On the Rapid Fluctuation of Voice Pitch*
Shizuo HIKI, Kunihiro SUGAWARA and Juro OIZUMI
(Research Institute of Electrical Communication, Tohoku University)

This paper describes a set of experiments aiming at 1) precise estimation of the small rapid fluctuation of the voice fundamental pitch in steady parts of voicing in utterances of various talkers under various conditions, 2) estimation of the contribution of the voice fluctuations to the subjective tonal quality, and 3) the same in the case of synthetic speech.

To measure fluctuations of the voice fundamental frequency, the peak interval of the waveform of the vibration of the anterior neck over the tracheal wall was extracted**. The interval were quantized to the nearest values with a 1/20 msec step, and the time series of the successive fundamental periods in each run of voicing was analyzed in detail with an electronic computer (NEAC 2230).

In the steady part of the voicing in utterances of Japanese vowels by normal talkers (32 males and 30 females, age 16 to 63), the range of fluctuations (percentage to the mean fundamental period) has been estimated at ±1.6% for mean of all talkers, ±3.2% for maximum and ±0.8% for minimum. Correlograms of the time series of the voice periods show that the fluctuation is comparatively smooth and has no apparent periodicity in rates above a few cps. Fluctuations of a larger extent and higher rates are observed superposed on the transient

Fig. 1 Some examples of the fluctuation of the time series of the successive pitch periods \( \{ T_n \} \) in the steady part of voicing of Japanese vowels pronounced by normal talker (N), with vibrate (V) and by pathologic talker (P).

---

* 声帯のビッケ周期のゆらぎの性質
比企静雄, 菅原邦宏, 大泉充郎(東北大・通研)
the fluctuation is much larger (for example, ±15%) and more random than in normal talkers. In the utterance with "vibrato" in singing, fluctuation is large and approximately sinusoidal, and has regular periodicity of several cps.

Different characteristics in the fluctuation for normal talkers, even though they are very small in extent and have no apparent periodicity, cause different tonal qualities in some cases. In synthetic speech, the perceptual quality is natural with a fluctuation of more than 1–2%, and becomes hoarse with a random fluctuation of more than 5–6%, or tremulous with a sinusoidal fluctuation of more than 8%.

Analysis of Speech by Single-Tuned Filter*
Shuji DOSHITA
(Department of Electrical Engineering, Kyoto University)

Appropriate selection of characteristics of filters is important in a design of speech analysis schemes. This paper describes some points about the response of single-tuned filters.

(1) Response to a formant-shaped signal

The response of a single-tuned filter having a center frequency $F_a$ and a bandwidth $B_a$, to an impulse fed through a simple resonant circuit having a center frequency $F_f$ and a bandwidth $B_f$ depends mainly on $|F_a - F_f|$ and $|B_a - B_f|$. For the case of $B_a = B_f$, the output waveform shows dips in envelope every $\frac{1}{|F_a - F_f|}$ sec. and decreases with exp $(-\pi B_a t)$. By using a condenser-input type CR-circuit with a fast rise time and a slow decay time as an envelope detector, the detected envelope can follow the build-up of the filter output and still smooth out the dips. The spectral response of a bank of such analysis channels is shown in Fig. 1, which shows that the components just matched to the resonant frequency $F_f$ grows with time more dominant compared with the other components.

(2) Analysis by single-tuned filter bank

The response of a bank of single-tuned filters to several types of signals including speech samples was experimentally examined. The spectrum analyzer consisted of a set of single-tuned filters, each having a selectable bandwidth of 33 cps, 67 cps, 100 cps or 167 cps and arranged with a 33 cps spacing and a choice of a CR-circuit or an LC-circuit as an envelope detector. The result of the analysis is displayed as a time-sampled spectrum section (Fig. 2b) and as a time-frequency pattern (Fig. 2a).

When a simple resonant circuit is excited by a pulse train, the outgoing signal approximates a vowel sound in a limited frequency range in the

 Fig. 1 Spectral response of the filter bank. The parameter is time after the impulse is applied

* 単共振フィルタによる音声の分析
堂下修司（京都大学・工学部）