temporal processing in elderly listeners," *J. Am. Acad. Audiol.*, vol. 92, pp. 183–189, 1996.
- [35] D. R. Frisina and R. D. Frisina, "Speech recognition in noise and presbycusis: Relations to possible neural mechanisms," *Hear. Res.*, vol. 106, pp. 95–104, 1997.
- [36] A. Starr, T. W. Picton, Y. Sinniger, L. J. Hood, and C. I. Berlin, "Auditory neuropathy," *Brain*, vol. 119, pp. 741–753, 1996.