صوت المتكلم تحت الضغط:
دراسة صوتية لحالات من المتهمين

Voice Under Stress:
Phonetic Investigation of Some Criminal Cases[١]

د/ إيمان محمد محمد قاسم

مدرس بقسم الصوتيات واللسانيات-كلية الآداب-جامعه الاسكندرية

Abstract

Tense voice is more likely to occur in situations of increased cognitive load. During the interrogation of the criminals under forensic circumstances, the offenders are exposed to some kind of stress (e.g., in the emotions like diffidence, anxiety, fear…) which are reflected in the production of their speech, and consequently, they are good indicators of tense vocalization. The aim of this study is to investigate the phonetic characteristics of stressed/unstressed voice of the crime offenders' speech under pressure and how can these characteristics encode information about the speaker: his emotions and -to some extent-his intentions. Various kinds of videos were collected from T.V interviews with offenders, victims, witnesses. Analyses of acoustic measures; F0, Jitter, Shimmer, H/N and formants were carried on selected items of stressed and unstressed utterances of each criminal. From such analyses, F0 seems to be the first factor in priority, followed respectively by the above-mentioned ones.

[١] **قُبِل ملخص هذا البحث في المؤتمر الدولي للصوتيات الجنائية IAFPA في إنجلترا عام ٢٠٢١, وعرض في شكل بوستر ولم يُنشر (مرفق في نهاية البحث).**

المقدمة:

في أثناء استجواب المتهمين فى قضية ما قضائيًا، يكون المتهم تحت ضغط نفسي وعبء ذهنى *cognitive load*، وعلى الرغم من أنه يحاول أن يخفي حقيقة حالته النفسية عند تعرضه لأسئلة قد تضطره إلى الاعتراف بجريمة ارتكبها أو اتُهم بارتكابها، أو حتى شاهدها، على الرغم من ذلك، تظهر عليه انفعالات وتصرفاتٌ مؤشرات، منها الخوف وعدم الثبات، أوعدم الثقة، والاهتزاز، والتلعثم فى كلامه، أو زلات اللسان *tongue slips*، ومثل هذه الانفعالات تبدو آثارها فى صوته، فتُعد دليلاً جيدًا ومصدرًا مهمًّا لمعلومات عن المتهم من خلال التغير فى صوته.

من هذا المنطلق نقول إن لتوصيف نوعية الأصوات أهمية كبيرة فى مجالات جديدة، منها الصوتيات الجنائية، ذلك المجال الذي غايته التعرف على صوت المتكلم سمعياً وتقنياً. ويفيد ذلك التوصيف كذلك في كشف الكذب والتضليل عندما يُقلد شخص صوت آخر، وفي اكتشاف صوت الشخص المشكوك فيه من خلال صفات ذلك الصوت فى المكالمات الهاتفية؛ إذ إنها تتأثر بفلاتر الهاتف التى يمر الكلام من خلالها، فتغير صفات بعض أصوات الكلام وصوت المتكلم. وفى المجال الجنائى تفيد دراسة نوعية الصوت فى التحقيقات الشرطية للتعرف على صوت المشتبه به، وتساعد فى تضييق دائرة البحث عن المتهم، كما تفيد في تحديد متهم بعينه *Speaker Identification* ، من خلال صوته - كما يحدث فى عملية التعرف على شخص عن طريق شاهد من شهود العيان.

هدف هذه الدراسة إذن التحقق من السمات الصوتية للمتهم أو الضحية أو الشاهد، في حالة التوتر أو البكاء أو التعبير عن الغضب بصوت المضغوط *stressed voice* ومقارنته بصوت الشخص نفسه في حالته الطبيعية التي لا يكون فيها واقعًا تحت ضغط *unstressed voice* دراسة تجريبية تحليلية، وتحاول الدراسة أن توضح كيف أن تغير نوعية صوت الشخص يمكن أن يمدنا بمعلومات عن المتكلم أثناء انفعاله – وإلى حدٍ ما عن قصده.

ا- موضوع الدراسة:

Subject of the study

يتعرض الإنسان أحياناً لمواقف تجعله تحت ضغط، وتتغير لها حالته الفيزيائية والنفسية والذهنية حين يتكلم، وهذه الحالات تتطلب التواصل بالصوت *voice communication* - في ظروف ضاغطة جدًّا- لطلب الإسعاف أوالشرطة أو رجال الإطفاء، وغير ذلك من حالات الطوارئ والمواقف، فى المجال الجنائي الشرعي *Forensic*، والمجال العسكرى *Military*، ومجال الأمن *Security*، وكذلك مجال الطب النفسي حين يتعلق الأمر بالمرضى النفسيين. وهذه المواقف تؤثر في طريقة نطق المتكلم للكلام عند إنتاجه - وقد يكون هذا عن غير وعي منه أحيانًا - ويشعر المستمع بهذا الاختلاف فى الصوت وطريقة الكلام. وقد يتمكن المتكلم من التحكم في الأصوات الصادرة عنه، الصامت منها والصائت لكن هناك ملامح بروزودية أخرى - كزيادة الزمن أو الشدة أو ارتفاع النغمة أوانخفاضها- تُعين المستمع على أن يدرك المعنى المقصود حسب الموقف.

هناك إذن سمات معينة للسلوك اللفظي *verbal behavior* وعلامات *markers* تبدو في الكلام، حين يكون المتكلم تحت ضغط، سواء أكانت تلك العلامات تصدر عن وعي من المتكلم أم عن غير وعي منه، ويدركها المستمع بوضوح (*Hansen ١٩٩٤*). وكذلك قد يركز المتكلم- في مثل تلك الظروف- على أجزاء بعينها من كلامه يراها ذات معلومة مهمة، ولا يركز على أجزاء أقل أهمية، ويعتمد التحكم فى شدة الصوت في تلك الحالة على ضبط المتغيرات فى الحنجرة وما فوقها (*Hicks and Hollien,1981*).

وذهب جون ليفر (Laver, 1991) *ص* ٢٤١ إلى أن هناك ثلاثة أوجه أساسية في الأداء الصوتي للمتكلم تؤدي وظائف لغوية و/ أوغير لغوية، وتمثل خلفية إدراكية للمستمع، تتيح له تصورًا صوتيًا كاملاً للمتكلم *Vocal Image of a Speaker* ، وسنقتصر هنا على الملامح غير اللغوية على النحو الآتي:

١- ملامح غير لغوية *Extralinguistics* وهي تلك الناشئة عن التركيب التشريحي (الفسيولوجي) لجهاز المتكلم الصوتي *Anatomical*، *كما خلقه الله؛* لذلك هي خارج إرادة المتكلم وتحكمه، وتؤدي وظيفة إخبارية (معلوماتية) *Informative* . بالإضافة إلى سمات تتصل بعادات المتكلم الكلامية *Habitual*، التي اكتسبها من المجتمع الذى نشأ فيه، ولها وظيفة إخبارية واتصالية.

٢ - ملامح ما وراء لغوية *Paralinguistic* (يترجمها بعض الباحثين، هكذا:" القريبة من اللغوية" أو "معلغوية")، تصنعها نغمة الصوت *Tone of voice*، وتعبر عن شخصية المتكلم، وعن حالته الانفعالية *Attitude*، كما أنها سمات يمكن التحكم فيها، ويمكن تعلمها، وتؤدى وظيفة تواصلية *Communicative*.

أ) **توصيف نوعية الصوت نطقياً**

شرح جون ليفر فى كتابه ١٩٨٠ نوعية الصوت *Voice Quality* غير الطبيعية المتنوعة، التي يمكن أن تصدر عن الأوتار الصوتية، مثل: الصوت الأجش، والصوت النَفَسي، والصوت المحشرج، والخشن، والصوت المخنوق، والصوت المهموس ، ووصف كلاً منها، ورصد الفروق بينها بقياس العناصر الفيزيائية الناتجة عنها. كما أوضح الآثار في المكونات الرنينية الناتجة عن تكبير تجاويف ما فوق الحنجرة *Supra-laryngeal cavities* التي قد تتغير قيم ترددداتها فى حالات الصوت المضغوط - الناتج عن الخوف والبكاء وغيرهما- عن مواقعها فى النطق الطبيعي .

ومما ذكره من التفاصيل عن نوعية الصوت (p.141) ،أن الصوت المشدود *tense voice* يتسم بالعلو من حيث الشدة وارتفاع في الحدة *pitch* ومن النوع الأجش، مع ارتفاع ضغط الهواء أسفل الأوتار الصوتية :

"Tense voice was characterized as a loud, high-pitched, harsh voice with high subglottal pressure ….. "(Laver 1980:p 141)

واستكمالاً لتوصيف الصوت المضغوط أو المشدود ،عرَّف هانسن (*Hansen* ١٩٨٩) *وموری*(*Murry et.al* ١٩٩٦) أنواع الضغوط Stressors والعناصر الفيزيائية – الأكوستية المرتبطة بها ،كالآتي:

ب) توصيف الضغوط *Stressors*

1-ضغوط فسيولوجية *physiological stressors* وتعني زيادة المجهود الصوتي فى الأوتار الصوتية.

٢- ضغوط فيزيائية *physical stressors* وتشير إلى رد الفعل العضلي في أثناء النطق، وتأثير العضلات المسئولة عن الحركات النطقية فى الجهاز الصوتي التي تؤدي إلى شد العضلات أو تشنجها، وكذلك زيادة

ضغط الهواء أسفل *الحنجرة subglottal air pressure* ، ما يؤدي إلى زيادة معدل فتح الأوتار الصوتية وغلقها، ومن ثم زيادة معدل التردد الأساسى F_0 فيؤدي إلى ارتفاع الحدة/النغمة *pitch* .

٣-ضغوط نفسية *psychological stressors* تتمثل في الانفعالات وما يصحبها من سمات وراء لغوية *paralinguistic*

٤-ضغوط إدراكية *perceptual stressors* تتمثل فى الأوامر العصبية - العضلية

وأوضح موري(*Murry et.al* ١٩٩٦) تأثير الضغوط في عملية الكلام، ذاهبًا إلى أن الفعل العضلى *muscular action* يجعل ضغط الهواء يتزايد في الحنجرة أسفل الأوتار الصوتية، مثلما فى حالة الغضب، وينتج عنه أسلوب الكلام بصوت *عالٍ loud style* لأن الشد فى الأوتار الصوتية يرفع قِيم التردد الأساسي F_0 وكذلك المكونات الناتجة عن عمليات التكبير فى تجاويف ما فوق الحنجرة وأيد كريمان (*Kreiman, et.al. 2011*) هذا، وأضاف أن المكونات الرنينية الأولى F_1، F_2 تتغير أماكنها تغيراً ملحوظاً فى الصوت العالي عنه فى المنخفض.

وشرح هانسن (*Hansen 2007*) كيف تؤثر الضغوط في عملية إنتاج الكلام موضحاً أن: " الذي يتكلم تحت أي شكل من أشكال الضغط ، ينتج عن هذا الضغط تحور ما، وتغيير فى عملية إنتاج الكلام "، وأن الضغوط حالة نفسعضوية؛ إذ تصحبها عادة انفعالات معينة، مثل: الخوف ،والتردد، والتلعثم، وتكرار زلات اللسان ، والمواقف التى يتعرض لها ، فتؤثر في الحالة الفيزيائية والحالة الذهنية *cognitive* للمتكلم.

ج- نوعية الصوت من الناحية الفيزيائية - التحليلية:

وبوجه عام، ركز العلماء عند التحليل الأكوستي للكلام على العناصر الفيزيائية، كالتردد الأساسي *Fundamental frequency(F_0)* والشِّدة *Intensity*، والزمن *Duration* ، وأماكن المكونات *Formant locations* ، وميول المنحنى النغمي للكلام *Spectral slope*.

وأكدت معظم الدراسات ***أن التردد الأساسي F_0*** من العناصرالمهمة، التي تثمر مراقبتها نتائج واضحة عند تحليل أسلوب الكلام تحت الضغط، وكذلك متوسط المنحنى النغمي للكلام *mean and median of F_0 contour* .وأضاف موري١٩٩٦ تأثير الأفعال العضلية *muscular action* التى تزيد ضغط الهواء تحت

الأوتار الصوتية، فيزيد الشد في الأوتار الصوتية *tension* بالإضافة إلى عدم انتظام حركة تلك الأوتار *irregular glottal pulses*. (*Kreiman, et.al. 2006*)

كما وجد جاسن (Jessen et.al١٩٧٣) أن التردد الأساسي مقياس مهم للتغيرات فى صوت الشخص نفسه فى حالاته المختلفة *Intra- speaker* ،ويشير التردد الأساسي ومستوياته إلى تفرد كل متكلم بصفات خاصة ،لاسيما معدل F_0 mean and median.

وأضاف كنزل(Kunzel ١٩٩٥) أن التردد الأساسي F_0 يقع في الصدارة من حيث الأهمية بين العناصر الفيزيائية، وهو العنصر الأقوى عند تحديد الاختلافات الدقيقة بين أصوات المتكلمين بشكل عام، وأن معدل التردد الأساسي يتأثر بالتغير في الانفعالات ، وأن سرعة الكلام *Tempo* تفيد في رصد اختلافات عدة، بين الأشخاص *Inter-speaker* ،ثم بين الشخص ونفسه *Intra- speaker* فى حالاته المختلفة.

وفي الجدول الآتي-رقم (١)- أمثلة لنتائج دراسة هانسن (*Hansen 2007*) فى اللغة الانجليزية، ودراسة فشل (*Fashal 1990*) في العربية العامية، ودراسة إبراهيم *(Ibrahim 2016)* فى العربية الفصيحة ؛ لقياس متوسط قيم التردد الأساسي F0 للانفعالات المختلفة فى الكلام، ومنها حالة الغضب التي تمثل الصوت المشدود ومقارنة ذلك بحالة المتكلم العادية.

Author	Normal/ mean F_0	Anger/ mean F_0	Male/Female
Hansen	124 Hz	283 Hz	Male
Fashal	170 Hz	224 Hz	Male
	220 Hz	255 Hz	Female
Ibrahim	177 Hz	213 Hz	Male

الجدول (١):متوسط قِيم التردد الأساسي فى الانجليزية والعربية العامية والعربية الفصيحة، في الكلامين المشدود والطبيعي.

اتفقت نتائج الدراسات الثلاث السابقة أيضًا على أن **متوسط سرعة الكلام** في أثناء الغضب والصوت العالي- بالمللي ثانية- أعلى منه في الكلام العادي. وأضاف هانسن (Hansen et.al.1989) أن زمن الكلمة - بحركاتها وسواكنها- تلعب دوراً مهماً فى مقدرة المستمع على إدراك محتوى المعلومة التي تصله من قِبل المتكلم.

أما من ناحية **شدة الصوت المضغوط** ، فإن قِيم الشدة- كما هو متوقع- تكون عالية في حالات الغضب و الصوت العالي، كما أن الشدة والزمن علاقتهما مترابطة فى حالات الغضب والصوت العالي المضغوط، (كما أظهرت الدراسات الثلاث المذكورة أعلاه).

ومن ناحية ***المكونات الترددية*** *Formants* التي تشكلها الممرات الصوتية فوق الحنجرة *vocal tract*، أظهرت نتائج دراسة هانسن (Hansen et.al.1989) أنه فى حالة الغضب والصوت العالي تحت تأثيرالضغط يظهر اختلاف ملحوظ في المكونين الأول F_1 والثاني F_2 بالزيادة في قيم الترددات ، وتتغير معها سعة المكون *Bandwidth* ويقل التغير فى المكونين الثالث F_3 والرابع F_4.

أما ***نوع التصويت الحنجرى*** *phonation* ،فشرحه جون ليفر1996 ,(*Laver ١٩٨٠*) تفصيلياً، وقال إن هناك بعض القياسات الفيزيائية المهمة التي تفرق بين الصوت الأجش، والنفسي، والمهموس، والمبحوح وغيرها. وهي تقيس قيم التردد والشدة فى شكل التغيرات فى الأوتار الصوتية *glottal pulse shaping* ،خاصة عدم انتظامها في الفتح والغلق ،وذلك عن طريق قياس قِيم جيتر*Jitter* وشيمر *Shimmer* .وهذه القياسات تفيد فى إظهار كمية- وكيفية -الاهتزازات غير المنتظمة في التردد والشدة، إما في حالات البكاء ،وإما بحة الصوت، أو أي شكل من أشكال الصوت غير الطبيعي.

وشرحت دراسة هايوارد (Hayward 2000) شكل التصويت الحنجري *glottal pulse shaping*، وتغيراته في حالة الصوت المشدود، وقياس مستوى منحناه الطيفي وميوله، وأنه يتميز بأعلى نسبة انحدار *steeper slope* في المنحنى الطيفي، إذا ما قورن بالصوت غير المضغوط أو الصوت النَفَسي. كما أوضحت دراسة كلاري وهانسن (Clary and Hansen 1992) عدم انتظام اهتزازات الأوتار الصوتية *pulses' irregularity* بسبب التغير المفاجئ، وعدم التساوي بين مرحلة الغلق ومرحلة فتح الأوتار الصوتية في أثناء عملية التصويت الخاص بالصوت العالي والمضغوط مع زيادة الطاقة وشدة الصوت ، عنه فى الصوت المنخفض. وأضافت نتائج دراسة إبراهيم (*Ibrahim2016*) أن الاختلاف فى قياس *Jitter , Shimmer* -الذي يرصد قِيم عدم انتظام تردد الأوتار الصوتية- كانت أعلى في الصوت العالي منه في الصوت العادي.

ومن الناحية الإدراكية، أوضحت بعض الدراسات قدرة ***المستمع*** على استيعاب محتوى المعلومات الصادرة من المتكلم، وأنها تعتمد على عناصر عِدة؛ فالمستمع يدرك الإشارات الكلامية لا بمجرد الاعتماد على

الدلالة الفيزيائية للصوامت والصوائت المتتابعة فى الكلام ، لكن أيضاً بالاعتماد على العناصر البروزودية *paralinguistics* التى تلازم اللغوية . ولذلك، و تحت ظروف الضغط، يضيف المتكلم عناصر فوقطعية، مثل: الزمن، والشدة، وتغيير النغمة؛ ليعين المستمع بها على فهم الرسالة الصوتية، وفي المقابل، يستطيع المستمع رصد علامات الضغط *pressure markers* فى الرسالة الكلامية.*(Hollien et.al 1991)*

وناقش *هولن (Hollien 1980)* فى مقالته (الدلائل التصويتية للضغط النفسي) *"Vocal indicators of psychological stress"* العلاقة بين الضغط والكذب- في حالة ما إذا استعمل العلماء العناصر الفيزيائية السابق ذكرها في عمل أجهزة لكشف الكذب وتقييمه- وأثبت أن هذه الأجهزة ***لم تنجح لأن نتائجها تعتمد على عناصر إحصائية صرف***، أي: نسبة الصواب المئوية التى تحققت في العناصر ونسبة الخطأ ، وأن نتائجها ليست لها قيمة ذات بال في التفريق بين الكذب والضغط.

هذا مجمل الدراسات السابقة وأهم نتائجها ، أما عن هذا البحث فيعتمد على الدراسة التجريبية التحليلية المقارنة لصوت المتهم المضغوط وهو متوتر وصوته فى حالته العادية ، ثم ما قد ينتهي إليه من نتائج ، قد تتفق وقد تختلف مع نتائج الدراسات السابقة – أو بعضها – على النحو الآتى:

II- إجراءات البحث ومنهجه

أ) تجميع المادة المسجلة:

جُمعت مقاطع فيديو من لقاءات تلفازية بين المذيع والمتهم، أو الضحية، أو الشاهد فى قضايا جنائية متنوعة، وأنواع من القضايا التى دار الحوار حولها: قضايا خاصة بتجارة الأعضاء ، و قضايا قتل ، وقضايا خيانة زوجية ، وقضايا اختطاف أطفال ، وقضايا تتعلق بمشاجرات بين العامة فى الشارع.

(Figure 1) Unstressed voice (left) : when the speaker introduces himself.

Stressed voice (right): when the speaker is speaking in tense voice; with crying, fear, or anxiety.

وللتوضيح ***فالصوت المضغوط*** هنا في حالة انفعال المتكلم في أثناء توجيه أسئلة تدينه، أو تحثه على الاعتراف، فيحاول الدفاع عن نفسه، ويظهر من طريقة كلامه ضرب من الانفعال والصوت العالى أو البكاء بصوت مبحوح أو محشرج (انظر الصورة التي إلى اليمين *Figure* 1)

والصوت الطبيعى غير المضغوط للمتكلم نفسه في بداية اللقاء معه، وسؤاله عن اسمه ومهنته أو أية معلومات عادية أخرى، فيتكلم دون انفعال (انظر الصورة التي إلى الشمال*Figure1*).

ب) تصنيف المادة المسجلة:

اختير خمسة عشر متكلمًا متنوعًا، ما بين متهم وضحية وشاهد، ثمانية رجال وسبع نساء ، ثم استُخلص ٤٦ من المنطوقات فى حالات الانفعال (الصوت المضغوط) ، و ٣٧ من منطوقات (الصوت غير المضغوط) .

وحُفظ كل نوع منهما فى مجلد/فايل، حمل كل منطوق مصنف بداخله رقمًا، ومصنف كل صوت من النساء أو الرجال وكذلك نوع الجريمة (انظر الجدول رقم ٢)

الجدول (٢) : اسم الملف ومحتواه: نوع الجريمة ، رقم المتكلم، نوع المتكلم، رقم المنطوق، بالاضافة إلى قياسات المعايير التى تتعلق بحالة الأوتار الصوتية فى الصوت المضغوط وغير المضغوط للمتكلم. واستُخدم برات سكريبت لقياس وتجميع عدد كبير من المادة الصوتية التى تم تحليلها

Sound_Name	Speaker	Gender	F0	Jitter (local)	Shimmer (local)	Jitter (absolute)	Shimmer (dB)	Harmonic to noise ratio	Noise to harmonic ratio
KidnapCrime1_Sp5_M_18	Sp5	M	157.5	1.59%	12.31%	0.01%	113.20%	9.94	0.159095
KidnapCrime1_Sp5_M_18	Sp5	M	239	1.35%	19.46%	0.01%	174.00%	3.714	0.439026
KidnapCrime1_Sp5_M_18	Sp5	M	228.1	4.30%	16.16%	0.02%	140.10%	7.523	0.297198
KidnapCrime1_Sp5_M_18	Sp5	M	161.9	3.29%	18.04%	0.02%	152.50%	8.783	0.184671
KidnapCrime1_Sp5_M_18	Sp5	M	176.5	6.54%	17.15%	0.04%	159.50%	3.555	0.498252
KidnapCrime1_Sp5_M_18	Sp5	M	136.6	2.23%	14.18%	0.02%	133.50%	10.293	0.131708
KidnapCrime1_Sp5_M_18	Sp5	M	134.5	1.92%	17.31%	0.01%	151.10%	8.795	0.207548
KidnapCrime1_Sp5_M_18	Sp5	M	191.1	3.08%	15.54%	0.02%	129.10%	9.968	0.156379
KidnapCrime1_Sp5_M_18	Sp5	M	274.1	4.58%	12.64%	0.02%	140.90%	9.115	0.252959
KidnapCrime1_Sp5_M_18	Sp5	M	227.9	3.37%	14.55%	0.02%	133.90%	9.754	0.209935
KidnapCrime1_Sp5_M_18	Sp5	M	254.6	3.93%	21.11%	0.02%	176.20%	6.284	0.331426
KidnapCrime1_Sp5_M_18	Sp5	M	294.6	3.32%	8.48%	0.01%	107.80%	11.191	0.144166
KidnapCrime1_Sp5_M_18	Sp5	M	243.9	3.82%	10.92%	0.02%	108.00%	8.947	0.258277
KidnapCrime1_Sp5_M_18	Sp5	M	231.1	1.00%	8.15%	0.00%	99.00%	11.487	0.128335
KidnapCrime1_Sp5_M_21	Sp5	M	121.3	2.79%	35.12%	0.02%	296.40%	3.158	0.54562
KidnapCrime1_Sp5_M_21	Sp5	M	192.7	2.21%	11.29%	0.01%	97.10%	11.012	0.174593

ج) **تجهيز المادة المسجلة وتحليلها:**

تمت تنقية المادة المسجلة من الضوضاء، ومن الخلفية الموسيقية عن طريق برنامج *Audacity version 6.3* واستُخدم برنامج برات *praat* وجهاز *CSL ٤٥٠٠* لقياس الآتي:

١-مدى التغير فى حركة الأوتار الصوتية للمنطوقات المضغوطة والمنطوقات غير المضغوطة،وذلك عن طريق قياس F_0 , *Jitter* , *Shimmer* , *HNR*

يوضح Figure (3) الآتى مدى التغير/عدم الانتظام بين الموجة والتى تليها من ناحية التردد jitter كما هو موضح بالسهم الأفقى ،و قَيم عدم الانتظام بين الموجات من ناحية الشدة shimmer كما هو موضح بالسهم الرأسي.

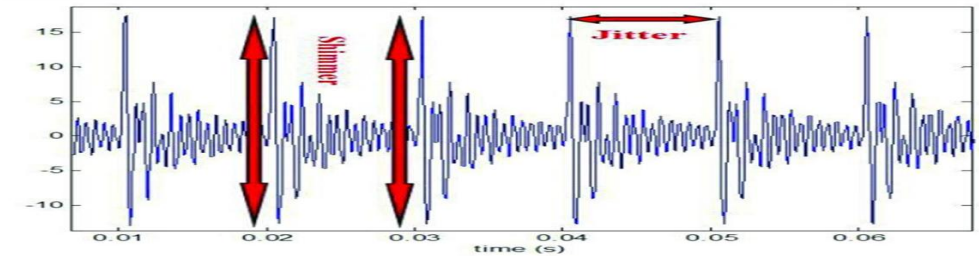


Figure (3),represents an Cycle-to-cycle frequency perturbation(Jitter/horizontally), and cycle-to-cycle intensity perturbation (shimmer/vertically).

٢- قياس مدى التغير فى حركة التجاويف التي فوق الحنجرة في أثناء إنتاج الصوتين المضغوط وغير المضغوط، وذلك عن طريق قياس المكونات الرنينية الثلاثة الأولى F_1 , F_2 , F_3 لبعض الحركات *vowels*

-واستُخدم فى تمثيل النتائج الإحصائية البرامج الآتية : *PRAAT Script and GG plot 2 package on R Studio*

د) التقييم السمعي *perceptual judgment*:

أوضحت دراسة هايوارد *(Hayward 2000)* أن التجربة في المجال الصوتي التجريبي تقوم على جانبين يكمل أحدهما الآخر،هما:الجانب الانطباعي impressionistic ،الجانب التحليلي الفيزيائي، وأنه يمكن استخلاص نتائج متكاملة ودقيقة بالاعتماد عليهما معًا. كما أضاف كيتنج (*Keating et. al. 2006*) أن التقييم الإدراكي للتفريق بين نوعيات الصوت قد تختلف من شخص لآخر، وترجع إلى طبيعة إدراك العناصر البروزودية المتنوعة.

وفي هذا البحث تم عمل تقييم سمعي للنماذج المختارة فى الحالتين: الصوت المضغوط والصوت غير المضغوط، وعُرض الملفان على عدد من المستمعين لتأكيد صحة اختيار المنطوقات فى الحالتين، قبل الشروع في التحليل الفيزيائي للكلام الطبيعي والمضغوط.

III - نتائج البحث

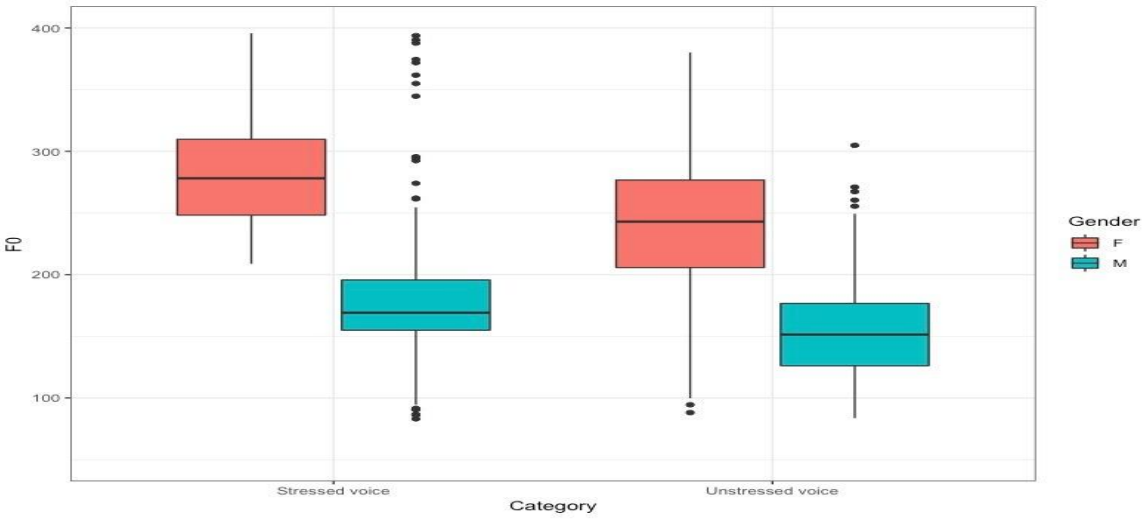
أولاً: نتائج عناصر التصويت الحنجري

Phonation Parameters

١ - التردد الأساسي F_0

بين عدد كبير من الدراسات فى هذا المجال ، في لغات عِدة، نجد اتفاقًا على أن التردد الأساسي F_0 يزيد بزيادة الشد فى الحنجرة والأوتار الصوتية، وأنه عنصر معتمد على المتكلم *a speaker dependent (Kunzle, et. al.1995)*. وقد أوضح جاسن وآخرون (Jessen et.al 1973) أن التردد الأساسي مقياس مهم للتغير والفروق بين أصوات الشخص الواحد في حالاته المختلفة *intra-speaker* .

وقد اتضح من نتائج هذا البحث أن قِيم متوسط التردد الأساسي يلعب دوراً مهماً فى تمييز الصوت المشدود/المضغوط من الصوت العادي غير المضغوط ،عند كل من النساء والرجال وأن متوسط قِيم F_0 في الصوت المضغوط *Stressed* كان أعلى من قيم الصوت غير المضغوط عند كل من النساء والرجال ، وأن صوت النساء فى الحالتين – بطبيعة الحال - أعلى من الرجال ، كما يتضح في الشكل الإحصائي (١).

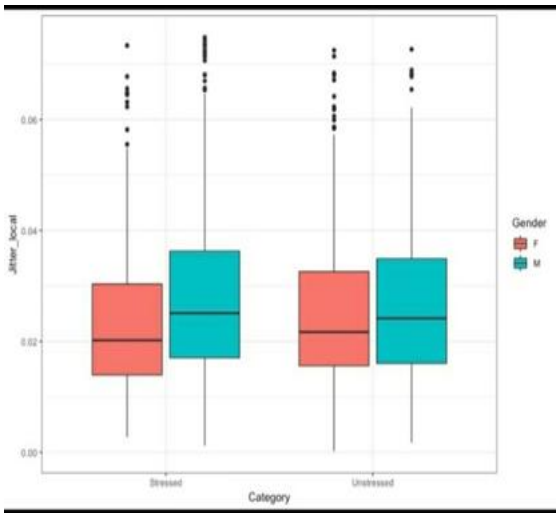


الرسم الإ:حصائى (١) نسب التردد الأساسي لكل من النساء (باللون الأحمر) والرجال (باللون الأزرق) لنوعي الصوت المضغوط وغير المضغوط.

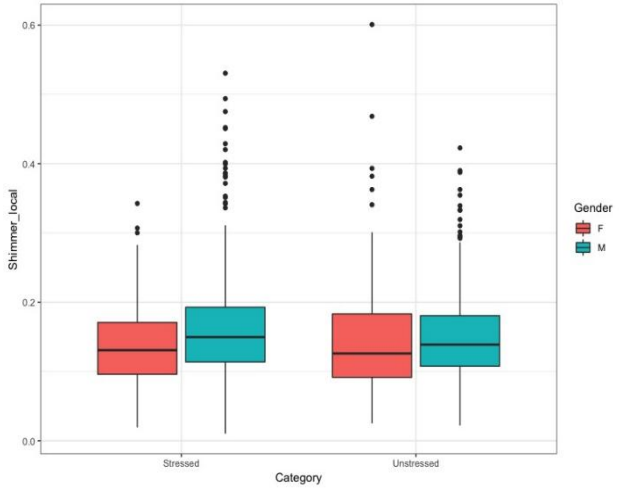
2- قيم عدم الانتظام فى التردد:

تُقاس نسبة عدم الانتظام فى كل حالة غلق/فتح والتي تليها فى **تردد** الأوتار الصوتية بال "*Jitter*" ،وهو يرصد نسب الاختلاف بدقة . كما يقاس عدم الانتظام فى **شدة** كل غلق / فتح في الأوتار الصوتية والتي تليها بال " *Shimmer* " وتسمى فى بعض الدراسات بالاضطراب *perturbation* فى العضلة *cricoid muscle* التى يتم فيها الشد فتحرك الأوتار الصوتية بشكل غير منتظم. *(Brenner and Shipp, 1998) , (Kreiman and Sidtis, 2011)*

وتُظهر نتائج هذه الدراسة أن متوسط قِيم ال "*Jitter*" تبدو متشابهة في الحالتين:الصوت المضغوط وغير المضغوط عند الرجال ، بينما متوسط قيمها أعلى عند النساء فى حالة الصوت غير المضغوط عنه في الصوت المضغوط، لكن الفرق غير مؤثر(كما هو مبين فى الشكل الإحصائي (٢).كما أظهرت قياسات ال " *Shimmer* " فى حالتي الصوت المضغوط وغيرالمضغوط عند الرجال والنساء ، نتائج غير مؤثرة ، (كما هو مبين فى الشكل الإحصائى (٣).



الرسم الإحصائى(٢): Jitter النساء (باللون الأحمر) والرجال (باللون الأزرق) لنوعي الصوت المضغوط وغير المضغوط.



الرسم الاحصائى (3): Shimmer لكل من النساء (باللون الأحمر) والرجال (باللون الأزرق) لنوعي الصوت المضغوط وغير المضغوط.

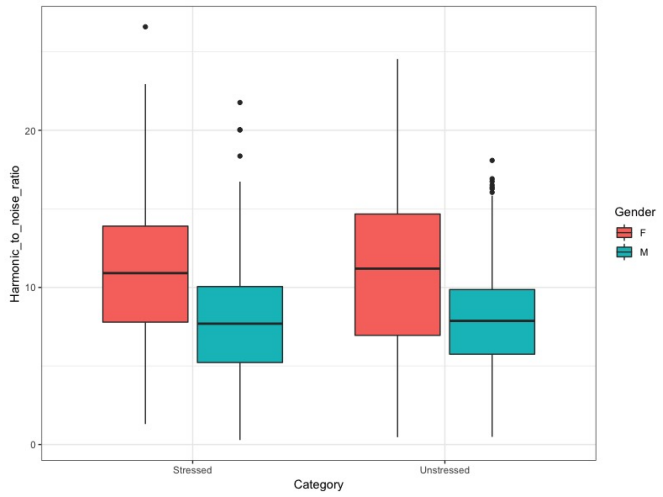
3- **معدل التردد المنتظم إلى غير المنتظم:**

Harmonic-to-Noise Ratio

قد يظهر عدم انتظام فى الأوتار الصوتية بسبب تغير مفاجئ مرده إلى علو الصوت أو وجود بحة فيه بسبب البكاء بصوت مخنوق أو صوت عاليُ *(Clary and Hansen 1992)*. ويُستخدم هذا النوع من القياسات فى علاج أمراض الكلام أحيانًا.

وتظهر أهمية هذا النوع من القياسات فى قياس نسبة التشوش *noise* الموجودة بين فتح / غلق الأوتار الصوتية نتيجة خروج كمية من الهواء في أثناء فتح الأوتار الصوتية وغلقها والتي يمكن أن تظهر أحيانًا مع الصوت العالي، أو المشدود، أو الصوت النَفَسي؛ فكلما كانت نسبة الهواء المتداخل مع التردد أقل يكون الصوت طبيعيًا *clear voice (Patel et al., 2011)* .

وتظهر نتائج هذه الدراسة أن معدل *Harmonic- to - noise ratios* *فى الصوت المضغوط أعلى منه في الصوت غير المضغوط عند النساء والرجال، بقيم تكاد تكون غير مؤثرة/فعالة Insignificant كما هو موضح فى الشكل رقم (٤).*



الرسم الإحصائي (٤) HNR معدل نقاء الصوت من الضوضاء لكل من النساء (باللون الأحمر) والرجال (باللون الأزرق) لنوعي الصوت المضغوط وغير المضغوط.

ثانياً: نتائج تكبيرات ما فوق الحنجرة

Subglottal Resonances

يتغيرشكل تجاويف ما فوق الحنجرة وحجمها طبقاً لتغير وطأة الحالة الانفعالية، لذلك فإن المكونات الثلاثة الأولى F_1, F_2, F_3 التى تمثل تكبير الصوت فى التجاويف تتغير عند أداء المتكلم في الحالات الانفعالية عنها فى الأداء الطبيعي،وقد قارن جولدن وهانسن *Godin and Hansen* فى دراستهما ٢٠٠٨ بين الكلام في أثناء الانفعالات بنوع من العبء الفيزيائي -كما سماه- *physical load* والكلام الطبيعي الحيادي -كما سماه- *neutral speech* ، وكانت نتيجة هذه الدراسة أن الكلام الانفعالي يؤثر في المكونات F_1, F_2، ووجد أن المكون الأول F_1 يعتمد على المتكلم.وأضاف كرايمان *Kreiman et.al 2011* أن لكل حركة *vowel* في أثناء الكلام مكوناتها الخاصة المميزة لها، وتُحدد هذه المكونات حسب أوضاع أعضاء النطق وأماكنها ، ويؤثر في هذا التغير عنصران أساسيان: الأول عملية النطق المشترك *coarticulation* عند الانتقال من (الصامت / الصائت) وإليه، والعنصر الثانى هو الضغط stress الذي له تأثير في كل مرحلة في النطق.

وأظهرت نتائج الدراسة الحالية أن F_1 فى أغلب المتحدثين - فى الصوت المضغوط- إما أعلى في قِيم التردد عن الصوت غير المضغوط وإما أنهم متساوون وذلك فى حالة الذكور، بينما الإناث على العكس،والصوت غير المضغوط أعلى فى قيم تردد المكونات . وكذلك بالنسبة إلى F_2 فى الصوت المضغوط أعلى منه في الصوت غير المضغوط عند الرجال، بينما عند النساء ، قيم F_2أعلى فى الصوت غير المضغوط(انظر الشكل ٥)

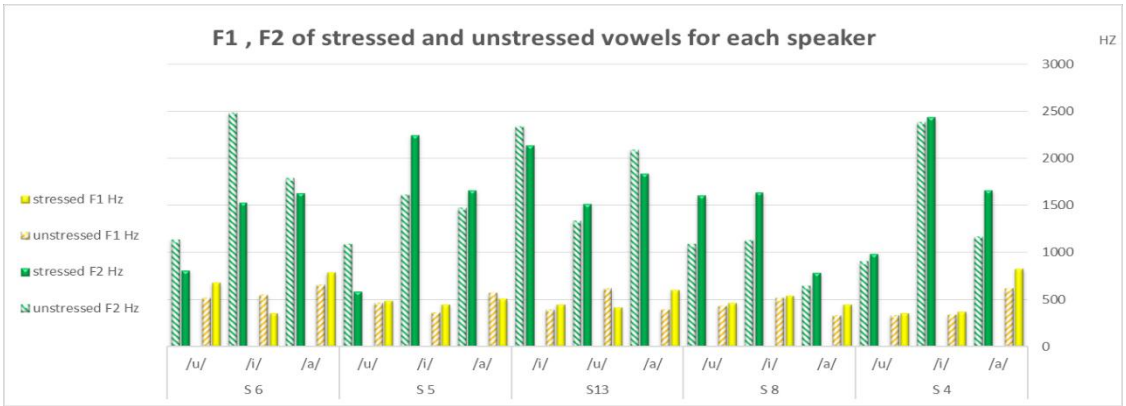


Figure ٥): متوسط القِيَم الترددية للمكونات F_1, F_2 للحركات فى المنطوقات المضغوطة وغير المضغوطةلنموذج من المتكلمين

٧- استنتاجات البحث:

Conclusions:

اهتمت هذه الدراسة بالمقارنة بين حالتين لصوت المتكلم نفسه *intra-speaker* : الحالة الحيادية في أثناء الكلام بشكل طبيعي *neutral state* والحالة الانفعالية في أثناء الكلام بصوت مضغوط *stressed* ، وكان الهدف من هذا الاختيار التطبيق على حالات cases من المتهمين أو الشهود أو الضحايا في أثناء الاستجواب في قضايا حقيقية لكي نتحقق من الصوت فى حالتيه *(stressed/unstressed)*بشكل طبيعي وبنطق تلقائي. وقد اختيرت مجموعة من المنطوقات (٣٧ منطوقًا بالصوت المضغوط) ومجموعة من (٤٦ منطوقًا بالصوت غير المضغوط) ، ثم عرضت المجموعتان على عدد من المستمعين لتأكيد اختيار المنطوقات فى الحالتين. ثم أجريت تحليلات للعناصر الفيزيائية F_0، *Jitter* ، *Shimmer* ، *HNR* ، وكذلك المكونات الرنينية الثلاثة الأولى F_1، F_2، F_3 ، وأسفر ذلك عن الآتي:

- أن التردد الأساسي F_0 كان العنصر الأهم والأول، وأن متوسط القيم - بشكل عام - يعتمد على المتكلم.، فكانت قيم التردد الأساسي F_0 فى الصوت المضغوط أعلى من نظيرتها في الصوت العادي، وذلك لكلٍ من أصوات الرجال وأصوات النساء. وهو ما أكدته دراسات سابقة فى لغات عدة، مثل: *(Kunzel 1995, Laver 1980 and Jassem et. al. 1973).*

- أن متوسط قِيم F_2 كانت أعلى في الصوت المضغوط عنه في الصوت غير المضغوط لكلٍ من الرجال والنساء، وذلك بنسب أعلى فى الترددات عند النساء وهو المنطقي. أما متوسط قِيم F_1 فى الحالتين المضغوط وغير المضغوط عند الرجال والنساء يرجع إلى أنه عنصر يعتمد *على المتكلم*.

- أن معدل الموجات المنتظمة إلى الضوضاء المتداخل مع الموجة HNR ، أعلى في الصوت المضغوط عنه فى الصوت غير المضغوط وذلك عند الرجال والنساء ،لكن بقيم تكاد تكون غير مؤثرة/ فعالة Insignificant

- أن نسبة الاضطراب فى تردد الموجة *Jitter* كانت قيمها شبه متساوية فى أصوات الرجال (المضغوط وغير المضغوط) ، بينما النسب أعلى فى صوت المرأة غير المضغوط عنه فى المضغوط لكن الفرق غير مؤثر. والشق الفيزيائي للاضطراب فى شدة الموجة *Shimmer*عنصر غير مؤثر في المقارنة بين الصوت المضغوط وغير المضغوط عند الجنسين.

V- References:

Baldwin J., French P. (1990) Forensic phonetics. Pinter, London.

Bou-Ghazale, S.E.(1996): Analysis, Modeling, and Perturbation of Speech Under Stress with Applications to Synthesis and Recognition. PhD thesis, Robust Speech Processing Laboratory, Duke Univ. Dept. of Electrical Engineering.In Hansen, J.H.L. and Patil, S. (2007): Speech Under Stress: Analysis, Modelling, and Recognition USA.

Bou-Ghazale, S.E., Hansen, J.H.L.(2000): A Comparative Study of Traditional and Newly Proposed Features for Recognition of Speech Under Stress. IEEE Transactions on Speech & Audio Processing 8(4), 429-442.

Brenner, M., Shipp T., Doherty E. T., Morrissey P. (1985). "Voice measures of psychological stress: laboratory and field data," in Vocal Fold Physiology, Biomechanics, Acoustics, and Phonatory Control, eds. Titze J. R., Scherer R. C. (Denver, CO: The Denver Center for the Performing Arts), 239-248.

Cairns, D.A., Hansen, J.H.L. (1994): Nonlinear Analysis and Detection of Speech under Stressed Conditions. Journal of the Acoustic Society of America 96(6), 3392-3400.

Catford, J.C. (1964) "Phonation Types",pp. 26-37 in Abercrombie, D et.al.(eds), In Honor of Daniel Jones, Longmans, London.

Clary, G., Hansen, J.H.L.(1992): A Novel Speech Recognizer for Keyword Spotting. In: Proceedings of the International Conference of Spoken Language Processing (ICSLP/02), Alberta, vol. 1, PP. 13-16.

Eriksson, A., Drygajlo, A.(2005): Forensic Speech Science. In: Tutorial, 9th European Conference on Speech Communication and Technology (Interspeech 05 - Eurospeech).

Fashal, M.M. (1998) "Some of the vocal emotional states: An acoustic study" Faculty of Arts, Alexandria University.

Hansen, J.H.L., Clements, M.A.(1987): "Evaluation of Speech under Stress and Emotional Conditions" 82(S1), 7-8.

Hansen, J.H.L.(1988) Analysis and Compensation of Stressed and Noisy Speech with Application to Robust Automatic Recognition. PhD thesis, School of Electrical Engineering, Georgia Institute of Technology, Atlanta.

Hansen, J.H.L. (1989). Evaluation of Acoustic Correlates of Speech Under Stress for Robust Speech Recognition. In: IEEE Proceedings of the 15th Northeast Bioengineering Conference, Boston, pp. 3.

Hansen, J.H.L. and Patil, S. (2007): Speech Under Stress: Analysis, Modelling, and Recognition. Center of Robust Speech System, University of Texas. USA.

Hayward, Katrina, (2000) Experimental Phonetics, Routledge, London and New York.

Hecker, M.H.L., Stevens, K.N., von Bismark, G., Williams, C.E.(1968): "Manifestations of Task Induced Stress in the Acoustic Speech Signal. Journal of the Acoustic Society of America 44, 993-1001.

Hicks, J.W., Hollien, H. (1981**)**: The Reflection of Stress in Voice-1: Understanding the Basic Correlates. In: Proceedings of the 1991 Carnahan Conference on Crime Countermeasures, pp. 189-195.

Hollien, H. (1980). "Vocal Indicators of Psychological Stress", Annual New York Academy of Sciences.

Hollien, H., Hicks, J.W.(1991): The Reflection of Stress in Voice-2: The Special Case of Psychological Stress Evaluators. In: Proceedings of the 1991 Carnahan Conference on Crime Countermeasures, pp. 196-197.

Ibrahim, O.A. (2016) "Building Audio-Visual Phonetically Annotated MSA Corpus for Speech Processing Purposes" M.A.Degree Dissertation, Faculty of Arts, Alexandria University.

Keating, A.P and Esposito, C.(2006), 'Linguistic Voice Quality', University of California, Los Angeles, USA,UCLA. Working Papers in Phonetics, No. 105, 85-91.

Koster, O., koster,J.P. (2004): The Auditory-Perceptual Evaluation of Voice Quality in Forensic Speaker Recognition. The Phonetician 89, 9-37.

Kramer, E. (1963): 'Judgement of personal characteristics and emotions from non-verbal properties of speech', Psychological Bulletin, 408-20.

Kreiman, J.,Bruce, R. (2005) "Perception of aperiodicity in pathological voice" JASA. 117 (4).

Kreiman J., Sidtis D. (2011). Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception. Hoboken, NJ: John Wiley & Son.

Kunzel, H.and Koster, M. (1995) "The Relation between Speech Tempo, Loudness and Fundamental Frequency: An important issue in forensic speaker recognition" Science and Justice,35 (4): 291-295.

Laver, J. (1968) 'Voice quality and indexical information', British Journal of Disorders of Communication 3, 43-54.

Laver, J. (1980) The Phonetic Description of Voice Quality, Cambridge University Press, Cambridge, England. **Laver J.** (1991). The Gift of Speech, Papers in the analysis of speech and voice, Edinburgh University Press.

Murray, I.R., Baber, C., South, A. (1996). Towards a Definition and Working Model of Stress and its Effects on Speech. Speech Communication 20, 3–12.

Patel, S., Scherer K. R., Björkner E., Sundberg J. (2011). Mapping emotions into acoustic space: the role of voice production. Biol. Psychol. 87 93-98.

Simonov, P.V., Frolov, M.V.(1977): Analysis of the Human Voice as a Method of Controlling Emotional State: Achievements and Goals. Aviation, Space, and Environmental Sciences 23-25.

Streeter, L.A, MacDonald, N.H., Apple., Krauss, R.M., Galotti, K.M.(1988): "Acoustic and Perceptual Indicators of Emotional Stress". Journal of the Acoustic Society of America 73(3), 917-928.

Teager, H.M., and Teager, S.M.(1982): "A Phenomenological Model for Vowel Production in the Vocal Tract". In: Speech Science: Recent Advances, pp. 72-100. In Speech Under Stress: Analysis, Modeling and Recognition 135.

Wagner I. (1995). A new jitter-algorithm to quantify hoarseness: an exploratory study. Forensic Linguist 2(1):18–27.

Williams, C.E., Stevens, K.N.(1972): Emotions and Speech: Some Acoustic Correlates, Journal of the Acoustic Society of America 52(4), 1238-1250.

Phonetic Investigation of Voice Under Stress:

Real Criminal Cases with Egyptian Arabic speakers

Eman Mohamed Qasem

Alexandria University, Phonetics and Linguistics Department, Egypt

eman.qasem@alexu.edu.eg ,

Introduction:

Tense voice is more likely to occur in situations of increased cognitive load. During the interrogation of the criminals under forensic circumstances, the offenders are exposed to some kind of stress (e.g., in the emotions like diffidence, anxiety, fear…) which are reflected in the production of their speech, and consequently, they are good indicators of tense vocalization.

The aim of this study is to investigate the phonetic characteristics of stressed/unstressed voice of the crime offenders' speech under pressure and how can these characteristics encode information about the speaker: his emotions and -to some extent-his intentions.

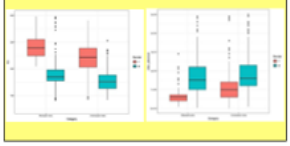
Materials and methods

Procedures:

- Data collection: various kinds of videos were collected from T.V interviews with offenders, victims, witnesses.
- kinds of crimes: -organs trading crime, - kidnap crime, killing crime, Fighting crime, and Adultery crime
- Data classification: about 15 speakers (8 males and 7 females) were chosen from the videos, their speech utterances were classified in two files: stressed voice and unstressed one.

Unstressed voice : when the speaker introduces himself.
Stressed voice: when the speaker is speaking in tense voice; with crying, fear or anxiety.(organs trading crime)

-Figures (1& 2): Represent examples of plots of stressed and unstressed variables for males and females.



Results:

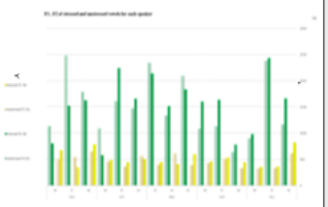
1- Voice quality parameters:

- F0 is the most significant parameter , its frequency values of stressed are higher than unstressed voice
- Jitter (absolute/ local) is slightly higher values in unstressed than stressed, especially by females.
- Shimmer (dB/local) seems to be insignificant parameter.
- Harmonic to noise ratios of stressed voice are higher than unstressed by both males and females; females values, however are higher.

2- Resonance parameters:

- Formants are speaker dependent. F1 and F2 are higher in stressed than unstressed by males than females

- In respect to vowel formants it is found that F1 of stressed voice is either higher in frequency values than unstressed or somewhat equal by males. Females are the opposite; unstressed voice are higher.
- F2 of stressed voice also are higher than unstressed by male speakers, while most of females are the opposite.



Conclusions:

•In this research, some examples of actual crimes have been selected from live interviews broadcasted by TV programs. The purpose of this selection is to investigate the performance of the criminals in the case of their emotional voice under pressure and their natural voice in ordinary situation. The change of the voice quality neutral to perturbated is reflected on the phonation voice F0, Jitter/ Shimmer, HNR and resonance parameters F1, F2,F3.

The most prominent conclusions of the study could be summarized as follows:

✓F0 is the first factor in priority followed respectively by the other voice quality factors.

✓HNR is the second feature that differentiate between stressed / unstressed voice especially of females

✓Jitter (absolute/local) were influence factors; their values were higher in unstressed than stressed, especially by females..

✓Shimmer (dB/local) seems to be insignificant. Parameter.

✓Formants are speaker dependent. F1 and F2 are higher in stressed than unstressed by males than females.

Acoustic analysis measures:

- Using PRAAT program version 3, and CSL program version, the following parameters were analyzed:
- F0,Jitter, Shimmer, H/N ratio of the chosen stressed and unstressed utterances of each speaker voice, in addition to the first three formants of stressed and unstressed vowels.

References:

•Hansen, J.H.L.(1989). Evaluation of Acoustic Correlates of Speech Under Stress for Robust Speech Recognition. In: IEEE Proceedings of the 15th Northeast Bioengineering Conference, Boston, pp. 3.

•Hollen, H.(1980). Vocal Indicators of Psychological Stress. In Forensic Psychology and Psychiatry. New York Academy of Sciences.pp:47-72.

•Laver, J. (1980).The phonetic description of voice quality. Cambridge: Cambridge University Press.