

# SPEECH NONLINEARITIES, MODULATIONS, AND ENERGY OPERATORS

Petros Maragos\*, Thomas F. Quatieri†, and James F. Kaiser‡

\*Division of Applied Sciences, Harvard University, Cambridge, MA 02138, USA

†MIT Lincoln Laboratory, 244 Wood St., Lexington, MA 02173, USA

‡Bellcore, 445 South St., Morristown, NJ 07962-1910, USA.

**ABSTRACT:** In this paper, we investigate an AM-FM model for representing modulations in speech resonances. Specifically, we propose a frequency modulation (FM) model for the time-varying formants whose amplitude varies as the envelope of an amplitude-modulated (AM) signal. To detect the modulations we apply the energy operator  $\Psi(x) = (\dot{x})^2 - x\ddot{x}$  and its discrete counterpart. We found that  $\Psi$  can approximately track the envelope of AM signals, the instantaneous frequency of FM signals, and the product of these two functions in the general case of AM-FM signals. Several experiments are reported on the application of this AM-FM modeling to speech signals, bandpass filtered via Gabor filtering.

## 1 Introduction

In his work on nonlinear modeling of speech production, Teager [1, 2] used the nonlinear operator

$$\Psi_d[x(n)] = x^2(n) - x(n-1)x(n+1) \quad (1)$$

on speech-related discrete-time signals  $x(n)$ . Kaiser [3] analyzed  $\Psi_d$  and showed that it can detect the frequency of single sinusoids and chirp signals, and has many useful properties; e.g.,

$$\Psi_d[Ar^n \cos(\Omega_0 n + \phi)] = A^2 r^{2n} \sin^2(\Omega_0) \quad (2)$$

Kaiser [4] recently introduced an operator closely related (see Section 3) to  $\Psi_d$  for continuous-time signals  $x(t)$ :

$$\Psi_c[x(t)] = [\dot{x}(t)]^2 - x(t)\ddot{x}(t) \quad (3)$$

where  $\dot{x} = dx/dt$ , and investigated several properties of  $\Psi_c$ ; e.g.,

$$\Psi_c[Ae^{rt} \cos(\omega_0 t + \theta)] = A^2 e^{2rt} \omega_0^2 \quad (4)$$

$$\Psi_c[x(t)y(t)] = x^2(t)\Psi_c[y(t)] + y^2(t)\Psi_c[x(t)] \quad (5)$$

$\Psi_c$  was originally derived to track the energy of a linear undamped oscillator. Namely, when  $\Psi_c$  is applied to the oscillation signal  $A \cos(\omega_0 t)$ , its output is  $(A\omega_0)^2$  and hence proportional to the energy of the source producing the oscillation. Thus  $\Psi_c$  can be viewed as an *energy operator*.

Teager applied  $\Psi_d$  to signals resulting from bandpass filtering speech vowels in the vicinity of their formants. If the formant

were due to a linear resonance, then by (4) the operator's output would be a decaying exponential. Teager observed, on the other hand, several "energy" pulses per pitch period, which he viewed as indicating modulation of formants caused by nonlinear phenomena such as rapidly varying separated air flow in the vocal tract. In our work we interpret these energy pulses by using an *AM-FM model*, i.e., a frequency modulation (FM) model for the variation of the center frequency of time-varying formants whose amplitude varies like the envelope of an amplitude-modulated (AM) signal. For pure AM and FM signals we found that the energy operators  $\Psi_c$  and  $\Psi_d$  can approximately track the envelope of AM signals, the instantaneous frequency of FM signals, and the product of these two functions in the case of AM-FM signals. Our coverage of these interesting results is brief and focuses on a few special cases; more details and general cases are given in [5]. We have also obtained promising experimental results from applying this AM-FM modeling to speech signals, bandpass filtered via Gabor filtering.

## 2 Continuous-time AM and FM

Consider a general AM signal

$$X_{AM}(t) = e(t) \cos(\omega_c t + \theta) \quad (6)$$

where  $e(t)$  is an envelope more slowly varying than the carrier. Henceforth, we shall drop the subscripts  $c, d$  from  $\Psi$ , since it will be clear from the context whether we refer to continuous or discrete time. By (5) and (4),

$$\begin{aligned} \Psi[X_{AM}(t)] &= e^2 \omega_c^2 + \cos^2(\omega_c t + \theta) \Psi(e) = [\omega_c e(t)]^2 (1 + \text{error}) \\ &\approx [\omega_c e(t)]^2, \text{ if } \Psi(e) \ll (\omega_c e)^2 \end{aligned} \quad (7)$$

where  $\approx$  is meant as "approximated by the dominant term". The order of the approximation error in (6) (where by *order* we mean the order of maximum value that a signal or a quantity can assume) is  $O[\Psi(e)/(e\omega_c)^2]$ . In [5] it was shown that this error order is  $\ll 1$  if  $e(t)$  is any bandlimited signal whose highest frequency  $\omega_a$  is  $\ll \omega_c$ . Then  $\Psi$  acts as an *envelope detector*, because  $\sqrt{\Psi[e(t) \cos(\omega_c t + \theta)]} \propto |e(t)|$ . Two special cases for  $e(t)$  are: (i) (AM with carrier)  $e(t) = 1 + ma(t)$ , where  $a(t)$  is the AM information signal and  $m \leq 1$  is the modulation index. (ii) (AM with suppressed carrier)  $e(t) = a(t)$ . For simplicity let  $e(t) = \cos(\omega_a t)$  with  $\omega_a \ll \omega_c$  (as standard to assume in AM); then

$$\Psi[\cos(\omega_a t) \cos(\omega_c t + \theta)] \approx [\omega_c \cos(\omega_a t)]^2 \quad (8)$$

This work was supported in part by the ARO under Grant DAALO3-86-K-0171 (Center for Intelligent Control Systems), in part by the NSF under Grant MIPS-86-58150 with matching funds from Bellcore, DEC, and Xerox, and in part by the Department of the Air Force.





The approximation error in (8) has order  $O[(\omega_a/\omega_c)^2]$ ; e.g., it is  $\leq 10\%$  if  $\omega_a/\omega_c \leq 1/3$ , or  $\leq 1\%$  if  $\omega_a/\omega_c \leq 0.1$ .

Consider now the general FM signal

$$X_{FM}(t) = \cos[\phi(t)] = \cos(\omega_c t + \Delta \int_0^t f(\tau) d\tau + \theta) \quad (9)$$

where  $\phi(t)$  is the instantaneous phase,  $\omega_c$  is the carrier frequency,  $\omega_i(t) = d\phi/dt = \omega_c + \Delta f(t)$  is the instantaneous frequency,  $f(t)$  is the FM information signal and varies more slowly than  $\cos(\omega_c t)$ , and  $\Delta$  is the maximum frequency deviation. It is assumed that  $|f(t)| \leq 1$  and  $\Delta \leq \omega_c$ . Applying  $\Psi$  to  $X_{FM}$  yields

$$\begin{aligned} \Psi[X_{FM}(t)] &= (\dot{\phi})^2 + \ddot{\phi} \frac{\sin(2\phi)}{2} = (\dot{\phi})^2 (1 + \text{error}) \\ &\approx (\dot{\phi})^2 = [\omega_i(t)]^2, \text{ if } \ddot{\phi} \ll 2(\dot{\phi})^2 \end{aligned} \quad (10)$$

The error in (10) has order  $O[\ddot{\phi}/2(\dot{\phi})^2]$ ; if this is  $\ll 1$ , then  $\Psi$  tracks well the instantaneous FM frequency. For the special case of an FM-chirp signal,  $f(t) = st$  and  $\omega_i(t) = \omega_c + \Delta st$ ; then the error order is  $\ll 1$  if  $\Delta s \ll \omega_c^2$ . For bandlimited signals  $f$  with highest frequency  $\omega_f$  the error order is  $\ll 1$  if  $\Delta \omega_f \ll 2(\omega_c)^2$  [5]. Such a special case is the FM-sine signal, where  $f(t) = \cos(\omega_f t)$ ,  $\beta = \Delta/\omega_f$  is the FM modulation index, and

$$\Psi[\cos(\omega_c t + \beta \sin(\omega_f t) + \theta)] \approx \omega_i^2(t) = [\omega_c + \Delta \cos(\omega_f t)]^2 \quad (11)$$

The error has order  $\ll 1$  if either  $\omega_f \ll \omega_c$  (a standard assumption in FM), or  $\Delta \ll \omega_c$  (small frequency deviation), or both. For example, the error will be  $\leq 10\%$  if  $(\Delta/\omega_c) \cdot (\omega_f/\omega_c) \leq 1/5$ .

Consider now a general AM-FM signal, i.e., an FM signal whose amplitude is the envelope  $e(t)$  of an AM signal:

$$X_{AFM}(t) = e(t) \cos[\phi(t)] = e(t) \cos(\omega_c t + \Delta \int_0^t f(\tau) d\tau + \theta) \quad (12)$$

Then, by (5) and (10),

$$\begin{aligned} \Psi[X_{AFM}(t)] &= e^2(t) [(\dot{\phi})^2 + \ddot{\phi} \frac{\sin(2\phi)}{2}] + \cos^2(\phi) \Psi(e) \\ &= [e(t) \omega_i(t)]^2 (1 + \text{error}) \\ &\approx [e(t) \omega_i(t)]^2, \text{ if } \ddot{\phi} \ll 2(\dot{\phi})^2 \text{ and } \Psi(e) \ll (e\dot{\phi})^2 \end{aligned} \quad (13)$$

The approximation error in (13) is of the order

$$O(\text{error}) = \max \left( O \left[ \frac{\ddot{\phi}}{2(\dot{\phi})^2} \right], O \left[ \frac{\Psi(e)}{(e\dot{\phi})^2} \right] \right) \quad (14)$$

This error is  $\ll 1$  if  $e(t), f(t)$  are bandlimited signals whose highest frequency is  $\ll \omega_c$  [5]. Then  $\sqrt{\Psi}[e(t) \cos(\int \omega_i(\tau) d\tau)] \approx |e(t)| \omega_i(t)$  and thus the  $\sqrt{\Psi}$ 's output is the product of two parts: an FM instantaneous frequency  $\omega_i(t)$  and the AM envelope  $|e(t)|$ . This result generalizes the tracking ability of  $\sqrt{\Psi}$ , which for  $A \cos(\omega_0 t)$  signals yields  $A\omega_0$ , whereas for AM-FM signals the constant amplitude  $A$  and frequency  $\omega_0$  are replaced by the AM envelope and the FM instantaneous frequency.

As a special case of the AM-FM model, let  $e(t) = \cos(\omega_a t)$  and  $f(t) = \cos(\omega_f t)$ . Then

$$\sqrt{\Psi}[\cos(\omega_a t) \cos(\omega_c t + \beta \sin(\omega_f t) + \theta)] \approx |\cos(\omega_a t)| [\omega_c + \Delta \cos(\omega_f t)] \quad (15)$$

The error will be  $\ll 1$  if (i)  $\Delta \ll \omega_c$  or  $\omega_f \ll \omega_c$  and (ii)  $\omega_a \ll \omega_c$ . If  $\Delta \ll \omega_c$ , then the  $\omega_i(t)$  variations have much smaller amplitude than that of  $\cos(\omega_a t)$ , and AM dominates over FM, i.e., the  $\sqrt{\Psi}$  output follows the AM envelope signal. AM also wins over FM if  $\omega_f \ll 2\pi/L < \omega_a$  where  $L$  is the time duration of the analysis

window. If  $\omega_f > 2\pi/L \gg \omega_a$ , then FM wins over AM, i.e.,  $\sqrt{\Psi}$  tracks the FM instantaneous frequency.

A by-product of the AM-FM model is a better algorithm for FM detection. Taking the derivative of the FM signal  $\cos \phi(t)$ , gives an AM-FM signal  $y(t) = -\dot{\phi}(t) \sin \phi(t)$  whose envelope is the FM instantaneous frequency. Hence, applying (13) with  $e = \dot{\phi}$  gives  $\sqrt{\Psi}[y(t)] \approx [\omega_i(t)]^2$ . Thus, we can build an FM detector from a differentiator followed by the  $\Psi$  operator. By (10),  $\sqrt{\Psi}$  applied to an FM signal tracks its instantaneous frequency  $\omega_i(t)$ , but if  $\sqrt{\Psi}$  is instead applied to the FM derivative, then it tracks  $[\omega_i(t)]^2$ . Hence, the latter case will give a better FM tracking, since the square law will emphasize the oscillations of the instantaneous frequency.

Although all the results in Section 2 (for notational simplicity) referred to unit-amplitude cosines with no exponential decay, they can be easily extended to incorporate an amplitude  $A$  and/or an exponential decay  $e^{rt}$  in the input signal by just multiplying the energy operator's output with  $A^2 e^{2rt}$ , because  $\Psi[Ae^{rt}x(t)] = A^2 e^{2rt} \Psi[x(t)]$ .

### 3 Discrete-time AM-FM

By discretizing derivatives we can obtain from  $\Psi_c$  an expression closely related to  $\Psi_d$  and thus link the two operators. We examined several cases, e.g., the 2-sample backward difference

$$\dot{x}(t) \mapsto x(n) - x(n-1) \implies \Psi_c[x(t)] \mapsto \Psi_d[x(n-1)] \quad (16)$$

Likewise, the 2-sample forward difference  $\dot{x} \mapsto x(n+1) - x(n)$  gives  $\Psi_d[x(n+1)]$ . Thus both *asymmetric 2-sample* differences succeed to transform  $\Psi_c$  into  $\Psi_d$  (modulo one sample shift). However, 2-sample or 3-sample *symmetric* differences fail because they give more complicated expressions [5]. Next we apply  $\Psi_d$  to a few cases of discrete AM and FM signals.

Let  $X(n) = \cos(\Omega_a n) \cos(\Omega_c n + \theta)$  be an AM signal. From the general property

$$\Psi_d[x(n)y(n)] = x^2(n) \Psi_d[y(n)] + y^2(n) \Psi_d[x(n)] - \Psi_d[x(n)] \Psi_d[y(n)] \quad (17)$$

and by (2) it follows that  $\Psi[X(n)]$  is equal to  $\cos^2(\Omega_a n) \sin^2(\Omega_c) + [\cos^2(\Omega_c n + \theta) - \sin^2(\Omega_c)] \sin^2(\Omega_a)$ . Hence,

$$\Psi[\cos(\Omega_a n) \cos(\Omega_c n + \theta)] \approx [\sin(\Omega_c) \cos(\Omega_a n)]^2 \quad (18)$$

if  $\sin^2(\Omega_a) \ll \tan^2(\Omega_c)$  (which holds if  $\Omega_a \ll \Omega_c$ ).

Consider the discrete-time FM-sine signal

$$Y(n) = \cos[\phi(n)] = \cos[\Omega_c n + \beta \sin(\Omega_f n) + \theta] \quad (19)$$

where  $\beta = \Delta/\Omega_f$ , and the instantaneous frequency is  $\Omega_i(n) = d\phi(n)/dn = \Omega_c + \Delta \cos(\Omega_f n)$ . For applying  $\Psi$  to  $Y(n)$  note that if  $A = \Omega_c n + \beta \cos(\Omega_f) \sin(\Omega_f n) + \theta$  and  $B = \Omega_c + \beta \sin(\Omega_f) \cos(\Omega_f n)$ , then  $Y(n+1)Y(n-1) = (\cos 2A + \cos 2B)/2 = \cos^2(A) - \sin^2(B)$ . If  $\Omega_f$  is sufficiently small such that  $\cos(\Omega_f) \approx 1$  and  $\sin(\Omega_f) \approx \Omega_f$ , then  $\cos(A) \approx Y(n)$ ,  $B \approx \Omega_i(n)$ , and by (1)

$$\Psi[\cos(\phi(n))] \approx \sin^2[\Omega_c + \Delta \cos(\Omega_f n)] \quad (20)$$

All the results in Section 3 can be easily extended to incorporate an amplitude  $A \neq 1$  and/or an exponential decay  $r^n$  in the input signal by just multiplying the output of  $\Psi$  by  $A^2 r^{2n}$ .



We conducted many experiments on synthetic discrete AM and FM signals that verified the theoretical results and validated the approximate formulas. Figure 1 illustrates some of the above conclusions.

## 4 Modeling Speech Resonances

Teager's work in steady-state vowels provides evidence for resonances with time-varying formants and amplitudes [2]. In our work we model a single speech resonance by a damped AM-FM model

$$R(n) = Ar^n \cos(\Omega_a n) \cos(\Omega_c n + \beta \sin(\Omega_f n) + \theta) \quad (21)$$

where  $\Omega_c$  is the center frequency of the formant, the instantaneous FM frequency  $\Omega_i(n) = \Omega_c + \Delta \cos(\Omega_f n)$  models the time-varying formant, and  $\Delta = \beta \Omega_f$  controls the amount of FM. The AM envelope  $|\cos(\Omega_a n)|$  tracks the amplitude variations and  $r$  is the rate of energy dissipation. Then by (18) and (20)

$$\sqrt{\Psi}[R(n)] \approx |Ar^n \cos(\Omega_a n) \sin(\Omega_c + \Delta \cos(\Omega_f n))| \quad (22)$$

This approximation assumes that  $\Omega_f$  is small and that  $\Omega_a \ll \Omega_c$  (see the discrete AM and FM case). Thus,  $\sqrt{\Psi}[R(n)]$  is a (damped) product of the envelope and the (sine of the) instantaneous frequency of the resonance. This class of signals may serve as a model of energy pulses observed on actual speech waveforms. An alternative model for the FM frequency that is consistent with our previous analysis is the chirp  $\Omega_i(n) = \Omega_c + \Delta sn$ . In addition, an alternative envelope model would be  $e(n) = 1 + m \cos(\Omega_a n)$  where  $m$  measures the amount of AM; then  $\sqrt{\Psi}[e(n) \cos(\int \Omega_i(n)) \approx e(n) |\sin \Omega_i(n)|$  if  $\Omega_f$  is small and  $\Omega_a \ll \Omega_c$  or  $m \ll 1$  [5].

In our work we extract a single resonance by bandpass filtering the speech signal with a Gabor filter, whose impulse and frequency response are

$$h(t) = \exp(-a^2 t^2) \cdot \cos(\omega_c t) \quad (23)$$

$$H(\omega) = \frac{\sqrt{\pi}}{2a} \left( \exp\left[-\frac{(\omega - \omega_c)^2}{4a^2}\right] + \exp\left[-\frac{(\omega + \omega_c)^2}{4a^2}\right] \right) \quad (24)$$

The Gaussian shape of  $H(\omega)$  avoids producing side lobes that could produce false pulses in the  $\Psi$ 's output. The bandwidth (measured between the points at 10% of peak value) is about  $BW = 1.7a$  (in Hz). Our design of the discrete bandpass Gabor filter proceeds as follows: A center formant frequency  $F_c$  is manually selected from the short-time speech spectrum by visual inspection. A value of  $a$  is selected such that  $BW \in [0.5F_c, F_c]$ .  $h(t)$  is discretized by replacing  $t$  with  $nT$ , where  $T$  is sampling period, and truncating  $h(n)$  to a symmetric FIR filter,  $h(n) = \exp(-b^2 n^2) \cdot \cos(\Omega_c n)$ , with  $-N \leq n \leq N$ ,  $b = aT$ , and  $\Omega_c = 2\pi F_c T$ . Then the Gabor bandpass filtering is performed by convolving the truncated  $h(n)$  with the speech signal. The integer  $N$  is chosen to truncate the Gaussian envelope of  $h(n)$  essentially to zero; e.g.,  $N = 2.5/(aT)$  yields  $\exp(-b^2 N^2) = 0.002$ .

Fig. 2 shows (a) a segment of a speech vowel /e/ sampled at  $F_s = 30$  kHz and (b) the output from  $\sqrt{\Psi}$  when applied to a bandpass filtered version of (a) extracted around a formant at  $F_c = 3400$  Hz using a Gabor filter with  $b = 1000T$  and  $N = 75$ . Fig. 3 is similar to Fig. 2 but for a sustained vowel /a/ sampled at 10 kHz with  $F_c = 2620$  Hz,  $b = 1500T$ , and  $N = 50$ . There are present 2-3 pulses per pitch period, and the damped AM-

FM model (22) may approximately explain the shape of these measured energy pulses. There have been cases where we have observed only one major pulse per pitch period. This may be partially explained by a low percent of (AM or FM) modulation or by small  $\Omega_a, \Omega_f$ .

Equation (22) has generally both an AM and an FM component. The pulses could be due to both or just one of them. If  $\Delta \ll \Omega_c$  (or if  $\Omega_f \ll 2\pi F_0 < \Omega_a$  where  $F_0$  is the pitch frequency), then AM wins over FM and  $\sqrt{\Psi}$  follows essentially the envelope of the resonance; such a case is illustrated in Fig. 1c via a synthetic AM-FM signal. If however  $\Omega_a \ll 2\pi F_0 < \Omega_f$  (or if  $m \ll 1$  in the case of the  $1 + m \cos(\Omega_a n)$  envelope), then the FM dominates over AM. By testing ideas on synthetic AM-FM signals and by running zero-crossing FM detectors on speech resonances, we have seen in our speech experiments that AM tends to dominate FM. Also if the FM frequency deviation  $\Delta$  is small, then the exponential decay makes the FM component harder to detect. Thus additional effort is required to isolate and purify the FM component.

Finally, note that we have assumed in all the above analysis and experiments the presence of a *single* resonance in the vocal tract output. Actual speech vowels are quasi-periodic and may consist of multiple resonances. Both these phenomena introduce an additive component to the single resonance which may alter the output of the energy operator. Consider the case that arises when there are two formants closely spaced. Let's model this situation with the signal  $x(t) = \sin(\omega_1 t + 2\theta) + \sin(\omega_2 t)$ . Then  $x(t) = 2 \cos(\omega_a t + \theta) \sin(\omega_c t + \theta)$  is an AM signal whose carrier and envelope frequencies are  $\omega_c = (\omega_1 + \omega_2)/2$  and  $\omega_a = (\omega_1 - \omega_2)/2$ . By (7),  $\sqrt{\Psi}[x(t)] \approx \omega_c |\cos(\omega_a t + \theta)|$ , and hence  $\sqrt{\Psi}$  will track the envelope, if the approximation error order  $O[(1-d)^2/(1+d)^2]$  is  $\ll 1$ , where  $d = \omega_2/\omega_1 \leq 1$  (assume  $\omega_1 \geq \omega_2$ ). For this error to be  $< 10\%$ ,  $d > 0.5$ , i.e., the two formants must be less than an octave apart. Then we observe an AM modulation of one formant by the other. Consider also the case of two consecutive pitch harmonics falling within the resonance bandwidth and passing through the Gabor filter. Then the above model (consisting of two additive sines) holds and may predict a possible tracking of an AM envelope. However, this AM envelope varies with a frequency roughly equal to the pitch frequency and thus the modulation does not introduce additional pulses over a pitch period. Finally, in the time-domain, closely spaced responses from the vocal tract due to consecutive pitch pulses may also introduce fluctuations in  $\sqrt{\Psi}$ 's output which are not consequences of the AM or FM modulation of the resonance itself.

## 5 Conclusions

Our discussion motivates the following model for the vocal tract response (over a pitch period):

$$S(n) = \sum_{k=1}^K A_k r_k^n \cos(\Omega_{a,k} n) \cos(\Omega_{c,k} + \beta_k \sin(\Omega_{f,k} n) + \theta_k) \quad (25)$$

where  $K$  is the number of resonances. We also may include interaction between resonances by allowing coupling between the  $\Omega_a$ 's and the  $\Omega_f$ 's of the same resonance or among different resonances. It is important to emphasize that this is not an AM-FM model of speech production, but rather the AM-FM is a mathematical vehicle to model the acoustical consequences of some nonlinear mechanisms of speech production. One approach to



obtaining the parameters of such a model would be to 1) find the  $K$  center frequencies  $\Omega_c$  (e.g., from the short-time Fourier transform), 2) iteratively extract each frequency band and model it as an AM-FM signal by using the operator  $\Psi$ , 3) subtract the modeled AM-FM component from the total speech signal and model the remainder of the resonances.

## References

- [1] H. M. Teager, "Some Observations on Oral Air Flow During Phonation", *IEEE Trans. ASSP-28*, 599-601, Oct. 1980.
- [2] H. M. Teager and S. M. Teager, "Evidence for Nonlinear Production Mechanisms in the Vocal Tract", *NATO Adv. Study Inst. on Speech Production and Speech Modelling*, Bonas, France, 1989; Kluwer Acad. Publ., 1990.
- [3] J. F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal", in *Proc. ICASSP*, Albuquerque, NM, Apr 1990.
- [4] J. F. Kaiser, "On Teager's Energy Algorithm and Its Generalization to Continuous Signals", in *Proc. DSP Workshop*, New Paltz, NY, Sep. 1990.
- [5] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Detecting Amplitude and Frequency Modulations using Energy Operators", Tech. Report, Harvard Robotics Lab, Dec. 1990.

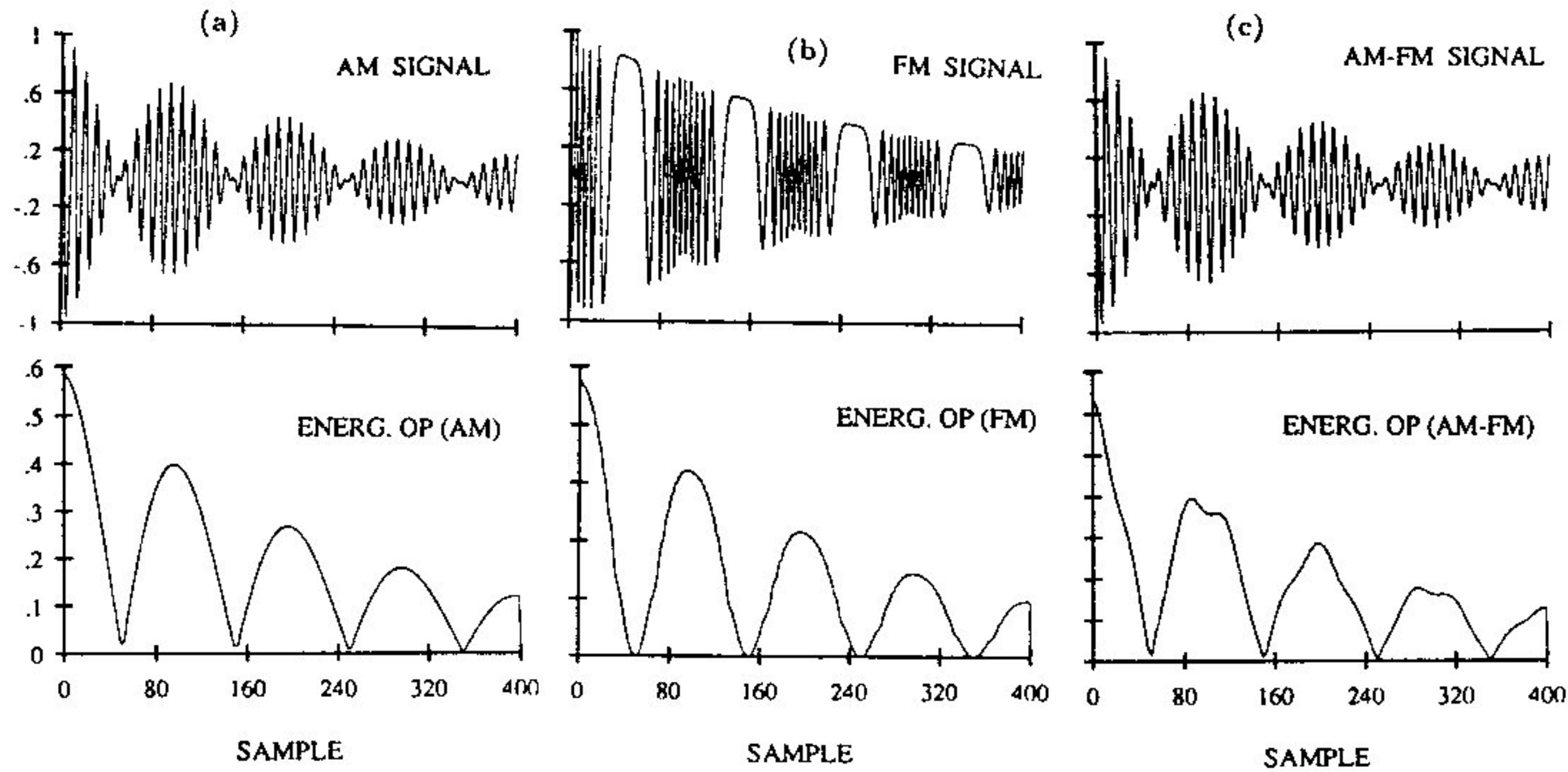


Figure 1. AM-FM signals  $X(n) = r^n \cos(\Omega_a n) \cos[\Omega_c n + (\Delta/\Omega_f) \sin(\Omega_f n)]$  and the energy operator outputs  $\sqrt{\Psi}[X(n)]$ , where  $\Omega_c = 0.2\pi$ ,  $r = 0.004$ : (a)  $\Omega_a = 0.01\pi$ ,  $\Omega_f = 0$ . (b)  $\Omega_a = 0$ ,  $\Omega_f = 0.02\pi$ ,  $\Delta = \Omega_c$ . (c)  $\Omega_a = 0.01\pi$ ,  $\Omega_f = 0.05\pi$ ,  $\Delta = \Omega_c/10$ .

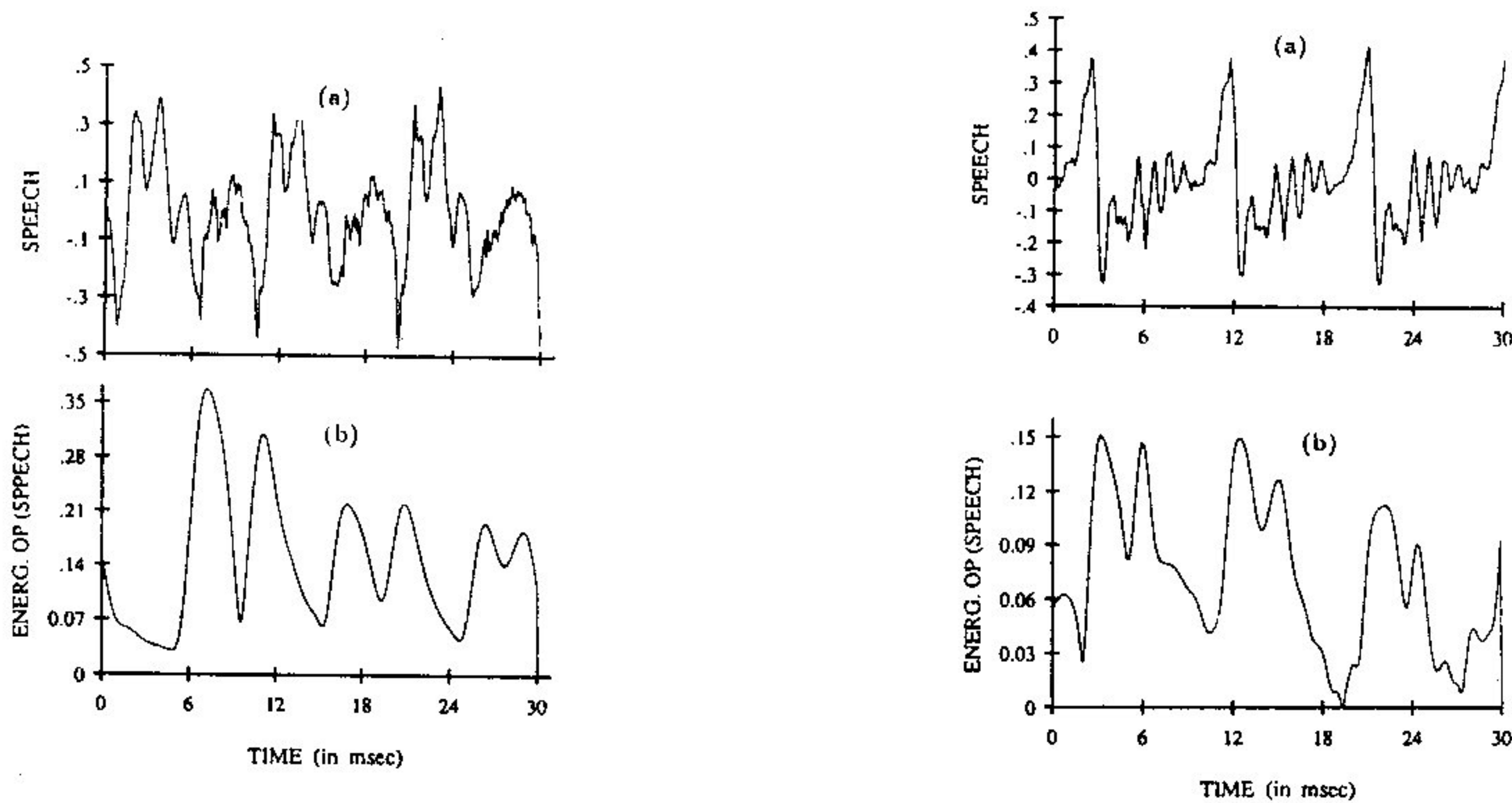


Figure 2. (a) Original speech vowel /e/. (b) Output of energy operator  $\sqrt{\Psi}$  when applied to a resonance of (a).

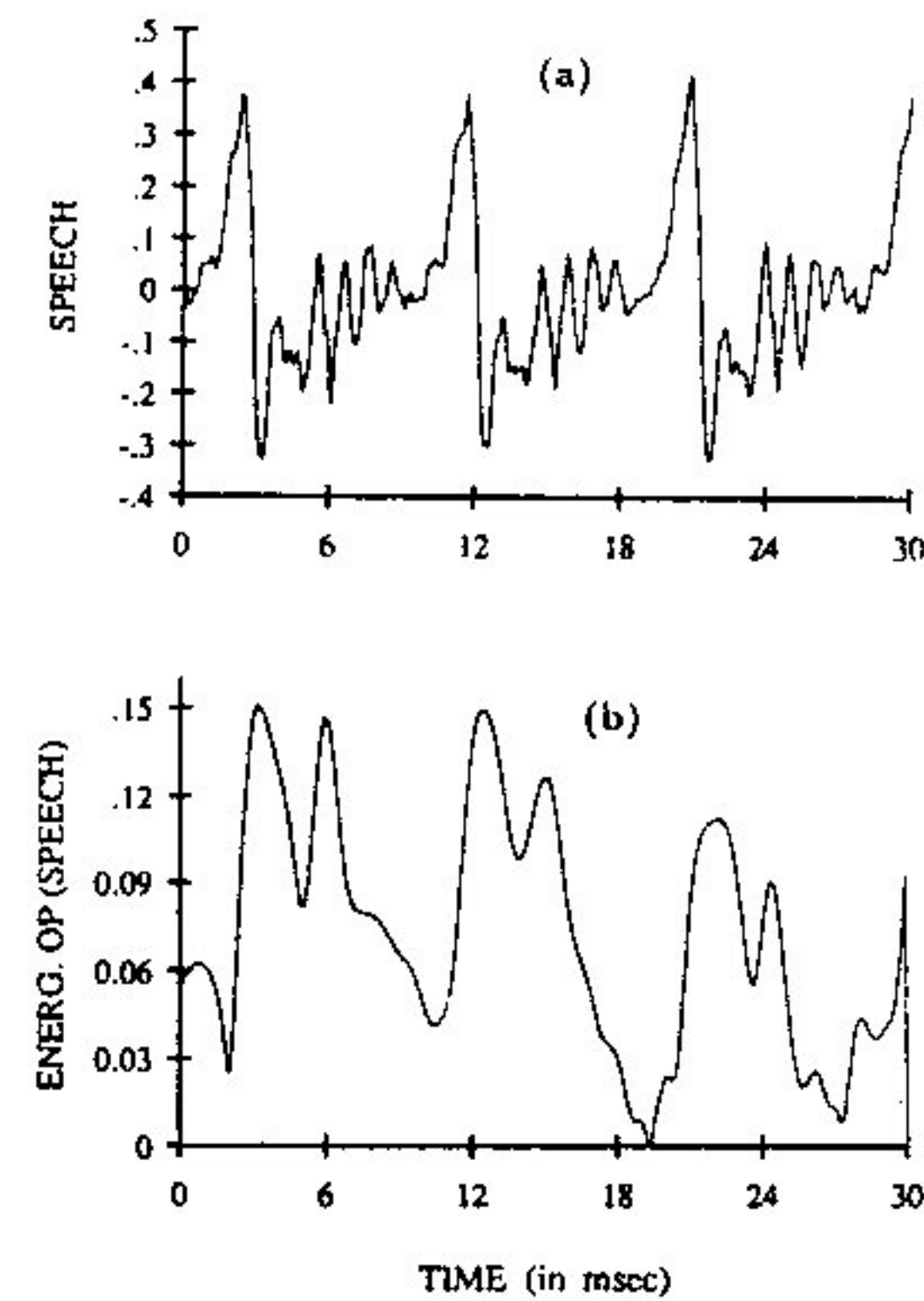


Figure 3. (a) Sustained speech vowel /a/. (b) Output of energy operator  $\sqrt{\Psi}$  when applied to a resonance of (a).