

## BÀI TẬP XỬ LÝ TIẾNG NÓI

1. Hàm truyền đạt của một bộ lọc số ở tần số formant  $F_k$  được cho bởi:

$$H_k(z) = \frac{1 - 2|z_k|\cos\theta_k + |z_k|^2}{1 - 2|z_k|\cos\theta_k z^{-1} + |z_k|^2 z^{-2}}$$

trong đó  $|z_k| = e^{-\sigma_k T}$ ,  $\theta_k = 2\pi F_k T$ ,  $T$ : chu kỳ lấy mẫu,  $2\sigma_k$ : dải thông.

- (a) Vẽ các điểm cực của  $H_k(z)$  trong mặt phẳng  $Z$
- (b) Viết phương trình sai phân mô tả quan hệ giữa tín hiệu ra  $y_k(n)$  và tín hiệu vào  $x_k(n)$
- (c) Vẽ sơ đồ khối của bộ lọc số này với 3 bộ nhân.
- (d) Bằng cách sắp xếp lại các số hạng của phương trình sai phân, vẽ sơ đồ khối của bộ lọc số chỉ có 2 bộ nhân

2. Hệ số phản xạ  $r_k$  được định nghĩa:  $r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k}$ . Chứng minh rằng  $r_k$  thỏa mãn :  
 $-1 \leq r_k \leq 1$

3. Một hệ thống phát hiện tần số cơ bản gồm có  **$M$  bộ lọc** thông dải giả thiết là lý tưởng có tần số cắt về phía tần số thấp của bộ lọc thứ  $k$  cho bởi:

$$F_k = 2^{k-1} F_1 \text{ với } k = 1, 2, \dots, M.$$

và tần số cắt về phía tần số cao cho bởi:

$$F_{k+1} = 2^k F_1 \text{ với } k = 1, 2, \dots, M.$$

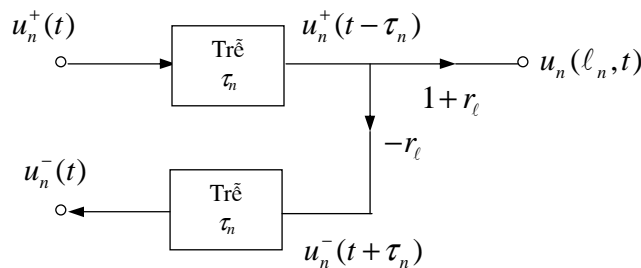
Các tần số cắt này được lựa chọn sao cho nếu đưa một tín hiệu tuần hoàn có **tần số cơ bản**  $F_0$  tới đầu vào của bộ lọc với  $F_k < F_0 < F_{k+1}$  thì đầu ra của các bộ lọc thứ 1 đến  $k-1$  sẽ có năng lượng rất yếu, đầu ra bộ lọc thứ  $k$  sẽ cho tần số cơ bản còn các bộ lọc từ thứ  $k+1$  đến  $M$  sẽ cho ra 1 hoặc nhiều hài của tần số cơ bản. Ở đầu ra của mỗi bộ lọc người ta sẽ có các bộ phát hiện và xác định tần số cơ bản.

- (a) Hãy xác định và giải thích cách lựa chọn  $F_1$  và  $M$  sao cho hệ thống này có thể đánh giá được các tần số cơ bản trong khoảng  $50 \text{ Hz} < F_0 < 800 \text{ Hz}$ .
- (b) Vẽ sơ đồ khối của hệ thống này và giải thích nguyên lý làm việc của nó.

4. Mô hình thực tế của trở kháng bức xạ của tuyến âm cho bởi

$$P_\ell(\ell, \Omega) = Z_\ell(\Omega) U(\ell, \Omega) \text{ với } Z_\ell(\Omega) = \frac{j\Omega L_r R_r}{R_r + j\Omega L_r}$$

trong đó  $P_\ell(\ell, \Omega)$  và  $U(\ell, \Omega)$  lần lượt là áp suất và thông lượng tại môi. Giả thiết rằng tiết diện của ống âm cơ bản ở đầu cuối của tuyến âm (về phía môi) là  $A_n$ . Ta có thể biểu diễn đầu cuối của tuyến âm theo sơ đồ như sau

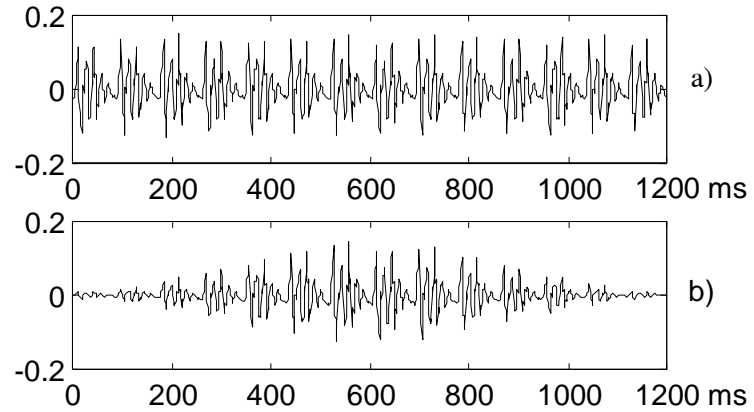


Hãy thiết lập quan hệ giữa các biến đổi Fourier của  $u_n^-(t + \tau_n)$  và  $u_n^+(t - \tau_n)$

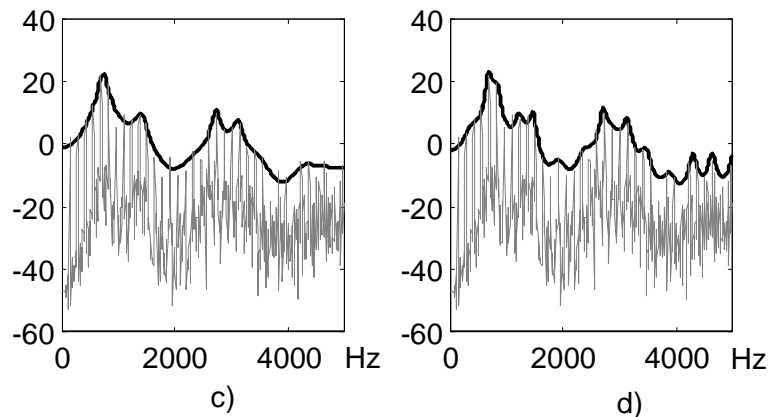
5.

(a) Hình vẽ a) là dạng tín hiệu tiếng nói. Hãy cho biết đây là tín hiệu ứng với âm hữu thanh hay vô thanh. Vì sao ?

(b) Trước khi phân tích phổ, dạng tín hiệu a) có thể được biến đổi thành b). Cách biến đổi như thế nào và biến đổi như vậy để làm gì ?



(c) Hình c) được tính toán xuất phát từ hình b). Hình c) gồm đường nét mảnh và đường nét đậm. Đường nét mảnh là gì và đường nét đậm là gì ? Những thông tin có được từ hình vẽ c) là gì ? Từ tín



hiệu hình b) có thể thực hiện biến đổi thế nào để có đường nét mảnh ? Cũng từ tín hiệu hình b) có bao nhiêu phương pháp để có đường nét đậm ? Trình bày các phương pháp đó.

(d) So sánh 2 hình vẽ c) và d) có thể thấy rằng đường nét mảnh của 2 hình là như nhau nhưng đường nét đậm của hình c) trơn hơn đường nét đậm của hình d). Từ các phương pháp đã nói ở mục 3), thay đổi tham số nào sẽ dẫn đến sự khác nhau này ?

6. Hàm truyền đạt của bộ lọc hiệu chỉnh dùng trong xử lý tiếng nói có dạng  $H(z) = 1 - az^{-1}$  với  $a$  là hằng số.

(a) Xác định biểu thức đáp ứng tần số của bộ lọc hiệu chỉnh

(b) Xác định và vẽ dạng đáp ứng biên độ của bộ lọc hiệu chỉnh với  $a = 0,98$ .

(c) Viết phương trình sai phân cho quan hệ vào-ra của bộ lọc này.

7. Tín hiệu tại thanh môn thường được mô phỏng dựa trên biểu thức sau:

$$g(n) = \begin{cases} na^n & n \geq 0 \\ 0 & n < 0 \end{cases}$$

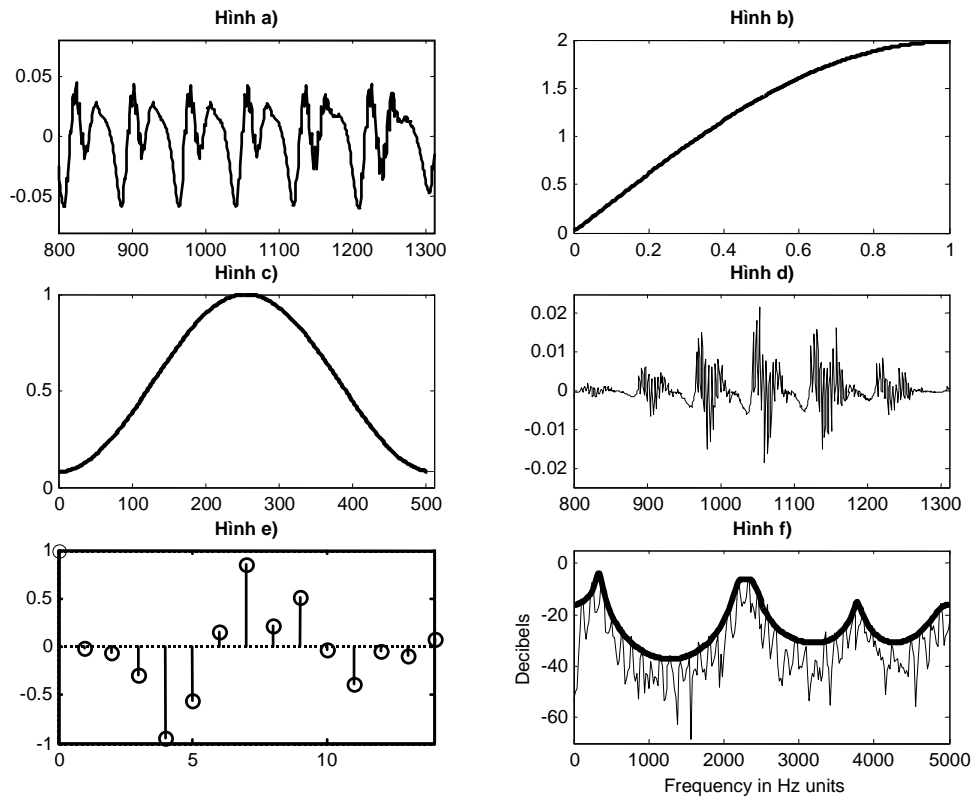
trong đó  $a$  là hằng số.

- (a) Hãy xác định biến đổi Z của  $g(n)$ .  
 (b) Xác định biên độ biến đổi Fourier của  $g(n)$ , tức là  $|G(e^{j\omega})|$ .  
 (c) Cần phải chọn  $a$  như thế nào để cho:

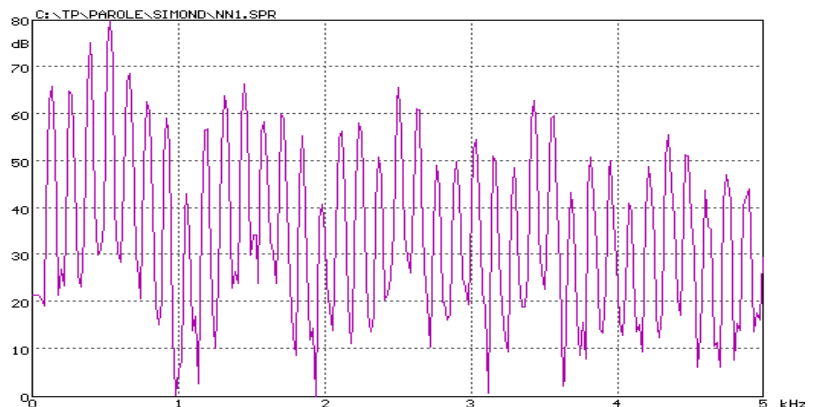
$$20 \lg |G(e^{j0})| - 20 \lg |G(e^{j\pi})| = 60 \text{ dB}$$

8. Các hình vẽ dưới đây liên quan đến các thao tác đối với tín hiệu tiếng nói ở hình a).

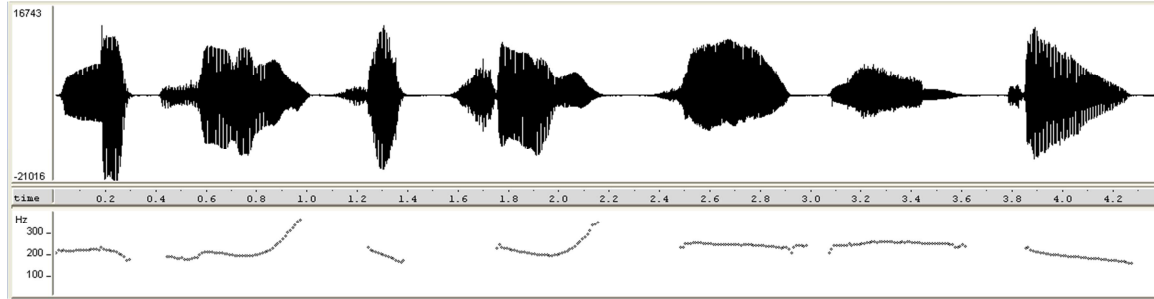
- (a) Hãy đưa ra các thông tin thu nhận từ các hình này.  
 (b) Thao tác nào liên quan đến các hình b), d), e), f). Giải thích các thao tác đó.



9. Phổ của một nguyên âm cho bởi hình vẽ bên. Biết rằng tần số lấy mẫu bằng 10 kHz và tuyến âm ứng với ống âm có tiết diện đều. Hãy xác định các tần số formant, tần số cơ bản và giải thích kết quả.



10. Hình vẽ dưới đây biểu diễn cho các âm tiết tiếng Việt: “phòng”, “học”, “xuống”, “bằng”, “đóng”, “trang”, “mục”. Hãy cho biết đây là các đại lượng nào của tiếng nói và chọn đoạn tín hiệu tương ứng với các âm tiết này đồng thời giải thích sự lựa chọn đó. (Chú ý: các âm tiết không xuất hiện trên hình theo đúng thứ tự như đã liệt kê)

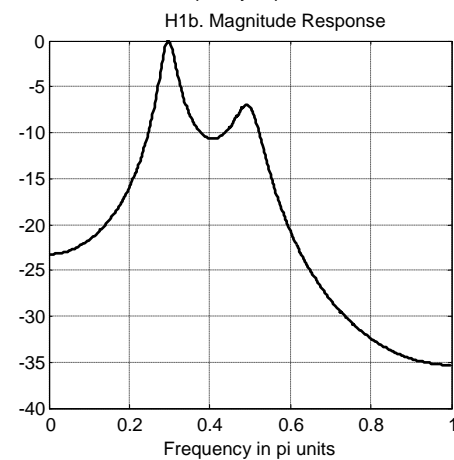
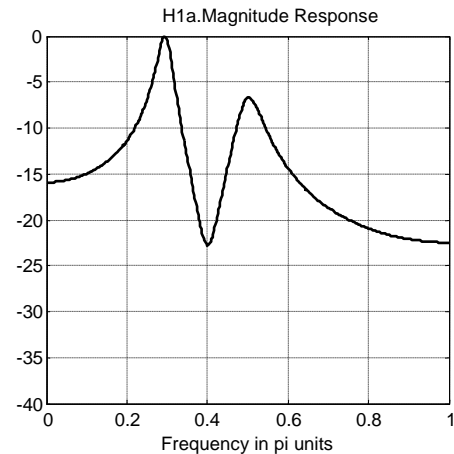


11. Xét hai hàm truyền đạt  $G(z)$  và  $H(z)$  như sau:

$$G(z) = \frac{1}{1 - 1,14z^{-1} + 0,9025z^{-2}}$$

$$H(z) = \frac{1}{1 + 0,81z^{-2}}$$

- Xác định các điểm cực, điểm không của hai hàm này và vẽ trên mặt phẳng  $z$ .
- Xét hai trường hợp ghép nối  $Y_1(z) = G(z)H(z)$ ,  $Y_2(z) = G(z) + H(z)$ . Tìm các điểm cực, điểm không của  $Y_1(z)$ ,  $Y_2(z)$  và vẽ trên mặt phẳng  $z$ . Nếu đây là hai trường hợp tổng hợp formant thì nói rõ đó là các trường hợp nào?
- Hãy chọn ra đáp ứng biên độ của  $Y_1(z)$ ,  $Y_2(z)$  từ H1a, H1b và giải thích sự lựa chọn đó. Trường hợp nào của H1a, H1b gần hơn với mô hình AR? Tại sao?



12. Cho tín hiệu tuần hoàn  $x(n)$

$$x(n) = \dots + 5\delta(n) + 3\delta(n-1) + 2\delta(n-2) + 5\delta(n-4) + 3\delta(n-5) + 2\delta(n-6) + 5\delta(n-8) + \dots$$

Hãy áp dụng phương pháp tự tương quan cải tiến và diễn giải từng bước để xác định chu kỳ  $T_0$  của tín hiệu này. Biết chu kỳ lấy mẫu bằng  $3 \text{ ms}$ .

**13.** Tín hiệu vào  $x(n)$  qua bộ lọc hiệu chỉnh sau đó nhân với cửa sổ Hamming. Bộ lọc hiệu chỉnh có hàm truyền đạt  $H(z) = 1 - 0,95z^{-1}$ . Biểu thức cửa sổ Hamming có dạng  $w(n) = 0,54 - 0,46 \cos\left(2\pi \frac{n}{N-1}\right)$ ,  $n = 0, 1, \dots, N-1$ ,  $N$  là độ rộng của sổ. Giả sử  $N = 16$ , tín hiệu vào bộ lọc hiệu chỉnh có dạng  $x(n) = u(n) - u(n - N)$ .

(a) Hãy tính tín hiệu ra bộ lọc hiệu chỉnh  $x_1(n)$  và tín hiệu sau cửa sổ Hamming  $x_2(n)$

(b) Giả thiết  $x_2(n)$  được đưa tới khối FFT. Hãy phát biểu tác dụng của cửa sổ Hamming bằng một câu duy nhất.

**14.** Đáp ứng biên độ của tuyến âm xác định được như sau:  
 $H(k) = 49\delta(k) + 52\delta(k-1) + 60\delta(k-2) + 50\delta(k-3) + 48\delta(k-4) + 57\delta(k-5) + 47\delta(k-6) + 40\delta(k-7) + 39\delta(k-8) + 44\delta(k-9) + 35\delta(k-10) + 26\delta(k-11) + 21\delta(k-12) + 18\delta(k-13) + 15\delta(k-14) + 14\delta(k-15)$

(a) Hãy nhận xét tuyến âm có bao nhiêu formant và các formant ứng với mẫu nào

(b) Trình bày giải thuật lập trình xác định các mẫu ứng với các formant này, kiểm tra xem giải thuật cho kết quả có phù hợp với nhận xét ở mục a) hay không

(c) Giả thiết tần số lấy mẫu  $F_s = 8 \text{ kHz}$ . Trình bày cách xác định giá trị các tần số formant ứng với các mẫu nói trên.

**15.** Bộ lọc cộng hưởng ở tần số formant có hàm truyền đạt  $H(z) = \frac{1}{1 - 1,4z^{-1} + 0,98z^{-2}}$   
 Biết tần số lấy mẫu bằng  $8 \text{ kHz}$ .

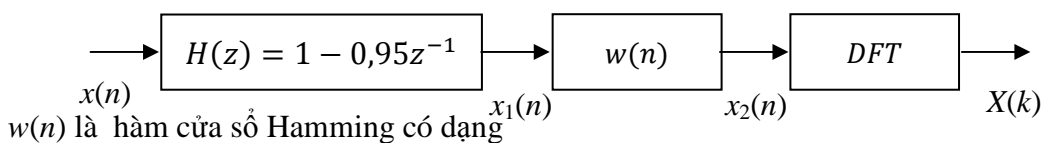
(a) Hãy xác định các điểm cực của  $H(z)$  và mô đun của  $H(z)$  ở các tần số  $0 \text{ Hz}$ ,  $1 \text{ kHz}$  và  $4 \text{ kHz}$ .

(b) Bộ lọc trên có thể được biến đổi bằng cách định nghĩa  $G(z) = H\left(\frac{z}{k}\right)$  trong đó  $0 < k < 1$ . Hãy xác định các điểm cực của  $G(z)$  và mô đun của  $G(z)$  ở các tần số  $0 \text{ Hz}$ ,  $1 \text{ kHz}$  và  $4 \text{ kHz}$  cho trường hợp  $k = 0,95$ .

(c) Vẽ dạng đáp ứng biên độ tương ứng với  $H(z)$  và  $G(z)$  trên cùng hình vẽ. Đáp ứng biên độ nào có dải thông rộng hơn và vì sao ?

**16.** Giả thiết tín hiệu  $x(n)$  có dạng  $x(n) = 2\delta(n) + 2\delta(n-1) + 2\delta(n-2) + 2\delta(n-3) + 2\delta(n-4)$

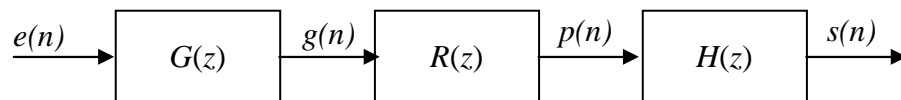
Tín hiệu  $x(n)$  được xử lý theo sơ đồ khối như sau:



$$w(n) = \begin{cases} 0,54 - 0,46 \cos\left(\frac{2\pi n}{M-1}\right), & 0 \leq n \leq M-1 \\ 0, & n \text{ còn lại} \end{cases}$$

$M$  là độ dài tín hiệu. Hãy tính và vẽ các tín hiệu  $x_1(n)$ ,  $x_2(n)$  và  $|X(k)|$ . Cho nhận xét về  $|X(k)|$ .

**17.** Tuyến âm có thể được mô phỏng theo sơ đồ khối như sau

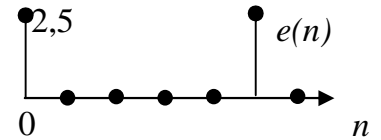


$$G(z) = \frac{1}{1 - z^{-1} + 0,5z^{-2}}, R(z) = 1 - 0,7z^{-1}, H(z) = \frac{1}{1 - \sum_{i=1}^4 a_i z^{-i}}$$

$$a_1 = 0,3, a_2 = -0,2, a_3 = 0,1, a_4 = -0,05$$

Khung tín hiệu  $e(n)$  gồm 7 mẫu có dạng như hình bên.

Hãy trình bày cách tính toán và vẽ các tín hiệu  $g(n)$ ,  $p(n)$  và  $s(n)$  cho một khung tín hiệu  $e(n)$ .



**18.**

- (a) Khung tín hiệu tiếng nói có 240 mẫu với tần số lấy mẫu  $F_s = 8 \text{ kHz}$ . Tính biến đổi Fourier của khung này dùng 1024 điểm  $FFT$  ta được 1024 mẫu  $X(k)$  với  $k = 1, 2, \dots, N$ . Tần số tương ứng của mẫu thứ 55 của  $X(k)$  sẽ bằng bao nhiêu và tại sao ?
- (b) Tần số tương ứng bằng  $\text{Hz}$  sẽ là bao nhiêu đối với tần số góc chuẩn hóa bằng  $\pi/4$ . Biết tần số lấy mẫu  $F_s = 10 \text{ kHz}$ .

## Một số đặc điểm ngữ âm tiếng Việt

- Âm tắc: tiếng nổ, phát sinh do luồng khí từ phổi đi ra bị cản trở hoàn toàn, phải phá vỡ sự cản trở đó để thoát ra.
- Âm xát: tiếng cọ xát, phát sinh do luồng không khí đi ra bị cản trở không hoàn toàn (chỉ bị khó khăn), phải lách qua một khe hở nhỏ và trong khi thoát ra như vậy phải cọ xát vào thành của bộ máy phát âm.
- Phụ âm bên: đầu lưỡi tiếp xúc với lợi chặn lối thoát của không khí, buộc nó phải lách qua khe hở ở hai bên cạnh lưỡi tiếp giáp với má mà ra ngoài tạo nên tiếng xát nhẹ (l).
- Luồng không khí thoát ra ngoài bị cản trở, tạo nên tiếng xát hay tiếng nổ, dạng tín hiệu không tuần hoàn gọi là tiếng động (ồn).
- Trong khi phát âm một số phụ âm, dây thanh cũng hoạt động đồng thời tạo nên tiếng thanh.
- Phụ âm có tỉ lệ tiếng động lớn hơn gọi là phụ âm ồn.
- Phụ âm có tỉ lệ tiếng thanh lớn hơn gọi là phụ âm vang.

1

## Một số đặc điểm ngữ âm tiếng Việt

- Phân loại phụ âm theo tắc hay xát, hữu thanh hay vô thanh, mũi hóa

Vị trí cấu âm				Môi	Đầu lưỡi		Mặt lưỡi	Cuối lưỡi	Họng
Phương thức cấu âm					Răng	Vòm miệng			
Bật hơi					th				
Tắc	Ồn	Không bật hơi	Vô thanh	p	t	tr	ch	c,k,qu	
			Hữu thanh	b	đ				
			Vang mũi		m	n		nh	ng,ngh
	Xát	Ồn	Vô thanh	ph	x	s		kh	h
			Hữu thanh	v	d,gi	r		g	
Vang bên			l						

32