

CHẤT GIỌNG CỦA NGƯỜI VIỆT QUA CÁC CHỈ SỐ RUNG ĐỘNG DÂY THANH

VŨ THI HẢI HÀ*

Abstract: Vocal cord disease is a common disease not only in Vietnam but also globally. According to some studies, vocal cord diseases account for 58% of laryngeal diseases [2], [3], [6]. These diseases affect the quality of the voice, mainly causing the voice to be hoarse, unclear in tone, and over time lead to loss of voice, making it difficult for patients to communicate in social life. However, research related to vocal cord pathology is still very limited and is in need of research attention. In this article, we describe the vibration of the vocal cords through two parameters, the frequency disturbance parameter (Jitter) and the amplitude characteristic evaluation parameter (Shimmer) to provide acoustic parameters that describe the characteristics of the Vietnamese speech.

Keywords: *voice quality, vocal cords, phonation, Shimmer, Jitter*

1. Dẫn nhập

Có thể nói rằng bệnh lý dây thanh rất thường gặp ở Việt Nam cũng như trên thế giới. Theo một số nghiên cứu, các bệnh về dây thanh đứng hàng đầu chiếm 58% trong các bệnh lý về thanh quản [2], [3], [6]. Những tổn thương ở dây thanh được hình thành do hậu quả của những chấn thương trong quá trình phát âm, gặp ở nhiều lứa tuổi khác nhau, cả nam và nữ. Những bệnh dây thanh thường gặp như: Viêm thanh quản mãn, hạt dây thanh, polip dây thanh, sẹo dây thanh, dính dây thanh, lao thanh quản, ung thư thanh quản... đều để lại sự tổn thương dây thanh. Những bệnh này ảnh hưởng tới chất lượng của giọng nói, chủ yếu là giọng nói bị khàn, không rõ âm sắc, nhanh mệt và lâu dần dẫn đến mất giọng gây khó khăn cho bệnh nhân khi giao tiếp trong đời sống xã hội.

Tại Việt Nam, những số liệu nghiên cứu và những công trình nghiên cứu trong lĩnh vực này có thể kể đến một số công trình của Lê Sỹ Nhơn (1991) nhận xét về đặc điểm dịch tễ lâm sàng và kết quả điều trị 252 ca rối loạn giọng được điều trị tại Viện Tai Mũi Họng từ năm 1986 - 1990; Ngô Ngọc Liễn, Nguyễn Văn Lợi, Trần Công Hòa (2002) nghiên cứu sự biến đổi thanh điệu của giáo viên tiểu học bị bệnh thanh quản và bước đầu nhận định phân tích âm học trong chẩn đoán bệnh và giám định nghề. Phạm Thị Ngọc (2010) nghiên cứu bệnh giọng nghề nghiệp của giáo viên tiểu học huyện Đông Anh. Trần Duy Ninh (2011) nghiên cứu rối loạn giọng nói của giáo viên tiểu học thành phố Thái Nguyên cho thấy tỉ lệ mắc rối loạn giọng nói. Những nghiên cứu này trong cộng đồng còn rất khiêm tốn và cần được quan tâm nghiên cứu.

Những nghiên cứu về dây thanh có ý nghĩa khoa học và thực tiễn không chỉ đối với y học mà còn cả ngôn ngữ học. Vì vậy, trong bài báo này, chúng tôi tiến hành miêu tả chất giọng của người Việt thông qua hai thông số, thông số đánh giá sự nhiễu loạn về tần số (Jitter) và

* TS., Viện Ngôn ngữ học

chỉ số đánh giá đặc trưng về biên độ (shimmer) nhằm đưa ra những tham số âm học về đặc trưng chất giọng của người Việt.

2. Một số vấn đề lí thuyết

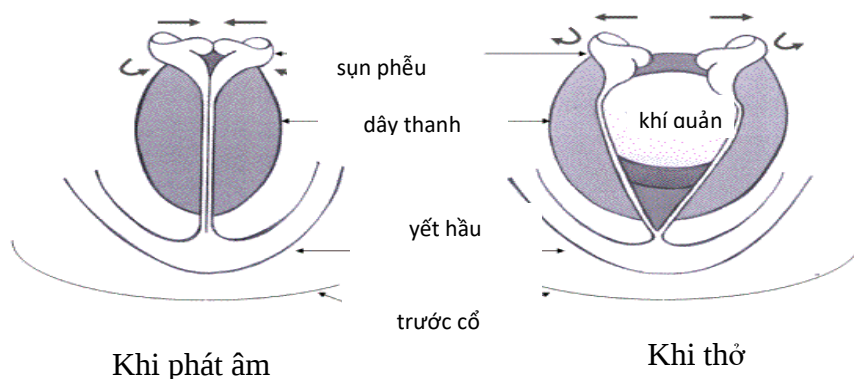
2.1. Dây thanh

Dây thanh (*vocal folds*) còn gọi là *dây thanh đôi*: chính là một cặp dây bằng màng nhầy được kéo căng từ sau ra trước, nằm bên trong thanh quản (larynx). Cặp dây gồm hai sợi ở cơ đàn hồi, hình nẹp có cấu trúc niêm mạc. Một đầu dây thanh nối liền với sụn giáp cổ định, một đầu kia nối với sụn phễu. Nhờ các cơ mà sụn phễu có thể cử động linh hoạt làm cho hai dây thanh có thể khép chặt để ngăn luồng hơi từ phổi đi ra hoặc mở ra theo những dạng khác nhau. Dây thanh rung động do sự đẩy hơi từ phổi để tạo ra âm thanh khi chúng ta nói, hát, la hét,... Khi lấy hơi vào, toàn bộ dây thanh sẽ mở ra để đưa không khí vào phổi và sẽ đóng lại khi chúng ta cố giữ hơi thở. Khi nói hoặc hát, dây thanh sẽ rung liên tục, số lần rung động tùy thuộc vào cao độ âm thanh.

Bên trên dây thanh là nắp thanh quản (epiglottis) có chức năng như một van đóng mở, khi nuốt thức ăn và đồ uống nắp thanh quản sẽ đóng lại để thức ăn và đồ uống không rơi vào thanh quản. Khi những thức ăn đồ uống đó đi sai đường, rơi vào thanh quản chính là lúc xảy ra hiện tượng ho, sặc.

Kích thước dây thanh của nam và nữ sẽ khác nhau. Đối với nam trưởng thành, giọng nói thường trầm hơn do dây thanh dài hơn và dày hơn. Dây thanh của nam thường dài khoảng 1.75 cm đến 2.5 cm, trong khi đối với nữ độ dài thường chỉ từ 1.25 đến 1.75 cm. Đối với trẻ em, dây thanh sẽ ngắn hơn nhiều so với cả nam và nữ trưởng thành. Sự khác nhau trong kích thước dây thanh dẫn đến sự khác nhau về cao độ trong giọng nói của mỗi người. Dây thanh dày lên theo tuổi tác, đặc biệt vào tuổi 14 -15, dây thanh dày lên rất nhanh, tạo nên hiện tượng gọi là “vỡ giọng” tuổi dậy thì.

Dây thanh giả (false vocal folds) là cặp màng nhầy nằm bên cạnh dây thanh có tác dụng bảo vệ, khi chúng ta la hét hay sử dụng quá nhiều lực, 2 dây thanh đôi sẽ thực hiện nhiệm vụ bảo vệ bằng cách đè lên phía trên để giảm tổn thương dây thanh do lực đẩy quá mạnh từ phổi.



Hình 1. Minh họa thanh quản, nhìn từ trên khi phát âm và khi thở
(<http://www.phys.unsw.edu.au/jw/voice.html>)

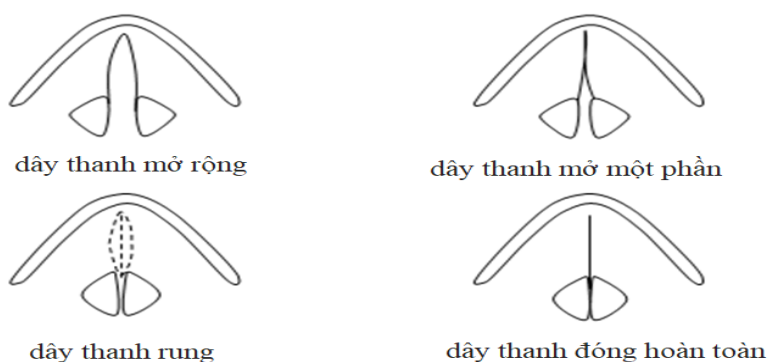
Khe thanh (glottis) là khe hở giữa hai dây thanh khi chúng chuyển động theo chiều ngang sang hai bên nhờ sụn phễu. Khi các cơ co lại, hai dây thanh ngấn đi và độ căng của nó cũng thay đổi. Khác với dây thanh thật, trong dây thanh giả không có cơ thịt, nó cấu tạo bằng vô số những lớp mô ngăn cách nhau bằng những tuyến nhờn nhỏ. Âm ban đầu được tạo ra ở hầu do sự rung của hai dây thanh. Nguyên tắc tạo âm đó được gọi là lí thuyết cơ đàn hồi (nyo- elastish) được đa số các nhà khoa học ủng hộ. Áp lực tăng lên đến mức làm khe thanh hé mở, không khí thoát ra ngoài làm áp lực giảm xuống, dây thanh lại trở về vị trí cũ. Cứ như vậy quá trình khép mở của hai dây thanh lặp đi lặp lại. Do tính chất đàn hồi của dây thanh, âm ban đầu được phát ra nhờ sự rung của dây thanh trong quá trình khép - mở. Dây thanh rung nhanh thì âm phát ra cao, dây thanh rung chậm thì âm phát ra trầm [22, tr.148].

2.2. Trạng thái của dây thanh và các thức tạo thanh

Tiêu chí thường được dùng để đối lập tiếng thanh trong ngôn ngữ là tính thanh: hữu thanh - vô thanh. Tính chất hữu thanh - vô thanh của âm phụ thuộc vào trạng thái của thanh môn khi phát âm. Nếu hai dây thanh khép lại gần nhau, khe thanh môn hẹp, không khí đi qua sẽ làm rung động dây thanh; nếu hai dây thanh tách xa nhau, khe thanh môn rộng, luồng không khí đi qua sẽ không thể làm rung động dây thanh.

Cả hai âm hữu thanh và vô thanh đều thuộc thức tạo thành **thường** (modal phonation). Bên cạnh đó còn hai thức tạo thanh khác được nhiều ngôn ngữ trên thế giới sử dụng: thức tạo thanh *thở* và thức tạo thanh *kẹt thanh*.

Trong thức tạo thanh **thở** (breathy phonation), các dây thanh gần nhau như khi tạo âm vô thanh, khi phát âm có rất ít hoặc không có luồng không khí đi qua thanh môn. Trong âm thức tạo thanh này, các nếp gấp thanh âm nằm tách biệt giữa các sụn ở phần dưới. Chúng vẫn có thể rung, nhưng đồng thời, một lượng lớn không khí đi qua thanh môn.



Hình 2. Bốn trạng thái khác nhau của dây thanh (theo Peter Roach)

Trong thức tạo thanh **kẹt thanh** (creaky voice), thanh môn bị thu ngắn lại do sụn phễu khép chặt. Hai dây thanh khép sát nhau ở một đầu nhưng hở ở đầu còn lại, do đó khi không khí đi qua thanh môn, dây thanh vẫn có thể rung động, nhưng rung động ở tần số thấp. Âm kẹt thanh còn được gọi bằng tên khác là âm thanh *quản hóa* (laryngealized).



Hình 3. Các trạng thái của dây thanh trong tạo âm thanh tiếng nói (theo [22, tr.150])

A (glottalized): thanh quản hóa

D (modal): thường

B (creaky): kẹt thanh

E (slack): chùng

C (stiff): căng

F (voiceless): vô thanh

2.3. Các chỉ số jitter, shimmer trong nghiên cứu sự rung động của dây thanh

2.3.1. Khái niệm Jitter

Khái niệm này đã được đề cập đến trong nghiên cứu của João Paulo Teixeira, Paula Odete Fer [13] "Jitter là tham số chỉ sự nhiễu loạn về tần số cho phép đo lường sự thay đổi của tần số thanh cơ bản khi chuyển từ một chu kỳ này sang một chu kỳ kế tiếp và đánh giá sự khác nhau về tần số giữa một chu kỳ và các chu kỳ kế tiếp. Jitter cho biết tính ổn định của cơ quan tạo thanh. Trong trường hợp người phát âm có các bệnh lý về giọng nói tham số sẽ Jitter thay đổi, điều này cho thấy cơ quan phát âm kém ổn định. Jitter tăng cho thấy sự rối loạn giọng nhưng không cho phép xác định nguyên nhân bệnh lý gây ra rối loạn này. Đơn vị đo jitter bằng ms hoặc tỉ lệ % của chiều dài chu kỳ". Jitter có thể được phân tích trên đoạn tín hiệu phát âm nguyên âm liên tục. Các jitter bị ảnh hưởng chủ yếu bởi sự thiếu kiểm soát rung động của dây thanh; tiếng nói của bệnh nhân thường có tỉ lệ jitter cao hơn. Hầu hết các nhà nghiên cứu coi là biến thể giá trị điển hình từ 0,5 đến 1,0% cho âm vị được duy trì ở người trẻ [8].

Các giá trị cho jitter có thể được đo bằng các tham số khác nhau. Jitter có thể được đưa ra bởi bốn tham số liên quan: jitter tuyệt đối (jitta) là nhiễu loạn tuyệt đối, jitter cục bộ hoặc tương đối (jitt), nhiễu loạn trung bình tương đối (rap) và nhiễu loạn thời gian năm điểm (ppq5).

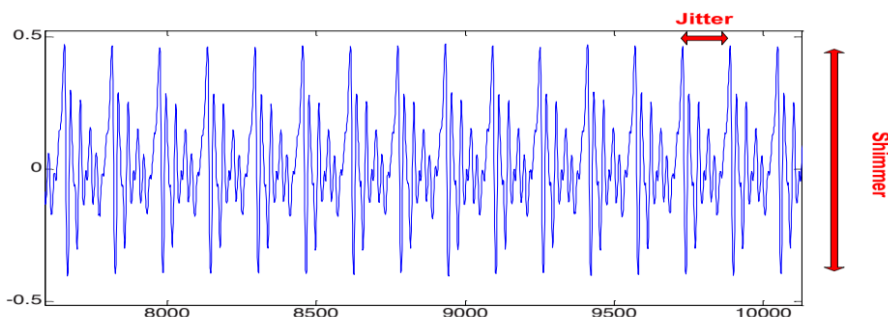
Jitter tuyệt đối là sự biến thiên theo chu kỳ của tần số cơ bản, nghĩa là chênh lệch tuyệt đối trung bình giữa các chu kỳ liên tiếp. Jitter tương đối hoặc Jitter cục bộ (jitt) là chênh lệch tuyệt đối trung bình giữa các giai đoạn liên tiếp, chia cho thời gian trung bình. Nó được biểu thị dưới dạng phần trăm. Jitter (rap) được hiểu là độ nhiễu loạn trung bình tương đối, sự khác biệt trung bình tuyệt đối giữa một chu kỳ và trung bình của nó với hai chu kỳ cạnh nó, chia cho thời gian trung bình. Nó được biểu thị dưới dạng phần trăm. Jitter (ppq5) là đơn vị đo độ nhiễu loạn theo chu kỳ năm điểm, được tính là trung bình của nó với bốn chu kỳ gần nhất, chia cho thời gian trung bình. Nó cũng được biểu thị dưới dạng phần trăm

2.3.2. Khái niệm shimmer

Là tham số chỉ sự nhiễu loạn về biên độ âm, nó xác định sự khác nhau về biên độ sóng âm giữa một chu kỳ với các chu kỳ liên tiếp, shimmer thay đổi trong trường hợp bệnh lý ở thanh quản, bệnh nhân rối loạn giọng thì shimmer tăng nhưng không có ý nghĩa quan trọng bằng jitter.

Shimmer (dB) được biểu thị bằng độ biến thiên của biên độ đỉnh cực đại tính bằng decibel, tức là logarit cơ số tuyệt đối trung bình 10 của chênh lệch giữa biên độ của các chu kỳ liên tiếp, nhân với 20. Hay nói cách khác, shimmer là tham số chỉ sự nhiễu loạn về biên độ âm, nó xác định sự khác nhau về biên độ sóng âm giữa một chu kỳ với các chu kỳ liên tiếp, shimmer thay đổi trong trường hợp bệnh lý ở thanh quản, bệnh nhân rối loạn giọng thì shimmer tăng.

Shimmer tương đối (relative) được xác định là sự chênh lệch tuyệt đối trung bình giữa các biên độ của các chu kì liên tiếp, chia cho biên độ trung bình, giá trị của shimmer tương đối được biểu thị bằng phần trăm. Shimmer (apq3) là đơn vị đo sự nhiễu loạn biên độ theo chu kì ba điểm. Sự chênh lệch trung bình tuyệt đối giữa biên độ của một chu kì và trung bình biên độ của các chu kì kế tiếp, chia cho biên độ trung bình. Nó được biểu thị bằng tỉ lệ phần trăm. Shimmer (apq5) được xác định là đơn vị đo sự nhiễu loạn biên độ theo chu kì năm điểm. Sự chênh lệch trung bình tuyệt đối giữa biên độ của một chu kì và biên độ trung bình của bốn chu kì liên tiếp gần nhất của nó, chia cho biên độ trung bình. Nó cũng được biểu thị bằng tỉ lệ phần trăm.



3. Phương pháp và tư liệu nghiên cứu

Bài viết sử dụng khí cụ phân tích ngữ âm nhằm miêu tả các đặc trưng âm học của âm tiết. Phương pháp phân tích, miêu tả ngữ âm âm vị học kết hợp với một số thủ pháp thống kê, so sánh cũng được sử dụng.

Tư liệu nghiên cứu bao gồm file ghi âm các âm tiết tiếng Việt do 6 CTV (3 nam, 3 nữ) giọng Bắc Bộ, tuổi từ 31 đến 65 tuổi. Cộng tác viên (CTV) được yêu cầu đọc một danh sách các âm tiết có chứa 3 nguyên âm /i, a, u/ (là những nguyên âm phổ biến ở các ngôn ngữ và ở ba đỉnh của tam giác nguyên âm) kết hợp với thanh điệu, phụ âm đầu, phụ âm cuối. Các nguyên âm được phát âm kéo dài hơn bình thường. Mỗi CTV đọc 3 lần. Nếu các CTV bị đọc vấp, đọc sai trong khi đọc các mẫu thì cho phép ghi âm lại các mẫu. Những người phát âm có thể lặp lại các mẫu cho đến khi nào đạt được một cách phát âm chính xác thì thôi.

Âm tiết mẫu bao gồm *xi, xa, xu, ti, ta, tu, mi, ma, mu, íp, áp, úp,...* Lời nói được đọc, và được ghi âm trực tiếp trên một máy tính xách tay qua chương trình Audacity tần số lấy mẫu ở tốc độ 22050Hz với hệ xử lí 16 bit, Mono. File âm thanh sau đó lưu ở định dạng *.wav. Mỗi người nói có tất cả 246 đoạn tín hiệu tiếng nói liên tục với độ dài trung bình 250ms/1 đoạn tín hiệu. Dữ liệu tiếng nói được ghi âm trong phòng kín nhiều nền thấp, không lẫn giọng nói khác và không lẫn tiếng động hay nhạc nền. Các nút ghi âm, nút dừng, nút chạy đều được người ghi âm điều khiển, vì vậy các CTV chỉ cần tập trung vào việc đọc văn bản trước mặt họ. Các CTV ngồi trước máy tính xách tay và micro kết hợp tai nghe được gắn lên đầu của họ, micro đặt cách miệng 7- 10 cm.

4. Kết quả nghiên cứu

Trong mục này, chúng tôi trình bày kết quả đo các thông số jitter, shimmer trên mẫu phát âm là các nguyên âm /i, a, u/ kết hợp với thanh điệu tiếng Việt. Các mẫu phát âm được phân tích qua phần mềm Praat. Sự khác biệt về giới tính cũng được đề cập đến. 9 thông số được miêu tả

bao gồm: jitter tuyệt đối (jitta) là nhiễu loạn tuyệt đối, jitter cục bộ hoặc tương đối (jitt), nhiễu loạn trung bình tương đối (rap), nhiễu loạn 5 điểm (ppq5), chênh lệch trung bình jitter giữa các giai đoạn liên tiếp, chia cho chu kì trung bình, shimmer cục bộ, decibe (shdB, shim), biên độ ba điểm thương số nhiễu (apq3), biên độ nhiễu năm điểm (apq5), chênh lệch tuyệt đối trung bình của chênh lệch liên tiếp giữa biên độ của các chu kì liên tiếp và tỉ lệ hài hòa âm (HNR).

4.1. Chỉ số đánh giá sự nhiễu loạn về tần số (jitter)

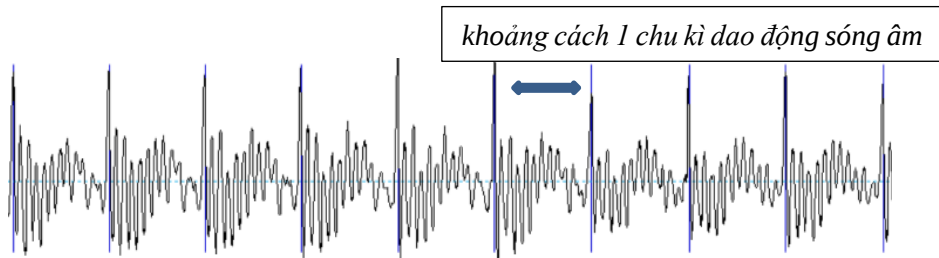
4.1.1. So sánh giá trị jitter giữa nam và nữ

Giá trị Jitter tương đối (Jitter local) hay giá trị chênh lệch trung bình tuyệt đối giữa các chu kì liên tiếp chia cho thời gian trung bình đối với nam giới và nữ giới người Việt khỏe mạnh trong các phát âm với thanh ngang, cao nằm trong khoảng 1,1 đến 2,4 %.

Có sự khác biệt về giới đối với các giá trị thu được. Sự khác biệt cho thấy jitter tương đối của giọng nữ cao hơn đáng kể so với giọng nam. Bảng thống kê trên cho biết các giá trị trung bình của jitter tương đối tăng dần ở nam giới đối với các nguyên âm /a, i, u/, trong khi phụ nữ cũng cho thấy sự gia tăng giá trị tương ứng đối với tham số này ở các nguyên âm /a, i, u/.

Giá trị biểu thị độ nhiễu loạn trung bình tương đối (jitter rap) của nam và nữ nằm trong khoảng 0,2 - 1,3 %, giá trị của nữ cao hơn nam, sự gia tăng giá trị cũng được thể hiện từ nguyên âm a đến u, trong đó a có giá trị thấp nhất, u có giá trị cao nhất.

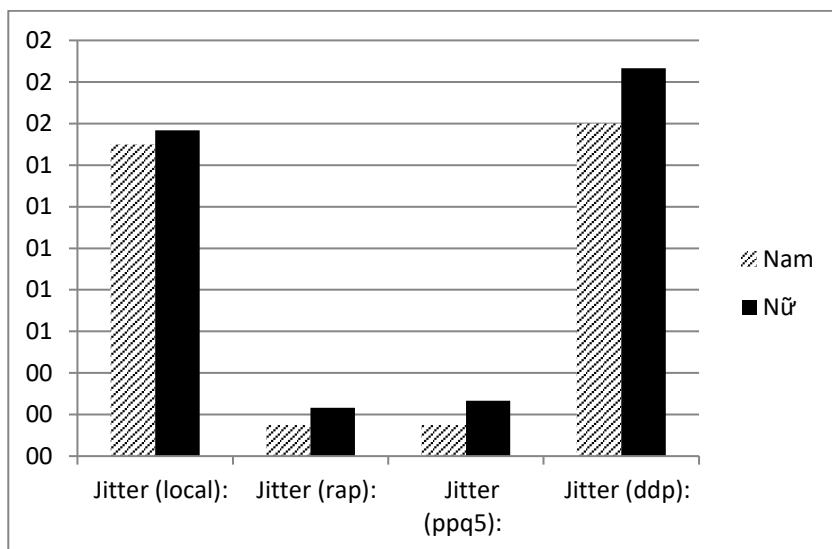
Cũng tương tự như jitter rap, giá trị đo độ nhiễu loạn theo chu kì năm điểm jitter (ppq5) cũng có sự gia tăng từ a đến u, nữ có xu hướng cao hơn nam. Giá trị này nằm trong khoảng 0,2 - 1,4%.



Hình 4. Dạng sóng và chu kì dao động sóng âm của nguyên âm /a/

Bảng 1: Giá trị của jitter đối với tín hiệu phát âm của nguyên âm /a, i, u/ với thanh ngang, cao, chất giọng thường

Jitter:	Nam			Nữ		
	a	i	u	a	i	u
Jitter (local): %	1,5	1,1	1,6	1,6	1,8	2,4
Jitta (μ s)	130	75	113	27,0	79,3	108,0
Jitter (rap): %	0,2	0,4	0,8	0,2	0,9	1,3
Jitter (ppq5): %	0,2	0,5	0,9	0,3	1,0	1,4
Jitter (ddp): %	1,6	1,3	2,6	1,9	2,8	3,9



Hình 5. Giá trị các jitter của nguyên âm /a/ giọng nam và nữ

Kết quả này có thể được giải thích vì giọng nam có tần số cơ bản thấp hơn và vì vậy thời gian glottal cao hơn giọng nữ. Do đó, trong một thời kì glottal dài hơn, nó được dự đoán sẽ gây nhiều loạn lâu hơn trong các thời kì glottal. Sự khác biệt về giới tính của F0 có thể được chứng minh do có sự khác biệt về mặt giải phẫu rõ rệt ở thanh quản của nam và nữ. Theo nhiều nghiên cứu, thanh quản của nam lớn hơn khoảng 40% so với nữ và nếp gấp thanh âm của nam là một khối dày hơn của nữ.

4.1.2. So sánh giá trị Jitter giữa các mẫu chứa thanh điệu

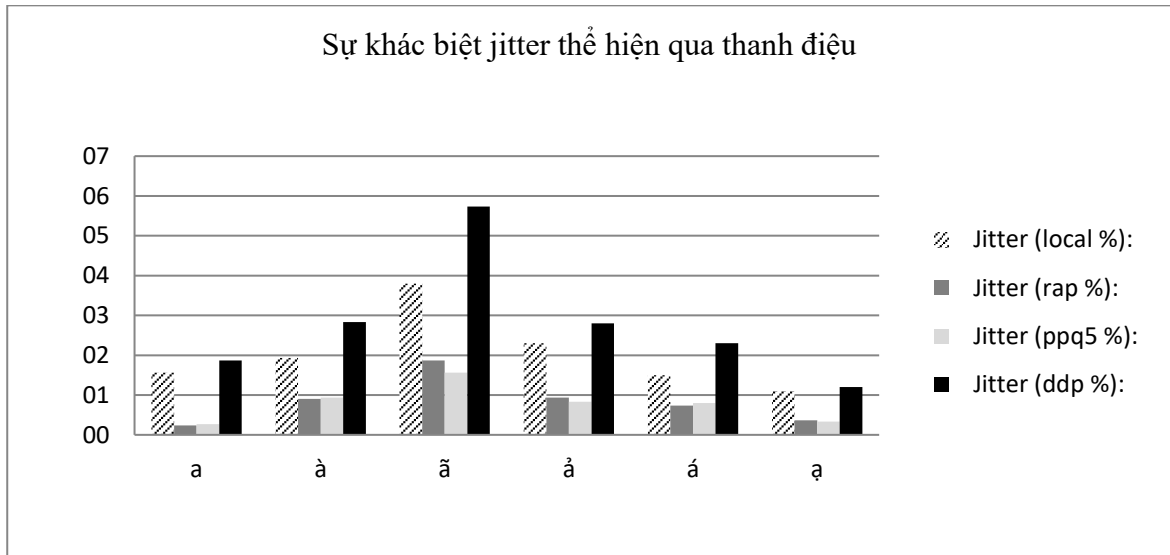
Những đặc trưng âm vị học của hệ thống thanh điệu tiếng Việt được thể hiện qua các tiêu chí về đường nét, âm vực và chất giọng. Giá trị jitter của các mẫu âm thanh chứa thanh điệu có sự khác biệt.

Bảng 2. Đặc trưng âm vị học hệ thanh điệu Bắc Bộ (theo Nguyễn Văn Lợi)

Thanh điệu	Cao độ		Chất giọng
	Đường nét	Âm vực	
Thanh Ngang	Ngang	Cao	Thường
Thanh Huyền	Xuống	Thấp	Thở (chùng)
Thanh Hỏi	Xuống-lên	Thấp	Thanh quản hoá
Thanh Ngã	Xuống-lên	Cao	Thanh quản hoá
Thanh Sắc	Lên	Cao	Thường
Thanh Nặng	Ngang (Xuống)	Thấp	Tắc thanh môn

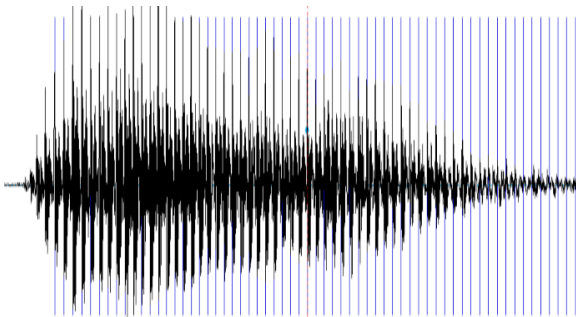
Bảng 3. Giá trị của jitter đối với tín hiệu phát âm của nguyên âm /a/ với 6 thanh điệu tiếng Việt ở nữ

	Nữ					
	a	à	ã	ả	á	ạ
Jitter (local %):	1,6	1,9	3,8	2,3	1,5	1,1
Jitter (rap %):	0,2	0,9	1,9	0,9	0,7	0,4
Jitter (ppq5 %):	0,3	0,9	1,6	0,8	0,8	0,3
Jitter (ddp %):	1,9	2,8	5,7	2,8	2,3	1,2

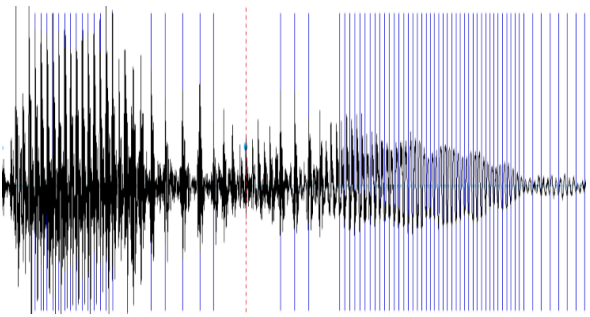


Hình 6. Giá trị các jitter của nguyên âm /a/ trong kết hợp với thanh điệu

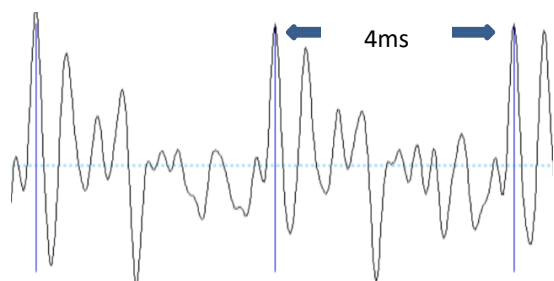
Có thể thấy, đối với những thanh có đường nét phức tạp, đổi hướng, chất giọng thanh quản hóa như ngã, hỏi các giá trị jitter có xu hướng cao hơn các thanh khác. Giá trị jitter tương đối (jitter local) cao nhất ở thanh ngã 3,8% thấp nhất ở thanh nặng 1,1%. Xu hướng này cũng được thể hiện ở các giá trị jitter rap, jitter ppq, ddp.



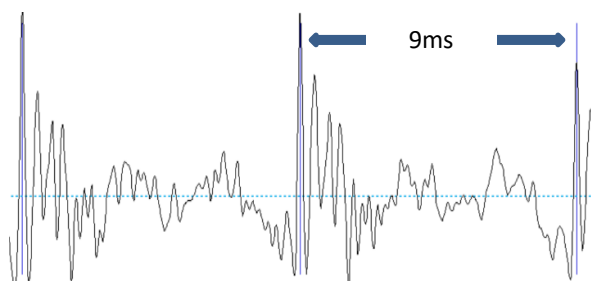
Hình 7. Sóng âm và các pulses tự động trong praat âm a



Hình 8. Sóng âm và các pulses tự động trong praat âm ả



Hình 9. Chu kì dao động sóng âm ở phần đầu âm ã



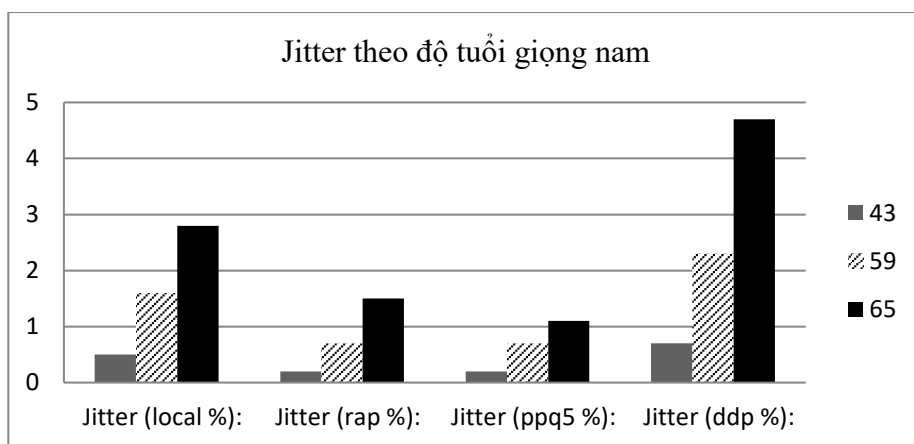
Hình 10. Chu kì dao động sóng âm ở phần giữa (đoạn thanh quản hóa) âm ã

Khi nhìn dạng sóng và các pulses của các mẫu a, ã cũng có thể nhận thấy trường độ dao động sóng âm của a theo những chu kì ổn định, đều đặn còn ã có chu kì dao động không đồng đều nhất là ở giai đoạn thanh quản hóa ở giữa. Điều này lí giải vì sao các giá trị Jitter ở các mẫu có thanh ngã luôn cao hơn các mẫu chứa các thanh khác.

4.1.3. So sánh giá trị jitter giữa các độ tuổi

Sự khác biệt tuổi tác còn được thể hiện trong các giá trị jitter tương đối. Nghiên cứu cho thấy sự khác biệt về tuổi có chiều hướng thống nhất ở cả nam và nữ. Jitter có xu hướng tỉ lệ thuận với số tuổi. Tuổi càng cao thì sự nhiễu loạn về tần số jitter sẽ càng tăng. Cụ thể: giá trị jitter local của nam 65 tuổi là 2,8% trong đó giá trị này của nam 45 tuổi là 0,5% thấp hơn 5,6 lần. Các giá trị jitter rap, jitter ppq5, ddp cũng có xu hướng tương tự.

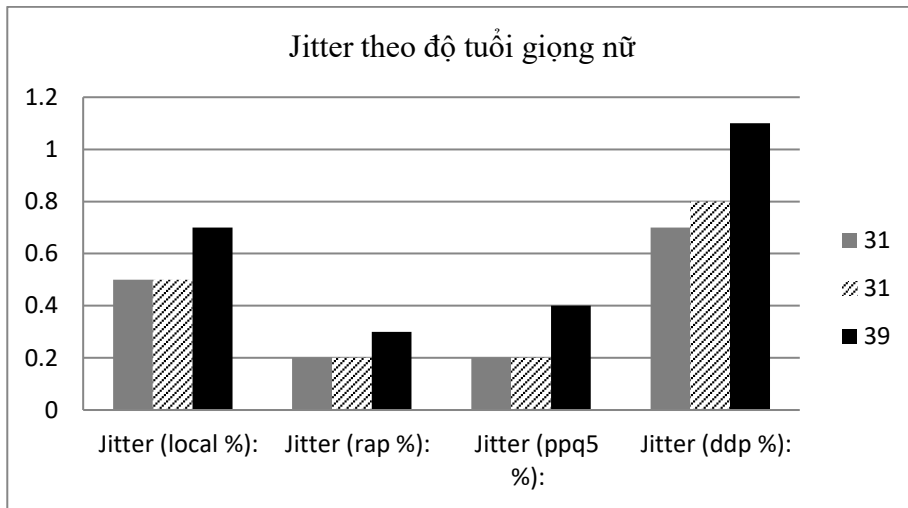
	Nam		
	43	59	65
Jitter (local %):	0,5	1,6	2,8
Jitter (rap %):	0,2	0,7	1,5
Jitter (ppq5 %):	0,2	0,7	1,1
Jitter (ddp %):	0,7	2,3	4,7



Hình 11. Giá trị các jitter theo độ tuổi giọng nam

Giá trị Jitter local của nữ 39 tuổi là 0,7 % trong đó giá trị này của nữ 31 tuổi là 0,5% thấp hơn 1,4 lần. Các giá trị Jitter rap, Jitter ppq5, ddp cũng có xu hướng tương tự.

	Nữ		
	31	31	39
Jitter (local %):	0,5	0,5	0,7
Jitter (rap %):	0,2	0,2	0,3
Jitter (ppq5 %):	0,2	0,2	0,4
Jitter (ddp %):	0,7	0,8	1,1



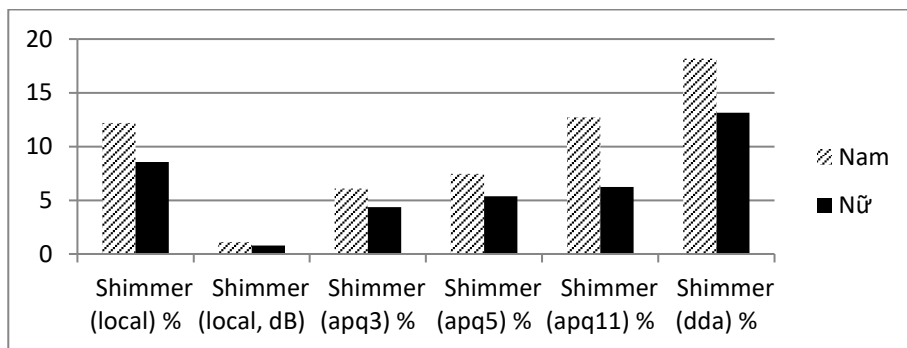
Hình 12. Giá trị các jitter theo độ tuổi giọng nữ

4.2. Chỉ số đánh giá đặc trưng về biên độ (shimmer)

Là tham số chỉ sự nhiễu loạn về biên độ âm, nó xác định sự khác nhau về biên độ sóng âm giữa một chu kì với các chu kì liên tiếp, shimmer thay đổi trong trường hợp bệnh lí ở thanh quản, bệnh nhân rối loạn giọng thì shimmer tăng.

4.2.1. So sánh giá trị shimmer giữa nam và nữ

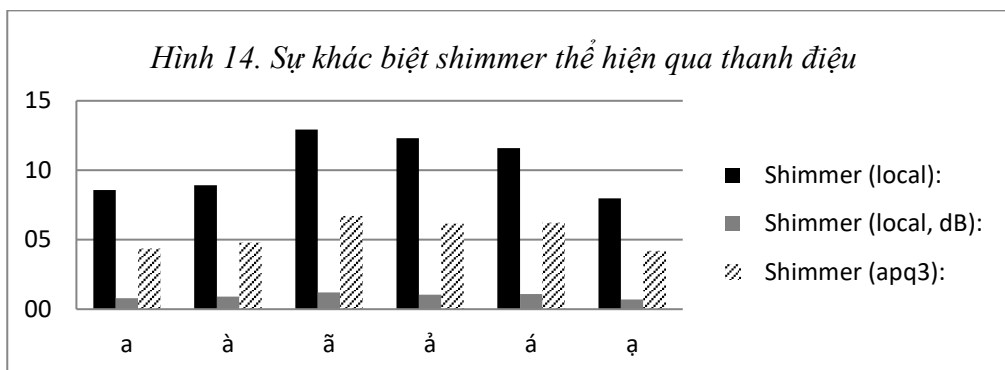
Shimmer:	Nam			Nữ		
	a	i	u	a	i	u
Shimmer (local):	12	6,5	8,0	8,6	10	9,7
Shimmer (local, dB):	1,1	0,6	0,7	0,8	1	1
Shimmer (apq3):	6,1	2,5	3,9	4,4	5,5	5,3
Shimmer (apq5):	7,5	4,0	5,6	5,4	6,5	6,7
Shimmer (apq11):	13	5,2	7,6	6,2	8,4	9,0
Shimmer (dda):	18	7,3	12	13	16	16



Hình 13. Giá trị các shimmer của nguyên âm /a/

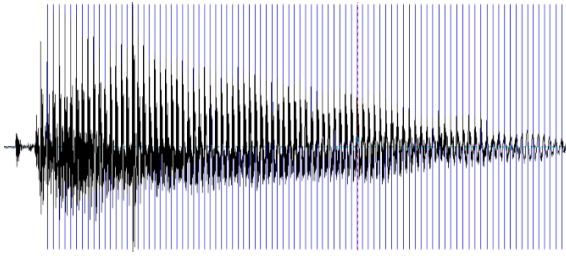
Các giá trị shimmer của nguyên âm a, i, u ở giọng nữ cao hơn giọng nam đối với các nguyên âm /i, u/. Chỉ số nhiễu loạn biên độ tương đối (shimmer local) nguyên âm /i, u/ của nữ nằm trong khoảng 9,7-10 %, của nam trong khoảng 6,5 -8%. Riêng trường hợp nguyên âm /a/ các chỉ số shimmer của nữ thấp hơn nam từ 0,5 đến 0,7%.

4.2.2. So sánh giá trị shimmer giữa các mẫu chứa thanh điệu

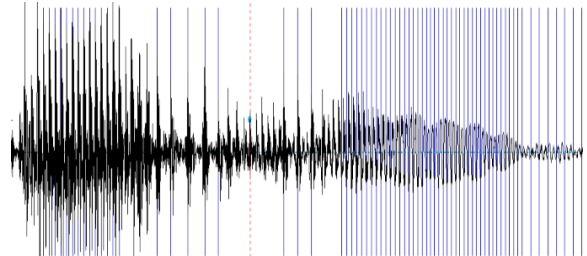


	Nữ					
	a	à	ã	ả	á	ạ
Shimmer (local):	8,6	8,9	12,9	12,3	11,6	8,0
Shimmer (local, dB):	0,8	0,9	1,2	1,0	1,1	0,7
Shimmer (apq3):	4,4	4,8	6,7	6,2	6,2	4,2
Shimmer (apq5):	5,4	6,3	8,7	7,4	8,0	4,3
Shimmer (apq11):	6,2	8,0	11,5	10,6	10,0	4,8
Shimmer (dda):	13,1	13,9	20,0	18,4	18,4	12,3

Đối với những thanh có đường nét phức tạp, đổi hướng, chất giọng thanh quán hóa như ngã, hỏi hay thanh có đường nét đi lên, cao độ cao, chất giọng thường các giá trị shimmer có xu hướng cao hơn các thanh khác. Giá trị shimmer tương đối (shimmer local) cao nhất ở thanh ngã 12,9% thấp nhất ở thanh nặng 8%. Xu hướng này cũng được thể hiện thống nhất ở các giá trị shimmer khác như: shimmer (local, dB), shimmer (apq3), shimmer (apq5), shimmer (dda).



Hình 15. Biên độ dao động sóng âm của mẫu a



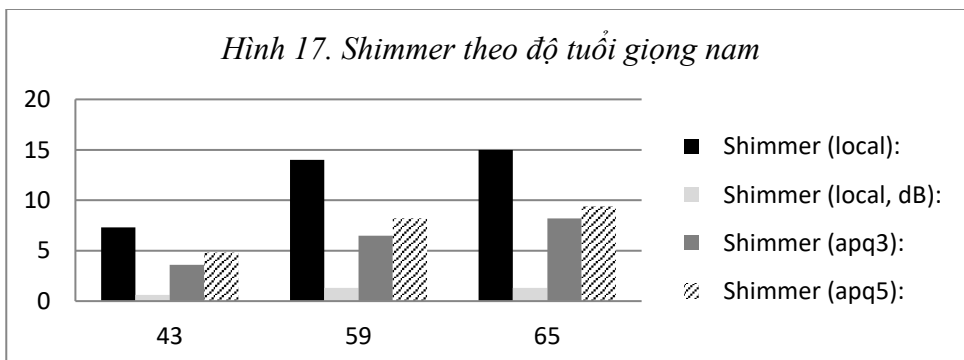
Hình 16. Biên độ dao động sóng âm của mẫu ã

Biên độ dao động sóng âm của các mẫu a, ã cũng cho thấy cường độ dao động sóng âm của a theo chiều hướng ổn định thấp dần đều. Biên độ dao động của ã có cường độ không đồng đều cường độ to ở phần đầu âm tiết, thấp nhất là ở giai đoạn thanh quản hóa ở giữa. Điều này lí giải vì sao các giá trị shimmer ở các mẫu có thanh ngã, thanh hỏi luôn cao hơn các mẫu chứa các thanh khác.

4.2.3. So sánh giá trị shimmer giữa các độ tuổi

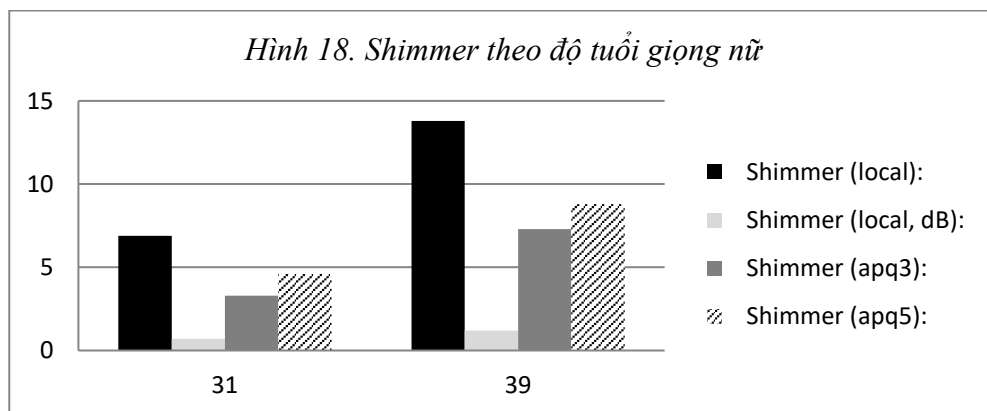
Ở những độ tuổi khác nhau thì sự khác nhau về biên độ sóng âm cũng được thể hiện một cách rõ ràng. Các giá trị nhiễu loạn biên độ Shimmer tỉ lệ thuận với độ tuổi của người nói. Người nói có tuổi càng cao thì sự khác nhau về biên độ sóng âm càng cao. Cụ thể: so sánh giữa tuổi 43 và 59 của nam. Giá trị shimmer ở độ tuổi 59 cao hơn độ tuổi 43 từ 1,7 đến 2,1 lần. Giá trị shimmer ở độ tuổi 65 cao hơn độ tuổi 59 từ 0,6 đến 1,2 lần. Giá trị shimmer ở độ tuổi 43 cao hơn độ tuổi 65 từ 1,9 đến 2,2 lần. Thống kê cho thấy tuổi càng trẻ thì giá trị nhiễu loạn biên độ càng ít, càng già thì xu hướng nhiễu loạn biên độ càng tăng.

	Nam		
	43	59	65
Shimmer (local):	7,3	14	15
Shimmer (local, dB):	0,60	1,30	1,30
Shimmer (apq3):	3,6	6,5	8,2
Shimmer (apq5):	4,8	8,2	9,4
Shimmer (apq11):	5,7	19,3	13
Shimmer (dda):	10,8	19,7	24



Đối với giọng nữ, so sánh giữa độ tuổi 31 và 39. Giá trị shimmer ở độ tuổi 39 cao hơn độ tuổi 31 từ 1,5 đến 2,2 lần. Thống kê cho thấy tuổi càng ít thì giá trị nhiễu loạn biên độ càng thấp, tuổi càng nhiều thì xu hướng nhiễu loạn biên độ càng cao.

	Nữ	
	31	39
Shimmer (local):	6,9	13,8
Shimmer (local, dB):	0,7	1,2
Shimmer (apq3):	3,3	7,3
Shimmer (apq5):	4,6	8,8
Shimmer (apq11):	6	9,2
Shimmer (dda):	9,9	21,9



5. Kết luận

Ở nước ngoài, các nghiên cứu chất thanh qua các tham số Jitter, Shimmer được nói đến trong nhận dạng người nói, trong nghiên cứu điều trị rối loạn phát âm. Bằng cách sử dụng các phần mềm và thuật toán phân tích người ta có thể dễ dàng tìm thấy sự khác biệt của jitter, shimmer giữa những người nói khác nhau.

Ở trong nước, những nghiên cứu về chất thanh trong giọng nói cũng đã bắt đầu được áp dụng bởi một số chuyên gia thuộc lĩnh vực y học. Với sự tồn tại của cơ sở dữ liệu mô tả chất lượng âm thanh hoặc sử dụng các công cụ phân tích kết hợp các thông số, có thể phân biệt giữa giọng nói bình thường và bệnh lí. Những công cụ này cho phép theo dõi quan điểm lâm sàng và giảm mức độ chủ quan của phân tích nhận thức. Các thông số âm thanh thường được sử dụng trong các ứng dụng phân tích âm thanh là tần số cơ bản (F0), jitter, shimmer và các dạng tần số. Trong đó:

Jitter là tham số chỉ sự nhiễu loạn về tần số cho phép đo lường sự thay đổi của tần số thanh cơ bản khi chuyển từ một chu kì này sang một chu kì kế tiếp và đánh giá sự khác nhau về tần số giữa

một chu kì và các chu kì kế tiếp. Jitter cho biết tính ổn định của cơ quan tạo thanh. Jitter có thể được đưa ra bởi bốn tham số liên quan: jitter tuyệt đối (jitta) là nhiễu loạn tuyệt đối, jitter cục bộ hoặc tương đối (jitt), nhiễu loạn trung bình tương đối (rap) và nhiễu loạn thời gian năm điểm (ppq5).

Shimmer là tham số chỉ sự nhiễu loạn về biên độ âm, nó xác định sự khác nhau về biên độ sóng âm giữa một chu kì với các chu kì liên tiếp, shimmer thay đổi trong trường hợp bệnh lí ở thanh quản, bệnh nhân rối loạn giọng thì shimmer tăng. Shimmer có thể được đưa ra bởi các tham số liên quan như: shimmer (dB), shimmer tương đối; shimmer (apq3) đơn vị đo sự nhiễu loạn biên độ theo chu kì ba điểm; shimmer (apq5) đơn vị đo sự nhiễu loạn biên độ theo chu kì năm điểm.

Jitter và shimmer là các tham số chỉ ra sự biến đổi xảy ra ở tần số cơ bản. Trong khi jitter chỉ ra sự thay đổi hoặc nhiễu loạn của tần số cơ bản, thì shimmer đề cập đến cùng một nhiễu loạn, nhưng nó có liên quan đến biên độ của sóng âm hoặc cường độ phát ra âm thanh. Jitter bị ảnh hưởng chủ yếu do thiếu kiểm soát rung động giọng nói và shimmer liên quan đến việc giảm sức đề kháng dây thanh và tổn thương khối trong nếp gấp thanh âm, có liên quan đến sự hiện diện của tiếng ồn khi phát ra hơi thở.

Về tham số jitter: qua nghiên cứu có thể thấy có sự khác biệt về giới đối với các giá trị thu được. Sự khác biệt cho thấy jitter tương đối của giọng nữ cao hơn đáng kể so với giọng nam. Giá trị trung bình của jitter tương đối tăng dần ở nam giới đối với các nguyên âm /a, i, u/, trong khi phụ nữ cũng cho thấy sự gia tăng giá trị tương ứng đối với tham số này ở các nguyên âm /a, i, u/. Giá trị biểu thị độ nhiễu loạn trung bình tương đối (jitter rap) của nam và nữ nằm trong khoảng 0,2 - 1,3 %, giá trị của nữ cao hơn nam, sự gia tăng giá trị cũng được thể hiện từ nguyên âm a đến u, trong đó a có giá trị thấp nhất, u có giá trị cao nhất. Cũng tương tự như jitter rap, giá trị đo độ nhiễu loạn theo chu kì năm điểm jitter (ppq5) cũng có sự gia tăng từ a đến u, nữ có xu hướng cao hơn nam.

Kết quả này có thể được giải thích vì giọng nam có tần số cơ bản thấp hơn và thời gian glottal cao hơn giọng nữ. Do đó, trong một thanh hầu dài hơn, nó được dự đoán sẽ gây nhiễu loạn lâu hơn. Sự khác biệt về giới tính của F0 có thể được chứng minh do có sự khác biệt về mặt giải phẫu rõ rệt ở thanh quản của nam và nữ. Theo nhiều nghiên cứu, thanh quản của nam lớn hơn khoảng 40% so với nữ và nếp gấp thanh âm của nam là một khối dày hơn của nữ.

Sự khác biệt ở các phát âm có thanh điệu, những thanh có đường nét phức tạp, đôi hướng, chất giọng thanh quản hóa như ngã, hỏi các giá trị Jitter có xu hướng cao hơn các thanh khác. Xu hướng này cũng được thể hiện ở các giá trị jitter rap, jitter ppq, dpp.

Sự khác biệt tuổi tác còn được thể hiện trong các giá trị jitter tương đối. Nghiên cứu cho thấy sự khác biệt về tuổi có chiều hướng thống nhất ở cả nam và nữ. Jitter có xu hướng tỉ lệ thuận với số tuổi. Tuổi càng cao thì sự nhiễu loạn về tần số jitter sẽ càng tăng.

Về tham số shimmer: Các giá trị shimmer của nguyên âm a, i, u ở giọng nữ cao hơn giọng nam đối với các nguyên âm /i, u/. Chỉ số nhiễu loạn biên độ tương đối (shimmer local) nguyên âm /i, u/ của nữ nằm trong khoảng 9,7-10 %, của nam trong khoảng 6,5 -8%. Riêng trường hợp nguyên âm /a/ các chỉ số shimmer của nữ thấp hơn nam từ 0,5 đến 0,7%.

Đối với những thanh có đường nét phức tạp, đổi hướng, chất giọng thanh quản hóa như ngã, hỏi hay thanh có đường nét đi lên, cao độ cao, chất giọng thường các giá trị shimmer có xu hướng cao hơn các thanh khác. Giá trị shimmer tương đối (shimmer local) cao nhất ở thanh ngã 12,9% thấp nhất ở thanh nặng 8%. Xu hướng này cũng được thể hiện thống nhất ở các giá trị shimmer khác như: shimmer (local, dB), shimmer (apq3), shimmer (apq5), shimmer (dda).

Các giá trị nhiễu loạn biên độ shimmer tỉ lệ thuận với độ tuổi của người nói. Người nói có độ tuổi càng cao thì sự khác nhau về biên độ sóng âm càng cao. Cụ thể: so sánh giữa độ tuổi 43 và 59 của nam. Giá trị shimmer ở độ tuổi 59 cao hơn độ tuổi 43 từ 1,7 đến 2,1 lần. Giá trị shimmer ở độ tuổi 65 cao hơn độ tuổi 59 từ 0,6 đến 1,2 lần. Giá trị shimmer ở độ tuổi 43 cao hơn độ tuổi 65 từ 1,9 đến 2,2 lần. Thống kê cho thấy tuổi càng trẻ thì giá trị nhiễu loạn biên độ càng ít, càng già thì xu hướng nhiễu loạn biên độ càng tăng.

Đây là những kết quả mà chúng tôi trình bày trong điều kiện tư liệu hiện tại, khi có thêm tư liệu từ nghiên cứu chất giọng của người nói ở các ngôn ngữ khác thì mới có thể so sánh, làm rõ hơn đặc trưng chất giọng của người Việt.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Nguyễn Duy Dương, Jonathan Livesay, *Bước đầu nghiên cứu các thông số rung động dây thanh trên người không có bệnh thanh quản*, Tạp chí Tai Mũi Họng, 2, tr.64-70, 2006.
2. Nguyễn Duy Dương, Ngô Ngọc Liễn, *Đặc điểm lâm sàng của rối loạn giọng kéo dài sau vi phẫu thanh quản*, Tạp chí Tai Mũi Họng Việt Nam, 2-5, tr.64-70, 2011.
3. Lê Thanh Hải, *Nghiên cứu hình thái lâm sàng và phân tích chất thanh của bệnh nhân viêm thanh quản do nấm*, Luận văn bác sĩ chuyên khoa cấp II, Đại học Y Hà Nội, 2014.
4. Lê Văn Lợi, *Thanh học và các bệnh về giọng nói, lời nói và ngôn ngữ*, Nhà xuất bản Y học, tr.15-88, 1999.
5. Nguyễn Văn Lợi, Jerold Edmondson, *Thanh điệu và chất giọng trong tiếng Việt hiện đại*, Tạp chí Ngôn ngữ, số 1, tr.1-16, 1997.
6. Nguyễn Văn Lợi, Ngô Ngọc Liễn, *Bệnh nghề nghiệp về thanh quản ảnh hưởng đến khả năng phát âm thanh điệu ở giáo viên tiểu học*, Những vấn đề ngôn ngữ học, Kỷ yếu hội nghị khoa học (Viện Ngôn ngữ học), tr.60-79, 2000.
7. Nguyễn Văn Lợi, *Đặc điểm ngữ âm - âm vị học của phụ âm tắc, hữu thanh, thở trong các ngôn ngữ ở Việt Nam và Đông Nam Á (trên bình diện đồng đại và lịch đại)*, Trong Những vấn đề ngôn ngữ học, Nxb KHXH, 2004.

Tiếng Anh

8. Boersma, Paul and Weenink, David. *Praat: doing phonetics by computer*, Phonetic Sciences, University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>
9. Brockmann-Bauser, *Improving jitter and shimmer measurements in normal voices*, Phd Thesis of Newcastle University, 2011.

10. E. Slyh, W. T. Nelson, and E. G. Hansen, *Analysis of mrate, shimmer, jitter, and F0 contour features across stress and speaking style in the SUSAS database*, presented at ICASSP, 1999.
11. Farrús, J. Hernando e P. Ejarque, *Jitter and shimmer measurements for speaker recognition*, Barcelona, 2007.
12. H. Wertzner, S. Schreiber, L. Amaro, *Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders*, Rev Bras Otorrinolaringologia, 71 (5) 2005, pp. 582-588, 2005.
13. J.P. Teixeira, C. Oliveira, C. Lopes, *Vocal acoustic analysis - jitter shimmer and HNR Parameters*, Procedia Technology, Elsevier, 9 (2013), pp. 1112-1122, 2013.
14. João Paulo Teixeira, Paula Odete Fer, *Jitter, Shimmer and HNR classification within gender, tones and vowels in healthy voices*, Procedia Technology 16 (2014) 1228 - 1237, 2014.
15. Meike Brockmann, Michael J. Drinnan, Claudio Storck, and Paul N. Carding, *Reliable jitter and shimmer measurements* in Voice Clinics: The Relevance of Vowel, Gender, Vocal Intensity, and Fundamental Frequency Effects in a Typical Clinical Task, Journal of Voice, Vol. 25, No. 1, pp. 44-53, 2011.
16. Michaelis, M. Fröhlich, H. W. Strube, E. Kruse, B. Story, and I. R. Titze, *Some simulations concerning jitter and shimmer measurement*, presented at 3rd International Workshop on Advances in Quantitative Laryngoscopy, Aachen, Germany, 1998.
17. Mireia Farrús, Javier Hernando, Pascual Ejarque, *Jitter and Shimmer Measurements for Speaker Recognition*, Praat software website: <http://www.fon.hum.uva.nl/praat/>.
18. Teixir Diagnostico da a, J. P; Oliveira, C. and Lopes, C., *Vocal acoustic analysis, jitter, shimmer and HNR Parameters*. Procedia Technology. Elsevier, Vol. 9, 2013; 1112-1122
19. Vasilakis M.; Stylianou, Y., *Spectral jitter modeling and estimation*, Biomedical Signal Processing and Control, 129, 2009.
20. Wertzner, H., Schreiber, S.; Amaro L., *Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders*, Rev Bras Otorrinolaringologia; 71, 5, 582-88, 2005.
21. Zwetsch, I., Fagundes, R., Russomano, T., Scolari, D., *Digital signal processing in the differential diagnosis of beningn larynx diseases*, Porto Alegre, 2006.
22. Ladefoged, Peter, *A course in phonetics*, 3rd edition, Harcourt, Brace, and Jovanovich, Fort Worth, TX, 1993.